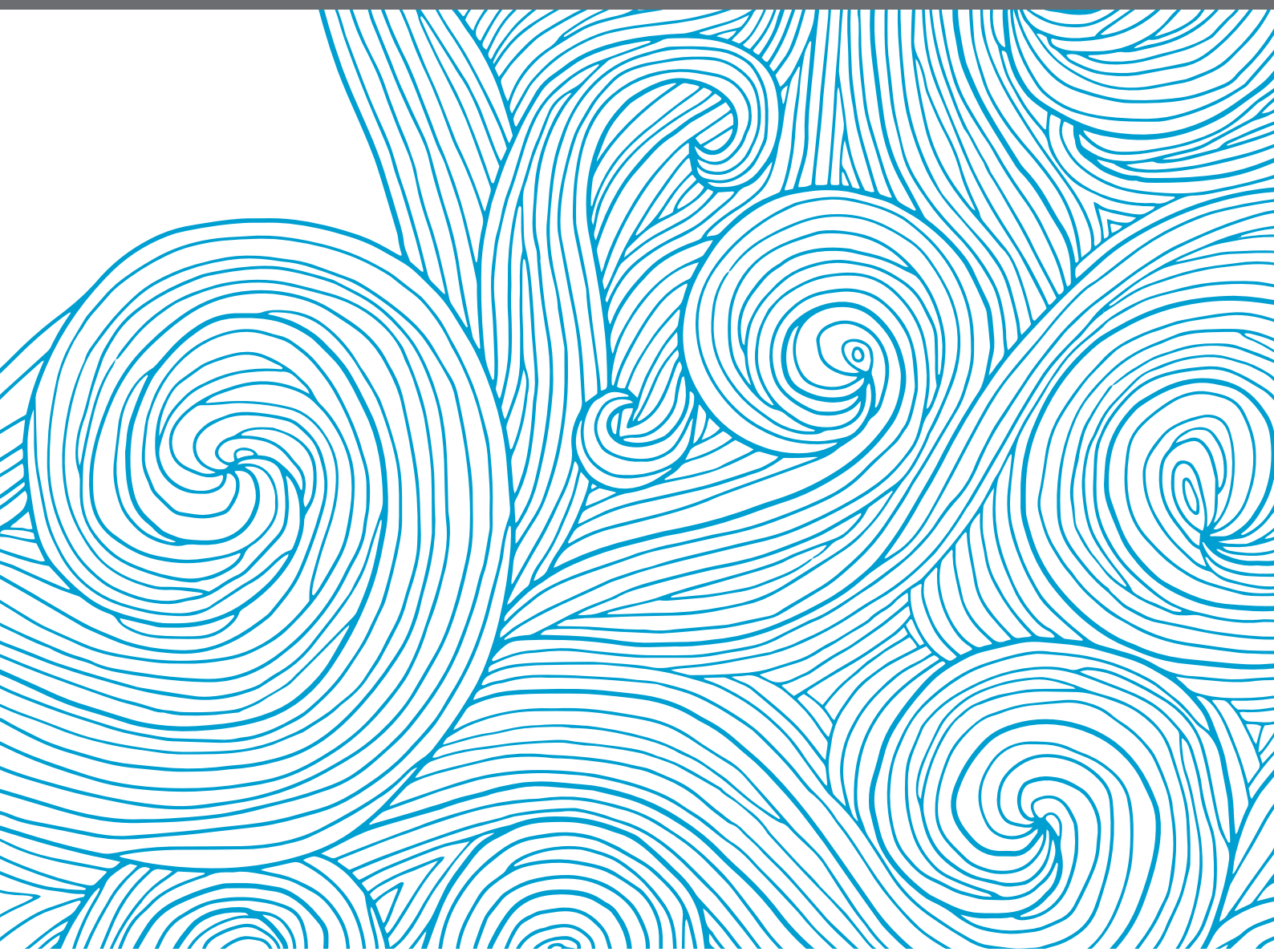
COLOUR AND LIGHT IN THE OCEAN

EDITED BY: Victor Martinez-Vicente, Astrid Bracher, Didier Ramon, Shubha Sathyendranath and Tiit Kutser

PUBLISHED IN: Frontiers in Marine Science



frontiers Research Topics

Frontiers eBook Copyright Statement

The copyright in the text of individual articles in this eBook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this eBook is the property of Frontiers.

Each article within this eBook, and the eBook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this eBook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or eBook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-88963-421-7
DOI 10.3389/978-2-88963-421-7

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

COLOUR AND LIGHT IN THE OCEAN

Topic Editors:
Victor Martinez-Vicente, Plymouth Marine Laboratory, United Kingdom
Astrid Bracher, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research (AWI), Germany
Didier Ramon, Hygeos, France
Shubha Sathyendranath, Plymouth Marine Laboratory, United Kingdom
Tiit Kutser, University of Tartu, Estonia

CLEO publications in *Frontiers in Marine Science*

Foreword
Josef Aschbacher,
Director of ESA's Earth Observation Programmes

Satellite data have drastically changed the view we have of the oceans. Covering about 70% of Earth's surface, oceans play a unique role for our planet and for our life – but large areas remain unexplored and are difficult to reach.

Since the 1980s, Earth-orbiting satellites have helped to observe what is happening at the ocean surface. Sensors like CZCS, AVHRR, SeaWifs and MODIS provided the first ocean colour data from space. Starting in 2002, ESA's Medium Resolution Imaging Spectrometer (MERIS) on-board the environmental satellite Envisat, provided detailed information on phytoplankton biomass and concentrations of other matter in the global oceans.

These satellite observations laid the groundwork for studying the marine environment and how it responds to climate change, and the research community has since delivered information on the variability of marine ecosystems.

Part of this work is reflected in this stunning collection of peer-reviewed publications presented at the workshop, Colour and Light in the Ocean from Earth Observation (CLEO), held at ESA's ESRIN site in Frascati, Italy, on 6–8 September 2016. The event attracted more than 160 participants from all over the world, including remote sensing experts, marine ecosystem modelers, in-situ observers and users of Earth observation data. Scientifically, the meeting covered applications in climate studies over primary productivity and ocean dynamics, to pools of carbon and phytoplankton diversity at global and regional scales. It also demonstrated the potential of Earth observation and its contribution to modern oceanography. Looking to the future, new satellites developed by ESA under the coordination of the European Commission will further our scientific and operational observations of the seas.

With Sentinel-3A in orbit and its twin Sentinel-3B following in 2017, there is a new category of data available for operational oceanographic applications and climate studies for years to come. These data are free and easy to access by anyone interested. Looking at the role of oceans in our daily lives, I am sure that this collection of scientific excellence will be valued by scientists of today and will inspire the next generation to carry these ideas into the future.

Citation: Martinez-Vicente, V., Bracher, A., Ramon, D., Sathyendranath, S., Kutser, T., eds. (2020). Colour and Light in the Ocean. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88963-421-7

Table of Contents

07 ***Using Optical Sensors on Gliders to Estimate Phytoplankton Carbon Concentrations and Chlorophyll-to-Carbon Ratios in the Southern Ocean***
Sandy J. Thomalla, A. Gilbert Ogunkoya, Marcello Vichi and Sebastiaan Swart

26 ***A Consumer's Guide to Satellite Remote Sensing of Multiple Phytoplankton Groups in the Global Ocean***
Colleen B. Mouw, Nick J. Hardman-Mountford, Séverine Alvain, Astrid Bracher, Robert J. W. Brewin, Annick Bricaud, Aurea M. Ciotti, Emmanuel Devred, Amane Fujiwara, Takafumi Hirata, Toru Hirawake, Tihomir S. Kostadinov, Shovonlal Roy and Julia Uitz

45 ***Obtaining Phytoplankton Diversity From Ocean Color: A Scientific Roadmap for Future Development***
Astrid Bracher, Heather A. Bouman, Robert J. W. Brewin, Annick Bricaud, Vanda Brotas, Aurea M. Ciotti, Lesley Clementson, Emmanuel Devred, Annalisa Di Cicco, Stephanie Dutkiewicz, Nick J. Hardman-Mountford, Anna E. Hickman, Martin Hieronymi, Takafumi Hirata, Svetlana N. Losa, Colleen B. Mouw, Emanuele Organelli, Dionysios E. Raitsos, Julia Uitz, Meike Vogt and Aleksandra Wolanin

60 ***Simulating PACE Global Ocean Radiances***
Watson W. Gregg and Cécile S. Rousseaux

79 ***Light Absorption by Phytoplankton in the Upper Mixed Layer of the Black Sea: Seasonality and Parametrization***
Tanya Churilova, Vyacheslav Suslin, Olga Krivenko, Tatiana Efimova, Nataliia Moiseeva, Vladimir Mukhanov and Liliya Smirnova

93 ***Uncertainty in Ocean-Color Estimates of Chlorophyll for Phytoplankton Groups***
Robert J. W. Brewin, Stefano Ciavatta, Shubha Sathyendranath, Thomas Jackson, Gavin Tilstone, Kieran Curran, Ruth L. Airs, Denise Cummings, Vanda Brotas, Emanuele Organelli, Giorgio Dall'Olmo and Dionysios E. Raitsos

115 ***Impact of El Niño Variability on Oceanic Phytoplankton***
Marie-Fanny Racault, Shubha Sathyendranath, Robert J. W. Brewin, Dionysios E. Raitsos, Thomas Jackson and Trevor Platt

130 ***The OLCI Neural Network Swarm (ONNS): A Bio-Geo-Optical Algorithm for Open Ocean and Coastal Waters***
Martin Hieronymi, Dagmar Müller and Roland Doerffer

148 ***Bio-Optical Properties of Two Neigboring Coastal Regions of Tropical Northern Australia: The Van Diemen Gulf and Darwin Harbour***
David Blondeau-Patissier, Thomas Schroeder, Lesley A. Clementson, Vittorio E. Brando, Diane Purcell, Phillip Ford, David K. Williams, David Doxaran, Janet Anstee, Nandika Thapar and Miguel Tovar-Valencia

175 ***Shellfish Aquaculture From Space: Potential of Sentinel2 to Monitor Tide-Driven Changes in Turbidity, Chlorophyll Concentration and Oyster Physiological Response at the Scale of an Oyster Farm***
Pierre Gernez, David Doxaran and Laurent Barillé

190 ***Regional Empirical Algorithms for an Improved Identification of Phytoplankton Functional Types and Size Classes in the Mediterranean Sea Using Satellite Data***
Annalisa Di Cicco, Michela Sammartino, Salvatore Marullo and Rosalia Santoleri

208 ***Extended Formulations and Analytic Solutions for Watercolumn Production Integrals***
Žarko Kovač, Trevor Platt, Suzana Antunović, Shubha Sathyendranath, Mira Morović and Charles Gallegos

224 ***Synergistic Exploitation of Hyper- and Multi-Spectral Precursor Sentinel Measurements to Determine Phytoplankton Functional Types (SynSenPFT)***
Svetlana N. Losa, Mariana A. Soppa, Tilman Dinter, Aleksandra Wolanin, Robert J. W. Brewin, Annick Bricaud, Julia Oelker, Ilka Peeken, Bernard Gentili, Vladimir Rozanov and Astrid Bracher

246 ***Reproduction of Spatio-Temporal Patterns of Major Mediterranean Phytoplankton Groups From Remote Sensing OC-CCI Data***
Gabriel Navarro, Pablo Almaraz, Isabel Caballero, Águeda Vázquez and Isabel E. Huertas

262 ***Validation and Intercomparison of Ocean Color Algorithms for Estimating Particulate Organic Carbon in the Oceans***
Hayley Evers-King, Victor Martinez-Vicente, Robert J. W. Brewin, Giorgio Dall'Olmo, Anna E. Hickman, Thomas Jackson, Tihomir S. Kostadinov, Hajo Krasemann, Hubert Loisel, Rüdiger Röttgers, Shovonlal Roy, Dariusz Stramski, Sandy Thomalla, Trevor Platt and Shubha Sathyendranath

282 ***Phytoplankton Group Identification Using Simulated and* In situ *Hyperspectral Remote Sensing Reflectance***
Hongyan Xi, Martin Hieronymi, Hajo Krasemann and Rüdiger Röttgers

295 ***An Exact Solution for Modeling Photoacclimation of the Carbon-to-Chlorophyll Ratio in Phytoplankton***
Thomas Jackson, Shubha Sathyendranath and Trevor Platt

305 ***Characterizing Spatial Variability of Ice Algal Chlorophyll* a *and Net Primary Production Between Sea Ice Habitats Using Horizontal Profiling Platforms***
Benjamin A. Lange, Christian Katlein, Giulia Castellani, Mar Fernández-Méndez, Marcel Nicolaus, Ilka Peeken and Hauke Flores

328 ***Primary Production: Sensitivity to Surface Irradiance and Implications for Archiving Data***
Trevor Platt, Shubha Sathyendranath, George N. White III, Thomas Jackson, Stéphane Saux Picart and Heather Bouman

337 ***Intercomparison of Ocean Color Algorithms for Picophytoplankton Carbon in the Ocean***
Víctor Martínez-Vicente, Hayley Evers-King, Shovonlal Roy, Tihomir S. Kostadinov, Glen A. Tarran, Jason R. Graff, Robert J. W. Brewin, Giorgio Dall'Olmo, Tom Jackson, Anna E. Hickman, Rüdiger Röttgers, Hajo Krasemann, Emilio Marañón, Trevor Platt and Shubha Sathyendranath

356 ***Comparison of Seasonal Cycles of Phytoplankton Chlorophyll, Aerosols, Winds and Sea-Surface Temperature off Somalia***
Muhammad Shafeeque, Shubha Sathyendranath, Grinson George, Alungal N. Balchand and Trevor Platt

371 ***Satellite Radiation Products for Ocean Biology and Biogeochemistry: Needs, State-of-the-Art, Gaps, Development Priorities, and Opportunities***
Robert Frouin, Didier Ramon, Emmanuel Boss, Dominique Jolivet, Mathieu Compiègne, Jing Tan, Heather Bouman, Thomas Jackson, Bryan Franz, Trevor Platt and Shubha Sathyendranath

391 ***Optical Classification of the Coastal Waters of the Northern Indian Ocean***
S. Monolisha, Trevor Platt, Shubha Sathyendranath, J. Jayasankar, Grinson George and Thomas Jackson

frontiers in Marine Science

ORIGINAL RESEARCH
published: 13 February 2017
doi: 10.3389/fmars.2017.00034

Using Optical Sensors on Gliders to Estimate Phytoplankton Carbon Concentrations and Chlorophyll-to-Carbon Ratios in the Southern Ocean

Sandy J. Thomalla[1,2]*, *A. Gilbert Ogunkoya*[3†], *Marcello Vichi*[2,3] and *Sebastiaan Swart*[1,2,4]

[1] Southern Ocean Carbon and Climate Observatory, CSIR - Ocean Systems and Climate, Cape Town, South Africa, [2] Department of Oceanography, University of Cape Town, Cape Town, South Africa, [3] Marine Research Institute (Ma-Re), University of Cape Town, Cape Town, South Africa, [4] Department of Marine Sciences, University of Gothenburg, Gothenburg, Sweden

One approach to deriving phytoplankton carbon biomass estimates (C_{phyto}) at appropriate scales is through optical products. This study uses a high-resolution glider data set in the Sub-Antarctic Zone (SAZ) of the Southern Ocean to compare four different methods of deriving C_{phyto} from particulate backscattering and fluorescence-derived chlorophyll (chl-a). A comparison of the methods showed that at low (<0.5 mg m^{-3}) chlorophyll concentrations (e.g., early spring and at depth), all four methods produced similar estimates of C_{phyto}, whereas when chlorophyll concentrations were elevated one method derived higher concentrations of C_{phyto} than the others. The use of methods derived from particulate backscattering rather than fluorescence can account for cellular adjustments in chl-a:C_{phyto} that are not driven by biomass alone. A comparison of the glider chl-a:C_{phyto} ratios from the different optical methods with ratios from laboratory cultures and cruise data found that some optical methods of deriving C_{phyto} performed better in the SAZ than others and that regionally derived methods may be unsuitable for application to the Southern Ocean. A comparison of the glider chl-a:C_{phyto} ratios with output from a complex biogeochemical model shows that although a ratio of 0.02 mg chl-a mg C^{-1} is an acceptable mean for SAZ phytoplankton (in spring-summer), the model misrepresents the seasonal cycle (with decreasing ratios from spring to summer and low sub-seasonal variability). As such, it is recommended that models expand their allowance for variable chl-a:C_{phyto} ratios that not only account for phytoplankton acclimation to low light conditions in spring but also to higher optimal chl-a:C_{phyto} ratios with increasing growth rates in summer.

Keywords: phytoplankton carbon, chlorophyll to carbon ratios, particulate backscattering, gliders, Southern Ocean

OPEN ACCESS

Edited by:
Victor Martinez-Vicente, Plymouth Marine Laboratory, UK

Reviewed by:
Jason R. Graff, Oregon State University, USA
Marco Bellacicco, Centre National D'Etudes Spatiales, Italy

***Correspondence:**
Sandy J. Thomalla
sthomalla@csir.co.za

†***Present Address:***
A. Gilbert Ogunkoya, Department of Ecology, Montana State University, Bozeman, MT, USA

Specialty section:
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

Received: *10 October 2016*
Accepted: *27 January 2017*
Published: *13 February 2017*

Citation:
Thomalla SJ, Ogunkoya AG, Vichi M and Swart S (2017) Using Optical Sensors on Gliders to Estimate Phytoplankton Carbon Concentrations and Chlorophyll-to-Carbon Ratios in the Southern Ocean. Front. Mar. Sci. 4:34. doi: 10.3389/fmars.2017.00034

INTRODUCTION

Marine phytoplankton at the global scale have an average biomass turnover time of 1 week or less (Falkowski et al., 1998). Despite their temporary existence, these living organisms can absorb carbon at a rate of 40–50 Pg C y^{-1} and are responsible for roughly half the net primary production on Earth (Longhurst et al., 1995; Antoine et al., 1996; Field et al., 1998; Falkowski et al., 1998;

Westberry et al., 2008). The Southern Ocean is responsible for 40% of global CO_2 uptake (Gruber et al., 2009) and the sub-basin of the Sub-Antarctic Zone (SAZ) is recognized as one of the regions of higher carbon export and an effective atmospheric CO_2 sink (Metzl et al., 1999; McNeil et al., 2001; Trull et al., 2001; Wang et al., 2001). This region of the Southern Ocean is characterized by deep convective mixed layers (>500 m) during winter, which favors the injection of waters rich in inorganic carbon into the mixed layer (Key et al., 2004). Conversely during summer, seasonal heat flux shoals the mixed layer (Swart et al., 2015), which improves the overall availability of light for photosynthesis favoring the transformation of inorganic carbon to particulate organic carbon (POC) and the potential for an effective "biological carbon pump" (Broecker and Peng, 1982; Volk and Hoffert, 1985). The deep mixed layers formed in the SAZ during winter are subducted northwards as Sub-Antarctic Mode Water (SAMW) and Antarctic Intermediate Water (AAIW, McCartney, 1977) driving an effective solubility pump which in combination with the biological pump maintains a strong CO_2 sink year round (McNeil et al., 2007).

If researchers are to accurately reflect the seasonal cycle of phytoplankton production in predictive climate models and thereby improve our understanding of the sensitivities of the biological carbon pump to changes in climate forcing factors (both needed for predicting long term trends), the Southern Ocean ecosystem has to be investigated at the appropriate scales that link the physical drivers to the biogeochemistry (Lévy et al., 2001; Le Quéré et al., 2007; Klein et al., 2008; Doney et al., 2009; Thomalla et al., 2011; Racault et al., 2012; Joubert et al., 2014; Carranza and Gille, 2015; Swart et al., 2015). There is increasing evidence in the Southern Ocean that seasonal to sub-seasonal temporal scales and meso- to submeso- spatial scales play an important role in determining the response of primary producers to physical forcing (Boyd, 2002; Fauchereau et al., 2011; Thomalla et al., 2011, 2015; Lévy et al., 2012; Swart et al., 2015), which may in turn affect their sensitivity to climate change.

Satellite remote sensing provides an essential tool for investigating patterns of phytoplankton variability at high-sampling frequency and with good spatial resolution globally. However, remotely detected water-leaving radiances emanate from only the first optical depth and require assumptions about their representativeness of the vertical structure of the water column. The frontier in ocean observation is adequate and sustained spatial sampling of the sub-surface ocean (Rudnick et al., 2004) conducted at an appropriate frequency to define and understand the growth timescales of phytoplankton. Autonomous platforms (e.g., floats and gliders) are able to profile the water column (0–1000 m) and characterize vertical biogeochemistry at smaller scales, but also for sufficiently long periods that may help to reduce uncertainties associated with carbon budgets at longer time scales. In addition, the volume of information that a single glider mission retrieves, can be instrumental in developing and validating statistically robust parameterizations for numerical models, which are otherwise performed with oftentimes inadequate data sets generated from once-off or "classical" (low spatial and/or low temporal frequency) sampling techniques. As such, high-resolution sampling of phytoplankton biomass through the water column is key to reducing uncertainties associated with carbon budgets.

Net autotrophic primary production is ultimately a function of the standing population of phytoplankton biomass, which is a system state variable that is not easily observed in remote regions like the Southern Ocean. Phytoplankton biomass refers to the total quantity of phytoplankton in a given volume of water expressed here as weight in carbon (mg C m^{-3}). Satellite ocean color data can provide a proxy of phytoplankton biomass through empirical combinations of radiometric information to obtain estimates of carbon and/or chlorophyll concentration (e.g., Stramski et al., 1999; Gardner et al., 2006; Blondeau-Patissier et al., 2014). In addition, production models are available to turn this information into rates of primary production (e.g., Behrenfeld and Falkowski, 1997; Arrigo et al., 1998, 2008; Moore and Abbott, 2000; Carr et al., 2006). However, chlorophyll is a dynamic property of phytoplankton (MacIntyre et al., 2002) that is influenced by shifts in the physiology of the cells, with intracellular pigments being modulated in response to changes in growth conditions (e.g., temperature, nutrients, light) (e.g., Halsey and Jones, 2015; Behrenfeld et al., 2016). In addition, different pigment content is expressed by the phylogenetic evolution of phytoplankton species with changes in accessory pigments leading to differences in light absorption per unit chlorophyll (MacIntyre et al., 2002). As such, temporal changes in chlorophyll over large ocean regions can result form physiologically or community driven modifications in cellular chlorophyll-to-carbon ratios, rather than to changes in biomass (Behrenfeld et al., 2005, 2016; Westberry et al., 2008, 2016; Mignot et al., 2014; Bellacicco et al., 2016). Such instances will have strong implications on assessments of long term trends in primary production, ecosystem trophic dynamics, and carbon export (Behrenfeld et al., 2016).

Other optical observations such as backscattering are better correlated to carbon than chlorophyll (Antoine et al., 2011) and can provide independent measures of phytoplankton biomass in open ocean waters (away from regions with highly scattering inorganic material). In addition, they can be measured both *in situ* and with satellite remote sensing. Unlike chlorophyll, these measures are more likely to be insensitive to changes in the intracellular concentration of pigments (Behrenfeld and Boss, 2006; Behrenfeld et al., 2016; Bellacicco et al., 2016; Westberry et al., 2016). However, the use of optical proxies for total organic carbon is complicated by the highly variable relationships reported in the literature for POC vs. backscattering which can vary by a factor of five (Cetinić et al., 2012). Much of the variability between POC and backscattering is driven by differences inherent to the types of particles in the system. For example, the carbon density (or POC to volume ratio) varies between species with diatoms and *Phaeocystis* having typically lower POC to volume ratios (Cetinić et al., 2012). In addition, the ability to differentiate phytoplankton specific carbon (C_{phyto}) from other suspended particulate matter is a big challenge operationally (Lü et al., 2009). Methodological constraints result in the carbon biomass of phytoplankton being poorly identifiable (Martinez-Vicente et al., 2013) and not easy to distinguish from other types of carbon (Eppley et al., 1992; Oubelkheir et al.,

2005). Nonetheless, the strong relationship observed between bio-optical observations of carbon biomass allows the use of autonomous instruments (e.g., gliders and floats) to retrieve data at high spatial and temporal resolution, increasing the capability of capturing higher frequency changes in ocean biogeochemistry (e.g., Mignot et al., 2014; Thomalla et al., 2015). Despite their advantages however, the availability of *in situ* bio-optical data on both a regional and global scale is still sparse, highlighting the need to prioritize their collection on future research campaigns. Given their growing importance in the trajectory of ocean ecosystem understanding, it is important to develop methods of detecting biogeochemical properties from instruments on autonomous platforms; and carbon from optical sensors is certainly one of the first candidates, given the importance of phytoplankton production in driving the carbon sink (Boss et al., 2008).

Rates of carbon production can also be derived from mathematical models that link standing biomass and environmental conditions to growth rates. Since the very first extensive discussions on models (e.g., Cullen, 1990) chlorophyll and/or carbon have been considered necessary to model irradiance-based growth. Mathematical models (empirical, semi-analytical, and analytical) are used to reconstruct primary production from satellite estimates of chlorophyll or carbon (see Behrenfeld and Falkowski, 1997 for model summary), and are similarly used in global ocean biogeochemical models to determine rates of carbon production (e.g., Moore et al., 2002; Aumont and Bopp, 2006; Vichi et al., 2007). Some of the models include mechanistic representations of phytoplankton acclimation that consider varying amounts of carbon and chlorophyll in the modeled phytoplankton (largely relying on the Geider et al., 1997 formulation). The Southern Ocean is likely to experience light-saturated conditions for a rather limited period, and therefore, the acclimation process and associated changes in chlorophyll-to-carbon ratios in the period leading to the summer blooms are very likely. The majority of global ocean biogeochemical models and Earth System Models however overestimate the magnitude of the spring-summer bloom (e.g., Doney et al., 2009; Steinacher et al., 2009; Vichi and Masina, 2009), which is usually attributed to the coarse resolution global ocean models and their inability to simulate deep vertical mixing. However, McKiver et al. (2015) demonstrated that an increase in horizontal resolution down to 25 km (considered to be eddy-permitting and partly eddy-resolving in the Southern Ocean) did not help to improve the timing and magnitude of the bloom in the SAZ. Coarse resolution models with a 1 to 2° grid tend to have an early bloom and a much larger magnitude than observed, and the substantial increase in resolution only partly reduced these biases. They went on to suggest that this behavior might instead result from inaccurate parameterization of the chlorophyll-to-carbon ratios. This strengthens the need to obtain combined measures of chlorophyll and carbon from the world oceans and particularly from the Southern Ocean.

This work tests four different methods of deriving phytoplankton carbon from optical data collected by sensors on a glider deployed in the SAZ. Some of these methods can be adjusted using available data from the SAZ, while other relationships are specific to the region of original sampling and can only be applied to the SAZ as they are. By comparing the different methods we aim to elucidate the differences of choosing one method over another and their relative validity in the Southern Ocean. In addition, the chlorophyll-to-carbon (chl-a:C_{phyto}) ratios generated from the different methods are compared with *in situ* data from the literature and a model simulation to highlight their respective ranges and distribution.

MATERIALS AND METHODS

Glider Data

The data used for this research were collected in the framework of the Southern Ocean Carbon and Climate Observatory (SOCCO; http://www.socco.org.za) during the Southern Ocean Seasonal Cycle Experiment (SOSCEx; Swart et al., 2012). Seaglider SG573 was deployed south of Gough Island in the South-East Atlantic Ocean at 43.0°S, 11°W (**Figure 1**). The glider was deployed on 25 September 2012 and retrieved on 15 February 2013, resulting in a high-resolution transect of 143 days (∼5.5 months) covering a total distance of 1693 km (see Swart et al., 2015 for more detail). The glider measures a suite of parameters that includes conductivity (salinity), temperature, pressure, dissolved oxygen, chlorophyll-*a* fluorescence (proxy for phytoplankton concentration), Photosynthetically Active Radiation (PAR), and two wavelengths of optical backscattering by particles, $b_{bp(470)}$ and $b_{bp(700)}$.

The glider was programmed to profile between the surface and 1000 m continuously at a nominal vertical velocity of 10 cm s^{-1}. Each dive cycle (which includes a descent and ascent profile) took ∼5 h to complete and covered an average horizontal distance of 2.8 km, rendering a temporal resolution of 2.5 h and a spatial resolution of 1.4 km (between each water column profile). Glider data were linearly interpolated to a 6-hourly frequency in order to grid unevenly spaced profiles, which typically ranged between 4 and 6 hourly. At the deployment and retrieval site of the glider, ship-based CTD cross-calibration casts were carried out yielding two independent inter-calibrations between the gliders, CTD sensors and bottle samples (Swart et al., 2015).

Glider fluorescence from a WETLabs ECO puck™ (BB2Fl-470/700) was dark corrected by subtracting the median fluorescence below 300 m from all raw instrument counts. Fluorescence quenching was isolated by selecting all daylight profiles between local sunrise and sunset plus 2.5 h (as quenching was on occasion observed in the profile following sunset). Fluorescence quenching was subsequently corrected using optical backscattering (b_{bp}), an alternate proxy for phytoplankton biomass that is not susceptible to quenching, based on the methods described in Sackmann et al. (2008). This method assumes a constant chlorophyll-to-carbon ratio throughout the surface waters and no cellular changes in chlorophyll packaging with depth as a photo adaptive strategy to low light levels. When backscattering was unavailable (intermittent sensor malfunction), quenching was corrected by extrapolating the maximum fluorescence value within the mixed layer to the surface according to Xing et al. (2012). In this method, the maximum fluorescence yield within the mixed layer is

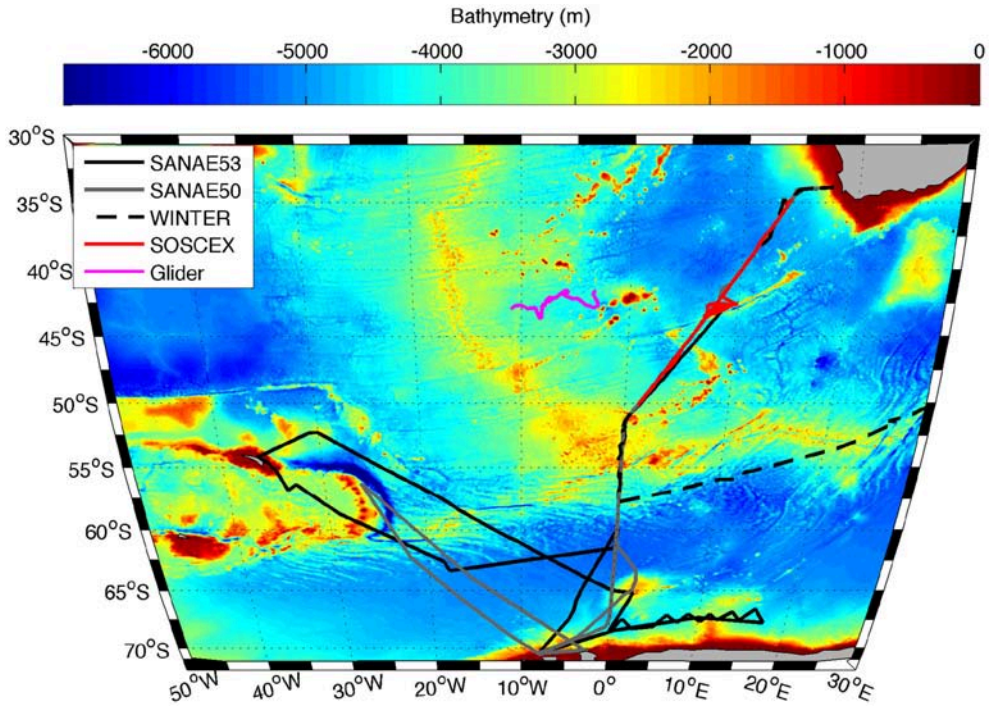


FIGURE 1 | Trajectory of the Seaglider (SG573) between 20 September 2012 and 15 February 2013 (pink line) together with cruise tracks of SANAE 53 (black solid line), SANAE 50 (grey line), WINTER (black dashed line), and SOSCEx (red line) overlaid on the bathymetry of the study region.

assumed to be representative of the mixed layer and implies homogeneity. This however is not necessarily the case as mixing and settling patterns between phytoplankton functional types are dynamic (Quéguiner, 2013). As such, when backscattering data are available, the Sackmann et al. (2008) method of quenching correction is preferred as it allows for small-scale variability in biomass with depth and reduces the possibility of overestimating surface chlorophyll values in the event of a subsurface chlorophyll maxima that is not biomass driven.

Fluorescence was converted to chlorophyll using a combination of the manufacturer's instrument specific chlorophyll conversion factor and *in situ* chlorophyll samples (250 ml filtered onto GF/F and extracted in 8 ml of 90% acetone for 24 h at −20°C) collected from CTD casts at glider intercept stations. This allowed all five gliders deployed during the SOSCEx experiment to be plotted simultaneously to form a statistically significant regression from 83 co-located glider chlorophyll and *in situ* chlorophyll samples (slope = 4.12, intercept= −0.21, $r^2 = 0.66$) with the manufacturer calibrated glider-based measurements being ∼4 times higher than the ship-based chlorophyll measurements. The slope of the regression was applied to all glider chlorophyll data to correct the manufacturer conversion to chlorophyll to one more suited to the regional characteristics of Southern Ocean chlorophyll (see Swart et al., 2015 for more detail).

Spikes from raw backscattering ($\lambda = 470$) were separated using a 7-point running median filter according to Briggs et al. (2011). Raw digital counts were then converted into particulate backscattering (b_{bp}) according to Dall'Olmo et al. (2009), following the equation:

$$b_{bp470} = 2\pi\ \chi_p\ [S\,(C - D) - \beta_{sw}] \quad (1)$$

where S is the instrument specific scaling factor; C are the raw digital counts and D are the dark counts (factory value); χ_p (equal to 1.1) is the factor used to convert the particulate volume scattering function into b_{bp} (Boss and Pegau, 2001) and β_{sw} is the volume scattering of pure water estimated using the models of Zhang and Hu (2009) and Zhang et al. (2009). Remaining spikes in particulate backscattering were removed with a threshold in shallow ($b_{bp470} > 0.048$) and deep ($b_{bp470} > 0.0025$) water. Profiles with high mean backscattering ($b_{bp} > 0.001$) in deep waters (> 150 m) were identified as bad profiles and discarded (see Thomalla et al., 2015 for more detail).

Mixed layer depth (MLD) was defined following the criterion of de Boyer Montégut et al. (2004) as the depth where the difference in temperature exceeds 0.2°C in reference to the temperature at 10 m ($\Delta T_{10m} = 0.2$°C). The temperature based criteria was chosen because (1) the salinity data (and hence density data) is contaminated in the final ∼5 weeks of the glider data due to Gooseneck barnacles bio-fouling the conductivity cell of the CTD sensor and (2) intermittent spiking and thermal lag errors (see Garau et al., 2011) of the salinity data resulted in intermittent false MLD determinations when using a density criterion alone to determine the MLD. An investigation into the two MLD methods used on the glider data shows that they match

each other closely (r= 0.86, $p \ll 0.01$) throughout the experiment (Swart et al., 2015).

Cruise Data

The *in situ* data used in this work were collected on four cruises to the Southern Ocean between austral winter 2012 and late summer 2013 (**Table 1**), which were typically separated into three legs: The GoodHope Line between Cape Town and Antarctica, the Buoy Run between Antarctica and South Georgia and the marginal ice zone along the continental margin (**Figure 1**). Data consisted of CTD casts and surface underway sampling for chlorophyll and POC collected routinely on SOCCO summer research cruises. CTD casts also provided $b_{bp_{470}}$ data that were co-located with Niskin bottle samples for POC. Chlorophyll samples (250 ml) were filtered onto Whatmann 25 mm GF/F glass fiber filters and extracted in 8 ml of 90% acetone at −20°C for 12–24 h. Fluorescence was measured on a Turner Trilogy Laboratory Fluorometer and converted to chlorophyll using a standard chlorophyll dilution calibration. POC samples (∼2 L) were filtered onto a pre-combusted 25 mm or 47 mm Whatmann GF/F filters and oven dried at 50°C. Filters were acidified by fuming with concentrated Hydrochloric Acid for 24 h to drive off inorganic carbon and re-dried in the oven. Filters were pelleted into 5 x 8 mm tin capsules and analyzed using a CHN analyser (Parsons et al., 1984; Knap et al., 1994). Blanks were interspersed every 6 to 20 samples (typically every 12). No replicate samples were analyzed.

Model Data

Numerical model results in the Sub-Antarctic zone were obtained from a simulation of the Biogeochemical Flux Model (BFM, http://bfm-community.eu) coupled with the NEMO ocean model (http://www.nemo-ocean.eu) at 0.25° resolution (PELAGOS025, McKiver et al., 2015). The BFM model (Vichi et al., 2007) allows computing phytoplankton dynamics in terms of stoichiometrically variable constituents, which also include variable chlorophyll to carbon ratios modified after Geider et al. (1997). All details of the model simulation are given in McKiver et al. (2015) and full descriptions of the biogeochemical model equations are available in Vichi et al. (2015). The variables extracted from the model results were total phytoplankton carbon and total phytoplankton chlorophyll, with the resulting dominant functional group in the region throughout the year being diatoms.

TABLE 1 | Names and dates of the various cruises indicating the different legs covered and the number of POC samples per cruise, with leg 1 between Cape Town and Antarctica along the GoodHope Line, leg 2 from Antarctica to South Georgia and leg 3 along marginal ice zone (Figure 1).

Cruise	Date	Leg	Number of POC samples
SANAE 50	02 Jan 2011–01 Feb 2011	1, 2 and 3	151
WINTER	09 July 2012–01 Aug 2012	1	132
SOSCEX	18 Feb 2013–10 Mar 2013	1	89
SANAE 53	28 Nov2013–11 Feb 2014	1, 2 and 3	149 (surface only)

Estimation of Phytoplankton Carbon

Four different methods of deriving the fraction of phytoplankton-specific carbon (C_{phyto}) from backscattering and chlorophyll were applied to the glider optical data. They represent examples of the different types of methods available in the scientific literature.

Linear Method (30%POC)

Linear relationships between POC and backscattering were first proposed by Stramski et al. (1999). The high correlation in open ocean waters was the result of the dominant organic particle concentration which controls changes in both POC and b_{bp} (Stramski and Morel, 1990; Gardner et al., 1993; Stramski and Reynolds, 1993; Loisel and Morel, 1998). In this method 221 POC samples from the cruises listed in **Table 1** (except SANAE53 where no CTD POC samples were available) were plotted against co-located $b_{bp_{470}}$ data (**Figure 2**). POC and $b_{bp_{470}}$ were linearly correlated with one another to provide a regression equation specific to the SAZ using a total least square method to account for the uncertainty in both the POC data and the backscattering:

$$POC = (39418 \pm 3000)\, b_{bp470} - (13 \pm 6)\,. \qquad (2)$$

We have assumed an uncertainty of 7% for both the input data to the total least square regression and reported the slope with all digits, as it is customary in the literature (**Table 2**), with the addition of the standard error. The relationships between b_{bp} and POC from various ocean basins in the literature (**Table 2**) are included on **Figure 2** for comparison.

POC however still needs to be converted into a fraction specific to phytoplankton (C_{phyto}). Behrenfeld et al. (2005) summarized ranges of field based C_{phyto}:POC ratios from different oceanic regions to an average phytoplankton contribution to total POC of ∼30%. Their 30% summary was derived from the studies of Eppley et al. (1992), Durand et al. (2001), Gundersen et al. (2001), and Oubelkheir (2001), where C_{phyto}:POC ratios ranged between 19 and 49% in regions that varied from eutrophic to oligotrophic. The 30%POC method as applied here converts b_{bp470} into POC using the linear regression (Equation 2) and then uses a constant 30% fraction to represent C_{phyto} as the phytoplankton contribution to total POC. A comparison of glider b_{bp470} with CTD b_{bp470} from five collocated profiles (at glider deployment, intercept and retrieval sites) gave a slope of 0.99 (n= 104) making it possible to apply the regression from Equation (2) to the glider b_{bp470} data to retrieve POC.

Behrenfeld et al. (2005) Method (B05)

Behrenfeld et al. (2005) derived a method of estimating C_{phyto} from remotely sensed backscattering data using a linear relationship between b_{bp440} and chlorophyll. In this method, chlorophyll and b_{bp} at 440 nm (m^{-1}) were estimated using the Garver-Siegel-Maritorena (GSM) semi-analytical algorithm (Garver and Siegel, 1997; Maritorena et al., 2002; Siegel et al., 2002) applied to monthly SeaWiFS data from September 1997 to January 2002. While the b_{bp} wavelength analyzed by Behrenfeld et al. (2005) (440 nm) differs to our study (B05) (470 nm) this discrepancy would only result in a small percentage difference in

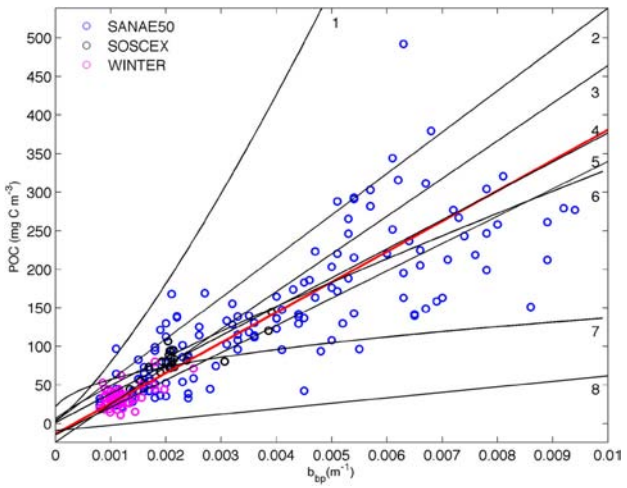


FIGURE 2 | Relationship between *in situ* POC and particulate backscattering (b_{bp} 470 nm) from cruise data (Table 1, SANAE53 was excluded as corresponding CTD backscattering data were not available) with the linear least squared regression ($r^2 = 0.93$) plotted in solid red. Overlain with fits (summarized in (Cetinić et al., 2012)) from (1) Stramski et al. (1999) from Ross Sea; (2) Stramski et al. (2008) from Atlantic/Pacific (excluding upwelling data); (3) Graff et al. (2015) from the North and South Atlantic and Equatorial Pacific; (4) Loisel et al. (2001) from Mediterranean; (5) Cetinić et al. (2012) from North Atlantic; (6) Stramski et al. (1999) from the Polar Frontal Zone (PFZ); (7) Balch et al. (2010) from north and south Atlantic; (8) Stramski et al. (2008) from Atlantic/Pacific (entire data set, including upwelling). See **Table 2** for regression coefficients.

b_{bp} values (Graff et al., 2015). As with Behrenfeld et al. (2005), our b_{bp470} was corrected for contributions from non-algal organisms other than phytoplankton (which also contribute to the optical backscattering signal) by subtracting a background value of 3.5 10^{-4} m^{-1}, with the assumption that this value represents the portion of non-algal particulate matter that does not co-vary with phytoplankton (Behrenfeld et al., 2005). Worth noting however is that the proportion of non-algal particulate matter is not constant as was shown by a study in the Mediterranean where the non-algal contribution to particulate backscattering varied both seasonally and regionally by more than one order of magnitude (Bellacicco et al., 2016). Only 0.02% of the glider backscattering data fell below this threshold and they were discarded. The corrected b_{bp470} was then converted into C_{phyto} through the equation

$$C_{phyto} = 13000\left(b_{bp470} - 3.5\ 10^{-4}\right). \quad (3)$$

The slope was chosen by Behrenfeld et al. (2005) to give satellite chl-a:C_{phyto} values within the range compiled by Behrenfeld et al. (2002) from laboratory experiments (average = 0.010, range = 0.001 to >0.06) and to give an average phytoplankton carbon contribution to total particulate organic carbon of ∼30% (range: 24 to 37%). Behrenfeld et al. (2005) did not provide any statistical assessment of the estimated parameters.

Martinez-Vicente et al. (2013) Method (M13)

Martinez-Vicente et al. (2013) derived phytoplankton carbon from the relationship between C_{phyto} and *in situ* backscattering

TABLE 2 | Comparison of POC vs. b_{bp} reported from literature for original wavelength [taken from Cetinić et al. (2012)] with the addition of results from Graff et al. (2015) and this study.

Author	Sample size	Region	POC vs. b_{bp}
Stramski et al., 2008	54	Pacific, Atlantic[a]	53607 b_{bp} + 2.5
Stramski et al., 2008	59	Pacific, Atlantic[b]	7085.01 b_{bp}–9.1
Stramski et al., 1999	33	APFZ[c]	17069*$b_{bp}^{0.859}$
Stramski et al., 1999	24	Ross Sea	476935.8*$b_{bp}^{1.277}$
Balch et al., 2010	18	N and S Atlantic	841*$b_{bp}^{0.395}$
Loisel et al., 2001		Mediterranean	37550b_{bp} + 1.3
Cetinić et al., 2012	321	North Atlantic	35422b_{bp} –14.4
Graff et al., 2015	53	N and S Atlantic, Equatorial Pacific	48811b_{bp}–24
This study	220	South Atlantic Southern Ocean	39418b_{bp}–13

[a]*excluding upwelling data,* [b]*for entire data set, including upwelling data and* [c]*Antarctic Polar Frontal Zone.*

b_{bp470} from the euphotic zone of a latitudinal study of the Atlantic ocean using a total least square linear regression. C_{phyto} was directly estimated from flow cytometry for 6 different groups of phytoplankton using information on phytoplankton abundances, cellular carbon per unit volume and mean cell volume. The total C_{phyto} concentration per sample was calculated as the sum of the contributions from each phytoplankton type. A significant linear relationship was found between b_{bp470} and C_{phyto}:

$$C_{phyto} = (30100 \pm 1100)\left[b_{bp470} - (7.6 \pm 0.4)\ 10^{-4}\right] \quad (4)$$

The linear regression was limited to $b_{bp470} < 0.003$ m^{-1} because samples above this threshold value (n= 8) exhibited a shift in the relationship that was not possible to describe with a single linear function. In this study we applied the M13 Equation (4) to all glider data, even those higher than 0.003 m^{-1} (these are all surface data representing about 3% of the total, with values up to 0.0048 m^{-1}). In addition, the form of Equation (4) implies that the domain does not comprise b_{bp} values lower than 7.6 10^{-4} m^{-1}. Since in this study most of the b_{bp470} values collected by the glider sensor deeper than 150 m were lower than this threshold (about 54% of the total data), they were discarded from the analysis during the application of the method.

Sathyendranath et al. (2009) Method (S09)

Sathyendranath et al. (2009) derived a method of estimating C_{phyto} using chlorophyll and POC observations from offshore regions of the North West Atlantic Ocean and the Arabian Sea. In their study, the authors considered that POC incorporates all types of particulate carbon in the system (including phytoplankton, bacteria, detritus, and viruses). The S09 method then assumes that for any given chlorophyll concentration, the minimum particulate carbon content represents an upper bound on phytoplankton carbon. In the S09 method, the authors log-transformed both POC and chlorophyll to linearize the relationship and to reduce the weight of the stations with high

values of POC and chlorophyll in the regression analysis. They then used a 1% quantile regression to represent the upper limit on carbon content from phytoplankton alone.

This method can be applied to the available data from the South Atlantic Southern Ocean, using co-located POC and chlorophyll data from both underway and profile CTD samples collected on all cruises listed in **Table 1** and depicted in **Figure 1**. The 1% quantile regression was fitted to the log- transformed data (**Figure 3**) to derive the following equation:

$$C_{phyto} = 42\ (chl)^{0.86} \quad (5)$$

The relationships derived by Buck et al. (1996) and Sathyendranath et al. (2009) are included in **Figure 3** for comparison and show that the SAZ data are characterized by lower carbon content per unit chlorophyll and that the values converge to the Sathyendranath et al. (2009) line at higher chlorophyll concentrations (**Figure 3**).

RESULTS

Seasonal Evolution of Chlorophyll and POC

The glider transect showed substantial seasonal changes in chlorophyll (**Figure 4A**) with relatively low chlorophyll concentrations (<0.5 mg m^{-3}) found at the beginning of the transect between late September and October, coinciding with the deepest MLDs (down to 200 m). Periods of enhanced chlorophyll were shown to be associated with submesoscale physical forcing of the MLD by Swart et al. (2015) and Thomalla et al. (2015). The spatial scales of the physical field are evident in Supplementary Figure 1, and investigated in more detail in a study by Du Plessis et al. (in review; see also their **Figure 6**), which shows small-scale excursions of mixed layer density and surface (100 m) stratification for the glider time series. These rapid changes in density are associated with submesoscale features that actively slump horizontal density gradients (Mahadevan et al., 2012) driving enhanced stratification during periods of lighter mixed layers and weaker stratification when mixed layers are dense (Supplementary Figure 1) (Du Plessis et al., in review). The shoaling of the MLD from ∼100 to 20 m through seasonal heat flux in late November saw a concomitant increase in surface chlorophyll (>0.55 mg m^{-3}) that was sustained throughout summer until the end of the sampling period in mid-February.

Overall, there is a rather good visual correspondence between the seasonal evolution of chlorophyll and POC (**Figure 4B**), which was attained through the application of equation 2 to the glider b_{bp470} data. The bulk of POC was located within the MLD, as it occurs for chlorophyll, with lowest POC concentrations (50–80 mg C m^{-3}) being observed in October, increasing in November (to about 80–120 mg C m^{-3}) and reaching maximum concentrations in December (with values up to 150 mg C m^{-3}) extending through to February. The more evident difference between the datasets is in late October where an increase in POC is not observed in chlorophyll and in early January, when a decrease in POC is not observed in chlorophyll.

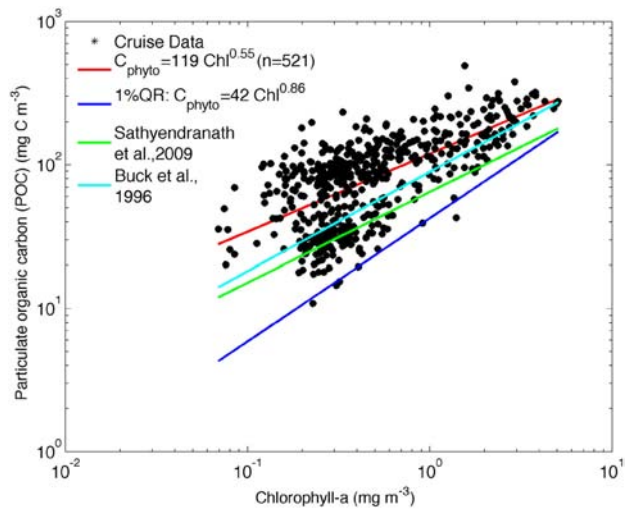


FIGURE 3 | POC against chlorophyll for all cruise data. The least-square fit to log-transformed data is shown in red [$\log_{10}$ POC = (2.08 +/− 0.01) + (0.55 +/− 0.03) $\log_{10}$ Chla], the C_{phyto} minimum carbon estimates by quantile regression [$\log_{10}$ POC = (1.63 +/− 0.05) + (0.86 +/− 0.09) $\log_{10}$ Chla, $p = 0.01$], in dark blue. The relationship between phytoplankton carbon and chlorophyll from Sathyendranath et al. (2009) (green line) and Buck et al. (1996) (light blue) are shown for comparison.

Phytoplankton Carbon Estimates

Figures 5B–D compares the results from the different C_{phyto} methods for selected depths. Corresponding chlorophyll concentrations are also shown (**Figure 5A**). C_{phyto} shows a very similar seasonal distribution to that of chlorophyll and very little difference between surface 10 and 40 m concentrations highlighting the homogeneity expected in well-mixed surface waters. When applied to the glider b_{bp470} and chlorophyll data, 3 of the methods of retrieving C_{phyto} described in Section Discussion showed similar results, despite differences in the equations used. The 30%POC and B05 methods (**Figures 5B–D**) showed almost identical distribution patterns with concentrations ranging between 0 and 45 mg C m^{-3}. This is not surprising since the slope of the B05 equation is about 1/3 of the one found in Equation (2), and by assuming a 30% contribution due to phytoplankton carbon the numbers become similar. These slopes are also very similar to the more recent equation developed by Graff et al. (2015) $C_{phyto} = 12128(b_{bp470} + 4.86 \times 10^{-5})$ (their **Figure 3**) that utilized direct measurements of C_{phyto} from the temperate South Atlantic, equatorial upwelling and oligotrophic gyres. In addition, the S09 method (**Figure 5D**), the only method that uses chlorophyll rather than particulate backscattering to retrieve C_{phyto}, showed remarkably similar results to the methods of B05 and 30%POC.

Values of C_{phyto} from the M13 method were overall much larger, particularly at 10 and 40 m (**Figures 5B,C**), with values up to 80 mg C m^{-3}. With this method, surface concentrations of C_{phyto} in October (∼30–40 mg C m^{-3}) were lower relative to the remaining transect but high when compared to the other three methods (∼20 mg C m^{-3}). Similarly, from November to February, high surface C_{phyto} from M13 ranged from ∼40–100 mg C m^{-3} compared with ∼ 20–50 mg C mg^{-3} in the other

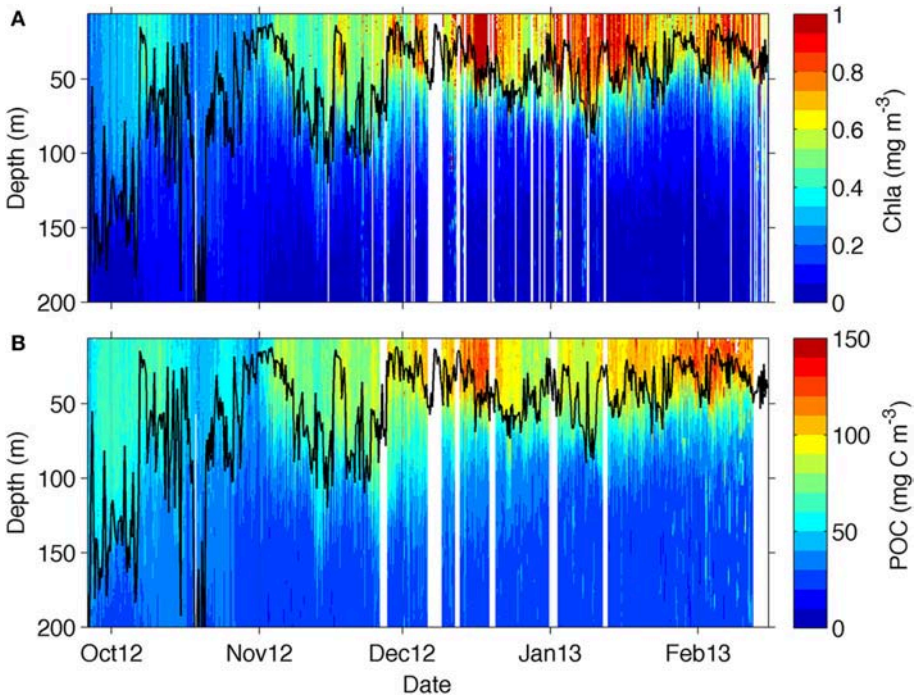


FIGURE 4 | Section plot from the glider transect data for (A) chlorophyll and **(B)** particulate organic carbon (POC) derived from backscattering data from equation (2). The mixed layer depth is marked as a solid black line.

three methods. The discrepancy between M13 and the other methods decreases with depth (80 m) and is less prominent at the beginning of the transect (October). This indicates that the difference between M13 and the other three methods is generally smaller when chlorophyll concentrations are low (<0.5 mg chl-a m^{-3}) in early spring and at depth (>80 m).

Seasonal Changes in Chlorophyll to Phytoplankton Carbon Ratios

In general, the surface ratios (**Figure 6**) appear lower in October (0.01–0.02) when mixed layers were deeper and chlorophyll concentrations were low, increasing toward mid-January (0.02–0.04) with decreasing mixed layers and increasing concentrations of phytoplankton biomass (**Figure 4**). The S09 method, despite being less variable through its dependence on chlorophyll alone, shows a similar background value to the chl-a:C_{phyto} ratios of the B05 and 30%POC methods, except at the beginning of the time series (last week of September through to the first week of October) where a decrease in chlorophyll is not mirrored by a corresponding decrease in C_{phyto} (from b_{bp470} computed methods) resulting in low chl-a:C_{phyto} ratios (from b_{bp470} computed methods). A similar deviation occurred in the last week of January to the first week of February where S09 chl-a:C_{phyto} ratios were noticeably higher than B05 and 30%POC but this time due to an increase in C_{phyto} (from b_{bp470} computed methods) that was not evidenced in chlorophyll. A final example of where chl-a:C_{phyto} ratios between fluorescence based and b_{bp470} based methods deviated is in the last week of October where a peak in the chl-a:C_{phyto} ratios from b_{bp470} computed methods (**Figure 6**) was the result of a distinct drop in carbon that was not evident in chlorophyll (see also **Figure 4B**).

The ratio obtained with the M13 method is about half that obtained from the other methods, with similar short term time variations (~1–7 days) to the other b_{bp470} derived methods but a different seasonal time evolution (i.e., no tendency for M13 chl-a:C_{phyto} ratios to increase from October to mid-January). The seasonal evolution of low to high chl-a:C_{phyto} ratios from spring to summer is evidenced in 30%POC, B05 and S09 methods.

The ratios obtained from the glider data in the SAZ are usually larger than the range obtained with global satellite data (Behrenfeld et al., 2005) as shown in **Figure 6**. This range is more in agreement with the M13 method and the ratio obtained by applying the original Sathyendranath et al. (2009) equation to the data, or by considering all POC as phytoplankton carbon. Interestingly, the model simulation by McKiver et al. (2015), which parameterizes a dynamical ratio, also shows a similar value even if the temporal trend is opposite.

DISCUSSION

This study uses a high-resolution glider data set from the sub-Antarctic Southern Ocean to compute C_{phyto} using four different methods available in the literature. The ability to get sound estimates of C_{phyto} is important as it provides a measure of phytoplankton carbon biomass that is core to many models of net primary productivity, a key indicator of the carbon

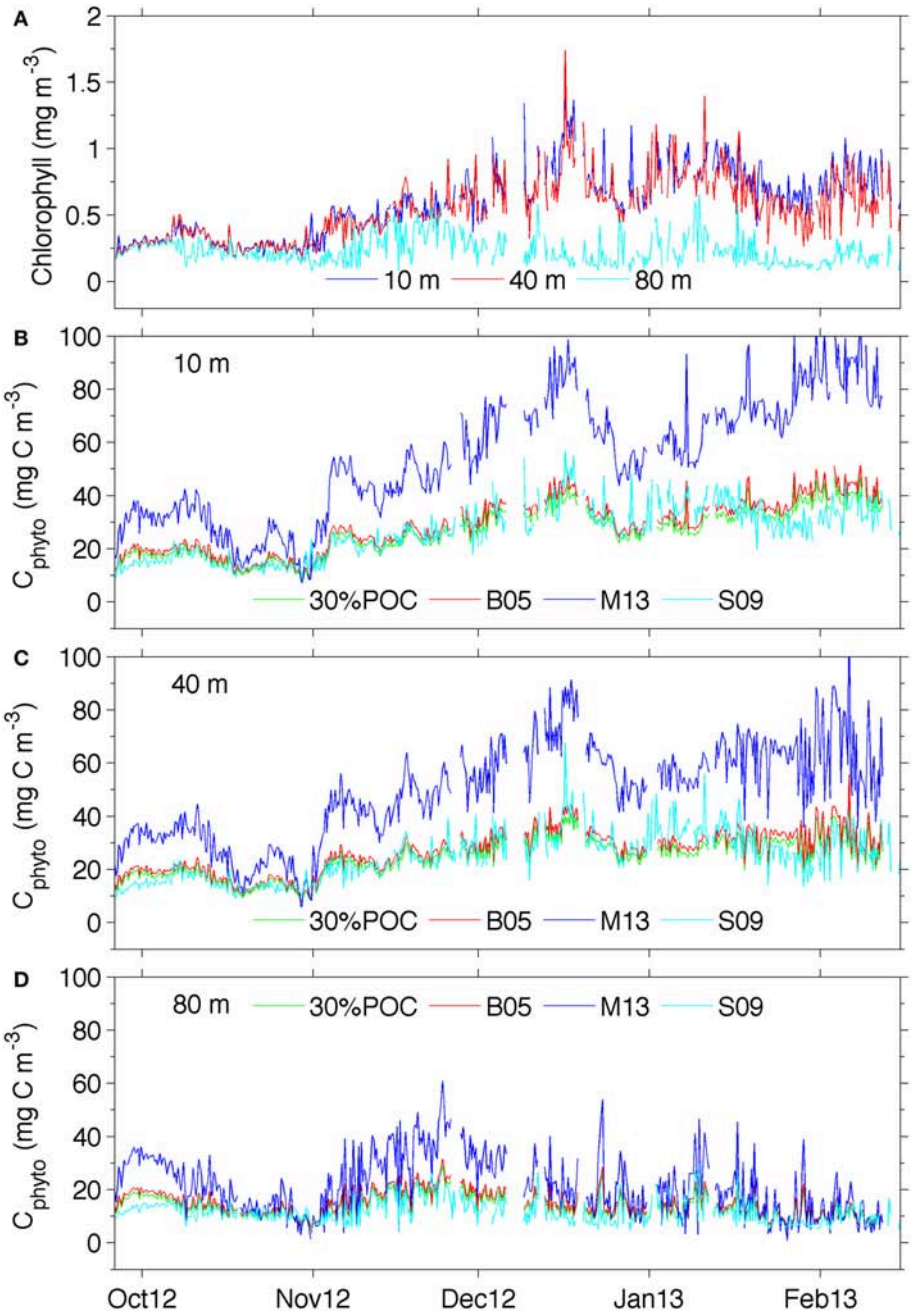


FIGURE 5 | (A) Temporal chlorophyll distribution at different depths and **(B)** temporal distribution of phytoplankton carbon at 10 m, **(C)** 40 m and **(D)** 80 m See supplementary Figure 3 for a version of panel **(B)** showing the ranges derived from using the standard error of the estimates for 30%POC equation (2), M13 equation (4), and S09 equation (5).

cycle. In addition, good measurements of C_{phyto} enable us to better understand changes in cellular chlorophyll-to-carbon ratios, which provides information on phytoplankton physiology (e.g., cellular adjustments to changes in light, temperature, and nutrients, Behrenfeld et al., 2005). As such, deriving good estimates of C_{phyto} and chl-a:C_{phyto} ratios will enable us to refine model parameterizations of phytoplankton dynamics (Sathyendranath et al., 2009). Since the biological pump in the Southern Ocean drives 33% of global organic carbon flux (Schlitzer, 2002; Lenton et al., 2013), it is of particular importance that we improve our understanding of the biological response of Southern Ocean phytoplankton to climate change.

Comparison of the Methods

Rather than assess the quality and merit of individual methods of deriving C_{phyto} against each other, which would require

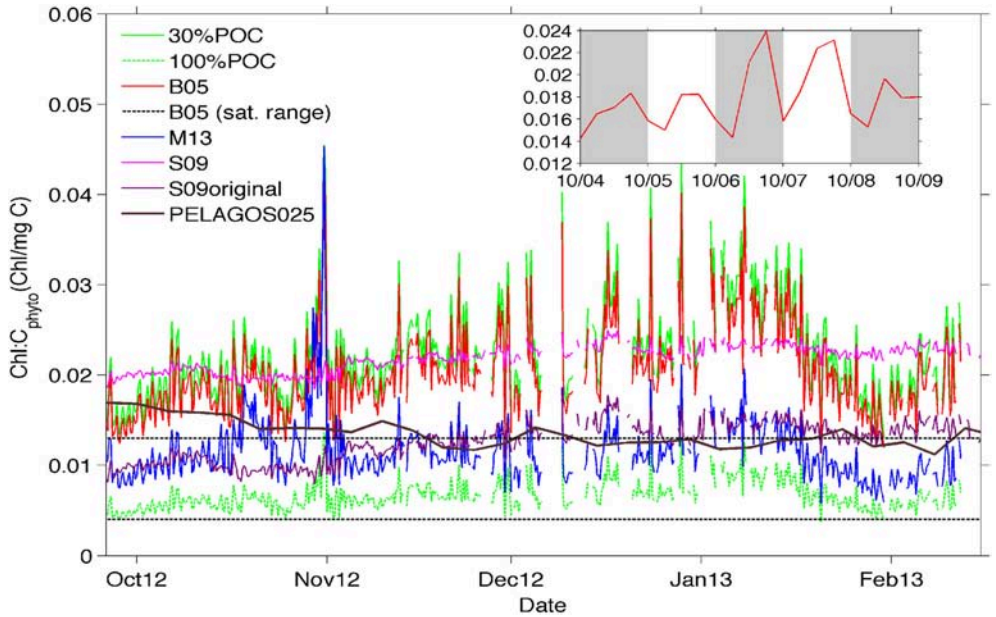


FIGURE 6 | Time-evolution of chl-a:C_{phyto} ratios at the surface (10 m) derived from the 30%POC method (solid lighter green top line), the B05 method (red line), the M13 method (blue line), and the S09 method (pink line). In addition, Chl-a:C_{phyto} ratios were calculated using the chl-a:POC ratio from the cruise data (which implies that all POC is phytoplankton specific) and is presented as 100%POC (darker green bottom dashed line). Included for comparision are (1) the chl-a:C_{phyto} ratios derived from the original equation from Sathyendranath et al. (2009) presented as S09original (purple line), (2) the satellite range of ratios from Behrenfeld et al. (2005) (black dotted lines) and (3) the ratios derived from the PELAGOS025 model (McKiver et al., 2015, extracted from the model for the same geographical co-ordinates as the glider transect in time but for a year 2011 simulation, solid black line). The inset shows a detail of the daily signal for the B05 method.

independent phytoplankton data that would inevitably be limited in time and space, this study aims to evaluate the differences of the various methods using high frequency optical glider data from the SAZ as a benchmark. It should also be noted that the methods applied in this research were developed using different data obtained from different oceanic regions. The M13 method is the only equation that was derived from *in situ* C_{phyto} data (computed from cell counts and assumptions of carbon content per cell) but from a latitudinal study of the Atlantic; the B05 equation is also non-specific to the Southern Ocean and derived from global satellite data; the 30%POC and S09 methods were recomputed using in *situ* POC and chlorophyll data from the Southern Ocean. To further extend the number of methods, the original equation from the S09 method developed using data from the NW Atlantic and Arabian Sea (Sathyendranath et al., 2009) has been included in **Figures 6**, **8** for comparison (presented as S09original).

The most evident result of the comparison is that three of the methods cluster together (30%POC, B05, and S09) with only one (M13) being substantially different in magnitude from the others (**Figure 5**). Despite their different origins, C_{phyto} from 30%POC, B05, and S09 (the only method that is derived from chlorophyll) all show similar results. The similarity in C_{phyto} derived using B05 (which made use of a global data set to convert b_{bp} to POC) and 30%POC (that used a regionally specific conversion based on Southern Ocean filtered samples) implies that the conversion from b_{bp} to POC used here is regionally robust. This is confirmed by a comparison of the POC vs. b_{bp} relationship from previously published values (**Figure 2**, **Table 2**), which shows the slope of the regression between POC and b_{bp} to fall well within the observed literature range from various ocean basins. The relationship from our data (from the south Atlantic Southern Ocean; **Figure 2**, **Table 2**) is very similar to that generated from the Mediterranean (Loisel et al., 2001), the North Atlantic Bloom Experiment (NABE, Cetinić et al., 2012) and the Polar Frontal Zone (PFZ, Stramski et al., 1999). Our POC vs. b_{bp} slope is however lower than that from the North and South Atlantic and Equatorial Pacific (Graff et al., 2015), the Ross Sea (Stramski et al., 1999), and the slope that is the basis of the NASA POC algorithm (Stramski et al., 2008), using data from the Pacific and Atlantic but excluding data from the upwelling regions (Cetinić et al., 2012). On the contrary, our slope is higher than that from the Atlantic and Pacific including data from the upwelling regions (Stramski et al., 2008) and yields higher POC (for b_{bp} >0.0027) when compared to the relationship from the north and south Atlantic (Balch et al., 2010).

The M13 method returns much higher concentrations of C_{phyto} with the difference being less when chlorophyll concentrations are below $< 0.5\ mg\ m^{-3}$) in the month of October (**Figure 5**) and at depth ($\sim$ 80 m). The reason for the higher C_{phyto} estimates using M13 is the steep slope (30100) of Equation (4), which is more than twice the slopes used in B05 (13000) and 30%POC (11825). The comparatively steep slope of M13 was noted in Martinez-Vicente et al. (2013) and the explanation proposed for the doubling of parameters between their equation and that of Behrenfeld et al. (2005) was due to "uncertainties

in satellite and *in situ* estimates of b_{bp} and/or differences in the spatio-temporal scales of the two studies." A big difference in the steepness of the slope was similarly noted between M13 (slope = 30100) and a study by Graff et al. (2015) in the North and South Atlantic and Equatorial Pacific (slope = 12128). They suggest that the difference may be due to variability in the assumed volume-based phytoplankton biomass conversions used in the M13 method to convert volume to C_{phyto}. This can yield an order of magnitude variability in resultant C_{phyto} estimates (Caron et al., 1995; Dall'Olmo et al., 2011).

Equation (4) from Martinez-Vicente et al. (2013) was generated from the typically low biomass region of the Atlantic, where *in situ* b_{bp} data were generally less than 0.003 m^{-1}. The eight data points that exceeded this threshold were removed from the regression by the authors, as they showed a shift in the relationship that could not be described with a single linear function. An inclusion of the 8 data points (where b_{bp} was > 0.003) would drive an even steeper slope, even higher concentrations of C_{phyto}, and a greater disagreement between the different methods used here. These 8 data points were characteristic of eutrophic conditions (as opposed to oligotrophic conditions) where larger cells characterized the phytoplankton community (more Nano- and Pico- eukaryotes and less *Prochlorochoccus*) with lower b_{bp}:C_{phyto} ratios (see Martinez-Vicente et al., 2013, their Supplementary Figure 1). Similarly, if microphytoplankton were included in their measurements (missed by flowcytometry), the slope would likely be even steeper. This would be consistent with optical theory that predicts a decrease in b_{bp}:C_{phyto} ratios with an increase in particle size, where C_{phyto} is proportional to volume (Boss et al., 2004; Martinez-Vicente et al., 2013). However, worth noting is that this theory represents phytoplankton cells as homogenous spheres that can overestimate the real refractive index for certain species (Vaillancourt et al., 2003). As described in Section Martinez-Vicente et al. (2013) method (M13), the M13 equation was applied to all glider b_{bp} data (ranging from 0.001 to 0.01 m^{-1}), with much of the data falling above the domain of applicability of Equation (4) (b_{bp} <0.003). It is thus possible that the overestimated C_{phyto} concentrations generated from the M13 method are the result of an unsuitable application of a regionally derived model from the low backscattering field of the Atlantic to the relatively high backscattering SAZ. This possibility is supported by all four methods generating similar C_{phyto} concentrations when biomass was below 0.5 mg m^{-3} (i.e., in early spring and at depth ∼ 80 m, **Figure 5**). As such, when chlorophyll concentrations are below 0.5 mg m^{-3} it appears to make little difference which method of estimating C_{phyto} is chosen by the user.

The S09 method showed similar C_{phyto} results in range and distribution to both B09 and 30%POC, highlighting its robustness in converting chlorophyll to C_{phyto}. This information is particularly useful for data sets that require conversion to phytoplankton specific carbon in the absence of backscattering measurements. However, when the chl-a:C_{phyto} ratios from the different methods are compared (**Figure 6**), the limitations of the S09 method are highlighted, namely the very low temporal variability in chl-a:C_{phyto} ratios. This can be attributed to the derivation of chl-a:C_{phyto} ratios from a monotonic function of chlorophyll concentration:

$$chla : C_{phyto} = \frac{1}{a}\left(chla\right)^{1-b} \quad (6)$$

where a= 42 and b= 0.86 are the coefficients derived from Equation (5). The implication of this is that the S09 method, though being non-linear, does not allow for a scenario where C_{phyto} increases or decreases without a corresponding change in chlorophyll. The remaining three methods on the other hand allow C_{phyto} to vary independently of chlorophyll, which accounts for the higher temporal variability observed in the chl-a:C_{phyto} ratios using the B05, 30%POC and M13 methods (**Figure 6**).

Seasonal and Sub-Seasonal Variations in the Chlorophyll-to-Carbon Ratio

How do surface chlorophyll-to-carbon ratios (chl-a:C_{phyto}) vary in time? In general chl-a:C_{phyto} ratios tend to be lower in spring (October), higher in summer (December – January), and decreasing again in late summer (February) in particular in the backscattering methods (30%POC, B05, and M15). The absence of a strong seasonal evolution in M13 is likely a consequence of the method which drives higher C_{phyto} (relative to the other methods) when chlorophyll concentrations are high, and more similar C_{phyto} when chlorophyll is below 0.5 mg m^{-3} thus dampening any seasonal driven signal in surface chl-a:C_{phyto} ratios.

According to the literature, chl-a:C_{phyto} ratios tend to be highest when larger diatoms are present and lowest when smaller species dominate (e.g., *Prochlorococcus*) (Sathyendranath et al., 2009). As such, some of the seasonal variability observed in the chl-a:C_{phyto} ratios could be the result of different sized species dominating the community at different times. Smaller cells in spring (October) when light is potentially limiting (mean MLD = 87 m and mean PAR in MLD = 0.01 μ E cm^{-2} s^{-1}) and late summer (January) when nutrients are potentially limiting (depleted reservoir), vs. larger cells in early summer (December) when light (Supplementary Figure 2) and nutrients are thought to be unrestricted (mean PAR in MLD = 0.03 μ E cm^{-2} s^{-1}). Previous studies have shown that the range of chl-a:C_{phyto} ratios observed in the Southern Ocean are the result of seasonal variation in the physiological response of phytoplankton to light and nutrient limitation (Behrenfeld et al., 2005). Phytoplankton chl-a:C_{phyto} ratios tend to decrease with increasing light conditions and decreasing temperature and nutrient concentrations (and vice versa) (Geider et al., 1997; Socal et al., 1997; Taylor et al., 1997; Behrenfeld et al., 2005; Lü et al., 2009; Halsey and Jones, 2015; Westberry et al., 2016). Increasing chl-a:C_{phyto} ratios as a physiological response of phytoplankton to low light conditions enables them to increase their light harvesting ability by increasing the volume of chlorophyll packed into their cells (Behrenfeld and Milligan, 2013; Halsey and Jones, 2015; Bellacicco et al., 2016). Indeed, photoacclimation to changes in the underwater light field as a bloom develops have been known to account for the entire range of seasonal variability

observed in bulk chlorophyll (e.g., Winn et al., 1995; Westberry et al., 2008). However, since Fe reservoirs are replenished in winter through deep mixing and depleted through biological uptake in the Spring Summer growing season (Boyd et al., 2010), nutrients are not considered the likely driver of the observed increase in chl-a:C_{phyto} ratios. Similarly, both light (through increased PAR Supplementary Figure 2 and decreased MLD **Figure 4**) and temperature (see Thomalla et al., 2015 their **Figure 2B**) tend to increase from October to December. It is thus unlikely that these processes (light, temperature or nutrients) are responsible for the seasonal increase in chl-a:C_{phyto} ratios observed during this period (**Figure 6**). However, since laboratory studies consistently show a decrease in chl-a:C_{phyto} ratios under Fe stress (Greene et al., 1992; Sunda and Huntsman, 1995), it is possible that limiting nutrients could contribute to the observed decrease in chl-a:C_{phyto} ratios in late summer (February). Similar results of declining chl-a:C_{phyto} ratios were observed by Behrenfeld et al. (2005) in regions dominated by large spring–summer blooms in phytoplankton abundance, where a decrease in chl-a:C_{phyto} ratio was observed prior to the seasonal biomass crash..

The physiological response of phytoplankton to relief from Fe stress has shown increases in chl-a:C_{phyto} ratios as growth rate increases (Greene et al., 1992; Geider et al., 1997; Sunda and Huntsman, 1997). This response is the result of upregulation of photosynthetic machinery and light harvesting capacity (Laws and Bannister, 1980). A measure of the growth rate estimated as the rate of change of MLD integrated chlorophyll (Supplementary Figure 2) shows an increase from 0 at the beginning of the time series (October) to ∼0.017 mg chl-a m^{-2} d^{-1} in mid-December. Although we are unable to quantify the role of Fe in this study, the increase in chl-a:C_{phyto} ratios over the same time period may be a physiological response to increased growth rates when neither Fe or light were considered limiting. This was a similar response to what was observed during bloom conditions by Westberry et al. (2013) in both natural and purposeful Fe addition experiments. In addition, phytoplankton must photoacclimate as the bloom develops to counter the effects of self-shading, further increasing chl-a:C_{phyto} ratios over the growing season.

Finally, an alternative explanation for the observed seasonal increase in chl-a:C_{phyto} ratios is that the ratio of C_{phyto} to total POC (and to b_{bp}) decreases as the seasonal bloom develops. In other words, as the season progresses a greater percentage of the particulate backscattering signal is due to non-phytoplankton specific carbon (e.g., heterotrophic bacteria, detritus, viruses, ciliates) (Christaki et al., 2011). This was observed in a study in the Mediterranean where the non-algal contribution to b_{bp470} was generally larger in the more productive regions that had elevated phytoplankton abundances (Bellacicco et al., 2016). Behrenfeld et al. (2005) reported that a compiled data set of field derived C_{phyto}:POC ratios spanning oligotrophic to eutrophic ocean regions ranged from 19 to 49%. These differences reflect system variability in producer-consumer dynamics, processes influencing the particle field and possible differences in export efficiency (Graff et al., 2015). If this ratio varies seasonally as much as regionally (i.e., by a factor of 2.5), it will have a substantial effect on optically derived chl-a:C_{phyto} ratios. Of the four methods applied here, all three b_{bp470} computed methods (B09, 30%POC and M15) apply a constant C_{phyto}:POC ratio with time. On the other hand, the S09 method may indirectly account for adjustments in C_{phyto}:POC through the non-linear relationship between chlorophyll and C_{phyto} that drives an increase in chl-a:C_{phyto} ratios with increasing chlorophyll. To elaborate; as the bloom develops, phytoplankton biomass increases with a concomitant decrease in the percentage of C_{phyto} relative to total POC. This may be reflected in the logarithmic relationship between chlorophyll and C_{phyto} (**Figure 3**), which would be observed as an increase in chl-a:C_{phyto} ratios with increasing chlorophyll concentrations in the S09 method (**Figure 6**).

Over and above the characteristic seasonal cycle in chl-a:C_{phyto} ratios, there is strong sub-seasonal and daily variability that are likely driven by phytoplankton community responses to smaller scale physical processes and day-night physiological adjustments. In the case of the M13 method, where the seasonal cycle is dampened, one can argue that the smaller sub-seasonal scales are in fact dominating the variability. Some distinct examples of smaller scale adjustments in the chl-a:C_{phyto} ratios are as follows:

1. The low chl-a:C_{phyto} ratios (from b_{bp470} derived methods) at the very beginning of the time series (late September/early October) (**Figure 6**) that are associated with deep MLD's (∼150 m, **Figure 4**) when the glider crosses the cold core of a mesoscale cyclone ∼200 km in diameter (Swart et al., 2015);
2. A peak in chl-a:C_{phyto} ratios (from b_{bp470} derived methods) in the last week of October (**Figure 6**) that coincides with an intense shoaling of the MLD to ∼25 m (**Figure 4**), the timescale of which relates to an oceanographic feature and an event that is mesoscale in nature (slumping of the lateral gradient in density) (Mahadevan et al., 2012; Swart et al., 2015); and
3. A distinct daily signal in chl-a:C_{phyto} ratios with higher ratios during the day that are arguably the result of daily physiological adjustments in phytoplankton chlorophyll content per unit carbon.

The drivers of sub-seasonal variability in chl-a:C_{phyto} ratios seen in late September (low) and late October (peak) are clearly linked to mesoscale features and associated adjustments in the depth of the mixed layer, which is deep in late September (∼150 m) and shallow in late October (∼25 m). However, the adjustments in chl-a:C_{phyto} ratios are not a physiological response of the phytoplankton to light. Were this the case, the opposite response would be true i.e., an increase in chl-a:C_{phyto} ratios would be observed when MLD's were deep and light was supposedly limiting. It is thus more likely that these small scale adjustments in chl-a:C_{phyto} ratios are the result of a shift in community structure from a small cell dominated community with low growth rates in late September to a population dominated by large cells with high growth rates under high light conditions in late October.

To test whether the diurnal variability in the chl-a:C_{phyto} signal (see insert in **Figure 6**) was driven by an artifact

of residual uncorrected solar quenching (or the quenching correction itself), we compared in **Figure 7** the linear regressions of midnight and midday C_{phyto} ($r^2 = 0.93$) to midnight and midday chlorophyll ($r^2 = 0.66$). Although the coefficient of determination for chlorophyll is lower, this is what can be expected from the effects of acclimation. The normal dispersion of the data however suggests that the variability in chlorophyll is not biased and hence implies that there is no systematic artificial error introduced through the quenching correction (**Figure 7**). As such, it is believed that the daily signal in chl-a:C_{phyto} ratios is a real response of the phytoplankton community. Interpreting this diurnal variability is difficult as it depends on numerous parameters that include phytoplankton concentration, composition and physiological status together with the detrital and small heterotroph concentration, all of which are typically not known and hence diurnal variability in chl-a:C_{phyto} ratios remains poorly understood. The likely drivers however include the balance between daytime production and night time degradation, changes in particle size distribution and changes in the refractive index driven by varying internal concentrations of organic compounds (e.g., accumulation of intracellular carbon through photosynthesis) (Kheireddine and Antoine, 2014). Although this small-scale variability elicits further detailed investigations it remains outside of the scope of this manuscript.

Implications for Model Parameterizations

Some biogeochemical models apply a constant chl-a:C_{phyto} ratio globally (usually 0.02 mg chl-a mg C^{-1}) even if the constraints are not developed equally (i.e., some regions are data poor), while others have a dynamical function that should include acclimation to prevailing light conditions. It is entirely possible that the inadequate parameterization of this ratio is the reason for the seasonal biases found in biogeochemical models (Doney et al., 2009; Steinacher et al., 2009; Vichi and Masina, 2009). In **Figure 6** we compare the chl-a:C_{phyto} ratios generated from a Southern Ocean data set (using the different optical methods) with those constrained in satellite data, cruise data and from one medium resolution global model (the PELAGOS025 model analyzed by (McKiver et al., 2015), Model data). The chl-a:C_{phyto} data from the model was extracted from the same location as the glider transect and over the same dates (month and day) but from a simulation for the year 2011, due to a limitation in the availability of atmospheric forcing functions.

Since C_{phyto} cannot exceed POC, the chl-a:POC ratio from the cruise data (**Table 1**) was used to set a minimum chl-a:C_{phyto} ratio assuming all POC was phytoplankton specific (presented as 100%POC, **Figure 6**). These ratios were understandably low (<0.01 mg chl-a mg C^{-1}) but above the minimum satellite range (0.004 mg chl-a mg C^{-1}) of Behrenfeld et al. (2005) and oftentimes close in magnitude to the ratios generated by the M15 method. The ∼50% lower chl-a:C_{phyto} ratios for M13 are driven by the high C_{phyto} values that this method generates relative to the other methods, particularly when biomass is high (December to February). It was proposed earlier that the low chl-a:C_{phyto} ratios were a possible result of an unsuitable application of a regionally derived model from the Atlantic to the SAZ. The S09 method allows us to do a direct comparison of the chl-a:C_{phyto} results generated from one model derived predominantly from data from the NW Atlantic (S09original) with the same model but derived from data from the Southern Ocean (S09, **Figure 6**). This comparison shows how the application of the S09original model to the SAZ data set results in much lower chl-a:C_{phyto} ratios that are more inline with those generated by the M13 method, that was similarly derived from cell counts from the Atlantic.

Both the M13 and S09original methods produce chl-a:C_{phyto} ratios that are very similar to the 100%POC method, which outputs the minimum ratios possible assuming all POC is phytoplankton specific. This is not surprising when one considers that the slope of the M13 method (30100 mg C m^{-2}) is 76%

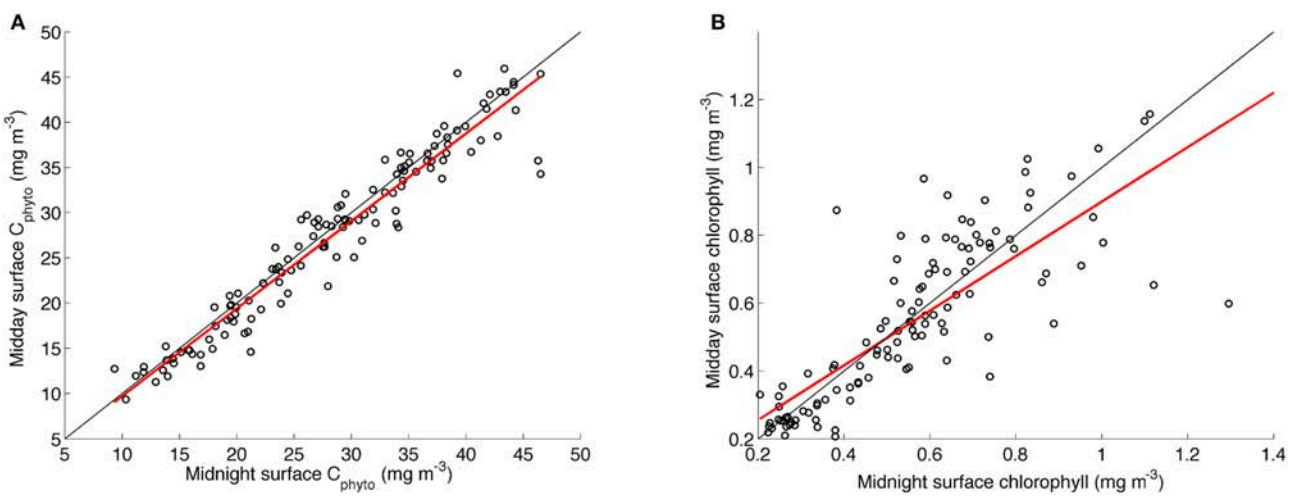


FIGURE 7 | Linear regressions of daily surface midnight vs. daily surface midday concentrations of (A) C_{phyto} from the glider transect using the B05 method (slope = 0.97, $r^2 = 0.93$, rmse = 2.49) and **(B)** chlorophyll (slope = 0.8, $r^2 = 0.66$, rmse = 0.15). The linear least square regression line appears in red with the one to one line in black.

of the POC: b_{bp} slope found in this study (39418 mg C m^{-2}) (**Figure 2**). This finding implies that when the M13 method is applied to the glider data set, there is very little non-algal POC that co-varies with C_{phyto}. This result is unlikely if one considers that the maximum C_{phyto}:POC ratio reported by Behrenfeld et al. (2005) is 49% and in the Equatorial Upwelling and temperate spring waters of the Graff et al. (2015) study rarely exceeded 40% (mean = 25%). Although phytoplankton has been known to dominate POC at a high biomass coastal upwelling site (Hobson et al., 1973) this observation is in contrast to the typical low contribution of C_{phyto} to POC in productive offshore waters (Hobson et al., 1973; Andersson and Rudehäll, 1993). As such, it is more likely that the slope of the M13 method (and S09original) is too high for this region and that the 30%POC, S09 and B05 methods are more appropriately applied to Southern Ocean data sets, supporting the argument of an unsuitable application of a regionally derived model from the Atlantic to the SAZ. The M13 method still resolves the shorter-term variations likely driven by photo-physiology but reduces the summer increase in the ratio explained by adjustments in nutrient driven growth rates and community structure.

A comparison of the results with the PELAGOS model (**Figure 6**) imply that an overall chl-a:C_{phyto} ratio of 0.02 mg chl-a mg C^{-1} seems adequate to represent the SAZ phytoplankton during spring and summer. However, the glider data indicate that the sub-seasonal and seasonal variations of this parameter are rather high and may play an important role in determining the magnitude of the bloom. In addition, the PELAGOS025 model shows an apparent misrepresentation of the seasonal cycle of chl-a:C_{phyto} ratios, with high ratios in winter and low ratios in summer (whereas the glider shows an increase in the chl-a:C_{phyto} ratios from winter through to summer). This is likely because PELAGOS025 is based on the Biogeochemical Flux Model (Vichi et al., 2007; Vichi and Masina, 2009) that uses a variation of the Geider et al. (1997) formulation for light acclimation and chlorophyll regulation. As such, the seasonal mismatch may be a result of the model assumption that all seasonal variability is purely due to acclimation to a prescribed high optimal chl-a:C_{phyto} ratio (deep mixed layers in winter driving increased chlorophyll synthesis through packaging responses to low light conditions). This may explain the observed seasonal biases of the models when compared with ocean color data reported in the literature. Rather, if our estimates of C_{phyto} based on backscattering are appropriate for the Southern Ocean (and this can only be validated with corresponding *in situ* estimates) then the glider data suggest that models need to account for low light adaptation in winter, which would dampen the acclimation response and allow for lower optimal chl-a:C_{phyto} ratios in winter. While in summer, the large increase in chlorophyll may need to coincide with a shift toward a community characterized by larger cells, relatively higher growth rates and higher optimal chl-a:C_{phyto} ratios. Such a mechanism is currently not implemented in any of the biogeochemical models because they usually consider one single group of diatoms. Numerical models, even the ones with a more sophisticated physiology like the BFM (or PISCES) may account for acclimation to light but are still dominated by the same kind of diatoms without any additional plasticity. The PELAGOS025 model does however do a better job in capturing the range of observed chl-a:C_{phyto} ratios, which the satellite data do not (**Figure 6**; see satellite range from Behrenfeld et al. (2005) for the Southern Ocean 0.004 – 0.013, their **Table 1**).

A Reference Chlorophyll-to-Carbon Ratio for the Sub-Antarctic Zone?

The data used for this research provides for the first time Cphyto obtained from a high-resolution glider data set to derive independent surface chl-a:Cphyto ratios from the SAZ. This is a much-needed parameter for biogeochemical models to improve the simulation of phytoplankton blooms in the region. To summarize the range of results obtained with the glider data from the different methods described in Section Materials and Methods, we compare in **Figure 8** their surface (10 m) chl-a:Cphyto ratios to several literature values and ranges (Montagnes et al., 1994; Sunda and Huntsman, 1995; Llewellyn and Gibb, 2000), as well as to the model results, using non-parametric distributions (mean, median and 5–25 percentiles).

Our analysis suggests that it is not possible to establish one single value for the conversion between chlorophyll and carbon. The large range of observed variability is driven by methodological uncertainties, regional differences and large seasonal variations. All chl-a:C_{phyto} ratios from the different methods applied to the glider data set (30%POC, B05, S09, M15) fell within the data set limits compiled by Behrenfeld et al. (2002), which for the global ocean range between 0.001 and >0.06 mg chl-a mg C^{-1} (Behrenfeld et al., 2005). Similarly, all methods generated chl-a:C_{phyto} ratios that fell within the range of laboratory culture measurements collated here (Supplementary Table 1) and within the range of ratios generated when one applies the 30% mean and 19–49% range in C_{phyto}:POC ratios (Behrenfeld et al., 2005) to the cruise POC data. Worth noting here is that although the Behrenfeld et al. (2005) range of 19–49% is presented for comparison (derived from the following references: (Eppley et al., 1992; Durand et al., 2001; Gundersen et al., 2001, and Oubelkheir, 2001), this range has been shown to extend to a minimum of 14% by Oubelkheir et al. (2005) and as high as 75% by Martinez-Vicente et al. (2013).

The smaller range (0.02–0.025 mg chl-a mg C^{-1}) of distribution in chl-a:C_{phyto} ratios from the S09 method (which does not allow for independent adjustments in C_{phyto} relative to chlorophyll) is clear (as in **Figure 6**). This range increases (0.008 – 0.018 mg chl-a mg C^{-1}) when applying the S09original model, which is characterized by a lower phytoplankton carbon per unit chlorophyll (**Figure 3**), likely due to the different region of origin of the analyzed samples. A much greater spread is evidenced in the cruise data (0.002–0.17 mg chl-a mg C^{-1}) relative to the glider data (0.006–0.05 mg chl-a mg C^{-1}). This is likely driven by the large regional coverage of the cruises relative to the glider (**Figure 1**) with the cruises likely sampling a much wider range of communities exposed to more varied growth conditions. In line with this argument, it follows that the range of variability of the cruise data is on a similar scale to that found in the laboratory culture experiments (0.0001–0.05), which are from a large variety of phytoplankton species and growing conditions.

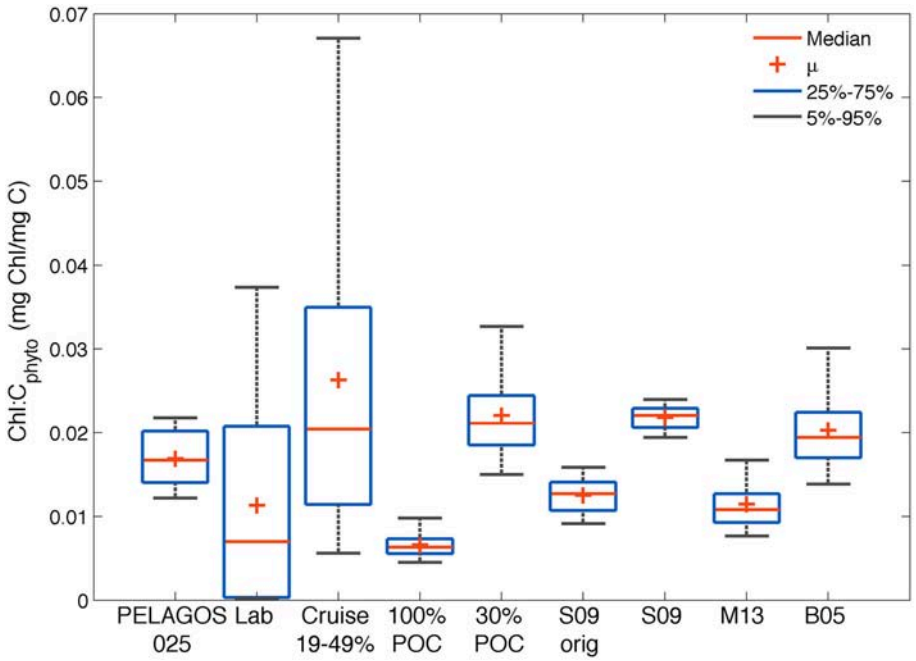


FIGURE 8 | Comparison of the distribution of chl-a:Cphyto ratios derived from (1) the four glider estimates in the SAZ (30%POC, S09, M13 and B05); (2) the PELAGOS025 model for the glider transect (0–10 m); (3) a collation of laboratory culture experiments from the literature [Lab: *n*= 127, Supplementary Table 1 (Montagnes et al., 1994; Sunda and Huntsman, 1995; Llewellyn and Gibb, 2000)]; (4) from Southern Ocean cruise data (Cruise, 0–20 m) assuming a phytoplankton carbon fraction of 19, 30, and 49%; (5) minimum possible ratios derived from the glider data if all POC was phytoplankton specific (100%POC) and 6) from the original Sathyendranath et al. (2009) model developed from data from the NW Atlantic and Arabian Sea (S09orig).

The high variability in observed chl-a:C_{phyto} ratios illustrates the potential error associated with predicting C_{phyto} concentrations based on chlorophyll concentrations alone.

If you round the median of the different data sets off to two decimal places, they fall into two distinct groups. Those with low median chl-a:C_{phyto} ratios (∼0.01 mg chl-a mg C^{-1}) from the laboratory cultures and glider data when the 100%POC, M13 and S09original methods are applied, and those with higher median chl-a:C_{phyto} ratios (∼0.02 mg chl-a mg C^{-1}) from the cruise data, PELAGOS025 model and the glider data with the 30%POC, B05 and S09 methods applied. Interestingly, another recent study from the north and south Atlantic, using direct measurements of C_{phyto}, had the same median chl-a:C_{phyto} ratio (0.01 range 0.029–0.002) as the M13 and S09original methods that were similarly all derived from the Atlantic (Graff et al., 2015). These results suggest that if reality is a low median chl-a:C_{phyto} ratio (∼0.01), then the M13 and S09original methods would produce the best results when applied to the glider data set from the SAZ. On the other hand, if reality is a higher median chl-a:C_{phyto} ratio (∼0.02) then the 30%POC, S09, and B05 methods better represent reality in the SAZ.

Nevertheless, even assuming that the values with a median of 0.02 mg chl-a mg C^{-1} are more realistic, the spread is sensibly large, and apparently not purely driven by acclimation given the phase discrepancy between the model results and the glider data (**Figure 6**). Further insights on the merit of one method over another can only be obtained with the aid of concurrent data on the phytoplankton community composition and their specific chlorophyll and carbon content. To this end, recent advances in technology such as sorting flow-cytometry and high sensitivity elemental analysis can allow for quantitative assessment of C_{phyto} (e.g., Graff et al., 2015) which will contribute to a more extensive set of field data for evaluating and validating optical methods of determining C_{phyto} concentrations.

CONCLUSIONS

This study used optical data from a high-resolution glider transect in the SAZ to compare four different empirical estimates of phytoplankton carbon (three based on particulate backscattering and one on chlorophyll). The chl-a:C_{phyto} ratios generated from the different methods were compared with *in situ* data from the literature and a model simulation to inform on their comparative range and distribution. Out of the four methods used, three (30%POC, B05, and S09) showed similar C_{phyto} concentrations, despite their different origins, in particular when chlorophyll concentrations were below 0.5 mg m^{-3}. The fourth (M13) derived higher concentrations of C_{phyto} when chlorophyll concentrations were high (>0.5 mg m-3). The S09 method produced very similar C_{phyto} concentrations to B09 and 30%POC, highlighting its robustness in converting chlorophyll to C_{phyto} in the absence of particulate backscattering. However, the S09 method does not allow for adjustments in C_{phyto} without proportional changes in chlorophyll and hence cannot account

for cellular adjustments in chl-a:C_{phyto} ratios that are not driven by biomass. Methods derived from b_{bp} on the other hand, showed high seasonal and sub-seasonal variability in chl-a:C_{phyto} ratios that can be attributed to adjustments in dominant species composition, physiological adaptations to varying light and nutrient regimes, changes in growth rates and variability in the ratio of C_{phyto} to total POC.

All chl-a:C_{phyto} ratios generated from the four different methods fell within the literature range compiled by Behrenfeld et al. (2002), the range of collated laboratory culture measurements and the cruise POC data (when a 19–49% range in C_{phyto}:POC ratio was applied). Nonetheless, the methods derived from the North Atlantic (M13 and S09original) were shown to generate mean chl-a:C_{phyto} ratios that were comparatively low (0.01 mg chl-a mg C^{-1}) and did not allow for sufficient variation of non-algal POC with $C_{phyto.}$ The 30%POC, S09, and B05 methods on the other hand generated higher mean chl-a:C_{phyto} ratios (0.02 mg chl-a mg C^{-1}) that were considered more appropriate for application to the SAZ. This highlights the potential for unsuitable application of regionally derived methods to the Southern Ocean.

Model simulations tend to overestimate the magnitude and miss the timing of the Southern Ocean spring-summer bloom, which McKiver et al. (2015) suggested was a possible result of inaccurate parameterization of the chl-a:C_{phyto} ratio. A comparison of the glider surface chl-a:C_{phyto} ratios with those from laboratory culture experiments from the literature, Southern Ocean cruise data (assuming a C_{phyto}:POC range of 19–49%) and the same model output used by McKiver et al. (2015) show that although an overall chl-a:C_{phyto} ratio of 0.02 mg chl-a mg C^{-1} could adequately represent the SAZ phytoplankton during spring and summer, the seasonal variation of this parameter is high and may play an important role in characterizing the bloom. It is proposed that the seasonal mismatch between the model and the glider data may result from the models assumption that all seasonal variability is simply due to acclimation. If indeed the seasonal ramp in chl-a:C_{phyto} ratios observed with the glider is related to higher growth rates (and not to a decrease in % contribution of C_{phyto} to b_{bp} as the seasonal bloom develops) then these results suggest that models need to accommodate a variable chl-a:C_{phyto} ratio that accounts for phytoplankton adaptation to low light conditions in spring (low optimal chl-a:C_{phyto} ratio) and higher optimal chl-a:C_{phyto} ratios with species-specific increasing growth rates in summer. Another option, as suggested by the work of Bellacicco et al. (2016), is that methods converting backscattering to C_{phyto} need to take into account the space-time variability of non-algal contributions to particulate backscattering, which can vary by more than one order of magnitude. To further our understanding of the merits of different optical methods of determining C_{phyto}, additional data is required on concurrent community composition and specific chlorophyll and carbon content.

AUTHOR CONTRIBUTIONS

All authors contributed to the final manuscript: ST designed the research, processed the data, and contributed to the writing of the manuscript. AO prepared all the figures and the first draft of the manuscript, MV contributed to the analysis of results and the writing of the manuscript, SS provided and processed the glider data. All authors discussed the results and drew the conclusions.

ACKNOWLEDGMENTS

This work was undertaken and supported through the Southern Ocean Carbon and Climate Observatory (SOCCO) Programme. This work was partially funded by the European Commission 7th framework programme through the GreenSeas Collaborative Project (265294 FP7-ENV-2010). This work was supported by CSIR's Parliamentary Grant funding (SNA2011112600001) and NRF SANAP grants (SNA2011120800004 and SNA14071475720). In addition MV was funded by the SANAP grant TRAIN-SOPP (SNA14072880912, grant no. 93089), University of Cape Town. We thank both the South African Maritime Safety Authority (SAMSA), the South African National Antarctic Programme (SANAP), the captain and officers of the MV SA Agulhas, and the RV SA Agulhas II for the successful completion of the voyages pertaining to this research. We acknowledge the dedication and professionalism shown by the oceanographic engineers of Sea Technology Services. Thanks are extended to the IOP team at the Applied Physics Laboratory, University of Washington and to Dr. Thomas Ryan-Keogh for his assistance in collating the laboratory culture chl-a:C_{phyto} ratios. We would like to thank the editor and our reviewers for their valuable contribution to the improvement of this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmars.2017.00034/full#supplementary-material

REFERENCES

Andersson, A., and Rudehäll, A. (1993). Proportion of plankton biomass in particulate organic carbon in the northern Baltic Sea. *Mar. Ecol. Prog. Ser.* 95, 133–139. doi: 10.3354/meps095133

Antoine, D., André, J., and Morel, A. (1996). Oceanic primary production: 2. Estimation at global scale from satellite (Coastal Zone Color Scanner) chlorophyll. *Global Biogeochem. Cycles* 10, 57–69. doi: 10.1029/95GB02832

Antoine, D., Siegel, D. A., Kostadinov, T., Maritorena, S., Nelson, N. B., Gentili, B., et al. (2011). Variability in optical particle backscattering in contrasting bio-optical oceanic regimes. *Limnol. Oceanogr.* 56, 955–973. doi: 10.4319/lo.2011.56.3.0955

Arrigo, K. R., van Dijken, G. L., and Bushinsky, S. (2008). Primary production in the Southern Ocean, 1997-2006. *J. Geophys. Res. Oceans* 113:C08004. doi: 10.1029/2007JC004551

Arrigo, K. R., Worthen, D. L., Schnell, A., and Lizotte, M. P. (1998). Primary production in Southern Ocean waters. *J. Geophys. Res. Oceans* 103, 587–600.

Aumont, O., and Bopp, L. (2006). Globalizing results from ocean *in situ* iron fertilization studies. *Global Biogeochem. Cycles* 20, GB2017. doi: 10.1029/2005GB002591

Balch, W. M., Bowler, B. C., Drapeau, D. T., Poulton, A. J., and Holligan, P. M. (2010). Biominerals and the vertical flux of particulate organic carbon from the surface ocean. *Geophys. Res. Lett.* 37:L22605. doi: 10.1029/2010GL044640

Behrenfeld, M. J., and Boss, E. (2006). Beam attenuation and chlorophyll concentration as alternative optical indices of phytoplankton biomass. *J. Mar. Res.* 64, 431–451. doi: 10.1357/002224006778189563

Behrenfeld, M. J., and Falkowski, P. G. (1997). Photosynthetic rates derived from satellite-based chlorophyll concentration. *Limnol. Oceanogr.* 42, 1–20. doi: 10.4319/lo.1997.42.1.0001

Behrenfeld, M. J., and Milligan, A. J. (2013). Photophysiological expressions of iron stress in phytoplankton. *Ann. Rev. Mar. Sci.* 5, 217–246. doi: 10.1146/annurev-marine-121211-172356

Behrenfeld, M. J., Boss, E., Siegel, D. A., and Shea, D. M. (2005). Carbon-based ocean productivity and phytoplankton physiology from space. *Global Biogeochem. Cycles* 19, GB1006. doi: 10.1029/2004GB002299

Behrenfeld, M. J., Marañón, E., Siegel, D. A., and Hooker, S. B. (2002). Photoacclimation and nutrient-based model of light-saturated photosynthesis for quantifying oceanic primary production. *Mar. Ecol. Prog. Ser.* 228, 103–117. doi: 10.3354/meps228103

Behrenfeld, M. J., O'Malley, R. T., Boss, E. S., Westberry, T. K., Graff, J. R., Halsey, K. H., et al. (2016). Revaluating ocean warming impacts on global phytoplankton. *Nat. Climate Change*. 6, 323–330. doi: 10.1038/NCLIMATE2838

Bellacicco, M., Volpe, G., Colella, S., Pitarch, J., and Santoleri, R. (2016). Influence of photoacclimation on the phytoplankton seasonal cycle in the Mediterranean Sea as seen by satellite. *Remote Sens. Environ.* 184, 595–604. doi: 10.1016/j.rse.2016.08.004

Blondeau-Patissier, D., Gower, J. F., Dekker, A. G., Phinn, S. R., and Brando, V. E. (2014). A review of ocean color remote sensing methods and statistical techniques for the detection, mapping and analysis of phytoplankton blooms in coastal and open oceans. *Prog. Oceanogr.* 123, 123–144. doi: 10.1016/j.pocean.2013.12.008

Boss, E., and Pegau, W. S. (2001). Relationship of light scattering at an angle in the backward direction to the backscattering coefficient. *Appl. Opt.* 40, 5503–5507. doi: 10.1364/AO.40.005503

Boss, E., Pegau, W. S., Lee, M., Twardowski, M. S., Shybanov, E., Korotaev, G., et al. (2004). Particulate backscattering ratio at LEO 15 and its use to study particles composition and distribution. *J. Geophys. Res. Oceans* 109, C01014. doi: 10.1029/2002JC001514

Boss, E., Swift, D., Taylor, L., Brickley, P., Zaneveld, R., Riser, S., et al. (2008). Observations of pigment and particle distributions in the western North Atlantic from an autonomous float and ocean color satellite. *Limnol. Oceanogr.* 53, 2112–2122. doi: 10.4319/lo.2008.53.5_part_2.2112

Boyd, P. W. (2002). Review environmental factors controlling phytoplankton processes in the Southern Ocean. *J. Phycol.* 861, 844–861. doi: 10.1046/j.1529-8817.2002.t01-1-01203.x

Boyd, P. W., Strzepek, R., Fu, F., and Hutchins, D. (2010). Environmental control of open-ocean phytoplankton groups: now and in the future. *Limnol. Oceanogr.* 55, 1353–1376. doi: 10.4319/lo.2010.55.3.1353

Briggs, N., Perry, M. J., Cetinic, I., Lee, C., D'Asaro, E., Gray, A. M., et al. (2011). High-resolution observations of aggregate flux during a sub-polar North Atlantic spring bloom. *Deep Sea Res. I Oceanogr. Res. Pap.* 58, 1031–1039. doi: 10.1016/j.dsr.2011.07.007

Broecker, W. S., and Peng, T. H. (1982). *Tracers in the Sea.* New York, NY: Eldigio Press; Columbia University.

Buck, K. R., Chavez, F. P., and Campbell, L. (1996). Basin-wide distributions of living carbon components and the inverted trophic pyramid of the central gyre of the North Atlantic Ocean, summer 1993. *Aquat. Microb. Ecol.* 10, 283–298. doi: 10.3354/ame010283

Caron, D. A., Dam, H. G., Kremer, P., Lessard, E. J., Madin, L. P., Malone, T. C., et al. (1995). The contribution of microorganisms to particulate carbon and nitrogen in surface waters of the Sargasso Sea near Bermuda. *Deep Sea Res. I Oceanogr. Res. Pap.* 42, 943–972. doi: 10.1016/0967-0637(95)00027-4

Carr, M. E., Friedrichs, M. A., Schmeltz, M., Aita, M. N., Antoine, D., Arrigo, K. R., et al. (2006). A comparison of global estimates of marine primary production from ocean color. *Deep Sea Res. II Top. Stud. Oceanogr.* 53, 741–770. doi: 10.1016/j.dsr2.2006.01.028

Carranza, M. M., and Gille, S. T. (2015). Southern Ocean wind-driven entrainment enhances statellite chlorophyll-a through the summer. *J. Geophys. Res. Oceans* 120, 304–323. doi: 10.1002/2014JC010203

Cetinić, I., Perry, M. J., Briggs, N. T., Kallin, E., D'Asaro, E. A., and Lee, C. M. (2012). Particlulate organic carbon and inherent optical properties during 2008 North Atlantic Bloom Experiment. *J. Geophys. Res. Oceans* 117:C06028. doi: 10.1029/2011JC007771

Christaki, U., Wambeke, F. V., Lefevre, D., Lagaria, A., Prieur, L., Pujo-Pay, M., et al. (2011). Microbial food webs and metabolic state across oligotrophic waters of the Mediterranean Sea during summer. *Biogeosciences* 8, 1839–1852. doi: 10.5194/bg-8-1839-2011

Cullen, J. J. (1990). On models of growth and photosynthesis in phytoplankton. *Deep Sea Res. A Oceanogr. Res. Pap.* 37, 667–683. doi: 10.1016/0198-0149(90)90097-F

Dall'Olmo, G., Boss, E., Behrenfeld, M. J., Westberry, T. K., Courties, C., Prieur, L., et al. (2011). Inferring phytoplnakton carbon and eco-physiological rates from diel cycles of spectra particulate beam-attenuation coefficient. *Biogeosciences* 8, 3423–3440. doi: 10.5194/bg-8-3423-2011

Dall'Olmo, G., Westberry, T. K., Behrenfeld, M. J., Boss, E., and Slade, W. H. (2009). Significant contribution of large particles to optical backscattering in the open ocean. *Biogeosciences* 6, 947–967. doi: 10.5194/bg-6-947-2009

de Boyer Montégut, C., Madec, G., Fischer, A. S., Lazar, A., and Iudicone, D. (2004). Mixed layer depth over the global ocean: an examination of profile data and a profile-based climatology. *J. Geophys. Res. Oceans* 109, C12003. doi: 10.1029/2004JC002378

Doney, S. C., Lima, I., Moore, J. K., Lindsay, K., Behrenfeld, M. J., Westberry, T. K., et al. (2009). Skill metrics for confronting global upper ocean ecosystem-biogeochemistry models against field and remote sensing data. *J. Mar. Syst.* 76, 95–112. doi: 10.1016/j.jmarsys.2008.05.015

Durand, M. D., Olson, R. J., and Chisholm, S. W. (2001). Phytoplankton population dynamics at the Bermuda Atlantic Time-series station in the Sargasso Sea. *Deep. Res. Part II Top. Stud. Oceanogr.* 48, 1983–2003. doi: 10.1016/S0967-0645(00)00166-1

Eppley, R. W., Chavez, F. P., and Barber, R. T. (1992). Standing stocks of particulate carbon and nitrogen in the equatorial Pacific at 150°W. *J. Geophys. Res. Oceans* 97, 655–661. doi: 10.1029/91JC01386

Falkowski, P. G., Barber, R. T., and Smetacek, V. (1998). Biogeochemical controls and feedbacks on ocean primary production. *Science* 281:200. doi: 10.1126/science.281.5374.200

Fauchereau, N., Tagliabue, A., Bopp, L., and Monteiro, P. M. S. (2011). The response of phytoplankton biomass to transient mixing events in the Southern Ocean. *Geophys. Res. Lett.* 38, L17601. doi: 10.1029/2011GL048498

Field, C. B., Behrenfeld, M. J., Randerson, J. T., and Falkowski, P. (1998). Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* 281, 237–240. doi: 10.1126/science.281.5374.237

Garau, B., Ruiz, S., Zang, G. W., Heslop, E., Kerfoot, J., Pascual, A., et al. (2011). Thermal lag correction on Slocum CTD glider data. *J. Atmos. Ocean. Technol.* 28, 1065–1071. doi: 10.1175/JTECH-D-10-05030.1

Gardner, W. D., Mishonov, A. V., and Richardson, M. J. (2006). Global POC concentrations from *in-situ* and satellite data. *Deep Sea Res. II Top. Stud. Oceanogr.* 53, 718–740. doi: 10.1016/j.dsr2.2006.01.029

Gardner, W. D., Walsh, I. D., and Richardson, M. J. (1993). Biophysical forcing of particle-production and distribution during a spring bloom in the North-Atlantic. *Deep Sea Res. II Top. Stud. Oceanogr.* 40, 171–195. doi: 10.1016/0967-0645(93)90012-C

Garver, S. A., and Siegel, D. A. (1997). Inherent optical property inversion of ocean color spectra and its biogeochemical interpretation: I. Time series from the Sargasso Sea. *J. Geophys. Res. Oceans*. 102, 607–625. doi: 10.1029/96JC03243

Geider, R. J., MacIntyre, H. L., and Kana, T. M. (1997). Dynamic model of phytoplankton growth and acclimation: responses of the balanced growth rate and the chlorophyll a:carbon ratio to light, nutrient-limitation and temperature. *Mar. Ecol. Prog. Ser.* 148, 187–200. doi: 10.3354/meps148187

Graff, J. R., Westberry, T. K., Milligam, A. J., Brown, M. B., Dall'Olmo, G., van Dongen-Vogels, V., et al. (2015). Analytical phytoplankton carbon measurement spanning divers ecosystems. *Deep Sea Res. I* 102, 16–25. doi: 10.1016/j.dsr.2015.04.006

Greene, R. M., Geider, R. J., Kolber, Z., and Falkowski, P. G. (1992). Iron-induced changes in light harvesting and photochemical energy conversion processes in eukaryotic marine algae. *Plant Physiol.* 100, 565–575. doi: 10.1104/pp.100.2.565

Gruber, N., Gloor, M., Mikaloff Fletcher, S. E., Doney, S. C., Dutkiewicz, S., Follows, M. J., et al. (2009). Oceanic sources, sinks, and transport of atmospheric CO_2. *Global Biogeochem. Cycles* 23, GB1005. doi: 10.1029/2008GB003349

Gundersen, K., Orcutt, K. M., Purdie, D. A., Michaels, A. F., and Knap, A. H. (2001). Particulate organic carbon mass distribution at the Bermuda Atlantic Time-series Study (BATS) site. *Deep. Res. Part II Top. Stud. Oceanogr.* 48, 1697–1718. doi: 10.1016/S0967-0645(00)00156-9

Halsey, K. H., and Jones, B. M. (2015). Phytoplankton strategies for photosynthetic energy allocation. *Ann. Rev. Mar. Sci.* 7, 265–297. doi: 10.1146/annurev-marine-010814-015813

Hobson, L., Menzel, D., and Barber, R. (1973). Primary productivity and sizes of pools of organic carbon in the mixed layer of the ocean. *Mar. Biol.* 19, 298–306. doi: 10.1007/BF00348898

Joubert, W. R., Swart, S., Tagliabue, A., Thomalla, S. J., and Monteiro, P. M. S. (2014). The sensitivity of primary productivity to intra-seasonal mixed layer variability in the sub-Antarctic Zone of the Atlantic Ocean. *Biogeosciences Discuss.* 11, 4335–4358. doi: 10.5194/bgd-11-4335-2014

Key, R. M., Kozyr, A., Sabine, C. L., Lee, K., Wanninkhof, R., Bullister, J. L., et al. (2004). A global ocean carbon climatology: results from Global Data Analysis Project (GLODAP). *Global Biogeochem. Cycles*, 18, 1–23. doi: 10.1029/2004GB002247

Kheireddine, M., and Antoine, D. (2014). Diel variability of the beam attenuation and backscattering coefficients in the northwestern Mediterranean Sea (BOUSSOLE site). *J. Geophys. Res. Oceans* 119, 5465–5482. doi: 10.1002/2014jc010007

Klein, P., Hua, B. L., Lapeyre, G., Capet, X., Gentil, S. L., and Sasaki, H. (2008). Upper ocean turbulence from high-resolution 3D simulation. *J. Phys. Oceanogr.* 38, 1748–1763. doi: 10.1175/2007JPO3773.1

Knap, A. H., Michaels, A. F., and Close, A. (1994). *The JGOFS Protocols Intergovernmental Oceanographic Commission*.

Laws, E. A., and Bannister, T. T. (1980). Nutrient limited and light limited growth of Thalassiosira-Fluviatilis in continuous culture with implications for phytoplankton growth in the ocean. *Limnol. Oceanogr.* 25, 457–473. doi: 10.4319/lo.1980.25.3.0457

Le Quéré, C., Rödenbeck, C., Buitenhuis, E. T., Conway, T. J., Langenfelds, R., Gomez, A., et al. (2007). Saturation of the Southern Ocean CO_2 sink due to recent climate change. *Science* 316, 1735–1738. doi: 10.1126/science.1136188

Lenton, A., Tilbrook, B., Law, R. M., Bakker, D., Doney, S. C., Gruber, N., et al. (2013). Sea-air CO_2 fluxes in the Southern Ocean for the period 1990-2009. *Biogeosciences* 10, 4037–4054. doi: 10.5194/bg-10-4037-2013

Lévy, M., Iovino, D., Resplandy, L., Klein, P., Treguier, A.-M., Madec, G., et al. (2012). Large-scale impacts of submesoscale dynamics on phytoplankton: local and remote effects. *Ocean Model.* 43–44, 77–93. doi: 10.1016/j.ocemod.2011.12.003

Lévy, M., Klein, P., and Treguier, A.-M. (2001). Impact of sub-mesosclae physics on production and subduction of phytoplankton in an oligotrophic regime. *J. Mar. Res.* 59, 5355–5565. doi: 10.1357/002224001762842181

Llewellyn, C. A., and Gibb, S. W. (2000). Intra-class variability in the carbon, pigment and biomineral content of prymnesiophytes and diatoms. *Mar. Ecol. Prog. Ser.* 193, 33–44. doi: 10.3354/meps193033

Loisel, H., Bosc, E., Stramski, D., Oubelkheir, K., and Deschamps P.-Y. (2001). Seasonal variability of the backscattering coefficient in the Mediterranean Sea based on satellite SeaWiFS imagery. *Geophys. Res. Lett.* 28, 4203–4206. doi: 10.1029/2001GL013863

Loisel, H., and Morel, A. (1998). Light scattering and chlorophyll concentration in case 1 waters: a reexamination. *Limnol. Oceanogr.* 43, 847–858. doi: 10.4319/lo.1998.43.5.0847

Longhurst, A., Sathyendranath, S., Platt, T., and Caverhill, C. (1995). An estimate of global primary production in the ocean from satellite radiometer data. *J. Plankton Res.* 17, 1245–1271. doi: 10.1093/plankt/17.6.1245

Lü, S., Wang, X., and Han, B. (2009). A field study on the conversion ratio of phytoplankton biomass carbon to chlorophyll-a in Jiaozhou Bay, China. *Chinese J. Oceanol. Limnol.* 27, 793–805. doi: 10.1007/s00343-009-9221-0

MacIntyre, H. L., Kana, T. M., Anning, T., and Geider, R. J. (2002). Photoacclimation of photosynthesis irradiance response curves and photosynthetic pigments in microalgae and cyanobacteria. *J. Phycol.* 38, 17–38. doi: 10.1046/j.1529-8817.2002.00094.x

Mahadevan, A., D'Asaro, E., Lee, C., and Perry, M. J. (2012). Eddy-driven stratification initiates North Atlantic spring phytoplankton blooms. *Science* 337, 54. doi: 10.1126/science.1218740

Maritorena, S., Siegel, D. A., and Peterson, A. R. (2002). Optimization of a semianalytical ocean color model for global-scale applications, *Appl. Opt.* 41, 2705–2714. doi: 10.1364/AO.41.002705

Martinez-Vicente, V., Dall'Olmo, G., Tarran, G., Boss, E., and Sathyendranath, S. (2013). Optical backscattering is correlated with phytoplankton carbon across the Atlantic Ocean. *Geophys. Res. Lett.* 40, 1154–1158. doi: 10.1002/grl.50252

McCartney, M. S. (1977). "Subantarctic mode water," in *A voyage of Discovery: George Deacon 70th Anniversary Volume*, *Supplement to Deep-Sea Research*, ed M. V. Angel (New York, NY; London: Pergamon Press), 103–119.

McKiver, W. J., Vichi, M., Lovato, T., Storto, A., and Masina, S. (2015). Impact of increased grid resolution on global marine biogeochemistry. *J. Mar. Syst.* 147, 153–168. doi: 10.1016/j.jmarsys.2014.10.003

McNeil, B. I., Metzl, N., Key, R. M., Matear, R. J., and Corbiere, A. (2007). An empirical estimate of the Southern Ocean air-sea CO_2 flux. *Global Biogeochem. Cycles* 21, GB3011. doi: 10.1029/2007GB002991

McNeil, B. I., Tilbrook, B., and Matear, R. J. (2001). Accumulation and uptake of anthropogenic CO2 in the Southern Ocean, South of Australia between 1968 and 1996. *J. Geophys. Res. Oceans* 106, 31431–31445. doi: 10.1029/2000JC000331

Metzl, N., Tilbrook, B., and Poisson, A. (1999). The annual fCO_2 cycle and the air-sea CO_2 flux in the sub-Antarctic Ocean. *Tellus B* 51, 849–861. doi: 10.1034/j.1600-0889.1999.t01-3-00008.x

Mignot, A., Claustre, H., Uitz, J., Poteau, A., D'Ortenzio, F., and Xing, X. (2014). Understanding the seasonal dynamics of phytoplankton biomass and the deep chlorophyll maximum in oligotrophic environments: a Bio-Argo float investigation. *Global Biogeochem. Cycles* 28, 856–876. doi: 10.1002/2013GB004781

Montagnes, D. J., Berges, J. A., Harrison, P. J., and Taylor, F. (1994). Estimating carbon, nitrogen, protein, and chlorophyll a from volume in marine phytoplankton. *Limnol. Oceanogr.* 39, 1044–1060. doi: 10.4319/lo.1994.39.5.1044

Moore, J. K., and Abbott, M. R. (2000). Phytoplankton chlorophyll distributions and primary production in the Southern Ocean. *J. Geophys. Res. Ocean* 105, 28709–28722. doi: 10.1029/1999JC000043

Moore, J. K., Doney, S. C., Kleypas, J. A., Glover, D. M., and Fung, I. Y. (2002). An intermediate complexity marine ecosystem model for the global domain. *Deep. Res. Part II Top. Stud. Oceanogr.* 49, 403–462. doi: 10.1016/S0967-0645(01)00108-4

Oubelkheir, K. (2001). *Biogeochemical Characterization of Various Oceanic Provinces through Optical Indicators Over Various Space and Time Scales*. Ph.D. thesis, Univ. de la Me'diter./CNRS, Marseille.

Oubelkheir, K., Claustre, H., Sciandra, A., and Babin, M. (2005). Bio-optical and biogeochemical properties of different trophic regimes in oceanic waters. *Limnol. Oceanogr.* 50, 1795–1809. doi: 10.4319/lo.2005.50.6.1795

Parsons, T. R., Maita, Y., and Lalli, C. M. (1984). *A Manual of Chemical and Biological Methods for Seawater Analysis*. Oxford: Pergamon Press.

Quéguiner, B. (2013). Iron frertilization and the structure of planktonic communities in high nutrients regions of the Southern Ocean. *Deep Sea Res. II.* 90, 43–54. doi: 10.1016/j.dsr2.2012.07.024

Racault, M.-F., Le Quéré, C., Buitenhuis, E., Sathyendranath, S., and Platt, T. (2012). Phytoplankton phenology in the global ocean. *Ecol. Indic.* 14, 152–163. doi: 10.1016/j.ecolind.2011.07.010

Rudnick, D. L., Davis, R. E., Eriksen, C. C., Fratantoni, D. M., and Perry, M. J. (2004). Underwater gliders for ocean research, *Mar. Technol. Soc. J.* 38, 73–84, doi: 10.4031/002533204787522703

Sackmann, B. S., Perry, M. J., and Eriksen, C. C. (2008). Seaglider observations of variability in daytime fluorescence quenching of chlorophyll-a in Northeastern Pacific coastal waters. *Biogeosciences Discuss.* 5, 2839–2865. doi: 10.5194/bgd-5-2839-2008

Sathyendranath, S., Stuart, V., Nair, A., Oka, K., Nakane, T., Bouman, H., et al. (2009). Carbon-to-chlorophyll ratio and growth rate of phytoplankton in the sea. *Mar. Ecol. Prog. Ser.* 383, 73–84. doi: 10.3354/meps07998

Schlitzer, R. (2002). Carbon export fluxes in the Southern Ocean: results from inverse modeling and comparison with satellite-based estimates. *Deep Sea Res. Part II Top. Stud. Oceanogr.* 49, 1623–1644. doi: 10.1016/S0967-0645(02)00004-8

Siegel, D. A., Maritorena, S., and Nelson, N. B. (2002). Global distribution and dynamics of colored dissolved and detrital organic materials. *J. Geophys. Res. Oceans* 107, 3228. doi: 10.1029/2001JC000965

Socal, G., Nöthig, E. M., Bianchi, F., Boldrin, A., Mathot, S., and Rabitti, S. (1997). Phytoplankton and particulate matter at the Weddell/Scotia confluence (47°W) in summer 1989, as a final step of a temporal succession (EPOS project). *Polar Biol.* 18, 1–9. doi: 10.1007/s003000050152

Steinacher, M., Joos, F., Frölicher, T. L., Bopp, L., Cadule, P., Doney, S. C., et al. (2009). Projected 21st century decrease in marine productivity: a multi-model analysis. *Biogeosciences Discuss.* 6, 7933–7981. doi: 10.5194/bgd-6-7933-2009

Stramski, D., and Morel, A. (1990). Optical properties of photosynthetic picoplankton in different physiological states as affected by growth irradiance. *Deep Sea Res. A Oceanogr. Res. Pap.* 37, 245–266. doi: 10.1016/0198-0149(90)90126-G

Stramski, D., and Reynolds, R. A. (1993). Diel variations in the optical properties of a marine diatom. *Limnol. Oceanogr.* 38, 1347–1364. doi: 10.1016/j.fsi.2011.01.004

Stramski, D. R., Reynolds, A., Kahru, M., and Mitchell, B. G. (1999). Estimation of particulate organic carbon in the ocean from satellite remote sensing. *Science* 285, 239–242. doi: 10.1126/science.285.5425.239

Stramski, D., Reynolds, R. A., Babin, M., Kaczmarek, S., Lewis, M. R., Rottgers, R., et al. (2008). Relationships between the surface concentration of particulate organic carbon and optical properties in the eastern South Pacific and eastern Atlantic Oceans. *Biogeosciences* 5, 171–201. doi: 10.5194/bg-5-171-2008

Sunda, W. G., and Huntsman, S. A. (1995). Iron uptake and growth limitation in oceanic and coastal phytoplankton. *Mar. Chem.* 50, 189–206. doi: 10.1016/0304-4203(95)00035-p

Sunda, W. G., and Huntsman, S. A. (1997). Interrelated influence of iron, light and cell size on marine phytoplankton growth. *Nature* 390, 389–392. doi: 10.1038/37093

Swart, S., Chang, N., Fauchereau, N., Joubert, W., Lucas, M., Mtshali, T., et al. (2012). Southern Ocean seasonal cycle experiment 2012: seasonal scale climate and carbon links. *S. Afr. J. Sci.* 108, 1–3. doi: 10.4102/sajs.v108i3/4.1089

Swart, S., Thomalla, S. J., and Monteiro, P. M. S. (2015). The seasonal cycle of mixed layer dynamics and phytoplankton biomass in the Sub-Antarctic Zone: a high-resolution glider experiment. *J. Mar. Syst.* 147, 103–115. doi: 10.1016/j.jmarsys.2014.06.002

Taylor, A. H., Geider, R. J., and Gilbert, F. J. H. (1997). Seasonal and latitudinal dependencies of phytoplankton carbon-to-chlorophyll a ratios: results of a modelling study. *Mar. Ecol. Prog. Ser.* 152, 51–66. doi: 10.3354/meps152051

Thomalla, S. J., Fauchereau, N., Swart, S., and Monteiro, P. M. S. (2011). Regional scale characteristics of the seasonal cycle of chlorophyll in the Southern Ocean. *Biogeosciences* 8, 4763–4804. doi: 10.5194/bgd-8-4763-2011

Thomalla, S. J., Racault, M.-F., Swart, S., and Monteiro, P. M. S. (2015). High-resolution view of the spring bloom initiation and net community production in the SubAntarctic Southern Ocean using glider data. *ICES J. Mar. Sci.* 72, 1999–2020. doi: 10.1093/icesjms/fsv105

Trull, T. W., Bray, S. G., Manganini, S. J., Honjo, S., and François, R. (2001). Moored sediment trap measurements of carbon export in the Subantarctic and Polar Frontal zones of the Southern Ocean, south of Australia. *J. Geophys. Res. Oceans* 106, 31489–31509. doi: 10.1029/2000JC000308

Vaillancourt, R. D., Brown, C. W., Guillard, R. R. L., and Balch, W. M. (2003). Light backscattering properties of marine phytoplankton: relationships to cell size, chemical composition and taxonomy. *J. Plankton Res.* 26, 191–212. doi: 10.1093/plankt/fbh012

Vichi, M., and Masina, S. (2009). Skill assessment of the PELAGOS global ocean biogeochemistry model over the period 1980–2000. *Biogeosciences Discuss.* 6, 2333–2535. doi: 10.5194/bgd-6-3511-2009

Vichi, M., Lovato, T., Lazzari, P., Cossarini, G., Gutierrez Mlot, E., Mattia, G., et al. (2015). *The Biogeochemical Flux Model (BFM): Equation Description and User Manual*. BFM version 5.1. Release 1.1, BFM Report Series 1, Bologna, Available online at: http://bfm-community.eu

Vichi, M., Pinardi, N., and Masina, S. (2007). A generalized model of pelagic biogeochemistry for the global ocean ecosystem. Part I: theory. *J. Mar. Syst.* 64, 89–109. doi: 10.1016/j.jmarsys.2006.03.006

Volk, T., and Hoffert, M. I. (1985). "Oceanic carbon pumps: analysis of relative strengths and efficiencies in ocean-driven atmospheric CO_2," in *The Carbon Cycle and Atmospheric CO_2: Natural Variations Archean to Present, Vol. 32*, eds E. T. Sunquist and W. S. Broecker (Washington, DC: American Geophysical Union, Geophysical Monography), 99–110.

Wang, X., Matear, R. J., and Trull, T. W. (2001). Modeling seasonal phosphate export and resupply in the Subantarctic and Polar Frontal zones in the Australian sector of the Southern Ocean. *J. Geophys. Res. Oceans* 106, 31525–31541. doi: 10.1029/2000JC000645

Westberry, T. K., Behrenfeld, M. J., Milligan, A. J., and Doney, S. C. (2013). Retrospective satellite ocean color analysis of purposeful and natural ocean iron fertilization. *Deep. Res. Part I Oceanogr. Res. Pap.* 73, 1–16. doi: 10.1016/j.dsr.2012.11.010

Westberry, T. K., Schultz, P., Behrenfeld, M. J., Dunne, J. P., Hiscock, M. R., Maritorena, S., et al. (2016). Annual cycles of phytoplankton biomass in the subarctic Atlantic and Pacific Ocean. *Global Biogeochem. Cycles* 30, 175–190. doi: 10.1002/2015GB005276

Westberry, T., Behrenfeld, M. J., Siegel, D. A., and Boss, E. (2008). Carbon-based primary productivity modeling with vertically resolved photoacclimation. *Global Biogeochem. Cycles* 22, GB2024. doi: 10.1029/2007GB003078

Winn, C. D., Campbell, L., Christian, J. R., Letelier, R. M., Hebel, D. V., Dore, J. E., et al. (1995). Seasonal variability in the phytoplankton community of the North Pacific Subtropical Gyre. *Global Biogeochem. Cycles*, 9, 605–620. doi: 10.1029/95GB02149

Xing, X., Claustre, H., and Blain, S. (2012). Quenching correction for *in vivo* chlorophyll fluorescence acquired by autonomous platforms: a case study with instrumented elephant seals in the Kerguelen region (Southern Ocean). *Limnol. Oceangr.* 10, 483–495. doi: 10.4319/lom.2012.10.483

Zhang, X., and Hu, L. (2009). Scattering by pure seawater at high salinity. *Opt. Express* 17, 12685–12691. doi: 10.1364/OE.17.012685

Zhang, X., Hu, L., and He, M. X. (2009). Scattering by pure seawater: effect of salinity. *Opt. Express* 17, 5698–5710. doi: 10.1364/OE.17.005698

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Thomalla, Ogunkoya, Vichi and Swart. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Marine Science

REVIEW
published: 21 February 2017
doi: 10.3389/fmars.2017.00041

A Consumer's Guide to Satellite Remote Sensing of Multiple Phytoplankton Groups in the Global Ocean

Colleen B. Mouw[1], Nick J. Hardman-Mountford[2], Séverine Alvain[3], Astrid Bracher[4,5], Robert J. W. Brewin[6,7], Annick Bricaud[8], Aurea M. Ciotti[9], Emmanuel Devred[10], Amane Fujiwara[11], Takafumi Hirata[12,13], Toru Hirawake[14], Tihomir S. Kostadinov[15], Shovonlal Roy[16] and Julia Uitz[8]*

[1] Graduate School of Oceanography, University of Rhode Island, Narragansett, RI, USA, [2] CSIRO Oceans and Atmosphere, Perth, WA, Australia, [3] Laboratoire d'Océanologie et de Géosciences - UMR 8187 LOG, Centre National de la Recherche Scientifique, Université Lille Nord de France - ULCO, Wimereux, France, [4] Alfred-Wegener-Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany, [5] Institute of Environmental Physics, University Bremen, Bremen, Germany, [6] Plymouth Marine Laboratory, Plymouth, UK, [7] National Centre for Earth Observation, Plymouth Marine Laboratory, Plymouth, UK, [8] Laboratoire d'Océanographie de Villefranche, Observatoire Océanologique de Villefranche, Centre National de la Recherche Scientifique, Sorbonne Universités, UPMC-Université Paris-VI, Villefranche-sur-Mer, France, [9] Center for Marine Biology, University of São Paulo, São Paulo, Brazil, [10] Bedford Institute of Oceanography, Fisheries and Oceans Canada, Halifax, NS, Canada, [11] Institute of Arctic Climate and Environment Research, Japan Agency for Marine-Earth Science and Technology, Yokosuka, Japan, [12] Faculty of Environmental Earth Science, Hokkaido University, Sapporo, Japan, [13] CREST, Japan Science Technology Agency, Tokyo, Japan, [14] Faculty of Fisheries Sciences, Hokkaido University, Hakodate, Japan, [15] Department of Geography and the Environment, University of Richmond, Richmond, VA, USA, [16] Department of Geography and Environmental Science and School of Agriculture Policy and Development, University of Reading, Reading, UK

OPEN ACCESS

Edited by:
Chris Bowler,
École Normale Supérieure, France

Reviewed by:
Michael J. Behrenfeld,
Oregon State University, USA
Daniele Iudicone,
Stazione Zoologica Anton Dohrn, Italy

***Correspondence:**
Colleen B. Mouw
cmouw@uri.edu

Specialty section:
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

Received: *18 October 2016*
Accepted: *03 February 2017*
Published: *21 February 2017*

Citation:
Mouw CB, Hardman-Mountford NJ, Alvain S, Bracher A, Brewin RJW, Bricaud A, Ciotti AM, Devred E, Fujiwara A, Hirata T, Hirawake T, Kostadinov TS, Roy S and Uitz J (2017) A Consumer's Guide to Satellite Remote Sensing of Multiple Phytoplankton Groups in the Global Ocean. Front. Mar. Sci. 4:41. doi: 10.3389/fmars.2017.00041

Phytoplankton are composed of diverse taxonomical groups, which are manifested as distinct morphology, size, and pigment composition. These characteristics, modulated by their physiological state, impact their light absorption and scattering, allowing them to be detected with ocean color satellite radiometry. There is a growing volume of literature describing satellite algorithms to retrieve information on phytoplankton composition in the ocean. This synthesis provides a review of current methods and a simplified comparison of approaches. The aim is to provide an easily comprehensible resource for non-algorithm developers, who desire to use these products, thereby raising the level of awareness and use of these products and reducing the boundary of expert knowledge needed to make a pragmatic selection of output products with confidence. The satellite input and output products, their associated validation metrics, as well as assumptions, strengths, and limitations of the various algorithm types are described, providing a framework for algorithm organization to assist users and inspire new aspects of algorithm development capable of exploiting the higher spectral, spatial and temporal resolutions from the next generation of ocean color satellites.

Keywords: remote sensing, ocean color, optics, phytoplankton functional types, phytoplankton size classes, particle size distribution, phytoplankton taxonomic composition, bio-optical algorithms

INTRODUCTION

The determination of phytoplankton community structure using satellite remote sensing has evolved from an aspiration to a highly active area of research, with numerous published approaches available over the past decade. Prior work had focused on the discrimination of dominant single phytoplankton groups such as coccolithophores, *Trichodesmium* spp., diatoms, and other harmful species such as *Karenia brevis*, *Karenia mikimotoi*, *Nodularia*, and *Microcystis* (IOCCG, 2014; see chapter 3 and references therein). A variety of approaches have emerged that attempt to discriminate "phytoplankton functional types" (PFT), which include algorithms that retrieve phytoplankton size classes (PSC), phytoplankton taxonomic composition (PTC), or particle size distribution (PSD). In this way, a PFT is an aggregation of phytoplankton, where irrespective of their phylogeny, they share similar biogeochemical or ecological roles. This broad definition lacks specificity, with no universal interpretation (Reynolds et al., 2002). Here PSC, PTC, and PSD serve as a further refinement of PFTs, where the choice of the considered functional type depends on the question at hand. Surveying the existing algorithms, with their varying inputs and outputs, can be overwhelming for non-experts wishing to use the data products from such approaches and determine which algorithm output may be most applicable to their problem at hand. This guide serves as a synthesis of the existing methods with clear articulation of the underlying approach, satellite input and output products, assumptions, strengths, limitations, and validation metrics.

There are several recent reviews of research accomplishments of phytoplankton composition retrieval from satellite (Nair et al., 2008; Brewin R. J. et al., 2011; De Moraes Rudorff and Kampel, 2012; IOCCG, 2014). Nair et al. (2008) provide a review of single-species and multiple type retrievals, while De Moraes Rudorff and Kampel (2012) review various algorithm approaches (empirical, semi-analytical, analytic). Brewin R. J. et al. (2011) directly compare the performance of PFT and PSC algorithms. IOCCG (2014) provides a comprehensive report of PFT accomplishments to date, giving users detailed information on the various satellite PFT techniques. Yet, since the time of these reviews the literature has grown quickly. Building on the IOCCG report, the goal here is to provide a simple guide to current PFT techniques that is attractive to a broad audience of marine scientists. We provide a direct comparison of the assumptions, strengths, limitations, required satellite input and output products and performance metrics for the different approaches. The goal of this guide is to provide such a comparison in accessible form to reduce the barrier of expert knowledge needed for users to make a sound and confident selection of an algorithm or group of algorithms. To address a similar requirement for primary productivity models, Behrenfeld and Falkowski (1997) produced a "consumer's guide to primary productivity models"; this contribution seeks to address a similar need for the users of PFT satellite products. Given phytoplankton form the base of the aquatic food web and their composition impacts the structure, function, and sustainability of the whole food web, we anticipate a broad user community, including: numerical model developers, environmental, and fisheries management entities, those seeking to understand climate-related changes in marine ecosystems and the carbon cycle, and members of the satellite remote sensing community that are non-PFT algorithm developers. Observationalists wanting to provide information to the broadest community are often looking for guidance on what variables or types of measurements would be of the highest value, in addition to identifying tools to put their observations into a larger context. Satellite remote sensing adds valuable synoptic observations on spatio-temporal scales impossible to sample *in situ*. In addition, by summarizing the parameters utilized in algorithm development, as well as satellite inputs and outputs, we aim to motivate identification of non-exploited parameter space and new algorithm development for extended PFT capability into the future.

Here, we focus on global open ocean methods solely dependent on inputs from ocean color radiance or its derived products. Thus, we exclude ecologically based methods that require additional physical and spatio-temporal information (e.g., Raitsos et al., 2008; Palacz et al., 2013). We utilize all of the algorithms that Kostadinov et al. (2017) directly compare plus three additional algorithms (Hirata et al., 2008; Devred et al., 2011; Li et al., 2013).

Unlike the "consumer's guide to primary productivity models" (Behrenfeld and Falkowski, 1997), where net primary production was the single common output between all compared models, satellite PFT algorithms have a variety of phytoplankton classes, units, and satellite product outputs. This presents an additional layer of challenge, precluding direct comparison of algorithm performance and explicit "how to" instructions as found in Behrenfeld and Falkowski (1997). Instead, other metrics, such as phenological cycles, are being explored as a way to inter-compare PFT algorithms (Kostadinov et al., 2017). It is not our purpose here to inter-compare algorithm performance, rather we seek to provide users with a simplified "go to" reference to understand existing algorithm types, their associated strengths and limitations, input requirements and output products, to aid in selecting the satellite PFT model that may best fit their application.

ALGORITHM OVERVIEW

Here, we focus on the four algorithm types that derive PFTs that are classified according to their theoretical basis, and include abundance-, radiance-, absorption-, and scattering-based approaches (**Figure 1**). The underlying assumptions and basic constructs for each of these algorithm types are described. We begin with the satellite inputs, followed by the outputs, then describe how they were derived (algorithm basis) and how successful the algorithm has been shown so far at retrieving the desired products (validation). A summary of notation can be found in **Table 1**.

Understanding Satellite Data Inputs: Ocean Color Radiometry

A satellite ocean color radiometer measures light (radiance) at the top of the atmosphere. On the global scale, the atmosphere

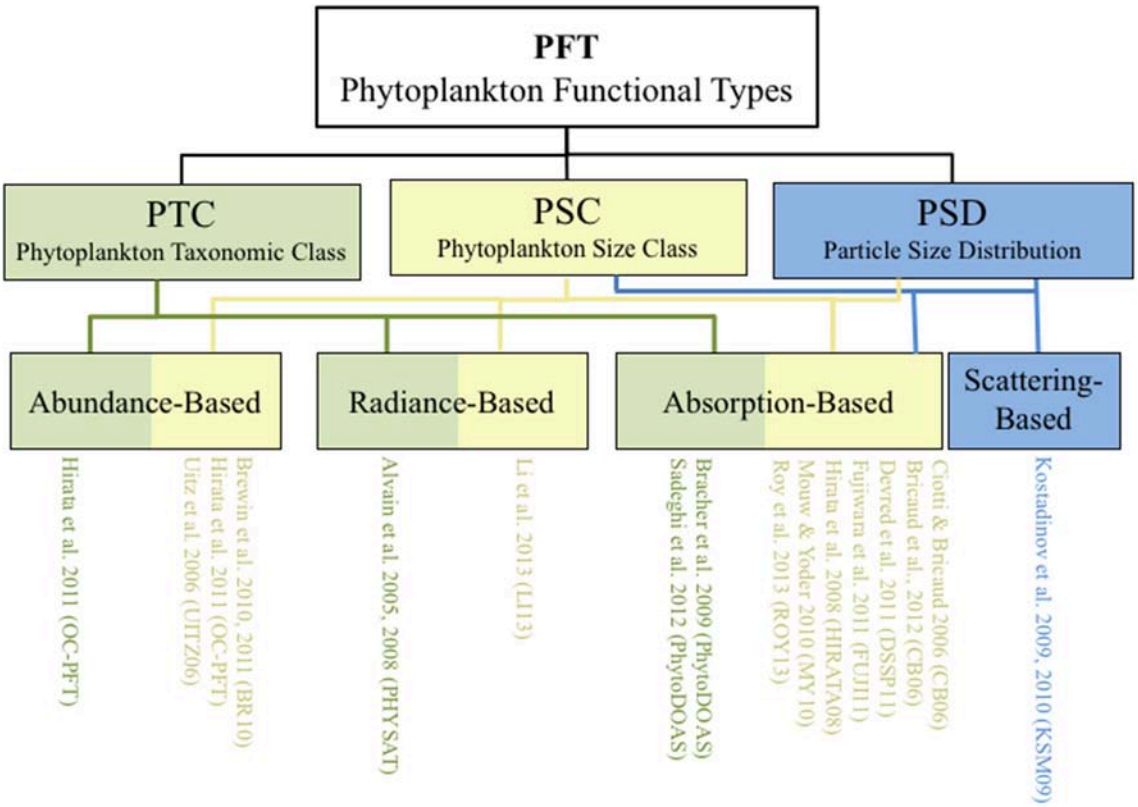


FIGURE 1 | Schematic of various phytoplankton functional type (PFT) algorithms grouped according to their output classification (PTC, PSC, or PSD) and algorithm development types (abundance-, radiance-, absorption-, and scattering-based). Color indicates the output classification of phytoplankton taxonomic class (PTC, green), phytoplankton size class (PSC, yellow) or particle size distribution (PSD, blue).

alone typically accounts for >90% of this signal (Mobley, 1994). After atmospheric correction, the primary measured variable is spectral remote sensing reflectance [$R_{rs}(\lambda)$] or normalized water-leaving radiance [$nLw(\lambda)$]. [Note that these variables are related via $R_{rs}(\lambda) = nLw(\lambda)/F_0(\lambda)$, where $F_0(\lambda)$ is the extraterrestrial solar irradiance centered at wavelength λ (Thuillier et al., 2003)]. In open ocean waters, the threshold of uncertainty acceptance for $R_{rs}(\lambda)$ is 5% (Bailey and Werdell, 2006). All other ocean color variables are estimated from $R_{rs}(\lambda)$ (**Figure 2**). This means that the inherent optical properties (IOPs, i.e., absorption and scattering/backscattering), which are independent from the ambient light field, as well as, biogeochemical variables such as chlorophyll-a concentration, [Chl], are estimated from $R_{rs}(\lambda)$, not measured directly from space. Approximate relationships between $R_{rs}(\lambda)$ and IOPs were presented by Gordon et al. (1988) so that:

$$R_{rs}(\lambda) = \Re \frac{f(\lambda)}{Q(\lambda)} \frac{b_b(\lambda)}{a(\lambda) + b_b(\lambda)} \quad (1)$$

where, $a(\lambda)$ is spectral total absorption coefficient and $b_b(\lambda)$ is spectral total backscattering coefficient, $\Re$ is a factor that accounts for reflection and refraction at the air-water interface, and f/Q accounts for the bidirectional nature of reflectance (Morel et al., 2002). The IOPs absorption and backscattering are functions of biological/biogeochemical variables. Phytoplankton abundance, composition and physiological status impact [Chl], PSD, light absorption, and backscattering, and thus $R_{rs}(\lambda)$. The algorithms that utilize absorption and backscattering satellite inputs obtain these IOP parameters from a variety of different semi-analytical inversion algorithms that are all fundamentally derived from the basic construct of Equation (1; Werdell et al., 2013; **Figure 2**). In contrary, the Phytoplankton Differential Optical Absorption Spectroscopy (PhytoDOAS) algorithm uses top of atmosphere satellite reflectance directly as input, to fit (and separate) simultaneously all absorbers in the atmosphere and ocean—accounting for atmospheric affects within the algorithm (Bracher et al., 2009; Sadeghi et al., 2012a).

Abundance-based algorithms use [Chl] as a satellite input (**Figure 2**). To date, all published abundance-based models utilize [Chl] derived by an empirical approach (O'Reilly et al., 1998),

$$log_{10}\left[Chl\right] = a_0 + \sum_{i=1}^{4} a_i \, log_{10}\left(\frac{R_{rs}(\lambda_{blue})}{R_{rs}(\lambda_{green})}\right)^i \quad (2)$$

where, a_0–a_4 are sensor-specific coefficients and $R_{rs}(\lambda_{blue})$ is the greatest of several input $R_{rs}(\lambda)$ values. However, within the constructs of the PFT algorithms, there is no reason why semi-analytically determined [Chl] could not

TABLE 1 | Summary of notation (units in parentheses, where applicable).

Optical parameters	
nLw(λ)	Normalized spectral water leaving radiance (mW cm^{-2} μm^{-1} sr^{-1})
$R_{rs}(\lambda)$	Spectral remote sensing reflectance (sr^{-1})
a(λ)	Total absorption (m^{-1})
a_{ph}	Phytoplankton absorption (m^{-1})
a_{cdm}	Absorption of colored dissolved and detrital material (m^{-1})
$b_b(\lambda)$	Total backscattering (m^{-1})
b_{bp}	Particulate backscattering coefficient (m^{-1})
η	Spectral slope of the particulate backscattering coefficient (unitless)
S	Spectral slope of colored dissolved and detrital material absorption (nm^{-1})
CDOM	Colored dissolved organic matter
NAP	Non-algal particles
Pigments	
HPLC	High precision liquid chromatography
DP	Diagnostic pigments
[Chl]	Chlorophyll-*a* concentration (mg m^{-3})
Phytoplankton	
PFT	Phytoplankton functional type
PSC	Phytoplankton size classes
PTC	Phytoplankton taxonomic composition
PSD	Particle size distribution
Micro	Microphytoplankton (>20 μm in diameter)
Nano	Nanophytoplankton (2–20 μm)
Pico	Picophytoplankton (0.2–2 μm)
Hapto	Haptophytes
Cocco	Coccolithophores
Dino	Dinoflagellates
Cyano	Cyanobacteria
Pro	*Prochlorococcus*
Syn	*Synechococcus*
Phaeo	*Phaeocystis*
S_f	Size parameter representing fractional picoplankton
S_{fm}	$1 - S_f$

be used in place of empirically determined [Chl]. The sensor-specific coefficients and bands are available at: http://oceancolor.gsfc.nasa.gov/cms/atbd/chlor_a. The level of acceptable uncertainty for [Chl] is 35% (Bailey and Werdell, 2006).

Within the portion of the satellite $R_{rs}(\lambda)$ signal that is attributed to phytoplankton (absorption by pigments and scattering by cellular material), pigment abundance is primarily responsible for first order magnitude variability in $R_{rs}(\lambda)$, while spectral shape differences associated with diversity in the taxonomic composition are secondary (Ciotti et al., 1999). Therefore, it is important to consider the overall phytoplankton contribution to total absorption and scattering budgets. Mouw et al. (2012) quantified this by looking at model output over the range of optical variability encountered in the global ocean considering scenarios where phytoplankton size did and did not vary. They find the magnitude of the [Chl] contribution to $R_{rs}(443)$ (443 nm is the wavelength where greatest phytoplankton absorption occurs) is much greater than the contribution of phytoplankton taxonomic composition to $R_{rs}(443)$ variability (see their Figures 6–8). This is due to the fact that chlorophyll-*a*, a pigment ubiquitous to all phytoplankton, has maximum absorption at 443 nm. PFT algorithms that exploit these second order characteristics, after accounting for the presence of colored dissolved organic matter (CDOM) and non-algal particles (NAP), are therefore subject to limitations due to relatively low signal-to-noise ratio of the residuals, that is, they operate near the limits of what is retrievable by the current state-of-the-art (e.g., Evers-King et al., 2014). Conversely, PFT algorithms that use the dominant abundance signal, such as [Chl], phytoplankton absorption, or particulate backscatter, are less impacted but have to face other limitations such as uncertainty in relationships between these properties and phytoplankton grouping.

Understanding Satellite PFT Outputs: PSC, PTC, and PSD

Here, we seek to summarize and simplify the satellite phytoplankton functional type algorithm products or outputs. The PSC output is most commonly grouped as pico- (0.2–2 μm), nano- (2–20 μm), and/or microplankton (>20 μm) following the size classification scheme proposed by Sieburth et al. (1978). However, a few models allow for multicomponent size classes not constrained by the traditional size groupings (Roy et al., 2013; Brewin et al., 2014b). The PSD satellite output (Kostadinov et al., 2009, 2010; Roy et al., 2013) can conform to the Sieburth et al. (1978) size classification. The PTC algorithms have a variety of outputs, dictated largely by the resolution of information available from *in situ* calibration and/or validation datasets. The PHYSAT approach (Alvain et al., 2005, 2008; Ben Mustapha et al., 2014) retrieves nanoeukaryotes, haptophytes (a major component of the nano-flagellates), *Prochlorococcus*, *Synechococcus*-like cyanobacteria, diatoms, coccolithophores, and *Phaeocystis*-like phytoplankton. The Hirata et al. (2011) approach retrieves pico-eukaryotes, prymnesiophytes (synonymous with haptophytes), diatoms, prokaryotes, green algae (chlorophytes), dinoflagellates, and *Prochlorococcus* sp., in addition to the main pico, nano, and micro size classes. The PhytoDOAS algorithm (Bracher et al., 2009; Sadeghi et al., 2012a) retrieves cyanobacteria, diatoms, coccolithophores, and dinoflagellates. We group similar classes together for clarity and simplicity. For example, haptophytes retrieved by Alvain et al. (2005, 2008) and Sadeghi et al. (2012a) are grouped with prymnesiophytes retrieved by Hirata et al. (2011). *Prochlorococcus* and *Synechococcus*, along with the broader prokaryotes class obtained by Hirata et al. (2011), are grouped as cyanobacteria (**Table 2**). Algorithm abbreviations follow those established by the algorithm's author(s), are consistent with those in Kostadinov et al. (2017), and are noted in **Figure 1** and **Table 2**.

These PFT output products are similar but are not identical and are defined by distinct units. These include dominance, [Chl] for each group (mg m^{-3}), fractional [Chl] (%), fractional

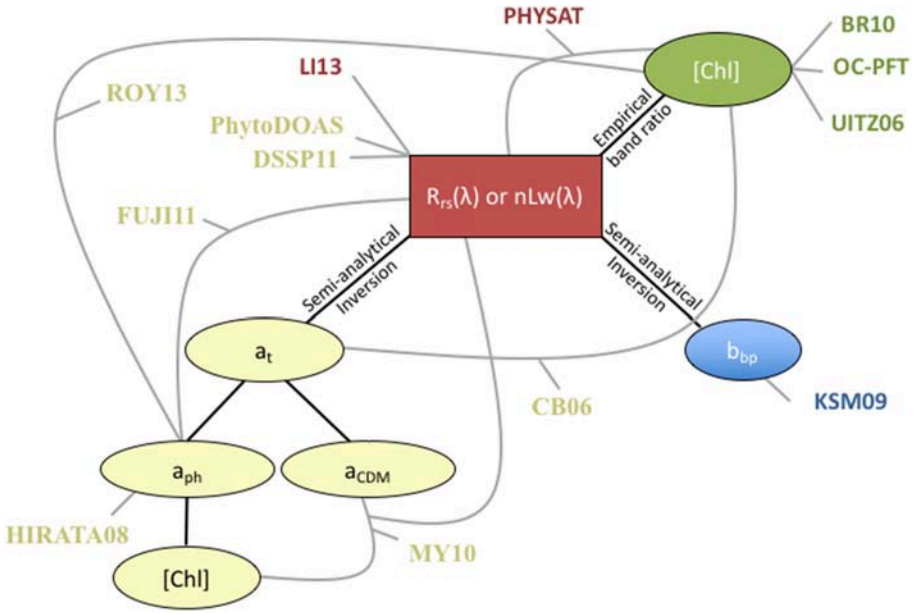


FIGURE 2 | Schematic of satellite product inputs utilized in each PFT algorithm. The red box indicates the $R_{rS}(\lambda)$ measured by a satellite radiometer. Ovals are derived satellite products and their connection to $R_{rS}(\lambda)$ is indicated as black lines. Gray lines indicate the connection of satellite input products used in the various PFT algorithms. The color of the algorithm abbreviation text indicates the algorithm type: abundance (green), radiance (red), absorption (yellow) and scattering (blue). Algorithm abbreviations are as in **Figure 1** and **Tables 2**, **3**.

biovolume (%), absorption (m^{-1}) of each group, and a continuous size parameter varying from 0 to 1 (see Equation 1 and **Table 3**). We also simplify output with regards to units. All phytoplankton groups or size classes, regardless of units, are grouped together in **Table 2** and **Figure 3**, which provide an overview of all algorithms. While users will most certainly require unit information, the overview table allows easy identification of the citations for the outputs of interest. For greater depth of information regarding units, a full list of output products, their validation source, and validation metrics are provided in **Table 3**.

An important consideration is the aspect of phytoplankton group dominance. Alvain et al. (2005, 2008) and Hirata et al. (2008) retrieve the dominant group for a given satellite image pixel. Alvain et al. (2005) define dominance as situations in which a given phytoplankton group is the major contributor to the radiance anomaly. This contribution is retrieved as dominant when the ratio (biomarker pigment concentration/[Chl]) value is at least equal to 50% of the value that will be observed if the phytoplankton group was alone in the sample. This approach allows an empirical relationship between radiances anomalies and *in situ* information. For this reason, PHYSAT interpretation needs to be carefully considered in terms of *in situ* data used to give a name to the remotely sensed signal. Alvain et al. (2005) classify daily images and compile monthly maps of the most frequent dominant phytoplankton group. The group present in more than half of the daily images is assigned as dominant in the monthly compilation. When no group remains dominant over the whole month, pixels are labeled as unidentified. Hirata et al. (2008) determine PSCs from diagnostic pigments and relate them to phytoplankton absorption at 443 nm [$a_{ph}(443)$] to retrieve PSCs from satellite imagery. In the development stage of relating diagnostic pigments to $a_{ph}(443)$ *in situ*, a PSC is defined as dominant if the marker pigment to diagnostic pigment ratio is >45%. However, in applying the approach to $a_{ph}(443)$ imagery, PSCs are determined based on threshold ranges of $a_{ph}(443)$, as such for a given pixel, only a single dominant type output is classified, regardless of temporal resolution of the satellite imagery. These are considerations users need to be aware of and can impact their interpretation and use. Further, when comparing satellite algorithms with biogeochemical model outputs, dominance (highest percentage of group) will vary whether one considers dominance of [Chl], $a_{ph}(\lambda)$, $b_{bp}(\lambda)$, or carbon—requiring care to ensure comparisons are done on the same terms.

Algorithm Basis

Abundance-based algorithms are based on the general observation that in the global open ocean a change in [Chl] is associated with a change in phytoplankton composition or size structure. The basis of this approach is that there is an upper limit of [Chl] in small cells imposed from genotypic and phenotypic constraints. Beyond this value, larger phytoplankton are responsible for an increase in [Chl] (Yentsch and Phinney, 1989; Chisholm, 1992).

Morel and Berthon (1989) suggested near surface [Chl] is related to water column-integrated chlorophyll content and its vertical distribution. Extending this work, Uitz et al. (2006) proposed quantitative relationships between the near surface [Chl] and (i) the water-column integrated chlorophyll content, (ii) its vertical distribution, and (iii) its community composition

TABLE 2 | Summary of satellite inputs and outputs.

Type	Algorithm references	Algorithm abbreviation	Development inputs							Satellite inputs						Satellite Outputs								
			nLw/Rrs	*[Chl]*	a_{ph}	a_{cdm}	b_{bp}	η	*S*	HPLC pigments	*nLw/Rrs*	*[Chl]*	*at*	a_{ph}	a_{cdm}	b_{bp}	*Micro*	*Nano*	*Pico*	*Hapto (cocco)*	*Dino*	*Cyano (Pro/Syn)*	*Diatom*	*Phaeo*
Abundance	Brewin et al., 2010	BR10		x						x		x					x	x	x					
	Brewin R. J. et al., 2011	BR10		x	x					x		x					x	x	x					
	Hirata et al., 2011	OC-PFT		x						x		x					x	x	x	x	x	x	x	
	Uitz et al., 2006	UITZ06								x		x					x	x	x					
Radiance	Alvain et al., 2005, 2008	PHYSAT	x	x						x	x	x						x		x		x	x	x
	Li et al., 2013	LI13	x							x	x						x	x	x					
Absorption	Bracher et al., 2009	PhytoDOAS			x						x											x	x	
	Sadeghi et al., 2012a	PhytoDOAS	x		x						x									x	x		x	
	Ciotti and Bricaud, 2006; Bricaud et al., 2012	CB06	x	x	x	x			x			x	x				(x)		x					
	Devred et al., 2011	DSSP11	x	x	x	x	x		x	x							x	x	x					
	Fujiwara et al., 2011	FUJI11	x	x	x			x			x			x			x		(x)					
	Hirata et al., 2008	HIRATA08		x	x					x				x			x	x	x					
	Mouw and Yoder, 2010a	MY10	x	x	x	x	x			x	x	x			x		x		(x)					
	Roy et al., 2013	ROY13		x	x					x		x		x			x	x	x					
Scattering	Kostadinov et al., 2009, 2010	KSM09					x	x			x					x	x	x	x					

The four algorithm types are indicated by color: abundance (green), radiance (red), absorption (yellow), scattering (blue). The development inputs, satellite inputs, and satellite outputs are indicated with "x" for each algorithm. Instances where other size classes could be inferred but are not directly retrieved are indicated with "(x)". Notation for column headers can be found in ***Table 1****.*

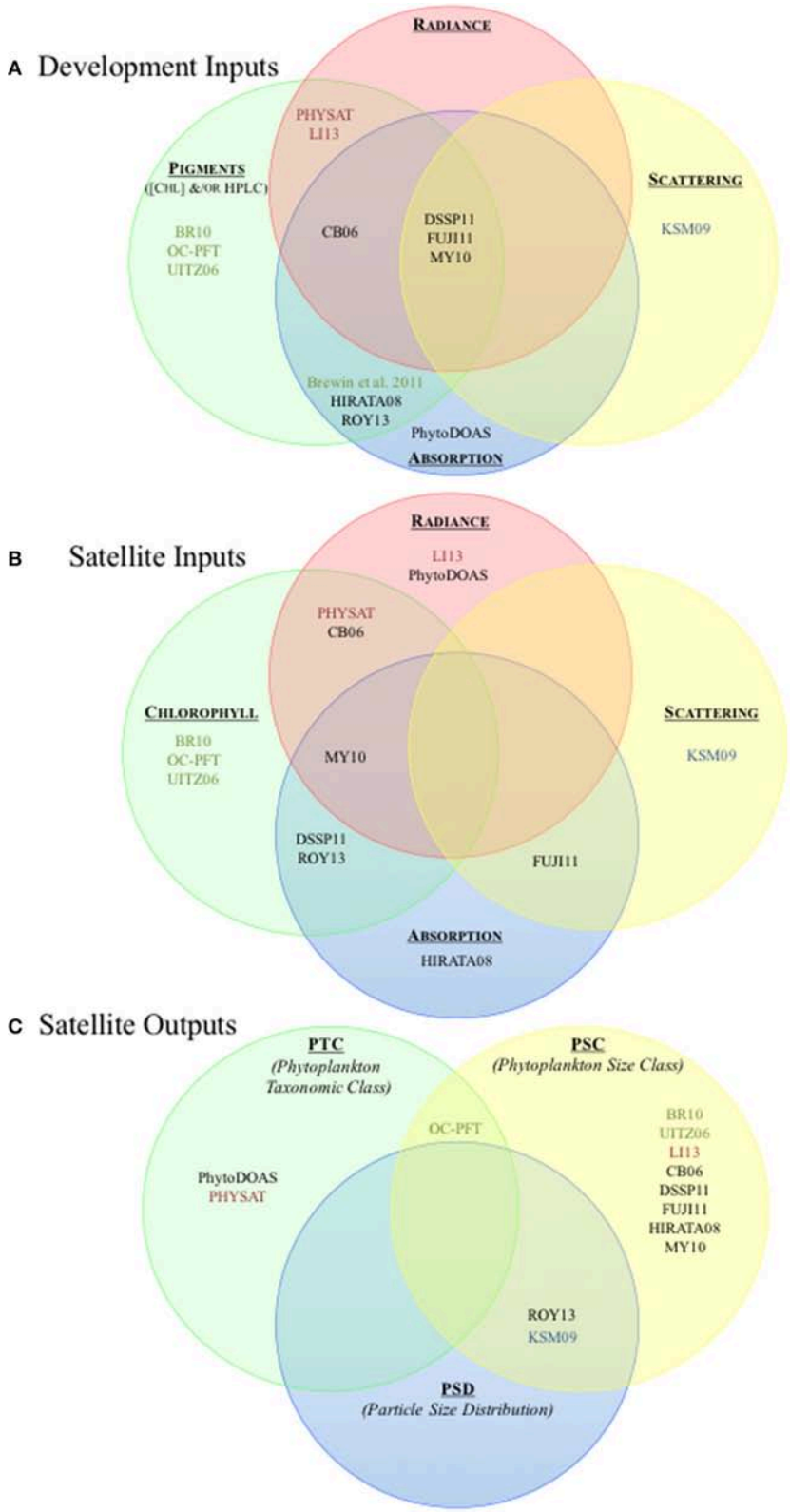


FIGURE 3 | Schematic of parameter space utilized in algorithm development, as well as satellite input and output products. (A) Overview of the parameters utilized in development of the four algorithm types. The primary optical data types are indicated with colored circles: pigments (green), radiance (red), absorption (blue), and scattering (yellow). **(B)** Overview of the satellite input products for the four algorithm types. Satellite input products are indicated by the colored

(Continued)

FIGURE 3 | Continued
circles: radiance (red), chlorophyll concentration (green), absorption (blue), and scattering (yellow). Overlapping circles indicate two or more satellite input products are utilized. **(C)** Overview algorithms satellite output by PFT types. The colored circles indicate the PFT type of the output products (phytoplankton taxonomic class (PTC, green), phytoplankton size class (PSC, yellow), and particle size distribution (PSD, blue). Overlapping circles indicate where a given algorithm produces two or more satellite output product types. The color of the text in all subplots indicates the algorithm type: abundance (green), radiance (red), absorption (black) and scattering (blue). Algorithm abbreviations are as in **Figure 1** and **Tables 2**, **3**.

in terms of three pigment-based PSC. The relationships were established from the analysis of a large high precision liquid chromatography (HPLC) pigment database, covering a broad range of trophic conditions in the global open ocean. Uitz et al. (2006) used a modified version of the diagnostic pigment indices of Vidussi et al. (2001) (described in the Algorithm Validation Section) to determine the depth-resolved contribution to the total chlorophyll biomass of three PSCs (pico-, nano-, and microphytoplankton). The resulting PSC-specific vertical profiles of [Chl] from stratified waters were discriminated from those sampled in well-mixed waters based on the ratio of the euphotic layer depth (calculated from the vertical [Chl] profile following Morel and Maritorena, 2001) and the mixed layer depth (extracted from a global monthly climatology). For the stratified and mixed waters, the [Chl] profiles of pico-, nano-, and microphytoplankton were sorted in trophic categories, defined by successive intervals of surface [Chl]. For each trophic category, average profiles of [Chl] associated with the pico-, nano, and microphytoplankton were calculated. The shape and magnitude of these profiles showed regular changes along the trophic gradient and, thus, could be parameterized as a function of surface [Chl]. Applied in a continuous manner to any given satellite-derived surface [Chl], the resulting empirical parameterization enables the ability to derive a vertical profile of [Chl] for each of the three pigment-based PSCs.

Hirata et al. (2011) estimate fractions of three PSCs and seven PTCs from empirical relationships between [Chl] and diagnostic pigments of various phytoplankton groups (see equations and coefficients in Hirata et al., 2011), based on global observations that abundance and composition of phytoplankton are not necessarily independent/de-coupled on synoptic scale. Brewin et al. (2010), extending the model proposed by Sathyendranath et al. (2001), describe the exponential functions that relate [Chl] to the fractional contribution of various PSCs,

$$[Chl]_{p,n} = C^{m}_{p,n}[1 - \exp(-S_{p,n}[Chl])] \quad (3a)$$

$$[Chl]_{p} = C^{m}_{p}[1 - \exp(-S_{p}[Chl])] \quad (3b)$$

$$[Chl]_{n} = [Chl]_{p,n} - [Chl]_{p} \quad (3c)$$

$$[Chl]_{m} = [Chl] - [Chl]_{p,n} \quad (3d)$$

where subscripts p, n, and m refer to pico- (>0.2–2 μm), nano- (>2–20 μm), and microplankton (>20 μm), respectively. $C^{m}_{p,n}$ and C^{m}_{p} are asymptotic maximum values for the associated size classes and $S_{p,n}$ and S_{p} determine the increase in size-fractionated [Chl] (parameter values can be found in Table 2 of Brewin et al., 2015), and have been found to vary with environmental conditions (Brewin et al., 2015; Ward, 2015). Both Brewin et al. (2010, 2012), Brewin R. J. W. et al. (2011) and Hirata et al. (2011) utilize the continuum of [Chl] (please see Figure 2 in Hirata et al., 2011 and Figure 4A in Brewin et al., 2010).

Radiance-based algorithms classify PFTs based on the shape and/or magnitude or the satellite-observed $R_{rs}(\lambda)$ or $nLw(\lambda)$. Radiance-based approaches assume that, after normalization, changes in radiance coincide with changes in PFT composition, as opposed to other in-water constituents such as CDOM or NAP that may or may not covary with the phytoplankton (e.g., Siegel et al., 2005). Alvain et al. (2005, 2008) normalize $R_{rs}(\lambda)$ to [Chl] and identify characteristic spectral bounds for several PTCs in terms of shape and amplitude (Ben Mustapha et al., 2014): nanoeukaryotes, *Prochlorococcus*, *Synechococcus*-like cyanobacteria, diatoms, *Phaeocystis*-like cells, and coccolithophores. More recently, based on theoretical relationships between radiance anomalies and specific phytoplankton groups, PHYSAT has been shown to potentially detect phytoplankton assemblages of several PTC as opposed to a single dominant one (Rêve et al., in revision). Alternatively, Li et al. (2013) consider a variety of spectral features on surface reflectance and use machine learning to select the most significant of these. They find continuum-removed and spectral curvature are the most significant spectral features with particular importance around 440–555 nm, which isolate absorption characteristics and measure non-linearity. They utilize these results with support vector regression to estimate PSCs.

Absorption-based algorithms comprise by far the majority of existing approaches. All of the approaches have some level of dependence on the spectral magnitude or shape of phytoplankton absorption [$a_{ph}(\lambda)$]. The magnitude of $a_{ph}(\lambda)$ is related to pigment composition and total pigment concentration, dominated by [Chl] at the peak wavelength (for oceanic waters) of 443 nm. Size information is contained in the absorption spectrum due to pigment packaging (e.g., Bricaud and Morel, 1986). Some of the approaches utilize chlorophyll-specific phytoplankton absorption in which phytoplankton absorption is normalized to [Chl] (Bracher et al., 2009; Mouw and Yoder, 2010a; Sadeghi et al., 2012a; Roy et al., 2013), either for a specific wavelength or to derive a spectral shape or slope that is related to second order signals including pigment composition and packaging. Several of the approaches (Ciotti and Bricaud, 2006; Mouw and Yoder, 2010a; Bricaud et al., 2012) stem from the theoretical underpinning of Ciotti et al. (2002) who identify that, despite the physiological and taxonomic variability, variation in $a_{ph}(\lambda)$ spectral shape can be defined by changes in the dominant size class. They determine chlorophyll-specific phytoplankton absorption (a^{*}_{ph}) as weighted between normalized mean pico- ($\bar{a}^{*}_{ph,pico}$)

TABLE 3 | Algorithm retrieval parameters and validation metrics.

Type	Algorithm	Type of retreival	Retreival parameters and units	Validation data source	Information source within reference	Validation measure	Strategy
Abundance	Brewin et al., 2010—BR10	PSC	Chla (mg m^{-3}): micro, nano, pico	HPLC pigments	Figure 6	pico: ME = 0.039 nano: ME = 0.076 micro: ME = 0.149	Semi-empirical
	Brewin R. J. et al., 2011—BR10	PSC	Absorption (m^{-1}): micro, nano, pico	Absorption	Figure 7	All at 443 nm: RMSE = 52.5% (size fractions not validated)	Empirical
	Hirata et al., 2011—OC-PFT	PSC, PTC	% Chl: pico, nano, micro diatom, dino, green hapto, prok, pico-euk prochlor	HPLC pigments	Table 4; Figure 4	pico: r^2 = 0.72, RMSE = 7.12 nano: r^2 = 0.56, RMSE = 8.55 micro: r^2 = 0.72, RMSE = 8.28 diatom: r^2 = 0.73, RMSE = 7.98 dino: r^2 = 0, RMSE = 1.87 green: r^2 = 0.40, RMSE = 4.71 hapto: r^2 = 0.37, RMSE = 10 prok: r^2 = 0.65, RMSE = 7.71 pico-euk: r^2 = 0.31, RMSE = 5.25 prochlor: r^2 = 0.72, RMSE = 6.25	Empirical
	Uitz et al., 2006—UITZ06	PSC	fractionated Chla (mg m^{-3}): micro, nano, pico	HPLC pigments	Figure 12A	log_{10} (predicted/measured) median = 0.02 mean = –0.012 std. dev. = 0.883	Empirical
Radiance	Alvain et al., 2005, 2008—PHYSAT	PTC	Dominance (presence over time): nanoeuk, prochlor, syn, diatom, phaeo	HPLC pigments	Figure 6 Alvain et al. (2008)	Classification success nanoeuc: 83% prochlor: 51% syn: 54% diatom: 57%	Empirical
	Li et al., 2013—LI13	PSC	Fractionated: pico, nano, micro	HPLC pigments	Figure 7	pico: r^2 = 0.587, RMSE = 15.2 nano: r^2 = 0.475, RMSE = 12 micro: r^2 = 0.617, RMSE = 17	Empirical Spectral features
Absorption	Bracher et al., 2009—PhytoDOAS	PTC	Chla (mg m^{-3}): cyano, diatom	HPLC pigments	Figure 8	diatom: r^2 = 0.92 cyano: r^2 = 0.81	Differential optical Absorption spectroscopy
	Sadeghi et al., 2012a—PhytoDOAS	PTC	Chla (mg m^{-3}): diatom, cocco, dino	Model and satellite	Figure 9 Product comparison*	diatom: – cocco: r^2 = 0.66 (MODIS PIC) dino: –	Differential optical Absorption spectroscopy

(Continued)

TABLE 3 | Continued

Type	Algorithm	Type of retreival	Retreival parameters and units	Validation data source	Information source within reference	Validation measure	Strategy
	Ciotti and Bricaud, 2006—CB06	PSC	S_f	Absorption	Table 4	RMSE all data = 17.2	Semi-analytical
							Non-linear optimization
	Bricaud et al., 2012—CB06	PSC	S_f	HPLC pigments	Table 1; Figure 4 in Brewin R. J. et al., 2011	52.0 ± 9.8% accuracy	Semi-analytical
	Validation in Brewin R. J. et al., 2011						Non-linear optimization
	Devred et al., 2011—DSSP11	PSC	% pico, % nano, % micro	HPLC pigments	Devred et al., 2011	pico: %diff = 3.4%, bias = 2.2%	Non-linear optimization
						nano: %diff = 11%, bias = −1.1%	
						micro: %diff = 12%, bias = −1.1%	
	Fujiwara et al., 2011—FUJI11	PSC	% Chla > 5μm	Size fractionated Chla	Figure 3	r^2 = 0.45, RMSE 22.7	Empirical
						Classification success = 69%	
	Hirata et al., 2008—HIRATA08	PSC	Dominance: pico, nano, micro	HPLC pigments	Table 3	Classification success	Empirical
						All data from AMT-07 = 73%	
	Mouw and Yoder, 2010b—MY10	PSC	S_{fm}	HPLC pigments	Table 2; Figure 13	micro: r^2 = 0.6, RMSE = 12.64	Look-up-table
	Roy et al., 2013—ROY13	PSD, PSC	PSD: exponent (unitless) and % Chl: micro, nano, pico	HPLC pigments absorption and error analysis	Figures 4, 5	pico: r^2 = 0.4	Semi-analytical
						nano: r^2 = 0.02	Non-linear optimization
						micro: r^2 = 0.42	
Scattering	Kostadinov et al., 2009—KSM09	PSD	PSD slope (unitless)	Coulter counter PSD	Figure 14	r^2 = 0.21 for PSD slope	Look-up-table
			PSD: pico, nano, micro [$\log_{10}(m^{-4})$]			r^2 = 0.26 for log10 (N_0)	
	Kostadinov et al., 2010—KSM09	PSC	% biovolume: pico, nano, micro	HPLC pigments	Figure 3	pico: r^2 = 0.34, RMSE = 24.1	Look-up-table
						nano: r^2 = 0.11, RMSE = 19.8	
						micro: r^2 = 0.42, RMSE = 17.1	

The four algorithm types are indicated by color: abundance (green), radiance (red), absorption (yellow), scattering (blue). The validation measure is as reported in the original algorithm publication and in units of the retrieval parameter, unless noted otherwise. Caution should be taken in comparing validation measures of differing units. ME is mean error, RMSE is root mean square error. A single asterisk indicates the validation data source is not a true validation; coccolithophores and diatoms are compared with numerical model output while coccolithophores are additionally compared with a satellite particulate inorganic carbon product and dinoflagellates are not compared. All algorithms with the exception of CB06 (Brazil continental shelf) and FUJI11 (Arctic and sub-Arctic) were developed for global extent. CB06 was later verified for global use by Bricaud et al. (2012).

and microplankton ($\bar{a}^*_{ph,\,micro}$) chlorophyll-specific absorption spectra.

$$a^*_{ph}(\lambda) = \left[S_f \times \bar{a}^*_{ph,\,pico}(\lambda)\right] + \left[\left(1 - S_f\right) \times \bar{a}^*_{ph,\,micro}(\lambda)\right] \quad (4)$$

where S_f is a dimensionless index constrained to vary between 0 and 1, specifying the relative contributions of microphytoplankton and picophytoplankton, respectively, to phytoplankton absorption. Equation (4) is based on the fact that the shape of the phytoplankton absorption spectrum flattens with increasing cell size. This relationship results from pigments being contained within particles (rather than in solution), known as the "discreetness effect," and secondarily how pigments are packaged within the cell, known as the "packaging effect" (Morel and Bricaud, 1981). Small cells have little cellular material between the chloroplast and cell wall making them highly efficient absorbers, resulting in higher magnitude and more peaked absorption. With large cells, light has to penetrate more cellular material to reach the chloroplast after passing through the cell wall, resulting in muted absorption affinity and in some cases shelf-shading (see Figure 7E in Ciotti et al., 2002). Note that the shape of the phytoplankton absorption spectrum (and therefore the S_f-value) can be affected by variations in pigment composition and intracellular pigment concentration resulting from photoacclimation, independent of cell size. Ciotti and Bricaud (2006) proposed a new $\bar{a}^*_{ph,pico}$ vector, based on an oceanic data set and Bricaud et al. (2012) utilize this relationship directly to retrieve S_f, absorption due to non-algal particles and colored dissolved organic matter at 443 nm [$a_{dg}(443)$] and the spectral slope of $a_{dg}(\lambda)$ through an inversion model. Mouw and Yoder (2010a) modify Equation (4) to vary with the percentage of microplankton (S_{fm}) rather than picoplankton. They develop an optical look-up-table (LUT) that contains ranges of S_{fm}, [Chl], and $a_{dg}(\lambda)$ from which $R_{rs}(\lambda)$ is calculated from radiative transfer. They utilize satellite [Chl] and $a_{dg}(443)$ to narrow the search space within the LUT, then find the closest match between satellite $R_{rs}(\lambda)$ and LUT $R_{rs}(\lambda)$ and retrieve the associated S_{fm} from the LUT.

Hirata et al. (2008) do not use the Ciotti et al. (2002) construct (i.e., Equation 4) that utilizes multiple wavelengths to characterize the spectral shape of $a_{ph}(\lambda)$. Instead, they identify a tight relationship between the magnitude of phytoplankton absorption at a single wavelength [$a_{ph}(443)$], related to [Chl], and the slope of $a_{ph}(443)$ to $a_{ph}(510)$, which is influenced by pigment packaging and composition. When this approach is applied to satellite data, it only uses $a_{ph}(443)$, and determines dominate size class using boundaries in $a_{ph}(443)$.

The approaches of Devred et al. (2011) and Brewin R. J. W. et al. (2011) are similar to that of Brewin et al. (2010), applying the constructs of Equations (3a–3d). Devred et al. (2011) use Equation (3) to derive chlorophyll-specific absorption coefficients for three PSCs. When this approach is applied to satellite data, it uses a semi-analytic inversion algorithm together with the derived chlorophyll-specific absorption coefficients to estimate size-fractionated [Chl], not using Equation (3) at this stage; hence this approach does not assume covariance between total [Chl] and size-fractionated [Chl] (as with an abundance-based approach). Fujiwara et al. (2011) is the only absorption-based approach that also uses backscatter as an input, which they determine empirically from $R_{rs}(\lambda)$ band ratios. They estimate PSC utilizing empirical relationships with phytoplankton absorption-spectra ratios and the particulate backscatter slope.

Roy et al. (2011) developed a semi-analytical algorithm based on phytoplankton absorption at a red wavelength (676 nm) to compute the equivalent spherical diameter of phytoplankton. Roy et al. (2013) further extended the algorithm to heterogeneous phytoplankton populations, where they utilized phytoplankton absorption at 676 nm to compute the PSD corresponding to the phytoplankton cells alone, and derived the power-law exponent/slope of the phytoplankton size spectrum. Knowing the slope of the phytoplankton cell-size distribution, the proportions of [Chl] within any diameter range of PSCs can be calculated.

The PhytoDOAS algorithm (Bracher et al., 2009; Sadeghi et al., 2012a) uses hyperspectral top of atmosphere reflectances to identify spectral features associated with PTCs. This approach requires hyperspectral satellite data and has been applied to the SCanning Imaging Absorption SpectroMeter for Atmospheric CHartographY (SCIAMACHY) onboard ENVISAT (more details in Bovensmann et al., 1999), which has a limited spatial coverage and resolution with 6-day revisit and 30 by 60 km pixel size. The differential optical absorption spectroscopy (DOAS) technique exploits sharp spectral features and when extended to phytoplankton, differentiates on spectral specific-absorption features of major PTCs. The DOAS method, utilizes observed backscattered radiation, normalized to the solar irradiance, at the top of the atmosphere and absorption cross sections (i.e., specific absorption coefficients, of all important absorbing constituents varying spectrally in the atmosphere-ocean system). The method uses non-linear optimization to fit these "differential absorption cross sections" of different phytoplankton groups, water vapor, and atmospheric trace gases: O_3, O_4, NO_2, glyoxal (CHOCHO), iodine oxide (IO), and spectral features caused by filling-in of Fraunhofer Lines due to Raman scattering. The contributions of broad-band scattering and absorption features, such as Mie- and Ray-leigh scattering in the atmosphere or NAP and CDOM in water, are approximated by a second-order polynomial in each fit. Bracher et al. (2009) adopted DOAS within 429–495 nm to retrieve absorption and biomass of cyanobacteria and diatoms independently. Sadeghi et al. (2012a) extended the method further to simultaneously retrieve diatoms, coccolithophores, and dinoflagellates over the 429–521 nm spectral range.

To date, there have only been two *scattering-based algorithms* published. Backscattering approaches retrieve information on all particles rather than just phytoplankton. Generally, the backscattering coefficient decreases according to a power law function with increasing wavelength. Smaller particles have a greater backscattering slope (η) than larger particles. Montes-Hugo et al. (2008) was the first to estimate phytoplankton size by considering the backscattering slope. They demonstrated their approach near the western shelf of the Antarctic Peninsula. Kostadinov et al. (2009) was the first to demonstrate the approach globally. They estimate spectral particulate backscattering

$[b_{bp}(\lambda)]$ from $R_{rs}(\lambda)$ and then calculate η from $b_{bp}(\lambda)$ based on Loisel et al. (2006). Using η, the PSD slope and reference abundance of particles are retrieved from a look-up-table that is constructed based on theoretical Mie scattering computations. These parameters are then used to estimate the number and volume concentrations for pico, nano, and micro sized particles. Assuming the relative proportions of biovolume are roughly constant across size classes, Kostadinov et al. (2010) validate the Kostadinov et al. (2009) approach with pigments and confirm pigment-based micro-, nano-, and pico-sized phytoplankton approximately represent micro-, nano-, and pico-sized particles derived from backscattering. Kostadinov et al. (2016a) further develop the KSM09 approach by using existing allometric relationships (Menden-Deuer and Lessard, 2000) to convert biovolume calculated from the PSD to phytoplankton carbon (C) in these three PSCs. These PSCs are the only carbon-based PFT retrievals available to date. The approach can be used to estimate phytoplankton carbon concentrations (absolute and fractional) in any size class in the 0.5–50 μm diameter range. It is desirable to express the PFTs in terms of carbon because: it is relatively insensitive to variations in phytoplankton physiological, unlike [Chl]; it is relevant to the carbon cycle and other biogeochemical cycles; and carbon is the unit used for PFTs in climate models (Hood et al., 2006).

Algorithm Validation

Nearly all algorithms are validated against estimates of phytoplankton size and composition estimates determined from *in situ* measurements of pigment concentrations with the HPLC technique (**Table 3**). The chemotaxonomic approach provides a means to quantify phytoplankton taxonomic composition utilizing a set of biomarker pigments (e.g., Jeffrey et al., 1997; Roy et al., 2011). Claustre (1994) and Vidussi et al. (2001) further proposed to utilize groupings of biomarker pigments to estimate phytoplankton size structure. They identified a set of seven diagnostic pigments specific to phytoplankton taxa, which were then assigned to one of the three size classes (micro-, nano-, and pico-) depending on the average cell size of the organisms. The diagnostic pigment-based approach enables estimating the contribution of the three phytoplankton size classes to the total chlorophyll *a* biomass as follows (Equation 5):

$$\Sigma DP = \sum_{i=1}^{7} W_i P_i \tag{5a}$$

$$f_{micro} = \frac{\sum_{i=1}^{2} W_i P_i}{\Sigma \text{ DP}} \tag{5b}$$

$$f_{nano} = \frac{\sum_{i=3}^{5} W_i P_i}{\Sigma \text{ DP}} \tag{5c}$$

$$f_{pico} = \frac{\sum_{i=6}^{7} W_i P_i}{\Sigma \text{ DP}} \tag{5d}$$

where, ΣDP is the sum of all the diagnostic pigments multiplied by the weight coefficients (W_i, values discussed below), f_{micro}, f_{nano}, and f_{pico} are the fractions of the micro-, nano- and pico-plankton size classes to [Chl], and P_i are the pigments' concentrations (P = {fucoxanthin; peridinin; 19′-hexanoyloxyfucoxanthin; 19′-butanoyloxyfucoxanthin; alloxanthin; chlorophyll-b and divinyl chlorophyll-b; zeaxanthin}). The most widely used coefficients are those proposed by Uitz et al. (2006) (W = {1.41; 1.41; 1.27; 0.35; 0.6; 1.01; 0.86}), which were derived from a global HPLC pigment database.

While diagnostic pigments have been widely used for validation due to the availability of extensive datasets of HPLC pigments across the global ocean (Peloquin et al., 2013), there are important limitations to consider. The diagnostic pigment-based approach does not necessarily reflect the true size structure of the phytoplankton communities because some taxonomic groups may spread over a broader size range (e.g., diatoms are typically found in the micro- but could also occur in the nano-size and sometimes in the pico-size classes) and some diagnostic pigments are shared by different taxonomic groups (e.g., fucoxanthin is the main carotenoid of diatoms but may also be found in prymnesiophytes). Recently modifications to the Vidussi et al. (2001) and Uitz et al. (2006) approach were proposed that account for the presence of some diagnostic pigments in more than one taxon. Hirata et al. (2011) and Devred et al. (2011) proposed further adjustments to the fucoxanthin pigment coefficient, to assign a portion of this pigment to nanoplankton. To address the analogous issue that prymnesiophytes predominate within the nanophytoplankton but can also be present in the pico-eukaryote population, Brewin et al. (2010) modified the 19′-hexanoyloxyfucoxanthin (19′-hex) coefficient to attribute a portion of this pigment to the picoplankton in low [Chl] waters. Furthermore, the DP for diatoms, fucoxanthin, is the precursor pigment for 19′-hex leading to some prymnesiophytes being classified as diatoms. For the algorithms that utilize HPLC pigments in their development, it should be noted that direct comparisons need to be considered carefully, as not all output products are developed and validated with the same set of diagnostic pigment coefficients. There are ongoing efforts to verify HPLC methods (i.e., Equation 5) through comparison with other techniques (e.g., Brewin et al., 2014a). As Nair et al. (2008) pointed out, any single method alone may not be entirely dependable, thus incorporating various methodologies leads to a more complete diagnosis of phytoplankton groups. Future efforts are necessary to complement HPLC methods with independent information on PFTs, for instance carbon-based size classes (Kostadinov et al., 2016a).

Not all validation approaches are based on HPLC pigment data, but rather use information on either absorption coefficients (Ciotti and Bricaud, 2006; Brewin R. J. W. et al., 2011) or size-fractionated [Chl] (Fujiwara et al., 2011). In addition, the study of Sadeghi et al. (2012b) does not perform a true validation, but rather compares numerical model results to other satellite products (**Table 3**).

Validation metrics are not reported using uniform metrics across algorithms causing an additional layer of complication when comparing algorithm performance. While it would be better to provide consistent validation measures across all algorithms, as mentioned previously, different satellite product outputs, units and use of variable development and validation

datasets/coefficients, preclude this ability. Reported validation measures are compiled in **Table 3**. Ideally, root mean square error (RMSE) (IOCCG, 2006) should be reported for matchups carried out according to the methods of Bailey and Werdell (2006) that specify the mean value of a five-by-five pixel box at the highest available pixel resolution measured by the sensor surrounding the location and within ±3 h of an *in situ* observation. In many cases, 9 km global area coverage satellite data are used to infer PFT classification. Thus, the spatial resolution is already coarse for validation matchups. However, algorithms can also be applied to full resolution (1 km) imagery improving validation efforts. In the case of PhytoDOAS, the input requires hyperspectral resolution and has been developed for use with SCIAMACHY, which has a resolution of 30 by 60 km. Spatial resolution differences between *in situ* point observations and the large SCIAMACHY pixels presents a limitation for validation using matchups (Bracher et al., 2009), since only very few *in situ* observations are within a homogeneous area of the size of a SCIAMACHY pixel. However, Aiken et al. (2007) point out that in the open ocean phytoplankton assemblages may be homogenously distributed over 50–100 km and smaller scales are possible for specific communities. In Sadeghi et al. (2012b), PhytoDOAS coccolithophore [Chl] was validated by comparison with satellite-derived particulate inorganic carbon (Balch et al., 2005).

ALGORITHM SELECTION

Users often select satellite products that most closely align with their application. When there are several satellite product choices for a given PFT type with varying facets and complexity, the optimal choice may not be clear. To help users determine what might be best suited for their purpose, in addition to the satellite inputs, outputs, and validation metrics described above, we compile a comparative list of assumptions, strengths, and limitations (**Table 4**). It is possible that merged products produce the best output beyond any individually selected algorithm (Palacz et al., 2013), yet an understanding of the underlying inputs into a merged product is always desirable.

Abundance-based algorithms assume a change in size and taxonomic structure with a change in chlorophyll. To the first order, and for large time and space scales, this holds true, but there are exceptions. Deviations from the mean state of the data in which the relationship is developed may occur (Hirata et al., 2011). This is particularly challenging at regional scales and in optically complex water where CDOM and NAP also complicate the retrieval of [Chl]. In a changing ocean, if shifts toward different phytoplankton assemblages with similar [Chl] occur, empirical relationships will require recalibration (Hirata et al., 2011). Abundance-based algorithms begin with uncertainty associated with the input satellite [Chl] product in addition to the uncertainty in relationships between [Chl] and phytoplankton grouping (**Figure 2**). Typically, band-ratio estimation of [Chl] (O'Reilly et al., 1998) has an accepted 35% uncertainty (Bailey and Werdell, 2006), which has recently been documented to be much less in the open ocean (16%) (Brewin et al., 2016), but becomes worse in coastal waters. Some semi-analytical inversions that retrieve [Chl], also have similar uncertainty across global scales (Brewin et al., 2015), but may maintain accuracy in coastal waters due to their ability to account for other in-water constituents contributing to the IOPs present that vary independently of each other. However, PFT approaches, which are broadly characterizing phytoplankton, may ultimately result in less uncertainty than the starting [Chl] product. The attractiveness of the abundance-based approaches is their ease of implementation and that they exploit the first-order signal in R_{rs}. [Chl] a primary biological variable that is routinely measured *in situ*, thus enabling extensive association of PFT fields with the abundance of *in situ* [Chl] that has accumulated across the globe. Once you know [Chl], PSC, or PTC estimates are a simple calculation.

Radiance-based approaches assume that after normalization to [Chl], changes in radiance coincide with changes in PFTs. They utilize $R_{rs}(\lambda)$ [or $nLw(\lambda)$], the fundamental parameter observed by a satellite radiometer and having uncertainty thresholds of 5% (Bailey and Werdell, 2006). Thus, the strength of radiance-based approaches is that they do not require or have limited dependence on products derived from $R_{rs}(\lambda)$. However, any normalization of the signal to derive the second-order relationships that tend to underpin these approaches will inevitably suffer from reduced signal to noise. Furthermore, when [Chl] is used in normalization (e.g., PHYSAT), the uncertainty associated with [Chl] is introduced (**Figure 2**). These algorithms are dependent on empirical relationships between radiance and PTCs or PSCs, thus as with empirical [Chl] dependencies described above, they require recalibration for long-term analyses. As with absorption- and scattering-based approaches and abundance-based approaches when using [Chl] determined from a semi-analytical model, radiance-based approaches allow for the ability to account for other optically active in water constituents (CDOM and NAP) as these also impact the spectral radiance (Alvain et al., 2012). This aspect allows potential development by users who have their own *in situ* datasets—it is possible to empirically associate a specific radiance anomaly to phytoplankton assemblages or specific composition (Alvain et al., 2012; Rêve et al., in revision). This highlights the importance of continued investment of detailed *in situ* databases to allow future development and use of remotely sensed phytoplankton groups. Radiance-based approaches are also influenced by physiological variability; however, the variability likely represents a larger proportion of the signal in normalized quantities.

PhytoDOAS (Bracher et al., 2009; Sadeghi et al., 2012a) has so far only been applied to a single sensor that has sufficient spectral resolution, precluding it from studies of phytoplankton composition where 30 km spatial resolution would be limiting. However, this is expected to improve in the near future: adaptations of the algorithm to similar high spectrally resolved satellite data with improved spatial coverage and resolution are currently ongoing. Ozone Monitoring Instrument (OMI) (since 2004) with 13 km by 24 km and TROPOMI (tropospheric OMI, to be launched in early 2017) with 3.5 km by 7 km global spatial resolution are, or will be, used with PhytoDOAS. In addition,

TABLE 4 | Comparative summary of algorithm assumptions, strengths and limitations.

	Assumptions	Strengths	Limitations
Abundance	• Change in size structure with change in [Chl] based on generalized relationships	• Easy to implement • Strong ecological basis	• Primarily empirical relationships with [Chl] that cannot detect regional deviations • Unable to distinguish mixed populations of similar abundance • Requires on-going recalibration as environmental change alters phytoplankton assemblages • Susceptible to physiological variability
Radiance	• After normalization to [Chl], changes in radiance are due primarily to variability in phytoplankton type	• Do not require or limited dependence on derived products • Input data (R_{rs}) has lower error than derived products	• Dependent on empirical relationships between radiance and pigments • Difficult to discriminate PFTs with similar normalized radiance signatures • Susceptible to physiological variability particularly normalized spectra
Absorption	• Variability largely the result of composition and pigment packaging • Primary variability in absorption is related to different PFTs	• Not directly dependent on concentration • Difficult to discriminate PFTs with similar absorption signatures	• Susceptible to physiological variability • Small deviations in spectral shape/magnitude can be difficult to retrieve
Scattering	• PSD and b_{bp} have a power-law shape • Relative proportions of biovolume to total particulate volume are roughly constant across classes	• Less sensitive to physiological variability	• Includes all particles, not just phytoplankton • Difficult to discriminate PFTs with similar scattering signatures

ocean color sensors are planned for the future with significantly increased spectral resolution (Mouw et al., 2015) that may allow a wider adoption of the PhytoDOAS method to even smaller spatial scales. For example, NASA's planned Plankton, Aerosol, Cloud, and ocean Ecosystem (PACE) mission with a hyperspectral ocean color sensor payload is expected to revolutionize the ability to use algorithms, such as PhytoDOAS, on more adequate spatio-temporal scales.

The number of existing absorption-based algorithms indicates the clear impact phytoplankton cell size and pigment composition have on the shape of the spectral absorption coefficient. These relationships have been reported in the literature for decades (e.g., Bricaud et al., 1988, 1995; Ciotti et al., 1999). The strengths of this type of algorithm include the ability to begin with inherent optical properties rather than [Chl] as the satellite input product, thus starting with reduced uncertainty at the onset. However, the assumed spectral shapes and coefficients utilized in semi-analytical approaches cannot fully capture natural variability across a variety of conditions resulting in uncertainties. These uncertainties are a balance of spectral accuracy and the accuracy of particular parameters over others (Werdell et al., 2013). As with R_{rs} approaches, those that require normalization by [Chl] inevitably reduce signal to noise and also reintroduce uncertainty associated with [Chl]. A limitation of absorption-based approaches is that they are sensitive to physiological variability associated with light and nutrient histories and these are likely to be of more influence when normalized quantities are used. Furthermore, small changes in the spectral shape of phytoplankton absorption can be difficult to retrieve from ocean-color (Garver et al., 1994; Wang et al., 2005), such that identifying and distinguishing different PFTs may not always be successful. Problems can also occur when trying to discriminate different phytoplankton groups with similar absorption signatures.

The scattering-based approaches presented here assume the PSD has a power-law shape and relative proportions of biovolume are roughly constant across size classes. Conversion to phytoplankton carbon for the carbon-based PFTs requires additional assumptions Kostadinov et al. (2016a). The models assume a relationship between the PSD and the spectral slope of $b_{bp}(\lambda)$. The use of $b_{bp}(\lambda)$ makes the approach less sensitive to physiological variability than other approaches. However, the particle size classes include all particles, not just phytoplankton and the relationship between $b_{bp}(\lambda)$ and phytoplankton cell size is still a matter of active debate (Stramski et al., 2004; Vaillancourt et al., 2004; Dall'Olmo et al., 2009; Whitmire et al., 2010). In addition, the sources of backscattering are still uncertain (Stramski et al., 2004) and applicability of Mie theory to particles and/or phytoplankton assemblages in seawater has its limitations (e.g., Dall'Olmo et al., 2009). It has been suggested that this approach represents phytoplankton carbon more closely—see Martinez-Vicente et al. (2013), and backscattering has been used to retrieve total phytoplankton carbon (Behrenfeld et al., 2005).

TABLE 5 | Online availability of permanently archived PFT algorithm products.

Algorithm	Location
BR10	http://catalogue.ceda.ac.uk/uuid/7201151d40683a4420da90c30640d4fa (Brewin et al., 2010; Brewin R. J. et al., 2011)
KSM09	https://doi.pangaea.de/10.1594/PANGAEA.859005 (PSD and Carbon-based size classes; Kostadinov et al., 2016b)
	http://hermes.acri.fr/index.php?class=demonstration_products (number-based size classes; Kostadinov et al., 2009)
MY2010	https://doi.pangaea.de/10.1594/PANGAEA.860474 (Mouw and Yoder, 2010b)
PHYSAT	http://log.univ-littoral.fr/Physat-2?lang=fr (Alvain et al., 2005, 2008) http://hermes.acri.fr/index.php?class=demonstration_products (Alvain et al., 2005, 2008)
PhytoDOAS	https://doi.org/10.1594/PANGAEA.870486 (Bracher et al., 2009, 2017; Sadeghi et al., 2012a)

Products were only available for SeaWiFS at time of writing for KSM09, BR10, PHYSAT, and MY2010 and for SCIAMACHY for PhytoDOAS.

Users are more focused on the satellite outputs, which they can use for various applications, rather than the intricacies of the type of algorithm used to produce the output. For ease of information identification, we have provided the validation metrics reported by algorithm type and satellite output types (**Table 3**). However, our purpose here is not to intercompare or validate algorithms. It is important to point out that the algorithms all use different approaches, datasets, and validation metrics. To be able to properly assess algorithm performance, one would have to carry out a comprehensive inter-comparison using the same validation data and consider errors of omission and commission (see Brewin R. J. et al., 2011), which is outside the scope of the present work. A validation effort is planned as part of the International Satellite Phytoplankton Functional Type Algorithm Inter-comparison Project (Hirata et al., 2012; http://pft.ees.hokudai.ac.jp/satellite/index.shtml) while an intercomparison based on phenology has been carried out by Kostadinov et al. (2017).

It is important to point out that many of these methods have been developed for the global open ocean. The optical complexity encountered in coastal waters is quite different from that found in the global datasets used to develop these algorithms. Additionally, the assumptions made by some are only valid for the global open ocean. The relationship between [Chl], CDOM absorption, and particulate backscatter is more variable in coastal water than the open ocean. For example, riverine sources, resuspension, and mixing may cause CDOM and NAP to vary independently of phytoplankton. For these reasons, band-ratio [Chl] estimates that utilize the blue and green region of the spectra are plagued with problems in coastal waters (Matthews, 2011). Thus, it is not advisable to apply open ocean abundance-based algorithms to coastal systems. Relationships would need to be assessed and likely redeveloped using a regionally specific dataset. Similar limitations would be expected for radiance-based methods. While the atmospheric correction can be a challenge over some coastal waters (Goyens et al., 2013), if $R_{rs}(\lambda)$ is accurately retrieved, the dynamic range of CDOM and NAP that impart a significant signal to $R_{rs}(\lambda)$ require empirical relationships and thresholds defined for various PFTs and PSCs to be reestablished. The approaches that build upon semi-analytic expressions that first retrieve IOPs from $R_{rs}(\lambda)$ and then PFTs from the retrieved IOPs, have the greatest ability to accommodate dynamic environments. These approaches parse the contributions of NAP, CDOM and phytoplankton before the phytoplankton IOPs are associated with a PFT. Similarly, this is done within the PhytoDOAS method by accounting for all relevant absorbers (from water and atmosphere) within the fitting of hyperspectral top of atmosphere reflectance. Accordingly, Brewin R. J. et al. (2011) find absorption-based approaches show an improvement over abundance-based approaches in coastal waters. However, the thresholds of detectability of approaches targeting optical signatures will not allow PFT retrieval in all cases.

The limitation of the ability to retrieve PFTs in some cases needs to be acknowledged. For example, Mouw and Yoder (2010a) are careful to consider the change in $R_{rs}(\lambda)$ produced by PSCs, [Chl] and CDOM absorption in relation to the radiometric sensitivity of the satellite senor. They find that when [Chl] or CDOM absorption were too high, the impact of size on $R_{rs}(\lambda)$ is masked. Likewise, when [Chl] is too low, the spectral response of $R_{rs}(\lambda)$ due to size is too small to differentiate from noise. Additionally, PHYSAT in its first version (Alvain et al., 2005) did not classify pixels where no phytoplankton group dominated due to the use of biomarker pigment threshold during the first empirical anomalies labeling steps. However, recent developments of PHYSAT have shown its capability to detect more than dominance cases utilizing detailed *in situ* data (Alvain et al., 2012; Ben Mustapha et al., 2014; Rêve et al., in revision).

The accessibility of products is another reason why users may select a given algorithm over another. Algorithms where simple calculations extracted from the publication can be quickly applied are far more likely to be utilized than those that require multiple complicated steps. The algorithm developer hosting the final output product for download by users has often remedied difficulty in this later situation. The PFT products that are currently accessible online are listed in **Table 5**. Further, the PFT products compiled for phenological comparison (Kostadinov et al., 2017) intend to be released in the near future. The availability of PFT product access is anticipated to grow substantially as future missions that have specified PFT products as part of their mission goals come online.

PFT algorithm development thus far has been focused on retrieving global distributions of PFTs. The next challenge is to detect change in these distributions over time. The temporal anomaly of PFTs can be a smaller signal than the bulk composition retrievals achieved thus far. The anomalies

are critical for understanding climate change issues and testing ecosystem model prediction. However, detecting change is confounded by inter- and intra-algorithm uncertainties and the relatively short record length of satellite data. Further, critical to this consideration are changes in phytoplankton physiology. Behrenfeld et al. (2016) show the importance of accounting for photoacclimation in temporal chlorophyll variability, as light-driven changes in chlorophyll can be associated with constant or increased photosynthesis. This finding of the necessity to account for physiological plasticity also directly impacts the PFT methods described here, most acutely for abundance-, radiance-, and absorption-based methods. While satellite PFT time series data have already been used to assess regional PFT variability and trends (Brewin et al., 2012; Sadeghi et al., 2012b; Alvain et al., 2013; Soppa et al., 2016). there is a need to characterize physiological plasticity in PFT retrievals to more accurately quantify phytoplankton compositional response to a changing ocean.

CONCLUSIONS

At the global scale, the current PFT algorithms demonstrate proof of concept in retrieving phytoplankton composition from satellite radiometry, opening the door for further development, and expand the use of satellite observations. While there are a variety of algorithm approaches, all agree on broad understanding of PFT distribution at large spatial-temporal scales, that are forced mainly by bathymetry and climatic regions. Larger cells and taxa tend to be found near coastal regions, especially under upwelling regimes, while smallest cells and taxa dominate in the center of oceans. Temperate regions are likely to present seasonal blooms of large cell sizes in spring and/or fall, while a less variable size distribution of phytoplankton is expected in tropical and subtropical areas and in the oligotrophic gyres.

Continual PFT algorithm development is anticipated, particularly with the expansion of sensor capability with future missions. Planned capability will expand spectral, spatial, and temporal resolution, in addition to radiometric sensitivity (Mouw et al., 2015). Increased spectral resolution will provide the ability to exploit more spectral signatures of PFTs (Isada et al., 2015; Wolanin et al., 2016). In addition to increased spectral resolution, increased spatial resolution may lend clarity to coastal processes and phytoplankton response to finer scale physical features. Improved temporal resolution on geostationary platforms will allow multiple views per day to investigate diurnal phytoplankton variability. Improved radiometric sensitivity will expand threshold detection required to detect the secondary impact of PFTs on radiometric variability. All of the potential capability in expanded satellite PFT products with the next generation of satellite sensors hinges on continued and increased investment in *in situ* observations to allow further algorithm development and validation. In addition to HPLC pigments that so many of these approaches are validated upon, training datasets also need to include unambiguous metrics of community composition that include particle size distribution and taxonomy (from imaging technologies) (Bracher et al., 2015). Exploiting compilations of abundance and biomass (Leblanc et al., 2012) and connections to genetically determined community composition (Malviya et al., 2016) are potentially rich resources for expanding training and validation datasets. In addition, coincident optical [i.e., $R_{rs}(\lambda)$ and IOPs] observations will be highly important to connect to the signals observed by satellite radiometers. The expanding optical sensors on Bio-Argo floats may also provide a valuable data stream for PFT development, particularly for vertical structure of phytoplankton communities (Mignot et al., 2014). It is important to expand the capability to measure phytoplankton carbon *in situ* (Graff et al., 2012, 2015) so future definitions of PFTs can be more carbon-relevant (Kostadinov et al., 2016a).

This document provides an overview of the primary components used in developing, implementing, and using satellite PFT products. While we do not provide direct recommendations for particular applications, our hope is that providing an accessible overview of the primary components of PFT algorithms will aid users in more confidently selecting products for a given application and ignite future conversations between satellite product developers and a variety of user communities. The satellite PFT literature is rapidly expanding and these tables and figures will require updating and the need to develop anew. In addition to the value we hope this brings to the user community, we equally hope this summary provides a framework for algorithm organization to inform where possible new approaches could be investigated into the future.

AUTHOR CONTRIBUTIONS

CM carried out the synthesis of algorithms, developed the organization, and prepared the manuscript with guidance from NH. CM prepared all figures and tables. All other co-authors contributed to ensuring accuracy of their algorithm description, overall synthesis, and editing.

FUNDING

The National Aeronautics and Space Administration (NASA) provided financial support for CM (NNX13AC34G) and TK (NNX13AC92G) for this effort. The contribution of AsB, RB, and AnB was partly funded via the ESA SEOM SY-4Sci Synergy project SynSenPFT.

ACKNOWLEDGMENTS

We acknowledge the International Satellite Functional Type Algorithm Intercomparison Project (http://pft.ees.hokudai.ac.jp/satellite/) for facilitation of a series of meetings and workshops that led to the development of this manuscript, with funding from JAXA and the UK National Centre for Earth Observation and support from the International Ocean Colour Coordination Group (IOCCG).

REFERENCES

Aiken, J., Fishwick, J., Lavender, S., Barlow, R., Moore, G., Sessions, H., et al. (2007). Validation of MERIS reflectance and chlorophyll during the BENCAL cruise October, 2002: preliminary validation of new products for phytoplankton function types and photosynthetic parameters. *Int. J. Remote Sens.* 28, 497–516. doi: 10.1080/01431160600821036

Alvain, S., Le Quéré, C., Bopp, L., Racault, M.-F., Beaugrand, G., Dessailly, D., et al. (2013). Rapid climatic driven shifts of diatoms at high latitudes. *Remote Sens. Environ.* 132, 195–201. doi: 10.1016/j.rse.2013.01.014

Alvain, S., Loisel, H., and Dessailly, D. (2012). Theoretical analysis of ocean color radiances anomalies and implications for phytoplankton groups detection in case 1 waters. *Opt. Express* 20, 1070–1083. doi: 10.1364/OE.20.001070

Alvain, S., Moulin, C., Dandonneau, Y., and Breon, F. (2005). Remote sensing of phytoplankton groups in case 1 waters for global SeaWiFS imagery. *Deep-Sea Res. I* 52, 1989–2004. doi: 10.1016/j.dsr.2005.06.015

Alvain, S., Moulin, C., Dandonneau, Y., and Loisel, H. (2008). Seasonal distribution and succession of dominant phytoplankton groups in the global ocean: a satellite view. *Global Biogeochem. Cycles* 22:GB3001. doi: 10.1029/2007GB003154

Bailey, S., and Werdell, P. (2006). A multi-sensor approach for the on-orbit validation of ocean color satellite data products. *Remote Sens. Environ.* 102, 12–23. doi: 10.1016/j.rse.2006.01.015

Balch, W. M., Gordon, H. R., Bowler, B. C., Drapeau, D. T., and Booth, E. S. (2005). Calcium carbonate measurements in the surface global ocean based on Moderate-Resolution Imaging Spectroradiometer data. *J. Geophys. Res.* 110, C07001. doi: 10.1029/2004JC002560

Behrenfeld, M. J., Boss, E., Siegel, D. A., and Shea, D. M. (2005). Carbon-based ocean productivity and phytoplankton physiology from space. *Global Biogeochem. Cycles* 19:GB1006. doi: 10.4319/lo.1997.42.7.1479

Behrenfeld, M. J., and Falkowski, P. G. (1997). A consumer's guide to phytoplankton primary productivity models. *Limnol. Oceanogr.* 42, 1479–1491.

Behrenfeld, M. J., O'Malley, R. T., Boss, E. S., Westberry, T. K., Graff, J. R., Halsey, K. H., et al. (2016). Revaluating ocean warming impacts on globalphytoplankton. *Nat. Clim. Chang.* 6, 323–330. doi: 10.1038/nclimate2838

Ben Mustapha, Z., Alvain, S., Jamet, C., Loisel, H., and Dessailly, D. (2014). Automatic classification of water-leaving radiance anomalies from global SeaWiFS imagery: application to the detection of phytoplankton groups in open ocean waters. *Remote Sens. Environ.* 146, 97–112. doi: 10.1016/j.rse.2013.08.046

Bovensmann, H., Burrows, J. P., Buchwitz, M., Frerick, J., Noël, S., Rozanov, V. V., et al. (1999). SCIAMACHY – mission objectives and measurement modes. *J. Atmos. Sci.* 56, 125–150. doi: 10.1175/1520-0469(1999)056<0127:SMOAMM>2.0.CO;2

Bracher, A., Dinter, T., Wolanin, A., Rozanov, V. V., Losa, S., and Soppa, M. A. (2017). Global monthly mean chlorophyll a surface concentrations from August 2002 to April 2012 for diatoms, coccolithophores and cyanobacteria from PhytoDOAS algorithm version 3.3 applied to SCIAMACHY data, link to NetCDF files in ZIP archive. doi: 10.1594/PANGAEA.870486 *Supplement to*: Losa, S., Soppa, A., Dinter, T., Wolanin, A., Rozanov, V. V., Brewin, R. J. W., Bricaud, A., Peeken, I. (in prep.). Synergistic exploitation of hyper- and multispectral Sentinel-measurements to determine Phytoplankton Functional Types at best spatial and temporal resolution (SynSenPFT). *Front. Mar. Sci.*

Bracher, A., Hardman-Mountford, N., Hirata, T., Bernard, S., Boss, E., Brewin, R. J. W., et al. (2015). *Phytoplankton Composition from Space: Towards a Validation Strategy for Satellite Algorithms*. NASA Technical Memorandum, 1–40.

Bracher, A., Vountas, M., Dinter, T., Burrows, J. P., Roettgers, R., and Peeken, I. (2009). Quantitative observation of cyanobacteria and diatoms from space using PhytoDOAS on SCIAMACHY data. *Biogeosciences* 6, 751–764. doi: 10.5194/bg-6-751-2009

Brewin, R. J., Hardman-Mountford, N. J., Lavender, S. J., Raitsos, D. E., Hirata, T., Uitz, J., et al. (2011). An intercomparison of bio-optical techniques for detecting dominant phytoplankton size class from satellite remote sensing. *Remote Sens. Environ.* 115, 325–339. doi: 10.1016/j.rse.2010.09.004

Brewin, R. J. W., Dall'Olmo, G., Pardo, S., van Dongen-Vogels, V., and Boss, E. S. (2016). Underway spectrophotometry along the Atlantic Meridional Transect reveals high performance in satellite chlorophyll retrievals. *Remote Sens. Environ.* 183, 82–97. doi: 10.1016/j.rse.2016.05.005

Brewin, R. J. W., Devred, E., Sathyendranath, S., Lavender, S., and Hardman-Mountford, N. (2011). Model of phytoplankton absorption based on three size classes. *Appl. Opt.* 50, 4535–4549. doi: 10.1364/AO.50.004535

Brewin, R. J. W., Hirata, T., Hardman-Mountford, N. J., Lavender, S. J., Sathyendranath, S., and Barlow, R. (2012). The influence of the Indian Ocean Dipole on interannual variations in phytoplankton size structure as revealed by Earth Observation. *Deep-Sea Res. II* 77–80, 117–127. doi: 10.1016/j.dsr2.2012.04.009

Brewin, R. J. W., Sathyendranath, S., Hirata, T., Lavender, S. J., Barciela, R. M., and Hardman-Mountford, N. J. (2010). A three-component model of phytoplankton size class for the Atlantic Ocean. *Ecol. Model.* 221, 1472–1483. doi: 10.1016/j.ecolmodel.2010.02.014

Brewin, R. J. W., Sathyendranath, S., Jackson, T., Barlow, R., Brotas, V., Airs, R., et al. (2015). Influence of light in the mixed-layer on the parameters of a three-component model of phytoplankton size class. *Remote Sens. Environ.* 168, 437–450. doi: 10.1016/j.rse.2015.07.004

Brewin, R. J. W., Sathyendranath, S., Lange, P. K., and Tilstone, G. (2014a). Comparison of two methods to derive the size-structure of natural populations of phytoplankton. *Deep Sea Res. I Oceanogr. Res. Pap.* 85, 72–79. doi: 10.1016/j.dsr.2013.11.007

Brewin, R. J. W., Sathyendranath, S., Tilstone, G., Lange, P. K., and Platt, T. (2014b). A multicomponent model of phytoplankton size structure. *J. Geophys. Res.* 119, 3478–3496. doi: 10.1002/2014JC009859

Bricaud, A., Babin, M., Morel, A., and Claustre, H. (1995). Variability in the chlorophyll-specific absorption-coefficients of natural phytoplankton: analysis and parameterization. *J. Geophys. Res.* 100, 13321–13332. doi: 10.1029/95JC00463

Bricaud, A., Bedhomme, A.-L., and Morel, A. (1988). Optical properties of diverse phytoplanktonic species: experimental results and theoretical interpretation. *J. Plankton Res.* 10, 851–873. doi: 10.1093/plankt/10.5.851

Bricaud, A., Ciotti, A. M., and Gentili, B. (2012). Spatial-temporal variations in phytoplankton size and colored detrital matter absorption at global and regional scales, as derived from twelve years of SeaWiFS data (1998–2009). *Global Biogeochem. Cycles* 26, 1. doi: 10.1029/2010GB003952

Bricaud, A., and Morel, A. (1986). Light attenuation and scattering by phytoplanktonic cells: a theoretical modeling. *Appl. Opt.* 25, 571–580. doi: 10.1364/AO.25.000571

Chisholm, S. (1992). "Phytoplankton Size," in *Primary Productivity and Biogeochemical Cycles in the Sea*, eds P. G. Falkowski and A. D. Woodhead (New York, NY: Springer), 213–237.

Ciotti, A., and Bricaud, A. (2006). Retrievals of a size parameter for phytoplankton and spectral light absorption by colored detrital matter from water-leaving radiances at SeaWiFS channels in a continental shelf region off Brazil. *Limnol. Oceanogr. Methods* 4, 237–253. doi: 10.4319/lom.2006.4.237

Ciotti, A., Cullen, J., and Lewis, M. (1999). A semi-analytic model of the influence of phytoplankton community structure on the relationship between light attenuation and ocean color. *J. Geophys. Res.* 104, 1559–1578. doi: 10.1029/1998JC900021

Ciotti, A., Lewis, M., and Cullen, J. (2002). Assessment of the relationship between dominant cell size in natural phytoplankton communities and the spectral shape of the absorption coefficient. *Limnol. Oceanogr.* 47, 404–417. doi: 10.4319/lo.2002.47.2.0404

Claustre, H. (1994). The trophic status of various oceanic provinces as revealed by phytoplankton pigment signatures. *Limnol. Oceanogr.* 39, 1206–1210. doi: 10.4319/lo.1994.39.5.1206

Dall'Olmo, G., Westberry, T. K., Behrenfeld, M. J., Boss, E., and Slade, W. H. (2009). Significant contribution of large particles to optical backscattering in the open ocean. *Biogeosciences* 6, 947–967. doi: 10.5194/bg-6-947-2009

De Moraes Rudorff, N., and Kampel, M. (2012). Orbital remote sensing of phytoplankton functional types: a new review. *Int. J. Remote Sens.* 33, 1967–1990. doi: 10.1080/01431161.2011.601343

Devred, E., Sathyendranath, S., Stuart, V., and Platt, T. (2011). A three component classification of phytoplankton absorption spectra: application to ocean-color data. *Remote Sens. Environ.* 115, 2255–2266. doi: 10.1016/j.rse.2011.04.025

Evers-King, H., Bernard, S., Robertson Lain, L., and Probyn, T. A. (2014). Sensitivity in reflectance attributed to phytoplankton cell size: forward

and inverse modelling approaches. *Opt. Express* 22, 11536–11551. doi: 10.1364/OE.22.011536

Fujiwara, A., Hirawake, T., Suzuki, K., and Saitoh, S.-I. (2011). Remote sensing of size structure of phytoplankton communities using optical properties of the Chukchi and Bering Sea shelf region. *Biogeosciences* 8, 3567–3580. doi: 10.5194/bg-8-3567-2011

Garver, S. A., Siegel, D. A., and Mitchell, B. G. (1994). Variability in near-surface particulate absorption spectra: what can a satellite ocean color imager see? *Limnol. Oceanogr.* 39, 1349−1367. doi: 10.4319/lo.1994.39.6.1349

Gordon, H., Brown, O., Evans, R., Brown, J., Smith, R., Baker, K., et al. (1988). A semianalytic radiance model of ocean color. *J. Geophys. Res.* D9, 10909–10924. doi: 10.1029/JD093iD09p10909

Goyens, C., Jamet, C., and Schroeder, T. (2013). Evaluation of four atmospheric correction algorithms for MODIS-AQUA images over contrasted coastal waters. *Remote Sens. Environ.* 131, 63–75. doi: 10.1016/j.rse.2012.12.006

Graff, J. R., Milligan, A. J., and Behrenfeld, M. J. (2012). The measurement of phytoplankton biomass using flow-cytometric sorting and elemental analysis of carbon. *Limnol. Oceanogr. Methods* 10, 910–920. doi: 10.4319/lom.2012.10.910

Graff, J. R., Westberry, T. K., Milligan, A. J., Brown, M. B., Dall'Olmo, G., Van Dongen- Vogels, V. K., et al. (2015). Analytical phytoplankton carbon measurements spanning diverse ecosystems. *Deep Sea Res. I* 102, 16–25. doi: 10.1016/j.dsr.2015.04.006

Hirata, T., Aiken, J., Hardman-Mountford, N., Smyth, T. J., and Barlow, R. G. (2008). An absorption model to determine phytoplankton size classes from satellite ocean colour. *Remote Sens. Environ.* 112, 3153–3159. doi: 10.1016/j.rse.2008.03.011

Hirata, T., Hardman-Mountford, N., and Brewin, R. J. W. (2012). Comparing satellite-based phytoplankton classification methods. *EOS Trans.* 93, 59–60. doi: 10.1029/2012eo060008

Hirata, T., Hardman-Mountford, N. J., Brewin, R. J. W., Aiken, J., Barlow, R., Suzuki, K., et al. (2011). Synoptic relationships between surface Chlorophyll-a and diagnostic pigments specific to phytoplankton functional types. *Biogeosciences* 8, 311–327. doi: 10.5194/bg-8-311-2011

Hood, R. A., Laws, E. A., Armstrong, R. A, Bates, N. R., Brown, C. W., Carlson, C. A., et al. (2006). Pelagic functional group modeling: progress, challenges and prospects. *Deep Sea Res. II* 53, 459–512. doi: 10.1016/j.dsr2.2006.01.025

IOCCG (2006). *Remote Sensing of Inherent Optical Properties: Fundamentals, Tests of Algorithms, and Applications*. Reports of the International Ocean-Colour Coordinating Group, No. 5. IOCCG, Dartmouth, NS.

IOCCG (2014). *Phytoplankton Functional Types from Space*. Reports of the International Ocean-Colour Coordinating Group, No. 15, IOCCG, Dartmouth, NS.

Isada, T., Hirawake, T., and Kobayashi, T. (2015). Hyperspectral optical discrimination of phytoplankton community structure in Funka Bay and its implications for ocean color remote sensing of diatoms. *Remote Sens. Environ.* 159, 134–151. doi: 10.1016/j.rse.2014.12.006

Jeffrey, S. W., Mantoura, R. F. C., and Wright, S. W. (eds.). (1997). *Phytoplankton Pigments in Oceanography: Guidelines to Modern Methods*. Paris: UNESCO.

Kostadinov, T. S., Cabré, A., Vedantham, H., Marinov, I., Bracher, A., Uitz, R., et al. (2017). Inter-comparison of phytoplankton functional type phenology metrics derived from ocean color algorithms and earth system models. *Remote Sens. Environ.* 190, 162–177. doi: 10.1016/j.rse.2016.11.014

Kostadinov, T., Siegel, D., and Maritorena, S. (2010). Global variability of phytoplankton functional types from space: assessment via the particle size distribution. *Biogeosciences* 7, 3239–3257. doi: 10.5194/bg-7-3239-2010

Kostadinov, T. S., Milutinovic, S., Marinov, I., and Cabré, A. (2016b). Size-partitioned phytoplankton carbon concentrations retrieved from ocean color data, links to data in netCDF format. doi: 10.1594/PANGAEA.859005. *Supplement to*: Kostadinov, T. S. et al. (2016): Carbon-based phytoplankton size classes retrieved via ocean color estimates of the particle size distribution. Ocean Sci. 12, 561Ű575. doi: 10.5194/os-12-561-2016

Kostadinov, T. S., Milutinović, S., Marinov, I., and Cabré, A. (2016a). Carbon-based phytoplankton size classes retrieved via ocean color estimates of the particle size distribution. *Ocean Sci.* 12, 561–575. doi: 10.5194/os-12-561-2016

Kostadinov, T. S., Siegel, D. A., and Maritorena, S. (2009). Retrieval of the particle size distribution from satellite ocean color observations. *J. Geophys. Res.* 114, C09015. doi: 10.1029/2009JC005303

Leblanc, K., Aristegui, J., Armand, L., Assmy, P., Beker, B., Bode, A., et al. (2012). A global diatom database – abundance, biovolume and biomass in the world ocean. *Earth Syst. Sci. Data* 4, 149–165. doi: 10.5194/essd-4-149-2012

Li, Z., Li, L., Song, K., and Cassar, N. (2013). Estimation of phytoplankton size fractions based on spectral features of remote sensing ocean color data. *J. Geophys. Res.* 118, 1445–1458. doi: 10.1002/jgrc.20137

Loisel, H., Nicolas, J.-M., Sciandra, A., Stramski, D., and Poteau, A. (2006). Spectral dependency of optical backscattering by marine particles from satellite remote sensing of the global ocean. *J. Geophys. Res.* 111, C09024. doi: 10.1029/2005JC003367

Malviya, S., Scalco, E., Audic, S., Vincent, F., Veluchamy, A., Poulain, J., et al. (2016). Insights into global diatom distribution and diversity in the world's ocean. *Proc. Natl. Acad. Sci. U.S.A.* 113, E1516–E1525. doi: 10.1073/pnas.1509523113

Martinez-Vicente, V., Dall'Olmo, G., Tarran, G., Boss, E., and Sathyendranath, S. (2013). Optical backscattering is correlated with phytoplankton carbon across the Atlantic Ocean. *Geophys. Res. Lett.* 40, 1154–1158. doi: 10.1002/grl.50252

Matthews, M. W. (2011). A current review of empirical procedures of remote sensing in inland and near-coastal transitional waters. *Int. J. Remote Sens.* 32, 1–45. doi: 10.1080/01431161.2010.512947

Menden-Deuer, S., and Lessard, E. J. (2000). Carbon to volume relationships for dinoflagellates, diatoms, and other protist plankton. *Limnol. Oceanogr.* 45, 569–579. doi: 10.4319/lo.2000.45.3.0569

Mignot, A., Claustre, H., Uitz, J., Poteau, A., D'Ortenzio, F., and Xing, X. (2014). Understanding the seasonal dynamics of phytoplankton biomass and the deep chlorophyll maximum in oligotrophic environments: a Bio-Argo float investigation. *Global Beiogeochem. Cycles* 28, 1–21. doi: 10.1002/2013gb004781

Mobley, C. D. (1994). *Light and Water*. San Diego, CA: Academic Press, Inc.

Montes-Hugo, M. A., Vernet, M., Smith, R., and Carder, K. (2008). Phytoplankton size-structure on the western shelf of the Antarctic Peninsula: a remote-sensing approach. *Int. J. Remote Sens.* 29, 801–829. doi: 10.1080/01431160701297615

Morel, A., Antoine, D., and Gentili, B. (2002). Bidirectional reflectance of oceanic waters: accounting for raman emission and varying particle scattering phase function. *Appl. Opt.* 41, 6289–6306. doi: 10.1364/AO.41.006289

Morel, A., and Berthon, J.-F. (1989). Surface pigments, algal biomass, and potential production of the euphotic layer: relationships reinvestigation in view of remote-sensing applications. *Limnol. Oceanogr.* 8, 1545–1562. doi: 10.4319/lo.1989.34.8.1545

Morel, A., and Bricaud, A. (1981). Theoretical results concerning light absorption in a discrete medium, and application to specific absorption of phytoplankton. *Deep Sea Res.* 11, 1375–1393. doi: 10.1016/0198-0149(81)90039-X

Morel, A., and Maritorena, S. (2001). Bio-optical properties of oceanic water: a reappraisal. *J. Geophys. Res.* 106, 7163–7180. doi: 10.1029/2000JC000319

Mouw, C. B., Greb, S., Aurin, D., DiGiacomo, P. M., Lee, Z., Twardowski, M., et al. (2015). Aquatic color radiometry remote sensing of coastal and inland waters: challenges and recommendations for future satellite missions. *Remote Sens. Environ.* 160, 15–30. doi: 10.1016/j.rse.2015.02.001

Mouw, C. B., and Yoder, J. A. (2010b). SeaWiFS (version R2010.0) Monthly Global 9 km Microplankton Fraction, links to NetCDF files. Supplement to: Mouw, CB; Yoder, JA (2010): Optical determination of phytoplankton size composition from global SeaWiFS imagery. *J. Geophys. Res.* 115. doi: 10.1029/2010JC006337 *Supplement to*: Mouw, C. B., Yoder, J. A. (2010). Optical determination of phytoplankton size composition from global SeaWiFS imagery. *J. Geophys. Res.* 115. doi: 10.1029/2010JC006337

Mouw, C. B., Yoder, J. A., and Doney, S. C. (2012). Impact of phytoplankton community size on a linked global ocean optical and ecosystem model. *J. Mar. Syst.* 89, 61–75. doi: 10.1016/j.jmarsys.2011.08.002

Mouw, C., and Yoder, J. (2010a). Optical determination of phytoplankton size composition from global SeaWiFS imagery. *J. Geophys. Res.* 115:C12018. doi: 10.1029/2010JC006337

Nair, A., Sathyendranath, S., Platt, T., Morales, J., Stuart, V., Forget, M., et al. (2008). Remote sensing of phytoplankton functional types. *Remote Sens. Environ.* 112, 3366–3375. doi: 10.1016/j.rse.2008.01.021

O'Reilly, J., Maritorena, S., Mitchell, B., Siegel, D., Carder, K., Garver, S., et al. (1998). Ocean color chlorophyll algorithms for SeaWiFS. *J. Geophys. Res.* 103, 24937–24953. doi: 10.1029/98JC02160

Palacz, A. P., St. John, M. A., Brewin, R. J. W., Hirata, T., and Gregg, W. W. (2013). Distribution of phytoplankton functional types in high-nitrate low-chlorophyll

waters in a new diagnostic ecological indicator model. *Biogeosciences* 10, 7553–7574. doi: 10.5194/bg-10-7553-2013

Peloquin, J., Swan, C., Gruber, N., Vogt, M., Claustre, H., Ras, J., et al. (2013). The MAREDAT global database of high performance liquid chromatography marine pigment measurements. *Earth Syst. Sci. Data* 5, 109–123. doi: 10.5194/essd-5-109-2013

Raitsos, D., Lavender, S., Maravelias, C., Haralabous, J., Richardson, A., and Reid, P. (2008). Identifying four phytoplankton functional types from space: an ecological approach. *Limnol. Oceangr.* 53, 605–613. doi: 10.4319/lo.2008.53.2.0605

Reynolds, C. S., Huszar, V., Kruk, C., Naselli-Flores, L., and Melo, S. (2002). Towards a functional classification of the freshwater phytoplankton. *J. Plankton Res.* 24, 417–428. doi: 10.1093/plankt/24.5.417

Roy, S., Sathyendranath, S., Bouman, H., and Platt, T. (2013). The global distribution of phytoplankton size spectrum and size classes from their light-absorption spectra derived from satellite data. *Remote Sens. Environ.* 139, 185–197. doi: 10.1016/j.rse.2013.08.004

Roy, S., Sathyendranath, S., and Platt, T. (2011). Retrieval of phytoplankton size from bio-optical measurements: theory and applications. *J. R. Soc. Interface* 8, 650–660. doi: 10.1098/rsif.2010.0503

Sadeghi, A., Dinter, T., Vountas, M., Taylor, B., Altenburg-Soppa, M., and Bracher, A. (2012b). Remote sensing of coccolithophore blooms in selected oceanic regions using the PhytoDOAS method applied to hyper-spectral satellite data. *Biogeosciences* 9, 2127–2143. doi: 10.5194/bg-9-2127-2012

Sadeghi, A., Dinter, T., Vountas, M., Taylor, B., Peeken, I., Altenburg Soppa, M., et al. (2012a). Improvements to the PhytoDOAS method for identification of coccolithophores using hyper-spectral satellite data. *Ocean Sci.* 8, 1055–1070. doi: 10.5194/os-8-1055-2012

Sathyendranath, S., Cota, G., Stuart, V., Maass, H., and Platt, T. (2001). Remote sening of phytoplankton pigments: a comparison of empirical and theoretical approaches. *Int. J. Remote Sens.* 22, 249–273. doi: 10.1080/014311601449925

Sieburth, J., Smetacek, V., and Lenz, J. (1978). Pelagic ecosystem structure: heterotrophic compartments of the plankton and their relationship to plankton size fractions. *Limnol. Oceangr.* 23, 1256–1263. doi: 10.4319/lo.1978.23.6.1256

Siegel, D., Maritorena, S., Nelson, N., and Behrenfeld, M. (2005). Independence and interdependencies among global ocean color properties: reassessing the bio-optical assumption. *J. Geophys. Res.* 110:C07011. doi: 10.1029/2004JC002527

Soppa, M. A., Völker, C., and Bracher, A. (2016). Diatom phenology in the southern ocean: mean patterns, trends and the role of climate oscillations. *Remote Sens.* 8, 420. doi: 10.3390/rs8050420

Stramski, D., Boss, E., Bogucki, D., and Voss, K. (2004). The role of seawater constituents in light backscattering in the ocean. *Prog. Oceanogr.* 61, 27–56. doi: 10.1016/j.pocean.2004.07.001

Thuillier, G., Herse, M., Labs, D., Foujols, T., Peetermans, W., Gillotay, D., et al. (2003). The solar spectral irradiance from 200 to 2400 nm as measured by the solspec spectrometer from the ATLAS and EURECA missions. *Solar Phys.* 214, 1–22. doi: 10.1023/A:1024048429145

Uitz, J., Claustre, H., Morel, A., and Hooker, S. (2006). Vertical distribution of phytoplankton communities in open ocean: an assessment based on surface chlorophyll. *J. Geophys. Res.* 111, C08005. doi: 10.1029/2005JC003207

Vaillancourt, R., Brown, C., Guillard, R., and Balch, W. (2004). Light backscattering properties of marine phytoplankton: relationships to cell size, chemical composition and taxonomy. *J. Plankton Res.* 26, 191–212. doi: 10.1093/plankt/fbh012

Vidussi, F., Claustre, H., Manca, B., Luchetta, A., and Marty, J.-C. (2001). Phytoplankton pigment distribution in relation to upper thermocline circulation in the eastern Mediterranean Sea during winter. *J. Geophys. Res.* 106, 19939–19956. doi: 10.1029/1999JC000308

Wang, P., Boss, E. S., and Roesler, C. (2005). Uncertainties of inherent optical properties obtained from semianalytical inversions of ocean color. *Appl. Opt.* 44, 4074–4085. doi: 10.1364/AO.44.004074

Ward, B. A. (2015). Temperature-correlated changes in phytoplankton community structure are restricted to polar waters. *PLoS ONE* 10:e0135581. doi: 10.1371/journal.pone.0135581

Werdell, P. J., Franz, B. A., Bailey, S. W., Feldman, G. C., Boss, E., Brando, V. E., et al. (2013). Generalized ocean color inversion model for retrieving marine inherent optical properties. *Appl. Opt.* 52, 2019–2037 doi: 10.1364/AO.52.002019

Whitmire, A., Pegau, W., Karp-Boss, L., Boss, E., and Cowles, T. (2010). Spectral backscattering properties of marine phytoplankton cultures. *Opt. Express* 18, 15073–15093. doi: 10.1364/OE.18.015073

Wolanin, A., Soppa, M. A., and Bracher, A. (2016). Investigation of spectral band requirements for improving retrievals of phytoplankton functional types. *Remote Sens.* 8:871. doi: 10.3390/rs8100871

Yentsch, C., and Phinney, D. (1989). A bridge between ocean optics and microbial ecology. *Limnol. Oceanogr.* 8, 1694–1705. doi: 10.4319/lo.1989.34.8.1694

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling Editor declared a shared affiliation, though no other collaboration, with several of the authors SA, AB, and JU and states that the process nevertheless met the standards of a fair and objective review.

Copyright © 2017 Mouw, Hardman-Mountford, Alvain, Bracher, Brewin, Bricaud, Ciotti, Devred, Fujiwara, Hirata, Hirawake, Kostadinov, Roy and Uitz. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Marine Science

REVIEW
published: 03 March 2017
doi: 10.3389/fmars.2017.00055

Obtaining Phytoplankton Diversity from Ocean Color: A Scientific Roadmap for Future Development

Astrid Bracher[1, 2], Heather A. Bouman[3], Robert J. W. Brewin[4, 5], Annick Bricaud[6, 7], Vanda Brotas[8], Aurea M. Ciotti[9], Lesley Clementson[10], Emmanuel Devred[11], Annalisa Di Cicco[12], Stephanie Dutkiewicz[13], Nick J. Hardman-Mountford[14], Anna E. Hickman[15], Martin Hieronymi[16], Takafumi Hirata[17, 18], Svetlana N. Losa[1], Colleen B. Mouw[19], Emanuele Organelli[4], Dionysios E. Raitsos[4], Julia Uitz[6, 7], Meike Vogt[20] and Aleksandra Wolanin[1, 2, 21]*

[1] Phytooptics Group, Climate Sciences, Alfred-Wegener-Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany, [2] Department of Physics and Electrical Engineering, Institute of Environmental Physics, University Bremen, Bremen, Germany, [3] Department of Earth Sciences, University of Oxford, Oxford, UK, [4] Remote Sensing Group, Plymouth Marine Laboratory, Plymouth, UK, [5] National Centre for Earth Observation, Plymouth Marine Laboratory, Plymouth, UK, [6] Sorbonne Universités, UPMC-Université Paris-VI, UMR 7093, LOV, Observatoire Océanologique, Villefranche/Mer, France, [7] Centre National de la Recherche Scientifique, UMR 7093, LOV, Observatoire Océanologique, Villefranche/Mer, France, [8] Faculdade de Ciencias da Universidade de Lisboa, MARE, Lisboa, Portugal, [9] CEBIMar, Universidade de São Paulo, São Paulo, Brazil, [10] CSIRO Oceans and Atmosphere, Hobart, TAS, Australia, [11] Department of Fisheries and Oceans, Bedford Institute of Oceanography, Dartmouth, NS, Canada, [12] Institute of Atmospheric Sciences and Climate, Italian National Research Council (CNR-ISAC), Rome, Italy, [13] Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA, [14] CSIRO Oceans and Atmosphere, Perth, WA, Australia, [15] Ocean and Earth Science, National Oceanography Centre Southampton, University of Southampton, Southampton, UK, [16] Department of Remote Sensing, Institute of Coastal Research, Helmholtz-Zentrum Geesthacht, Geesthacht, Germany, [17] Faculty of Environmental Earth Science, Hokkaido University, Sapporo, Japan, [18] CREST, Japan Science and Technology Agency, Tokyo, Japan, [19] Graduate School of Oceanography, University of Rhode Island, Narragansett, RI, USA, [20] Department of Environmental Systems Science, Institute for Biogeochemistry and Pollutant Dynamics, ETH Zürich, Zürich, Switzerland, [21] GeoForschungsZentrum Potsdam, Potsdam, Germany

OPEN ACCESS

Edited by:
Laura Lorenzoni,
University of South Florida, USA

Reviewed by:
Matthew J. Oliver,
University of Delaware, USA
Catherine Mitchell,
Bigelow Laboratory for Ocean Sciences, USA

***Correspondence:**
Astrid Bracher
astrid.bracher@awi.de

Specialty section:
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

Received: 30 November 2016
Accepted: 15 February 2017
Published: 03 March 2017

Citation:
Bracher A, Bouman HA, Brewin RJW, Bricaud A, Brotas V, Ciotti AM, Clementson L, Devred E, Di Cicco A, Dutkiewicz S, Hardman-Mountford NJ, Hickman AE, Hieronymi M, Hirata T, Losa SN, Mouw CB, Organelli E, Raitsos DE, Uitz J, Vogt M and Wolanin A (2017) Obtaining Phytoplankton Diversity from Ocean Color: A Scientific Roadmap for Future Development. Front. Mar. Sci. 4:55. doi: 10.3389/fmars.2017.00055

To improve our understanding of the role of phytoplankton for marine ecosystems and global biogeochemical cycles, information on the global distribution of major phytoplankton groups is essential. Although algorithms have been developed to assess phytoplankton diversity from space for over two decades, so far the application of these data sets has been limited. This scientific roadmap identifies user needs, summarizes the current state of the art, and pinpoints major gaps in long-term objectives to deliver space-derived phytoplankton diversity data that meets the user requirements. These major gaps in using ocean color to estimate phytoplankton community structure were identified as: (a) the mismatch between satellite, *in situ* and model data on phytoplankton composition, (b) the lack of quantitative uncertainty estimates provided with satellite data, (c) the spectral limitation of current sensors to enable the full exploitation of backscattered sunlight, and (d) the very limited applicability of satellite algorithms determining phytoplankton composition for regional, especially coastal or inland, waters. Recommendation for actions include but are not limited to: (i) an increased communication and round-robin exercises among and within the related expert groups, (ii) the launching of higher spectrally and spatially resolved sensors, (iii) the development of algorithms that exploit

hyperspectral information, and of (iv) techniques to merge and synergistically use the various streams of continuous information on phytoplankton diversity from various satellite sensors' and *in situ* data to ensure long-term monitoring of phytoplankton composition.

Keywords: ocean color, phytoplankton functional types, algorithms, satellite sensors, roadmap

USER NEEDS FOR PHYTOPLANKTON DIVERSITY FROM SPACE

Marine phytoplankton play an important role in the global carbon cycle via the biological carbon pump (e.g., IPCC, 2013) and contribute about 50% to the global primary production (Field et al., 1998). Over the past 30 years, ocean color remote sensing has revolutionized our understanding of marine ecosystems and biogeochemical processes by providing continuous global estimates of surface chlorophyll a concentration (chl-a, mg m^{-3}), a proxy for phytoplankton biomass (e.g., McClain, 2009). However, chl-a alone does not provide a full description of the complex nature of phytoplankton community structure and function. Phytoplankton have different morphological (size and shape) and physiological characteristics (growth and mortality rates, nutrient uptake kinetics, temperature, and light requirements) as well as different biogeochemical and ecological functions (e.g., silicification, calcification, nitrogen fixation, aggregation and sinking rates, lipid production, energy transfer; e.g., Le Quéré et al., 2005). Phytoplankton community structure is thus important to many fundamental biogeochemical processes, including: nutrient uptake and cycling, energy transfer through the marine food web, deep-ocean carbon export, and gas exchange with the atmosphere. Phytoplankton community composition also has important consequences for fisheries (e.g., fish recruitment) and specific species (Harmful Algal Blooms, HABs; a list of all abbreviations is given in **Table 1**) can directly impact human health (e.g., Cullen et al., 1997).

The ability to observe the spatial-temporal distribution (including phenology) and variability of different phytoplankton groups is a scientific priority for understanding the marine food web, and ultimately predicting the ocean's role in regulating climate and responding to climate change on various time scales. Thus, identifying the drivers of phytoplankton composition on global and regional scales is required to assess climate ecosystem interactions and to increase our understanding of the role of the ocean's biodiversity for marine ecosystem service provision. Coasts are especially vulnerable to major human threats caused by harmful algal blooms, eutrophication, hypoxia, and other processes deteriorating water quality. High resolution data on phytoplankton diversity is urgently needed for many socio-economic applications (e.g., fisheries, aquaculture, and coastal management, see IOCCG, 2009).

Some fishery models (e.g., Jennigs et al., 2008) already utilize information on phytoplankton biomass derived from ocean color satellites, however information on size and taxonomic composition from satellite is highly desirable to improve stock assessments (IOCCG, 2009). To better represent the variable biogeochemical state of the ocean, Earth System, and climate models (including those used in the IPCC assessments) have increasingly included a larger amount of biological complexity in their ocean biogeochemistry modules. To simplify the representation of the vast planktonic diversity, plankton have been grouped into plankton functional types according to their biogeochemical functions (e.g., Le Quéré et al., 2005). Biogeochemical models now commonly include 3–10 plankton functional types (e.g., Bopp et al., 2013; Laufkötter et al., 2015), with a few models including up to 100 or more types (Follows et al., 2007; Dutkiewicz et al., 2015; Masuda et al., 2017). Since *in situ* observations on plankton biogeography and abundance are scarce and many vast oceanic regions are too remote to be routinely monitored, biogeochemical modelers rely on surface

TABLE 1 | Abbreviations and acronyms used throughout the text.

AC	Atmospheric correction
AOP	Apparent optical property
chl-a	Chlorophyll *a* concentration
CDOM	Colored dissolved organic matter
EnMAP	Environmental Mapping and Analysis Program mission
HABs	Harmful Algal Blooms
HICO	Hyperspectral Imager for the Coastal Ocean
HPLC	High Performance Liquid Chromatography
HyspIRI	Hyperspectral InfraRred Imager
IOP	inherent optical property
MERIS	Medium Resolution Imaging Spectrometer
MODIS	Moderate Resolution Imaging Spectroradiometer
MSI	MultiSpectral Instrument
NASA	National Aeronautics and Space Administration
OC	Ocean color
OC-PFT	Algorithm of Hirata et al. (2011)
OLCI	Ocean and Land Colour Instrument
OMI	Ozone Monitoring Instrument
PACE	Pre-Aerosol, Clouds, and ocean Ecosystem
PhytoDOAS	Algorithm of Bracher et al. (2009), further adapted by Sadeghi et al. (2012a)
PFT	Phytoplankton functional types
PG	Phytoplankton groups
PSC	Phytoplankton size class
PT	Phytoplankton types
RTM	Radiative transfer model
SCIAMACHY	Scanning Imaging Absorption Spectrometers for Atmospheric Chartography
SeaWiFS	Sea-viewing Wide Field-of-view Sensor
S	Sentinel
TROPOMI	TROPOspheric Monitoring Instrument
UVN	Ultra-violet/Visible/Near-Infrared Instrument

ocean estimates of phytoplankton composition from satellite observations to evaluate model simulations and help to develop and validate their models. Increased biological realism in these models has been suggested as a mean to reduce the large uncertainty in future projections of net primary production, and carbon export (Bopp et al., 2013; Laufkötter et al., 2015). Information on global phytoplankton community composition from ocean color satellites is therefore highly desirable for Earth system model development and the quantification of key processes related to present and future global biogeochemical cycles. Particularly for the quantification of carbon fluxes in the world's ocean, high quality remote sensing data on phytoplankton community composition are a first priority (see science plan of the EXPORT project, Siegel et al., 2016).

Thus, continuous, global-scale, high-resolution satellite ocean color products that go beyond bulk chl-a and provide information on phytoplankton diversity is urgently needed to improve near-real time and forecasting models for marine services facilitating the above-mentioned applications. User requests for satellite data on phytoplankton diversity as an essential ocean/climate variable is providing impetus for its incorporation into international climate change initiatives and mission (capability) planning. In this article the current state of the art regarding algorithms, their validation and application is reviewed, then the gaps to meet user requirements are discussed, and finally detailed recommendations for future medium and long term actions are provided.

STATE OF THE ART

Diversity of phytoplankton, often represented by species richness and evenness, can be characterized in multiple dimensions (e.g., taxonomic, phylogenetic, morphological, or functional diversity, among others). This diversity is staggeringly large and even within a species there are often a large range of ecotypes with different environmental niches, life stages and/or morphological, and physiological characteristics (e.g., Bouman et al., 2006). For almost all purposes scientists tend to cluster species into groups specific to the purposes of their research. For instance, climate scientists and marine biogeochemists define phytoplankton functional types (PFT) based on their biogeochemical functions (e.g., diatoms as silicifying PFT). Based on satellite products, we here refer to any clustering of species (and ecotypes) as "Phytoplankton Groups" (PG). PG defined based on taxonomic criteria are referred to as phytoplankton types (PT), and PG defined based on their size range are referred to as phytoplankton size classes (PSC).

Satellite ocean-color remote sensing is unsurpassed in its ability to characterize the state of the surface ocean biosphere at high temporal and spatial scales. Beyond chl-a, increasing efforts have been invested internationally over the last two decades to develop ocean color algorithms to retrieve information on phytoplankton composition and size structure (see recent summary in IOCCG, 2014 and list of global approaches applied to satellite data in **Table 2**). These developments provide an opportunity to yield new operational satellite products. Ocean color algorithms to assess phytoplankton diversity make use of information originating from phytoplankton abundance, cell size, bio-optical properties (such as pigment composition, absorption, and backscattering characteristics) to differentiate PG (**Table 2**, **Figure 1** left). The abundance based approaches of Uitz et al. (2006), Brewin et al. (2010), Brewin et al. (2015), and Hirata et al. (2011) use satellite chl-a as input to derive PSC or PT based on empirical relationships linking *in situ* marker pigments to chl-a which are determined using high precision liquid chromatography (HPLC). Abundance-based approaches use satellite chl-a as input and by that exploit the largest signal in water leaving radiance to extract variability due to PG out of chl-a. This is then a simple calculation and can be applied easily to chl-a products from different sensors. However, they cannot predict atypical associations and may not hold in a future ocean.

Another class of algorithms relies on spectral features in reflectance, absorption, and/or backscattering spectra caused

TABLE 2 | A compilation of global algorithms to retrieve phytoplankton composition from satellite data.

Approach		Phytoplankton composition product		References
ABUNDANCE		Size classes		Uitz et al., 2006; Brewin et al., 2010, 2015
		Size classes and multiple taxa		Hirata et al., 2011
SPECTRAL	REFLECTANCE	Multiple taxa		Alvain et al., 2005, 2008; Li et al., 2013; Ben Mustapha et al., 2014
		Single taxon	Coccolithophores	Brown and Yoder, 1994; Moore et al., 2012
			Trichodesmium	Subramaniam et al., 2002; Westberry et al., 2005
	ABSORPTION	Size index		Ciotti and Bricaud, 2006; Mouw and Yoder, 2010; Bricaud et al., 2012
		Size classes		Devred et al., 2006, 2011; Hirata et al., 2008; Fujiwara et al., 2011; Roy et al., 2013
		Multiple taxa		Bracher et al., 2009; Sadeghi et al., 2012a; Werdell et al., 2014
	BACK-SCATTERING	Size classes		Kostadinov et al., 2009, 2016; Fujiwara et al., 2011
ECOLOGICAL		Taxonomic groups		Palacz et al., 2013

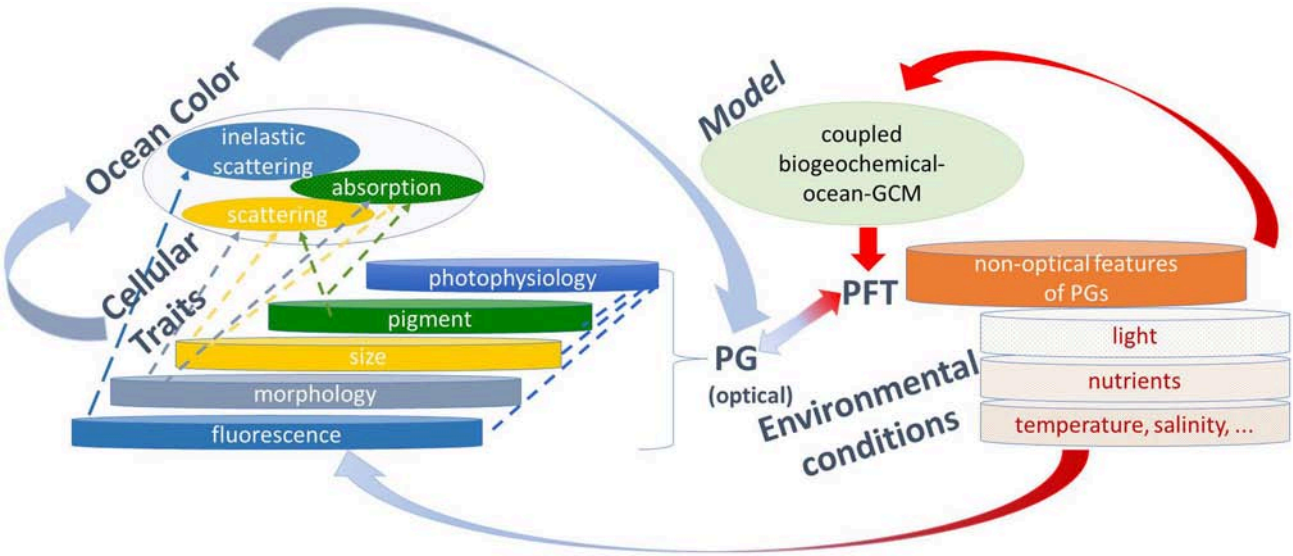


FIGURE 1 | Illustration of phytoplankton diversity as found in nature impacted by environmental conditions, and how it can be derived from observations and modeling. Through *in situ* measurements (which represent the most real conditions), phytoplankton are grouped according to cellular traits that influence their optical properties such as pigments, size, morphology, and fluorescence, all also responding to photophysiology, which are named optical features of phytoplankton groups (PG). In addition, inferences can be made about PG through non-optical features, such as nutrient requirements, stoichiometry, etc. The optical properties can be measured by ocean color and used to infer PG from remote sensing (highlighted by blue arrows). Coupled biogeochemical-ocean general circulation models (GCM) produce projections of phytoplankton functional types (PFT) which are, with PG classified according to functions, mainly incorporating non-optical and rarely optical properties (highlighted by red arrows). PG information from ocean color and ecosystem models can be combined (highlighted by blue-red arrows) to improve our knowledge. For instance, ocean-color PG can be used for model improvements and evaluation, and models could be re-developed to explicitly include optical properties of which the ocean-color PG use which will help to advance the application of ocean color PG.

by the variation in phytoplankton structure and pigment composition (Brown and Yoder, 1994; Subramaniam et al., 2002; Alvain et al., 2005, 2008; Westberry et al., 2005; Ciotti and Bricaud, 2006; Devred et al., 2006, 2011; Hirata et al., 2008; Bracher et al., 2009; Kostadinov et al., 2009, 2016; Mouw and Yoder, 2010; Fujiwara et al., 2011; Bricaud et al., 2012; Moore et al., 2012; Sadeghi et al., 2012a; Li et al., 2013; Roy et al., 2013; Ben Mustapha et al., 2014; Werdell et al., 2014). Spectral-based approaches exploit as much of the backscattered spectrum observed by satellite as necessary to extract the signatures of specific PG to ocean color. Generally, these methods are computationally much more expensive and require specific adaptations for each sensor. However, these algorithms rely on much less empirical relationships than the abundance based approaches and are based on physical principles (radiative transfer). Differences exist on the different satellite inputs (e.g., radiance, absorption, backscattering) and the underlying principles (for a comprehensive overview Mouw et al., 2017). Another approach incorporates various environmental parameters to predict PT based on their ecological preferences (Raitsos et al., 2008; Palacz et al., 2013). This method uses artificial neural networks to link the different biological and physical data sets. While the approach of Raitsos et al. (2008) was regionally developed for the North-Atlantic, the approach by Palacz et al. (2013) is not purely based on remote sensing data but also requires a coupling to a dynamic plankton model.

Products obtained from the PG algorithms (**Table 2**) are typically dominance (Brown and Yoder, 1994; Alvain et al., 2005; Moore et al., 2012; Ben Mustapha et al., 2014), presence or absence of a certain PT (Westberry et al., 2005; Werdell et al., 2014), fraction or concentration of chl-a of the three PSC (Devred et al., 2006, 2011; Uitz et al., 2006; Hirata et al., 2008, 2011; Kostadinov et al., 2009, 2016; Brewin et al., 2010, 2015; Fujiwara et al., 2011; Li et al., 2013; Roy et al., 2013) or a size factor characterizing the contribution of pico- (or micro-) phytoplankton to the phytoplankton community (Ciotti and Bricaud, 2006; Mouw and Yoder, 2010; Bricaud et al., 2012). Currently, only the products OC-PFT (Hirata et al., 2011) and PhytoDOAS (Bracher et al., 2009; Sadeghi et al., 2012a) enable the simultaneous determination of chl-a for several PT. PhytoDOAS retrieves the imprints of absorption characteristics of specific phytoplankton groups among all other atmospheric and oceanic absorbers from top of atmosphere data of the hyperspectral satellite sensor SCIAMACHY (Scanning Imaging Absorption Spectrometers for Atmospheric Chartography). All other satellite-based PG algorithms (**Table 2**) have been applied to water-leaving reflectance data from multispectral sensors [e.g., SeaWiFS (Sea-viewing Wide Field-of-view Sensor), MERIS (Medium Resolution Imaging Spectrometer), MODIS (Moderate Resolution Imaging Spectroradiometer)].

To be able to detect unexpected changes in phytoplankton community composition, satellite PG data based on exploiting the spectral signatures, and based on limited empirical assumptions are preferred. In the few last years, radiative transfer models (RTM) have been used to develop and assess the sensitivity of analytical (spectral) PG retrievals or to find

suitable spectral characteristics necessary for ocean color sensors to retrieve PG. Werdell et al. (2014) and Wolanin et al. (2016) used the GIOP (Generalized Inherent Optical Property) model software (Werdell et al., 2013) to invert reflectance spectra (either water-leaving or top of atmosphere), and Wolanin et al. (2015) used the coupled ocean-atmosphere RTM SCIATRAN (Rozanov et al., 2014) to test the sensitivity of a PT retrieval (PhytoDOAS). Evers-King et al. (2014) and Xi et al. (2015) used the ocean RTM HydroLight (Sequoia Scientific.) to specifically model the variation of composition of PSC or certain (dominant) PT, respectively, and assessed the potential of retrievals in different water types. Werdell et al. (2014) optimized the inversion scheme of GIOP to finally retrieve absence or presence of *Noctiluca miliaris* from MODIS data, while Wolanin et al. (2016) used this method to identify optimal band placements for multi- and hyper-spectral satellite data for successful retrievals of certain PT. The results of this study indicate that four additional bands (381, 473, 532, and 594 nm) for the Ocean and Land Colour Instrument (OLCI) would potentially enable absorption-based quantitative retrievals of diatoms, cyanobacteria, and coccolithophores. Recent methods have been developed to retrieve PG from *in situ* hyperspectral algal or particulate absorption coefficients, and validated using *in situ* measurements (Moisan et al., 2013; Organelli et al., 2013; Zhang et al., 2015). As absorption coefficients can be estimated from satellite measurements using inverse bio-optical models, this opens the way to applications of these methods to satellite data.

Some PG algorithms (most of the ones listed in **Table 2**) have been inter-compared at the global scale: firstly using *in situ* PSC (derived from HPLC pigments) in terms of dominance (Brewin et al., 2011) and secondly, under the 2nd Satellite PFT Algorithm International Intercomparison Project: http://pft.ees.hokudai.ac.jp/satellite/index.shtml. The initiative strengthens the links between algorithm developers at a global scale which will help also to guide modelers and policy makers on the specific assumptions underlying each product: the inter-comparison among most algorithms presented in **Table 2** and to an ensemble mean of Earth System Models is presented in Kostadinov et al. (2017). A user guide for application to open ocean waters on the most common algorithms (Mouw et al., 2017) explains the current global PG algorithms and their associated uncertainties and also includes a discussion on the advantages and disadvantages of these algorithms. A global *in situ* dataset of HPLC and optical properties is being developed to further evaluate these algorithms. This initiative organized and held breakout sessions at the International Ocean Color Symposia (IOCS) in 2013 and 2015 and at a specific expert International Ocean-Colour Coordinating Group (IOCCG) and National Aeronautics and Space Administration (NASA)-sponsored workshop in 2014, which focused on PG algorithms development, validation, and user needs. For each meeting the outcome resulted in a written summary of recommendation for community actions and the planning of future activities (see IOCS, 2013; Bracher et al., 2015a; IOCS, 2015) which also form the baseline for Section Recommendations toward Operational Products of Phytoplankton Diversity from Space.

To date, the majority of existing PG satellite retrieval approaches have relied on HPLC pigment data to derive *in situ* PG data: for developing and validating algorithms large *in situ* PT (e.g., Alvain et al., 2005; Hirata et al., 2011; Soppa et al., 2014; Swan et al., 2016) and PSC (e.g., Uitz et al., 2006; Brewin et al., 2010) data sets have been complied, complemented by the global pigment data set compiled under the MAREDAT project (Peloquin et al., 2013) and recent submissions to public data bases: e.g., SEABASS (http://seabass.gsfc.nasa.gov/), BODC (http://www.bodc.ac.uk), LTER Network Data Portal (https://portal.lternet.edu/nis/home.jsp), PANGAEA Data Publisher for Earth & Environmental Science (https://www.pangaea.de). Among all available *in situ* PG data sets, HPLC-phytoplankton pigment data contain the largest number of observations resulting in the greatest spatial coverage with standardized quality control protocols (Hooker et al., 2012). However, size fractionated *in situ* data of chl-a serve as a more direct validation data set for assessing satellite retrievals on PSC (e.g., Brewin et al., 2014). The long-term and spatially extended Continuous Plankton Recorder (CPR) data sets have been used for constructing and evaluating ecological algorithms focusing on larger phytoplankton (Raitsos et al., 2008). The CPRs, especially with the recent global data effort (Global Alliance of CPR Surveys, http://www.globalcpr.org/), may provide a unique platform on taxonomic information to modern satellite sensors for several oceanic regions around the globe. Inline (coupled) flow cytometry and microscopy techniques have been developed and enable a more precise classification of the phytoplankton groupings than HPLC marker pigments (e.g., Sosik and Olson, 2007). In addition, phytoplankton group specific Inherent Optical Properties (IOPs, i.e., absorption and backscattering) determined in the field have been used as algorithm inputs for several spectral approaches (Ciotti and Bricaud, 2006; Bracher et al., 2009; Mouw and Yoder, 2010; Fujiwara et al., 2011; Sadeghi et al., 2012a). Hyperspectral IOP measurements when obtained via continuous measurements (e.g., Boss et al., 2013) can help validating satellite-derived PG by increasing the number of match-ups, assessing variability within a satellite pixel, and quantifying the uncertainties in the two-step satellite methods (i.e., from water-leaving reflectance to IOP to PG).

Satellite PG time series data have already been used to assess regionally and globally the variability and trend of phytoplankton community composition (e.g., Brewin et al., 2012; Sadeghi et al., 2012b) and PG phenology (Alvain et al., 2013; Soppa et al., 2016a) also linking to environmental variables. In addition, satellite PG data were used to assess globally particulate organic carbon export (Mouw et al., 2016), for detection of regional HAB events (e.g., Kurekin et al., 2014), the estimation of recruitment of juvenile fish (Trzcinski et al., 2013) and for inferring globally oceanic emissions of volatile organic compounds (Arnold et al., 2009; Booge et al., 2016).

All ocean color data are limited in coverage to sun-light, cloud and ice-free conditions, and only deliver information on the surface ocean (first optical depth which is 4.6 times shallower than the euphotic depth). Therefore, for many applications, additional methods have to be used to resolve variability and trends of phytoplankton community structure and abundance.

Satellite-derived algorithms are increasingly compared not only to *in situ* data but to global model output from Earth System Models. Starting over two decades ago, biogeochemical models began incorporating multiple PT (e.g., Baretta et al., 1995; Le Quéré et al., 2005; Gregg and Casey, 2007) mainly to incorporate their biogeochemical relevance. As a first step models incorporated a "diatom" group given their importance in the silica cycle, but also given their potential important role in carbon export compared to other PT (e.g., Chai et al., 2002). As models became more sophisticated and started to simulate nitrogen biogeochemistry, many added a "diazotroph" class. Given the biogeochemical importance of these groups of phytoplankton, the modeling community refers to them as PFT (see **Figure 1** right). This classification is the closest to ocean color PT products as defined above. Though less common, models can also group phytoplankton in terms of size: the model of Ward et al. (2012) includes 25 size classes of phytoplankton. The advantage of such an approach is that it can use empirical allometric relationships of key growth parameters (e.g., maximum growth rates). Such a model output is more compatible with ocean color PSC products.

Since 2009, marine ecosystem modelers collaborate systematically with the remote sensing community in the MARine Ecosystem Model Intercomparison Project (MAREMIP). MAREMIP fosters the development of models based on PFT. Complementary to the Coupled Model Intercomparison Project Phase 5 and Phase 6 efforts (see http://cmip-pcmdi.llnl.gov/), MAREMP thus specifically targets the inter-comparison of the representation of current and future marine biology in global ocean models, and promotes the interactions between modelers and observationalists and the development of targeted observations. MAREMIP, as well as many single model studies conducted by marine ecosystem modelers worldwide (e.g., Ye et al., 2012; Dutkiewicz et al., 2015), have been using satellite-derived PT products for the evaluation of model performance in terms of plankton biogeography and global biogeochemical cycling (e.g., Hashioka et al., 2013; Vogt et al., 2013; Laufkötter et al., 2015). Initial studies have shown that models and satellite estimates of phytoplankton biogeography diverge, for example (a) in the timing of the phytoplankton bloom (Hashioka et al., 2013), (b) in phytoplankton dominance patterns and the global contribution of diatoms to total phytoplankton biomass (Vogt et al., 2013), and (c) in net primary production (Laufkötter et al., 2015).

GAP ANALYSIS

Current satellite data sets on phytoplankton composition (PG) are not generally available in a format readily adoptable by a wide user community. Some potential users (e.g., fishery managers) still use chl-a rather than satellite PG data, in part due to a lack of confidence in the PG products, and climate modelers use only a limited fraction of the currently existing products, due to the lack of uncertainty estimates associated with each product, and issues related to the compatibility between model and satellite output. In the following section, we detail the gaps, which need to be addressed if we want to respond to user needs and promote the use of a wider range of new remote sensing products (see summary in **Table 3**, left columns).

Gap 1: Information Mismatch between Satellite-Derived Phytoplankton Composition Products and User Group Target Variables

At present, there is a mismatch between the PG detected from satellite (and which differ between algorithms) and the groupings required by the user community. **Figure 1** illustrates PFT as they may be found in the environment, and which respond to environmental conditions based on the interplay of different variables (nutrients, temperature, salinity, light, and others). Optical (size, morphology, pigmentation, fluorescence) and non-optical (e.g., nutrient requirements, stoichiometry) properties of phytoplankton allow for distinctive groupings. The optical properties include photo-physiological responses which are driven by photoadaptation associated with certain P**G**, and photacclimation which is mostly independent of PG. From ocean color (**Figure 1** top-level left) absorption, scattering, and fluorescence properties of different PG can be derived. Coupled biogeochemical ocean models (**Figure 1** top-level right) often use groupings in terms functional groups (e.g., calcifiers, nitrogen fixers, etc.) which necessarily do not link to the optically based PG which are for example, either picophytoplankton (PSC algorithms), prokaryotic phytoplankton (e.g., Bracher et al., 2009), *Synechococcus* like cyanobacteria (e.g., Alvain et al., 2005), prochlorophytes (e.g., Hirata et al., 2011) or *Trichodesmium* only (e.g., Westberry et al., 2005) and not just nitrogen fixers. This highlights the need to enhance linkages between optical and functional PG to improve our knowledge. The algorithms listed above also provide results in different units (e.g., size factor, fraction of total chl-a, chl-a, dominance or just presence of a PG), which do not always match with users requirements (e.g., a numerical model might require carbon biomass). There are also substantial differences in the PG definitions among the users themselves (see IOCCG, 2014). While biogeochemical and RT models require a quantitative assessment of PT or PSC, end users for coastal environmental management need PG products as indicators for water quality, HAB presence, eutrophication and fisheries stock assessment. To help users selecting the appropriate PG data sets, the work already accomplished by inter-comparing (Kostadinov et al., 2017) and by setting up a user guide (Mouw et al., 2017) on global satellite PG needs to be extended to new algorithms and more explicit recommendations on which algorithm is best suited for specific users and science questions. The later can only be done when the uncertainties of these algorithms have been evaluated more consistently (see Gap 2). Improvements are also needed in terms of the representation of PG in the current generation of models to better constrain present and future projections of marine biogeochemistry. Furthermore, as the community is moving toward biogeochemical models of increased complexity, information on phytoplankton community composition from space including all PT, or other indices of biodiversity (pigments, size) will provide valuable resources for

TABLE 3 | Summary of gap analysis for phytoplankton composition from space: gap (left), status of existing work (second left), and recommendations for actions (right columns).

Gap	Status	Medium-term action	Long-term action
Mismatch	2nd PFT Intercomparison initiative (too little funding): inter-comparison, user guide, *in situ* data base	Workshops with users, *in situ* and algorithm developers, ecologists	Mechanistic frame-work to connect complementary different PG to PFT
	Some application of satellite PG for models have started	Extend Website and user guide on new PG products	
Uncertainties	No appropriate *in situ* data: Global HPLC-but not really PG	Round-robins on PG data format and quality standardization, method standardization	Curate existing *in situ* PG data bases and time series programs
	Other PG data require integration	Exploit additional *in situ* PG data	
	Difficult merging data bases	Support new sensor validation	Ensure *in situ* data acquisition and validation of new OC missions
	Spectral IOPs (esp. b_b) limited		Develop further inline and autonomous techniques
	Sparse *in situ* data can't resolve sub-pixel variability	Add AOP&IOP to PG *in situ* data	
	Deficient theoretical background: Inversion limited success, RTM lack	Use complementary data to constrain algorithms	Framework for clear traceability of uncertainties
	PG information for all water types Limited traceability of errors		
Satellite sensors	Multispectral sensors with limited PG: few bands do not resolve optical difference of PG	Develop atmospheric correction for hyperspectral sensors	Exploit adding bands to multispectral—OLCI
	Hyperspectral sensors: SCIAMACHY and OMI PG data but low spatial resolution	Adapt PT algorithms to Sentinels TROPOMI and UVN (high spatial coverage and resolution)	Merge all sensors' PG data for long term high coverage info
	No L_w, R_{RS} data (since no AC)	Develop synergistic PT products from hyper and multispectral data	Launch hyperspectral OC sensors (e.g., PACE)
	HICO: no PG, low coverage		
Regional capability	Satellite PG (mostly) only global	Exploit additional data to constrain algorithms	Launch high spatial resolution of multi-spectral (e.g., MSI) and hyperspectral (EnMAP, HyspIRI) sensors
	AC and standard OC products poor quality in complex waters	Round-robins for PG algorithm development and validation specific to regions	

Some actions are related to several gaps but are only stated at the medium term (second right) with agency supported activities embedded in long-term actions at international level (right); abbreviations in ***Table 1****.*

the next generation modelers. Thus, there is a need for on-going product development along with effective communication between remote-sensing scientists, biological oceanographers, and modelers to ensure future developments are consistent and comparable between parties and that ultimately improve climate predictions.

Gap 2: Lack of Traceability of Uncertainties in PG Algorithms

The quantitative assessment of uncertainty in PG satellite products is still insufficient. This is due to the above mentioned mismatch definition (see Gap 1), the limited theoretical background to connect optical signatures to diversity of phytoplankton communities across different environments and limitations in appropriate *in-situ* data.

At the cellular level, a detailed understanding of how pigment packaging (function of cell size and intracellular pigment concentration; Morel and Bricaud, 1981) and pigment composition that both govern the shape and magnitude of chl-a specific absorption (especially in the blue-green regions of the spectrum, which is commonly used in PG algorithms) requires further work. Both reconstruction (Bidigare et al., 1990) and decomposition (Hoepffner and Sathyendranath, 1993) methods are often applied separately to bio-optical datasets to explore the link between pigments and phytoplankton absorption. Reconstruction approaches conventionally apply a single pigment-specific absorption coefficient to a particular pigment or pigment type (e.g., photosynthetic and photoprotective carotenoids), often obtained from measurements of extracted pigments in solvent. Only a handful of studies have examined the absorptive properties of pigment-protein complexes (e.g., Johnsen and Sakshaug, 2007), yet differences in the spectral shape once pigments are embedded in proteins can be significant. Improved models on phytoplankton photoacclimation combined with new approaches in determining cell size should assist in improving our understanding of how pigment packaging

influences the spectral signature of natural phytoplankton assemblages. Efforts inverting hyperspectral reflectance and absorption spectra to obtain PG have shown limited success, leading to identification of certain PT with no quantification (Werdell et al., 2014; Kudela et al, 2015; Xi et al., 2015) of PSC fractions (Organelli et al., 2013) or quantification of accessory pigments in addition to chl-a (Chase et al., 2013; Moisan et al., 2013; Bracher et al., 2015b). PT specific absorption properties are available but large spectral variability is related to algal culturing and variations in size, pigment composition and pigment packaging due to physiological responses of PT. In contrast, due to high measuring uncertainties spectral scattering properties (including back-scattering and volume scattering function) are still less known (Tan et al., 2015; Harmel et al., 2016). Thus, PG related specific IOPs are not adequately represented in RTMs. This further limits tracing uncertainty in algorithms, pointing to the need for coincident IOP observations along with expended *in situ* datasets of phytoplankton composition.

Other errors in algorithms are also difficult to assess, for instance the accuracy of *in situ* data used as input or for validation of algorithms due to mostly non-standardized acquisition (see details below), the above mentioned mismatch definition (see Gap 1), and the spatial and temporal upscaling of specific PT and PSC signatures of diverse communities. Several studies have demonstrated that adding spectrally-resolved optics to biogeochemical models improves model skill (e.g., Dutkiewicz et al., 2015) as well as comparability to observed optical properties (e.g., Fujii et al., 2007; Baird et al., 2016). A minority of global numerical models resolve the bio-optical properties of different PG (Gregg and Casey, 2007; Dutkiewicz et al., 2015; Baird et al., 2016). These advancements may provide a way forward to investigate the biological realism of phytoplankton biogeography using a larger range of satellite PG products.

In addition, the reliance on HPLC for development and validation presents challenges to quantitatively assess the uncertainty of PG satellite products. However, inputs to HPLC PG datasets are (diagnostic) accessory pigment concentrations, which are only to a certain degree congruent with taxonomy or phytoplankton size. Size can vary considerably within certain functional or taxonomic groups, e.g., diatoms can range from 3 to 500 μm but are characterized by the same diagnostic pigment (fucoxanthin) across this size range. Similarly, grouping by accessory pigments can be problematic as there is substantial variability in pigment concentration as a function of physiological response to the environmental conditions and more importantly a given biomarker pigment is present in several PT (e.g., fucoxanthin in diatoms and haptophytes). Some PT, e.g., coccolithophores, cannot be inferred from HPLC pigments. In consideration of the expanding satellite sensor capabilities, there is a need for coordinated efforts to compile and generate comprehensive *in situ* datasets (not just HPLC) for assessing phytoplankton composition. There is also a need to provide best practice guidance to merge the different types of datasets (e.g., HPLC, microscopy, flow cytometry) into an integrated product that encompasses different ways of grouping phytoplankton species.

Gap 3: Missing Capabilities of Current Ocean Color Satellite Measurements

Differences among PT in their spectral absorption are small: many PT contain, despite specific marker pigments, the same suites of pigments or pigments of similar absorptive properties (note that besides the pigment absorption properties, spectral absorption is also ruled by the algal community size structure). Given the limited number of wavebands and the broad band resolution of current multi-spectral sensors can provide only limited information on the variability in phytoplankton spectral absorption caused by shifts in community structure (Bricaud et al., 2004; Organelli et al., 2011). This restricts all multispectral satellite phytoplankton composition products based on spectral principles to either indicating dominance, presence of PT or identifying major size class fractions within the total phytoplankton community to a high level of uncertainty.

Satellite instruments, with a very high spectral resolution (1 nm and better, originally designed for atmospheric applications), provide additional opportunities for distinguishing multiple PT based on their optical properties. The capability to retrieve quantitatively major PT groups based on their optical signature has been clearly shown with the PhytoDOAS method (Bracher et al., 2009; Sadeghi et al., 2012a) in the open ocean using hyperspectral satellite data from the atmospheric sensor SCIAMACHY. However, the exploitation of hyperspectral satellite data for ocean color has been so far very limited because hyperspectral sensors like SCIAMACHY (spectral resolution $<$0.5 nm) do not provide operational water-leaving radiance products and have very large foot-prints (30 by 60 km per pixel) and low global coverage (6 days). This provides a major constraint on assessing the retrieval's accuracy with *in situ* point measurements. It also limits the application of such PT satellite data sets. The difficulty of working with SCIAMACHY data is that one has to handle strong atmospheric absorbers (true for all hyperspectral satellite data) and the heterogeneity of big pixels; hence, the PhytoDOAS algorithm was designed to retrieve three PT directly from top of atmosphere radiances, by separating their high frequency absorptions from each other and relevant atmospheric absorbers, while accounting for broad band effects by using a low order polynomial. This method requires high spectral resolution ($<$1 nm). SCIAMACHY data acquisition ended with the lost contact to the ENVISAT satellite (April 2012). First results from adapting PhytoDOAS to the Ozone Monitoring Instrument (OMI) sensor (measuring since 2004) are very promising (Oelker et al., 2016) and will enable the extension of the spectrally derived PT data into the future with much improved global coverage (daily) and smaller foot print (13 $\times$ 24 km). OMI is also the precursor instrument to the in 2017 launched Sentinel-5-Precursor (S-5-P) with TROPOspheric Monitoring Instrument (TROPOMI) and in the 2020s launched Ultra-violet/Visible/Near-Infrared Instrument (UVN) instruments on Sentinel-4 and Sentinel-5 (all with a pixel size of 3.5 $\times$ 7 km).

The Hyperspectral Imager for the Coastal Ocean (HICO) provided data with high spatial (100 m) and spectral ($\sim$6 nm) resolution and limited coverage (only a restricted number of scenes globally). However, so far lack of robust atmospheric

correction for HICO (see current implementation in http://seadas.gsfc.nasa.gov/) has prevented the exploitation of the full spectrum. Eventually, not much more than standard phytoplankton information (chlorophyll, fluorescence line height) as for multispectral data was derived (Ryan et al., 2014). It is a big challenge to provide spectrally consistent high quality atmospheric correction for PG retrievals.

The new ocean-color sensor OLCI on Sentinel-3 already provides two more bands in the visible range than its predecessor MERIS. It is anticipated that the number of bands will further increase for future multispectral ocean-color sensors. In addition, hyperspectral missions like Pre-Aerosol, Clouds, and ocean Ecosystem (PACE; global, high coverage, 1 km pixels, launch 2022) and Environmental Mapping and Analysis Program mission (EnMAP; regional, low coverage, 30 m pixels, launch 2019) are planned for operating in the near future. However, hyperspectral instruments like ENMAP or PACE with 5 nm resolution are still very different from atmospheric instruments like SCIAMACHY. Hence, algorithms will have to be developed (or adapted) to retrieve the PT from these new instruments.

To monitor marine ecosystem and assess their vulnerability to future anthropogenic and climate change, beyond a good spatial-temporal resolution of existing data long-term time series data are needed to monitor trends in phytoplankton community structure, and its variability on inter-annual to decadal time scales. The average cloud-free repeat time per pixel for an ocean color sensor is only 100 observations per year for the temperate and tropical zones (Werdell et al., 2007), while it is much lower for high latitudes (e.g., 12 per year for the East Greenland Sea, Cherkasheva et al., 2014). Yet, merged ocean-color products significantly increase this temporal coverage (Maritorena et al., 2010; Racault et al., 2015). The development of long time series of PG satellite products, covering more than a decade, has just started (e.g., references given in Mouw et al., 2017). Such data sets are necessary to respond to user needs. Efforts have been taken to apply the multispectral PG algorithms not only to SeaWiFS but also to MODIS and MERIS. Synergistic use of multiple sensors will enable creating long-term time series moving from monthly to daily resolution, but also provides an opportunity to improve performance of individual retrievals. The ESA project SynSenPFT is an example for that where an algorithm was developed by synergistically using PT information from SCIAMACHY-PhytoDOAS and Ocean Color Climate Change Initiative chl-a-OC-PFT retrievals. This was done to obtain high spatially and temporally resolved PT chl data using their spectral imprints retrieved from high spectrally resolved satellite data and a global PT data set was developed, from 2002 to 2012 on 4 by 4 km daily resolution (Soppa et al., 2016b).

Gap 4: Lack of Regional Capability of PG Algorithms

Thus far, most PG algorithms work globally or some of them have been validated on restricted regions, but nearly all are limited to open ocean conditions. A spatial pattern matching between modeled- and satellite PG showed a relatively large discrepancy on smaller spatial scales than larger scales, especially around continental shelves (Hirata et al., 2013). However, PG satellite products retrieved are necessary especially for coastal areas and inland waters where water quality and HABs issues are most urgent. In these optically complex waters, optical constituents vary independently making ocean color retrievals challenging. In extremely high colored dissolved organic matter (CDOM) and low scattering waters, CDOM absorption dominates the whole visible spectrum resulting in very low water-leaving reflectance (<1%) and thus, the phytoplankton signal itself is weak. By contrast, the main problems in highly scattering waters are the masking of pigment absorption by non-algal (mineral) particle absorption and significant near infrared water reflectance (IOCCG, 2000). Successful results in these types of water are hampered by limited spatial and spectral resolution of sensors. This already makes it difficult to achieve accurate atmospheric correction and obtain reliable ocean color standard products. It also inhibits the observation of the patchy distribution of phytoplankton communities. To derive certain PT beyond size- and/or pigment-based discrimination of phytoplankton requires developing empirical methods that rely on covariation: Via the exploitation of additional data (light, temperature, nutrients, ...), retrievals and optical modeling for specific regions could be further constrained (and optimized), as for example in the study by Brewin et al. (2015) where information on ambient light field extracted from satellite information was combined with an abundance based PSC approach.

RECOMMENDATIONS TOWARD OPERATIONAL PRODUCTS OF PHYTOPLANKTON DIVERSITY FROM SPACE

In the following we give recommendations how to fill the gaps identified in the previous chapter. **Table 3** summarizes the mid- and long-term actions that are detailed below. Note, that several actions will address several gaps simultaneously. We recommend that the implementation of these actions is done in communication and collaboration between ocean color scientists, observationalists, numerical modelers, and other users. This will ensure that products are aligned to new *in situ* and satellite observational techniques and fulfill the ever changing needs of the wide range of user communities.

Improving Match between Satellite PG and Users' Needs

A mechanistic framework needs to be developed which draws the complementary use of the various PG data and links them to PFT (**Figure 1**). This will assure that users are aware of the actual specific groups in the different satellite products and how they compare to the groups they require in their specific application. Such a framework requires an international effort and funding, including experts in *in situ* measurements (HPLC, microscopy, flow cytometry, genetics, bio-optics), algorithm developers and representatives of the user communities (modeling, marine services). Certain medium-term actions should be taken:

– Regular workshops to improve communications between users (biogeochemical-, ecosystem-, RT-modelers, taxonomists, ecologists, fishery managers, HAB, and water quality experts) and algorithm developers should be held to achieve a common understanding and a consistent and comprehensive definition of the "groups" in PG algorithms, but also of their metrics [% vs. chl-a (or carbon) vs. dominance]. This could potentially lead in joint future proposals.
– A website informing on PG algorithms activities and development (user guide, algorithm inter-comparison and validation protocols, forum, ...) should be sustained.
– A "living" user guide (regularly updated) should describe available PG and PSC algorithms and satellite products including their definition, uncertainty, strengths, and limitations.

Curation and Acquisition of *In situ* Data for Improving and Assessing PG Retrievals

Within international cooperation of space agencies the curation of existing measurements of *in situ* PG abundance (HPLC, microscopy, flow cytometry, particle imaging, genomics, ...) and corresponding optical [IOPs, apparent optical properties (AOPs)] data needs to be secured:

– Specific comprehensive datasets should be selected that include coincident IOPs, AOPs, and phytoplankton composition that serve as a resource for PG algorithm development, refinement, and validation, and improve the ability to inter-compare validation metrics.
– Standardized data quality, nomenclature, and format among different data bases (e.g., SEABASS, LTER, PANGAEA, BODC, ...) should be assured to enable easy compilation and expansion.
– Pigment databases used for testing, validating and refining PT algorithms should be archived along with complete information on all detected pigments to allow rigorous inter-comparison studies and sensitivity analysis.
– Methods to convert from *in situ* data to PG biomass or fractions should be assessed and protocols for merging different datasets (e.g., HPLC, microscopy, ...) should be formulated.
– Existing time series sites of phytoplankton composition collected at fixed sites, covering a range of oceanic regimes, including the coastal ocean should be utilized to inter-compare algorithms and to assess their uncertainties considering seasonal and multiannual effects.
– *In situ* measurements need to be calibrated and standardized to advance the knowledge of phytoplankton composition *in situ* with defined uncertainties.
– Existing relevant (hyperspectral and/or PG related) IOP data from laboratory studies need to be curated, so they can be used as algorithm input and to develop the necessary theoretical background.
– Agency directed programs need to sustain the *in situ* data acquisition in order to secure assessment of accuracy of PG and related (e.g., atmospheric correction, CDOM, total suspended matter) satellite products. The following is recommended:
– The protocols of data acquisition need to be standardized by internationally run round-robin exercises and calibrations. Currently, various NASA WGs are updating the ocean optics protocols (Mueller et al., 2003), specifically for HPLC pigments and IOPs which should be supported also by other agencies.
– Methodological errors associated with different approaches of measuring phytoplankton absorption (filter pad method and methods measuring directly a water sample) need to be identified and quantified. This also includes the inter-comparison of *in situ* and bench top techniques over a range of environmental settings (optically complex to oligotrophic open ocean). For that, universal protocols for data collection and processing need to be established.
– Methodological errors associated with different approaches of measuring phytoplankton absorption (filter pad method and methods measuring directly a water sample) need to be identified and quantified which also include the inter-comparison of *in situ* and benchtop techniques over a range of environmental settings (optically complex to oligotrophic open ocean). For that universal protocols for data collection and processing need to be established, such as those currently prepared by NASA.
– Differences in various particle imaging and identification technologies (e.g., holography, flow cytometry, flow cam, etc.) need to be assessed.
– Following the launch of new ocean color sensors (particularly now OLCI on S-3), there is a need for validation activities to be funded across all AOPs and IOPs to PG products.
– Additional phytoplankton composition observational capability needs to be added to existing time series sites.
– Target locations for future field sampling (informed by existing products), uncertainty assessments, and potentially supported by modeling need to be identified.
– New technologies for in-line and autonomous measurements should be supported via the use of robotic platforms (e.g., profiling floats, autonomous surface water vehicles) to increase the spatial and vertical coverage of measurements. The development of new miniature sensors that can be deployed on these platforms to provide accurate measurement of phytoplankton community and carbon (e.g., miniature imaging flow cytometers, sensors for metagenomics hyperspectral IOPs and AOPs) should also be supported to ensure appropriate evaluation of satellite PG products performance on their spatial and temporal resolution.

Theoretical Background to Further Develop PG Retrievals and Assess Their Uncertainty

To fill Gap 2, the development of a framework for clear traceability of uncertainties in PG satellite products needs to be supported by the specific assessment of mismatch definition, *in situ* error, retrieval error, or errors due to the spatial and temporal upscaling of specific PG signatures in diverse communities. This requires the steps mentioned in Section Improving Match between Satellite PG and Users' Needs and

in Section Curation and Acquisition of *In situ* Data for Improving and Assessing PG Retrievals, but also steps linked to improving PG algorithms (see also Section Sustaining Long-Term PG Satellite Data) which require a solid theoretical background.

On one hand inverse modeling needs to be optimized by further developing the theoretical background to connect optical signatures to diversity of phytoplankton communities across different environments (especially in optically complex waters). The degree of independent information in hyperspectral wavelength signals will depend on the water type (e.g., waters optically dominated by phytoplankton alone, or other particles, or CDOM) and will determine whether different phytoplankton products can be independently derived from a given hyperspectral spectrum. This statistical-informational problem needs to be considered in the application of (global or regional) inversion algorithms. Measurements on spectral specific IOPs (in particular scattering properties) on natural and cultured samples will lead to a better description of optics in RTM. On the other hand, developing a mechanistic understanding of the spectral properties of PG to retrieve bio-optical indices of diversity requires a better utilization of expertise crossing a wide range of fields, including taxonomy and molecular ecology in connection with optically-derived PG. The usage of global numerical models which resolve the bio-optical properties of different PG will provide a way forward for connecting more specifically a larger range of satellite PG products (highlighted in **Figure 1** with red-blue arrow connecting optical PG with PFT). Models could group their "model phytoplankton-analogs" according to more dimensions of diversity (e.g., accessory pigments, scattering characteristics, etc.—see optical PG in **Figure 1**) that link closer to the satellite PG definitions than the more classical PFT designations (e.g., nitrogen fixers, silicifiers,...). Models that include spectrally-resolved optics and bio-optical properties of phytoplankton could also prove to be a powerful tool for exploring the inter-dependency and regionally varying skill of different satellite PG approaches.

Sustaining Long-Term PG Satellite Data

As outlined in Section Gap 3: Missing Capabilities of Current Ocean Color Satellite Measurements, long-term data sets of sufficient spatial and temporal resolution are needed to be established which are also adequate for regional applications (see Section Gap 4: Lack of Regional Capability of PG Algorithms).

At first, the exploitation of hyperspectral data needs to be intensified in order to base those data on deriving the spectral imprints of phytoplankton groups in ocean color:

- In preparation for the exploitation of future hyperspectral ocean color sensor [PACE, EnMAP or Hyperspectral InfraRred Imager (HyspIRI) and hopefully more] missions, much more effort needs to be put into the development of atmospheric correction for hyperspectral satellite data; methods should be developed over open ocean and complex waters, with the help of RTM, and considering multispectral atmospheric correction methods. Also current hyperspectral satellite data sets, such as SCIAMACHY, HICO, OMI (and from 2017 also TROPOMI), should be explored further.
- Hyperspectral and quantitative spectral PT and PSC algorithms should be adapted, extended and applied at global and regional scales to data with high spectral resolution from former, current and near-future "atmospheric" satellite sensors such as SCIAMACHY, OMI, TROPOMI and also, where possible, to the HICO data to set-up time series of hyperspectral PG data from 2002 onwards.
- Hyperspectral and quantitative spectral PT and PSC algorithms should be adapted, extended and applied at global and regional scales to data with high spectral resolution from former, current and near-future "atmospheric" satellite sensors such as SCIAMACHY, OMI, TROPOMI, and also, where possible, to the HICO data to set-up time series of hyperspectral PG data from 2002 onwards.
- The potential of spectral PG quantitative satellite retrievals for future satellite sensor should be further explored based on assessment via RTM, hyperspectral satellite, and *in situ* data to retrieve as many as possible of the PG requested by users from hyper- (e.g., like PACE, ENMAP) and multispectral data sets (e.g., like OLCI bands and a few additional bands).

In addition a framework is needed at an international level for integrating PG information from different sensors (hyper-/multispectral, global coverage/high spatial, and/or temporal resolution) to meet user requirements across different scales with special focus to regional applications (in order to fill Gap 4, see Section Gap 4: Lack of Regional Capability of PG Algorithms).

- At first, a workshop with experts across the different disciplines (*in situ*, algorithm, modeling) is necessary to define round-robin exercises for the regional (thematic) assessment of algorithms (validation, uncertainties) under specified protocols. When these regional algorithm inter-comparisons and assessments with different kinds of *in situ* optical and PG data are executed, best practice to use different kinds of PG data to obtain synergistic PG information should also be developed.
- PG data merged from all low spatial resolution hyperspectral available sources to date (incl. TROPOMI) should be used synergistically with multispectral derived PG products for building up high spatially resolved PG long-term global records from 1997 into the future.
- In addition, PG retrievals should be optimized on trigonal scale based on other environmental and ecological information (e.g., from remote sensing, climatology) such as sea surface temperature, available light, wind speed, mixed layer depth, nutrients fluorescence, and optically active water constituents (CDOM, total suspended matter) and better links to biogeography.

Based on the outcome of the activities to foster hyperspectral data exploitation, synergistic use of new satellite sensors

for detection of phytoplankton diversity across all oceanic, coastal, and inland water environments should allow for merged PG products. This will secure the prolongation of PG data time series as climate data records.

CONCLUSIONS

Synoptic observations on phytoplankton diversity, obtained from satellite ocean color data, have the potential to improve models for assessing and predicting climate change and for managing marine services, and they are currently the only means available for high resolution, long-term monitoring of changes in marine ecosystem structure at the regional to global scale. Yet, to meet the requirements of an essential ocean/climate variable (highly accurate and error-characterized) further scientific investment into existing and further developed methods is needed. In particular, the satellite phytoplankton group products should: (i) match those requested by the user communities; (ii) provide quantitative per-pixel uncertainty; (iii) exploit past, current and future hyperspectral remote-sensing; (iv) be tuned for regional applications (including coastal and inland-water regions); and (v) exploit better the various streams of satellite information, from the various sensors in space. Improved understanding of how the optical signatures (inherent optical properties) of phytoplankton groups vary will also aid algorithm development. These actions can only be achieved with coordinated and sustained investment, across national and international agencies, and through interdisciplinary co-operation between satellite algorithm developers, *in situ* experts and end users (e.g., ecosystem modelers).

AUTHOR CONTRIBUTIONS

AB organized the specific CLEO session, lead its discussion, and the writing team which set up the outline of this manuscript. She wrote the draft version and synoptically implemented the revisions by all other coauthors. HB, VB, SD, AH, MH, TH, SL, JU, MV, and AW planned and participated in the writing session after the workshop and substantially helped to revise several versions of the manuscript. In addition also several experts who could not participate in the discussion and writing sessions provided substantial comments and suggestions to improve the manuscript: ABri, RB, AC, LC, ED, AD, NH, CM, EO, and DR.

FUNDING

The ESA SEOM SY-4Sci Synergy project (No. 400112410/14/I-NB) funded partially the contribution of AB, ABri, RB, SL, and AW. SD acknowledges NASA (NNX13AC34G) for funding. CM acknowledges funding from NNX13AC34G for this effort.

ACKNOWLEDGMENTS

We thank ESA/ESRIN within the context of activities of the ESA Scientific Exploitation of Operational Missions (SEOM) for funding the "Phytoplankton Diversity at Global and Regional Scale" Session at the Colour and Light from Earth Observation (CLEO) workshop and International Satellite Functional Type Algorithm Intercomparison Project (http://pft.ees.hokudai.ac.jp/satellite/), with funding from JAXA, UK National Centre for Earth Observation, and the IOCCG for facilitation of a series of meetings that led to the development of this manuscript.

REFERENCES

Alvain, S., Le Quéré, C., Bopp, L., Racault, M. F., Beaugrand, G., Dessailly, D., et al. (2013). Rapid climatic driven shifts of diatoms at high latitudes. *Remote Sens. Environ.* 132, 195–201. doi: 10.1016/j.rse.2013.01.014

Alvain, S., Moulin, C., Dandonneau, Y., and Breon, F. (2005). Remote sensing of phytoplankton groups in case 1 waters for global SeaWiFS imagery. *Deep Sea Res. I* 52, 1989–2004. doi: 10.1016/j.dsr.2005.06.015

Alvain, S., Moulin, C., Dandonneau, Y., and Loisel, H. (2008). Seasonal distribution and succession of dominant phytoplankton groups in the global ocean: a satellite view. *Global Biogeochem. Cycles* 22:GB3001. doi: 10.1029/2007gb003154

Arnold, S. R., Spracklen, D. V., Williams, J., Yassaa, N., Sciare, J., Bonsang, B., et al. (2009). Evaluation of the global oceanic isoprene source and its impacts on marine organic carbon aerosol. *Atmos. Chem. Phys.* 9, 1253–1262. doi: 10.5194/acp-9-1253-2009

Baird, M. E., Cherukuru, N., Jones, E. M., Margvelashvili, N., Mongin, M., Oubelkheir, K., et al. (2016). Remote-sensing reflectance and true colour produced by a coupled hydrodynamic, optical, sediment, biogeochemical model of the Great Barrier Reef, Australia: comparison with remotely-sensed data. *Environ. Modell. Softw.* 78, 79–96. doi: 10.1016/j.envsoft.2015.11.025

Baretta, J. W., Ebenhöh, W., and Ruardij, P. (1995). The European regional seas ecosystem model, a complex marine ecosystem model. *Neth. J. Sea Res.* 33, 233–246. doi: 10.1016/0077-7579(95)90047-0

Ben Mustapha, Z., Alvain, S., Jamet, C., Loisel, H., and Dessailly, D. (2014). Automatic classification of water-leaving radiance anomalies from global SeaWiFS imagery: application to the detection of phytoplankton groups in open ocean waters. *Remote Sens. Environ.* 146, 97–112. doi: 10.1016/j.rse.2013.08.046

Bidigare, R. R., Ondnwk, M. E., Marrow, J. H., and Kiefer, D. A. (1990). *In vivo* absorption of algal pigments. *SPIE* 1302, 290–302.

Booge, D., Marandino, C. A., Schlundt, C., Palmer, P. I., Schlundt, M., Atlas, E. L., et al. (2016). Can simple models predict large scale surface ocean isoprene concentrations? *Atmos. Chem. Phys.* 16, 11807–11821. doi: 10.5194/acp-16-11807-2016

Bopp, L., Resplandy, L., Orr, J. C., Doney, S. C., Dunne, J. P., Gehlen, M., et al. (2013). Multiple stressors of ocean ecosystems in the 21st century: projections with CMIP5 models. *Biogeosciences* 10, 6225–6245. doi: 10.5194/bg-10-6225-2013

Boss, E., Picheral, M., Leeuw, T., Chase, A., Karsenti, E., Gorsky, G., et al. (2013). The characteristics of particulate absorption, scattering and attenuation coefficients in the surface ocean; Contribution of the Tara Oceans expedition. *Methods Oceanogr.* 7, 52–62. doi: 10.1016/j.mio.2013.11.002

Bouman, H. A., Ulloa, O., Scanlan, D. J., Zwirglmaier, K., Li, W. K., Platt, T., et al. (2006). Oceanographic basis of the global surface distribution of Prochlorococcus ecotypes. *Science* 312, 918–921. doi: 10.1126/science.1122692

Bracher, A., Hardman-Mountford, N., Hirata, T., Bernard, S., Brewin, R., Bricaud, A., et al. (2015a). *Report on IOCCG Workshop "Phytoplankton Composition from SPACE: towards a Validation Strategy for Satellite Algorithms*," NASA Tech. Memo 2015-217528_01-22-15, Portland, ME. Available online at: http://www.ioccg.org/groups/PFT-TM_2015-217528_01-22-15.pdf (Accessed October 25–26, 2014)

Bracher, A., Taylor, M. H., Taylor, B. B., Dinter, T., Röttgers, R., and Steinmetz, F. (2015b). Using empirical orthogonal functions derived from remote sensing

reflectance for the prediction of phytoplankton pigment concentrations. *Ocean Sci.* 11, 139–158. doi: 10.5194/os-11-139-2015

Bracher, A., Vountas, M., Dinter, T., Burrows, J. P., Röttgers, R., and Peeken, I. (2009). Quantitative observation of cyanobacteria and diatoms from space using PhytoDOAS on SCIAMACHY data. *Biogeosciences* 6, 751–764. doi: 10.5194/bg-6-751-2009

Brewin, R. J. W., Hardman-Mountford, N. J., Lavender, S. J., Raitsos, D. E., Hirata, T., Uitz, J., et al. (2011). An inter-comparison of bio-optical techniques for detecting dominant phytoplankton size class from satellite remote sensing. *Remote Sens. Environ.* 115, 325–339. doi: 10.1016/j.rse.2010.09.004

Brewin, R. J. W., Hirata, T., Hardman-Mountford, N. J., Lavender, S. J., Sathyendranath, S., and Barlow, R. (2012). The influence of the Indian Ocean Dipole on interannual variations in phytoplankton size structure as revealed by Earth Observation. *Deep Sea Res. II* 77–80, 117–127. doi: 10.1016/j.dsr2.2012.04.009

Brewin, R. J. W., Sathyendranath, S., Hirata, T., Lavender, S. J., Barciela, R. M., and Hardman-Mountford, N. J. (2010), A three-component model of phytoplankton size class for the Atlantic Ocean. *Ecol. Model.* 221, 1472–1483. doi: 10.1016/j.ecolmodel.2010.02.014

Brewin, R. J. W., Sathyendranath, S., Jackson, T., Barlow, R., Brotas, V., Airs, R., et al. (2015). Influence of light in the mixed-layer on the parameters of a three-component model of phytoplankton size class. *Remote Sens. Environ.* 168, 437–450. doi: 10.1016/j.rse.2015.07.004

Brewin, R. J. W., Sathyendranath, S., Lange, P. K., and Tilstone, G. (2014). Comparison of two methods to derive the size-structure of natural populations of phytoplankton. *Deep Sea Res. Part I* 85, 72–79. doi: 10.1016/j.dsr.2013.11.007

Bricaud, A., Ciotti, A. M., and Gentili, B. (2012), Spatial-temporal variations in phytoplankton size and colored detrital matter absorption at global and regional scales, as derived from twelve years of SeaWiFS data (1998–2009). *Global Biogeochem. Cycles* 26:GB1010. doi: 10.1029/2010GB003952

Bricaud, A., Claustre, H., Ras, J., and Oubelkheir, K. (2004). Natural variability of phytoplanktonic absorption in oceanic waters: influence of the size structure of algal populations. *J. Geophys. Res.* 109:C11010. doi: 10.1029/2004jc002419

Brown, C. W., and Yoder, J. A. (1994). Coccolithophorid blooms in the global ocean. *J. Geophys. Res.* 99, 7467–7482. doi: 10.1029/93JC02156

Chai, F., Dugdale, R. C., Peng, T.-H., Wilkerson, F. P., and Barber, R. T. (2002). One-dimensional ecosystem model of the equatorial Pacific upwelling system. Part I. Model development and silicon and nitrogen cycle. *Deep Sea Res. II* 49, 2713–2745. doi: 10.1016/S0967-0645(02)00055-3

Chase, A., Boss, E., Zaneveld, R., Bricaud, A., Claustre, H., Ras, J., et al. (2013). Decomposition of *in situ* particulate absorption spectra. *Methods Oceanogr.* 7, 110–124. doi: 10.1016/j.mio.2014.02.002

Cherkasheva, A., Bracher, A., Melsheimer, C., Köberle, C., Gerdes, R., Nöthig, E.-M., et al. (2014). Influence of the physical environment on phytoplankton blooms: a case study in the Fram Strait. *J. Mar. Syst.* 132, 196–207. doi: 10.1016/j.jmarsys.2013.11.008

Ciotti, A. M., and Bricaud, A. (2006). Retrievals of a size parameter for phytoplankton and spectral light absorption by colored detrital matter from water-leaving radiances at SeaWiFS channels in a continental shelf region off Brazil. *Limnol. Oceanogr. Methods*, 4, 237–253. doi: 10.4319/lom.2006.4.237

Cullen, J. J., Ciotti, A. M., Davis, R. F., and Lewis, M. R. (1997). Optical detection and assessment of algal blooms. *Limnol. Oceanogr.* 42, 1223–1239. doi: 10.4319/lo.1997.42.5_part_2.1223

Devred, E., Sathyendranath, S., and Stuart, V. (2006). A two-component model of phytoplankton absorption in the open ocean: theory and applications. *J. Geophys. Res.* 111:C03011. doi: 10.1029/2005jc002880

Devred, E., Sathyendranath, S., Stuart, V., and Platt, T. (2011). A three component classification of phytoplankton absorption spectra: application to ocean-color data. *Remote Sens. Environ.* 115, 2255–2266. doi: 10.1016/j.rse.2011.04.025

Dutkiewicz, S., Hickman, A. E., Jahn, O., Gregg, W. W., Mouw, C. B., and Follows, M. J. (2015). Capturing optically important constituents and properties in a marine biogeochemical and ecosystem model. *Biogeosciences* 12, 4447–4481. doi: 10.5194/bg-12-4447-2015

Evers-King, H., Bernard, S., Lain, L. R., and Probyn, T. A. (2014). Sensitivity in reflectance attributed to phytoplankton cell size: forward and inverse modelling approaches. *Opt. Express* 22, 11536–11551. doi: 10.1364/OE.22.011536

Field, C. B., Behrenfeld, M. J., Randerson, J. T., and Falkowski, P. G. (1998). Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* 281, 237–240. doi: 10.1126/science.281.5374.237

Follows, M. J., Dutkiewicz, S., Grant, S., and Chisholm, S. W. (2007). Emergent biogeography of microbial communities in a model ocean. *Science* 315, 1843–1846. doi: 10.1126/science.1138544

Fujii, M., Boss, E., and Chai, F. (2007). The value of adding optics to ecosystem models: a case study. *Biogeosciences* 4, 817–835. doi: 10.5194/bg-4-817-2007

Fujiwara, A., Hirawake, T., Suzuki, K., and Saitoh, S.-I. (2011). Remote sensing of size structure of phytoplankton communities using optical properties of the Chukchi and Bering Sea shelf region. *Biogeosciences* 8, 3567–3580. doi: 10.5194/bg-8-3567-2011

Gregg, W. W., and Casey, N. W. (2007). Modeling coccolithophores in the global ocean. *Deep Sea Res. II* 54, 447–477. doi: 10.1016/j.dsr2.2006.12.007

Harmel, T., Hieronymi, M., Slade, W., Röttgers, R., Roullier, F., and Chami, M. (2016). Laboratory experiments for inter-comparison of three volume scattering meters to measure angular scattering properties of hydrosols. *Opt. Express* 24, A234–A256. doi: 10.1364/OE.24.00A234

Hashioka, T., Vogt, M., Yamanaka, Y., Le Quéré, C., Noguchi Aita, M., Alvain, S., et al. (2013). Phytoplankton competition during the spring bloom in four Plankton Functional Type Models. *Biogeosciences* 10, 6833–6850. doi: 10.5194/bg-10-6833-2013

Hirata, T., Aiken, J., Hardman-Mountford, N., Smyth, T. J., and Barlow, R. (2008). An absorption model to determine phytoplankton size classes from satellite ocean color. *Remote Sens. Environ.* 112, 3153–3159. doi: 10.1016/j.rse.2008.03.011

Hirata, T., Hardman-Mountford, N. J., Brewin, R. J. W., Aiken, J., Barlow, R., Suzuki, K., et al. (2011). Synoptic relationships between surface Chlorophyll-a and diagnostic pigments specific to phytoplankton functional types. *Biogeosciences* 8, 311–327. doi: 10.5194/bg-8-311-2011

Hirata, T., Saux-Picart, S., Hashioka, T., Aita-Noguchi, M., Sumata, H., Shigemitsu, M., et al. (2013). A comparison between phytoplankton community structure derived from a global 3D ecosystem model and satellite observation. *J. Mar. Syst.* 109–110, 129–137. doi: 10.1016/j.jmarsys.2012.01.009

Hoepffner, N., and Sathyendranath, S. (1993). Determination of the major groups of phytoplankton pigments from the absorption spectra of total particulate matter. *J. Geophys. Res.* 98, 22784–22803. doi: 10.1029/93jc01273

Hooker, S. B., Clementson, L., Thomas, C. S., Schlüter, L., Allerup, M., Ras, J., et al. (2012). *The Fifth SeaWiFS HPLC Analysis Round-Robin Experiment (SeaHARRE-5)*, Vol. 98. Greenbelt, MD: NASA Tech. Memo 2012–217503; NASA Goddard Space Flight Center.

IOCCG (2000). "Remote sensing of ocean color in coastal, and other optically Complex, Waters," in *Reports of the International Ocean Color Coordinating Group, No. 3*, ed S. Sathyendranath (Dartmouth, NS: IOCCG), 140.

IOCCG (2009). "Remote sensing in fisheries and aquaculture: the societal benefits," in *Report of the International Ocean-Colour Coordinating Group, No. 8*, eds M.-H. Forget, V. Stuart, and T. Platt (Dartmouth, NS: IOCCG), 98.

IOCCG (2014). "Phytoplankton functional types from space," in *Reports of the International Ocean Color Coordinating Group, No. 15*, ed S. Sathyendranath (Dartmouth, NS: IOCCG), 156.

IOCS (2013). "Splinter session 10: phytoplankton community structure from ocean color," in *Proceedings of the International Ocean color Symposium (IOCS) 2013*, eds A. Bracher and T. Hirata (Darmstadt), 54–58. Available online at: http://iocs.ioccg.org/wp-content/uploads/report-iocs-2013-meeting.pdf

IOCS (2015). "Breakout session 1: remote sensing of phytoplankton composition - possibilities, applications and future needs," in *Proceedings of the International Ocean color Symposium (IOCS) 2015*, eds C. Mouw, A. Bracher, N. Hardman-Mountford (San Francisco, CA), 22–23. Available online at: http://iocs.ioccg.org/wp-content/uploads/2012/10/report-iocs-2015-meeting.pdf

IPCC (2013). "Climate change 2013: the physical science basis," in *Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change,* eds T. F. Stocker, D. Qin, G.-K. Plattner, M. M. B. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, P. M. Midgley (Cambridge, UK; New York, NY, USA: Cambridge University Press), 1535.

Jennigs, S., Melin, F., Blanchard, J. L., Forster, R. M., Dulvy, N. K., and Wilson, R. W. (2008). Global-scale predictions of community and ecosystem properties from simple ecological theory. *Proc. Biol. Sci.* 275, 1375–1383. doi: 10.1098/rspb.2008.0192

Johnsen, G., and Sakshaug, E. (2007). Biooptical characteristics of PSII and PSI in 33 species (13 pigment groups) of marine phytoplankton, and the relevance for pulse-amplitude-modulated and fast-repetition-rate fluorometry. *J. Phycol.* 43, 1236–1251. doi: 10.1111/j.1529-8817.2007.00422.x

Kostadinov, T. S., Siegel, D. A., and Maritorena, S. (2009). Retrieval of the particle size distribution from satellite ocean color observations. *J. Geophys. Res.* 114:C09015. doi: 10.1029/2009JC005303

Kostadinov, T., Cabré, A., Vedantham, H., Marinov, I., Bracher, A., Brewin, R., et al. (2017). Intercomparison of phytoplankton functional types derived from ocean color algorithms and earth system models: phenology. *Remote Sens. Environ.* 190, 162–177. doi: 10.1016/j.rse.2016.11.014

Kostadinov, T., Milutinovic, S., Marinov, I., and Cabré, A. (2016). Carbon-based phytoplankton size classes retrieved via ocean color estimates of the particle size distribution. *Ocean Sci.* 12, 561–575. doi: 10.5194/os-12-561-2016

Kudela, R., Palacios, S. L., Austerberry, D. C., Accorsi, E. K., Guild, L. S., and Torres-Perez, J. (2015). Application of hyperspectral remote sensing to cyanobacterial blooms in inland waters. *Remote Sens. Environ.* 167, 196–205. doi: 10.1016/j.rse.2015.01.025

Kurekin, A. A., Miller, P. I., and van der Woerd, H. J. (2014). Satellite discrimination of *Karenia mikimotoi* and Phaeocystis harmful algal blooms in European coastal waters: merged classification of ocean color data. *Harmf. Algae* 31, 163–176. doi: 10.1016/j.hal.2013.11.003

Laufkötter, C., Vogt, M., Gruber, N., Aita-Noguchi, M., Aumont, O., Bopp, L., et al. (2015). Drivers and uncertainties of future global marine primary production in marine ecosystem models. *Biogeosciences* 12, 6955–6984. doi: 10.5194/bg-12-6955-2015

Le Quéré, C., Harrison, S. P., Colin Prentice, I., Buitenhuis, E. T., Aumont, O., Bopp, L., et al. (2005). Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. *Glob. Chang. Biol.* 11, 2016–2040. doi: 10.1111/j.1365-2486.2005.01004.x

Li, Z., Li, L., Song, K., and Cassar, N. (2013). Estimation of phytoplankton size fractions based on spectral features of remote sensing ocean color data. *J. Geophys. Res. Oceans* 118, 1445–1458. doi: 10.1002/jgrc.20137

Maritorena, S., Fanton d'Andon, O. H., Mangin, A., and Siegel, D. A. (2010). Merged satellite ocean color data products using a bio-optical model: characteristics, benefits and issues. *Remote Sens. Environ.* 114, 1791–1804. doi: 10.1016/j.rse.2010.04.002

Masuda, Y., Yamanaka, Y., Hirata, T., and Nakano, H. (2017). Competition and community assemblage dynamics within a phytoplankton functional group: simulation using an eddy-resolving model to disentangle deterministic and random effects. *Ecol. Modell.* 343, 1–14. doi: 10.1016/j.ecolmodel.2016.10.015

McClain, C. R. (2009). A decade of satellite ocean color observations. *Ann. Rev. Mar. Sci.* 1, 19–42. doi: 10.1146/annurev.marine.010908.163650

Moisan, T. A. H., Moisan, J. R., Linkswiler, M. A., and Steinhardt, R. A. (2013). Algorithm development for predicting biodiversity based on phytoplankton absorption. *Cont. Shelf Res.* 55, 17–28. doi: 10.1016/j.csr.2012.12.011

Moore, T. S., Dowell, M. D., and Franz, B. A. (2012). Detection of coccolithophore blooms in ocean color satellite imagery: a generalized approach for use with multiple sensors. *Remote Sens. Environ.* 117, 249–263. doi: 10.1016/j.rse.2011.10.001

Morel, A., and Bricaud, A. (1981). Theoretical results concerning light absorption in a discrete medium, and application to specific absorption of phytoplankton. *Deep Sea Res.* 28, 1375–1393. doi: 10.1016/0198-0149(81)90039-X

Mouw, C. B., and Yoder, J. A. (2010). Optical determination of phytoplankton size composition from global SeaWiFS imagery. *J. Geophys. Res.* 115:C12018. doi: 10.1029/2010JC006337

Mouw, C. B., Barnett, A., McKinley, G. A., Gloege, L., and Pilcher, D. (2016). Phytoplankton size impact on export flux in the global ocean. *Global Biogeochem. Cycles* 30, 1542–1562. doi: 10.1002/2015GB005355

Mouw, C. B., Hardman-Montford, N., Alvain, S., Bracher, A., Brewin, R., Bricaud, A., et al. (2017). A Consumer's guide to satellite remote sensing of multiple phytoplankton groups in the global ocean. *Front. Mar. Sci.* 4:41. doi: 10.3389/fmars.2017.00041

Mueller, J. L., Fargion, G. S., and McClain, C. R. (2003). *Ocean Optics Protocols for Satellite Ocean Color Sensor Validation, Revision 4, Volume III: Radiometric Measurements and Data Analysis Protocols*. NASA/TM-2003-211621/Rev4-Vol. III, NASA Goddard Space Flight Center, Greenbelt, MD.

Oelker, J., Dinter, T., Richter, A., Burrows, J. P., and Bracher, A. (2016). "Towards improved spatial resolution of hyper-spectral phytoplankton functional type products," in *Oral presentation at 'Colour and Light in the Ocean from Earth Observation (CLEO) Relevance and Applications Products from Space and Perspectives from Models'* (Frascati: ESA-ESRIN). Available online at: http://esaconferencebureau.com/docs/default-source/16c12-presentations/d2_12-05_oelker.pdf?sfvrsn=2 (September 6–8, 2016).

Organelli, E., Bricaud, A., Antoine, D., and Uitz, J. (2013). Multivariate approach for the retrieval of phytoplankton size structure from measured light absorption spectra in the Mediterranean Sea (BOUSSOLE site). *Appl. Opt.* 52, 2257–2273. doi: 10.1364/AO.52.002257

Organelli, E., Nuccio, C., Melillo, C., and Massi, L. (2011). Relationships between phytoplankton light absorption, pigment composition and size structure in offshore areas of the Mediterranean Sea. *Adv. Oceanogr. Limnol.* 2, 107–123. doi: 10.4081/aiol.2011.5320

Palacz, A. P., St. John, M. A., Brewin, R. J. W., Hirata, T., and Gregg, W. W. (2013). Distribution of phytoplankton functional types in high-nitrate low-chlorophyll waters in a new diagnostic ecological indicator model. *Biogeosciences* 10, 7553–7574. doi: 10.5194/bg-10-7553-2013

Peloquin, J., Swan, C., Gruber, N., Vogt, M., Claustre, H., and Ras, J. (2013). The MAREDAT global database of high performance liquid chromatography marine pigment measurements (NetDCF) - Contribution to the MAREDAT World Ocean Atlas of Plankton Functional Types. *Earth Syst. Sci. Data* 5, 109–123. doi: 10.5194/essd-5-109-2013

Racault, M.-F., Raitsos, D. E., Berumen, M. L., Brewin, R. J. W., Platt, T., Sathyendranath, S., et al. (2015). Phytoplankton phenology indices in coral reef ecosystems: application to ocean-colour observations in the Red Sea. *Remote Sens. Environ.* 160, 222–234. doi: 10.1016/j.rse.2015.01.019

Raitsos, D. E., Lavender, S. J., Maravelias, C. D., Haralambous, J., Richardson, A. J., and Reid, P. C. (2008). Identifying phytoplankton functional groups from space: an ecological approach. *Limnol. Oceanogr.* 53, 605–613. doi: 10.4319/lo.2008.53.2.0605

Roy, S., Sathyendranath, S., Bouman, H., and Platt, T. (2013). The global distribution of phytoplankton size spectrum and size classes from their light-absorption spectra derived from satellite data. *Remote Sens. Environ.* 139, 185–197. doi: 10.1016/j.rse.2013.08.004

Rozanov, V. V., Rozanov, A., Kokhanovsky, A., and Burrows, J. P. (2014). Radiative transfer through atmosphere and ocean: software package SCIATRAN. *J. Quant. Spectrosc. Rad. Transf.* 133, 13–71. doi: 10.1016/j.jqsrt.2013.07.004

Ryan, J. P., Davis, C. O., Tufillaro, N. B., Kudela, R. M., and Gao, B.-C. (2014). Application of the hyperspectral imager for the coastal ocean to phytoplankton ecology studies in Monterey Bay, CA. *Remote Sens.* 6, 1007–1025. doi: 10.3390/rs6021007

Sadeghi, A., Dinter, T., Vountas, M., Taylor, B. B., Altenburg-Soppa, M., and Bracher, A. (2012a). Improvement to the PhytoDOAS method for identification of coccolithophores using hyper-spectral satellite data. *Ocean Sci.* 8, 1055–1070. doi: 10.5194/os-8-1055-2012

Sadeghi, A., Dinter, T., Vountas, M., Taylor, B., Altenburg-Soppa, M., Peeken, I., et al. (2012b). Remote sensing of coccolithophore blooms in selected oceanic regions using the PhytoDOAS method applied to hyper-spectral satellite data. *Biogeosciences* 9, 2127–2143. doi: 10.5194/bg-9-2127-2012

Siegel, D. A., Buesseler, K. O, Behrenfeld, M. J., Benitez-Nelson, C., Boss, E., Brzezinski, M. A., et al. (2016). Prediction of the export and fate of global ocean net primary production: the EXPORTS science plan. *Front. Mar. Sci.* 3:22. doi: 10.3389/fmars.2016.00022

Soppa, M. A., Hirata, T., Silva, B., Dinter, T., Peeken, I., Wiegmann, S., et al. (2014). Global retrieval of diatoms abundance based on phytoplankton pigments and satellite. *Remote Sens.* 6, 10089–10106. doi: 10.3390/rs61010089

Soppa, M. A., Völker, C., and Bracher, A. (2016a). Diatom phenology in the southern ocean: mean patterns, trends and the role of climate oscillations. *Remote Sens.* 8:420. doi: 10.3390/rs8050420

Soppa, M., Dinter, T., Losa, L., Wolanin, A., and Bracher, A. (2016b). *SY-4Sci Synergy R & D Study 4: Phytoplankton Functional Types (SynSenPFT), D1.2 Algorithm Theoretical Base Document (ATBD), Version 8.1, ESRIN Contract No.: 4000112410/14/I-NB*. Available online at: https://www.awi.de/fileadmin/user_upload/AWI/Forschung/Klimawissenschaft/Physikalische_

Ozeanographie_der_Polarmeere/SEOM-SynSenPFT-ATBD-D1.2_v8.1.pdf (Accessed September 15, 2016).

Sosik, H. M., and Olson, R. J. (2007). Automated taxonomic classification of phytoplankton sampled with imaging-in-flow cytometry. *Limnol. Oceanogr. Methods* 5, 204–216. doi: 10.4319/lom.2007.5.204

Subramaniam, A., Brown, C. W., Hood, R. R., Carpenter, E., and Capone, D. G. (2002). Detecting Trichodesmium blooms in SeaWiFS imagery. *Deep Sea Res. II* 49, 107–121. doi: 10.1016/S0967-0645(01)00096-0

Swan, C., Vogt, M., Gruber, N., and Laufkötter, C. (2016). A global seasonal surface ocean climatology of phytoplankton types based on CHEMTAX analysis of HPLC pigments. *Deep Sea Res. I* 109, 137–156. doi: 10.1016/j.dsr.2015.12.002

Tan, H., Oishi, T., Tanaka, A., and Doerffer, R. (2015). Accurate estimation of the backscattering coefficient by light scattering at two backward angles. *Appl. Opt.* 54, 7718–7733. doi: 10.1364/AO.54.007718

Trzcinski, M. K., Devred, E., Platt, T., and Sathyendranath, S. (2013). Variation in ocean color may help predict cod and haddock recruitment. *Mar. Ecol. Prog. Ser.* 491, 187–197. doi: 10.3354/meps10451

Uitz, J., Claustre, H., Morel, A., and Hooker, S. B. (2006). Vertical distribution of phytoplankton communities in open ocean: an assessment based on surface chlorophyll. *J. Geophys. Res.* 111:C08005. doi: 10.1029/2005JC003207

Vogt, M., Hashioka, T., Payne, M., and Buitenhuis, E. T. (2013). The distribution, dominance patterns and ecological niches of plankton functional groups in Dynamic Green Ocean Models and satellite estimates. *Biogeosci. Discuss.* 10, 17193–17247. doi: 10.5194/bgd-10-17193-2013

Ward, B. A., Dutkiewicz, S., Jahn, O., and Follows, M. J. (2012). A size structured food-web model for the global ocean. *Limnol. Oceanogr.* 57, 1877–1891. doi: 10.4319/lo.2012.57.6.1877

Werdell, J., Franz, B. A., Bailey, S. W., Harding, L. W. Jr., and Feldman, G. C. (2007). Approach for the long-term spatial and temporal evaluation of ocean color satellite data products in a coastal environment. *Proc. SPIE Int. Soc. Opt. Eng.* 6680:12. doi: 10.1117/12.732489

Werdell, P. J., Franz, B. A., Bailey, S. W., Feldman, G. C., Boss, E., and Brando, V. E. (2013). Generalized ocean color inversion model for retrieving marine inherent optical properties. *Appl. Opt.* 52:2019. doi: 10.1364/AO.52.002019

Werdell, P. J., Roesler, C. S., and Goes, J. I. (2014). Discrimination of phytoplankton functional groups using an ocean reflectance inversion model. *Appl. Opt.* 53:4833. doi: 10.1364/AO.53.004833

Westberry, T. K., Siegel, D. A., and Subramaniam, A. (2005). An improved bio-optical model for the remote sensing of *Trichodesmium* spp. blooms. *J. Geophys. Res.* 110:C06012. doi: 10.1029/2004JC002517

Wolanin, A., Dinter, T., and Bracher, A. (2015). *SY-4Sci Synergy R & D Study 4: Phytoplankton Functional Types (SynSenPFT), D2.1 - Report on Using Radiative Transfer Modelling to Develop a Correction Scheme and Investigate the Sensitivity of the Improved PhytoDOAS (Version 3.0) Algorithm*. Version 1, ESRIN Contract No.: 4000112410/14/I-NB, https://www.awi.de/fileadmin/user_upload/AWI/Forschung/Klimawissenschaft/Physikalische_Ozeanographie_der_Polarmeere/SEOM-SynSenPFT-D2.1_v1.pdf (Accessed October 30, 2015).

Wolanin, A., Soppa, M. A., and Bracher, A. (2016). Investigation of spectral band requirements for improving retrievals of phytoplankton functional types. *Remote Sens.* 8:871. doi: 10.3390/rs8100871

Xi, H., Hieronymi, M., Röttgers, R., Krasemann, H., and Qiu, Z. (2015). Hyperspectral Differentiation of phytoplankton taxonomic groups: a comparison between using remote sensing reflectance and absorption spectra. *Remote Sens.* 7, 14781–14805. doi: 10.3390/rs71114781

Ye, Y., Voelker, C., Bracher, A., Taylor, B., and Wolf-Gladrow, D. (2012). Environmental controls on N2 fixation by Trichodesmium in the tropical eastern North Atlantic. *Deep Sea Res. I* 64, 104–117. doi: 10.1016/j.dsr.2012.01.004

Zhang, X., Huot, Y., Bricaud, A., and Sosik, H. (2015). Inversion of spectral absorption coefficients to infer phytoplankton size classes, chlorophyll concentration, and detrital matter. *Appl. Opt.* 54, 5805–5815. doi: 10.1364/AO.54.005805

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Bracher, Bouman, Brewin, Bricaud, Brotas, Ciotti, Clementson, Devred, Di Cicco, Dutkiewicz, Hardman-Mountford, Hickman, Hieronymi, Hirata, Losa, Mouw, Organelli, Raitsos, Uitz, Vogt and Wolanin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Marine Science

ORIGINAL RESEARCH
published: 06 March 2017
doi: 10.3389/fmars.2017.00060

Simulating PACE Global Ocean Radiances

Watson W. Gregg and Cécile S. Rousseaux*

NASA Global Modeling and Assimilation Office, Greenbelt, MD, USA

The NASA PACE mission is a hyper-spectral radiometer planned for launch in the next decade. It is intended to provide new information on ocean biogeochemical constituents by parsing the details of high resolution spectral absorption and scattering. It is the first of its kind for global applications and as such, poses challenges for design and operation. To support pre-launch mission development and assess on-orbit capabilities, the NASA Global Modeling and Assimilation Office has developed a dynamic simulation of global water-leaving radiances, using an ocean model containing multiple ocean phytoplankton groups, particulate detritus, particulate inorganic carbon (PIC), and chromophoric dissolved organic carbon (CDOC) along with optical absorption and scattering processes at 1 nm spectral resolution. The purpose here is to assess the skill of the dynamic model and derived global radiances. Global bias, uncertainty, and correlation are derived using available modern satellite radiances at moderate spectral resolution. Total chlorophyll, PIC, and the absorption coefficient of CDOC (a_{CDOC}), are simultaneously assimilated to improve the fidelity of the optical constituent fields. A 5-year simulation showed statistically significant ($P < 0.05$) comparisons of chlorophyll ($r = 0.869$), PIC ($r = 0.868$), and a_{CDOC} ($r = 0.890$) with satellite data. Additionally, diatoms ($r = 0.890$), cyanobacteria ($r = 0.732$), and coccolithophores ($r = 0.716$) were significantly correlated with in situ data. Global assimilated distributions of optical constituents were coupled with a radiative transfer model (Ocean-Atmosphere Spectral Irradiance Model, OASIM) to estimate normalized water-leaving radiances at 1 nm for the spectral range 250–800 nm. These unassimilated radiances were within −0.074 mW cm^{-2} μm^{-1} sr^{-1} of MODIS-Aqua radiances at 412, 443, 488, 531, 547, and 667 nm. This difference represented a bias of −10.4% (model low). A mean correlation of 0.706 ($P < 0.05$) was found with global distributions of MODIS radiances. These results suggest skill in the global assimilated model and resulting radiances. The reported error characterization suggests that the global dynamical simulation can support some aspects of mission design and analysis. For example, the high spectral resolution of the simulation supports investigations of band selection. The global nature of the radiance representations supports investigations of satellite observing scenarios. Global radiances at bands not available in current and past missions support investigations of mission capability.

Keywords: PACE, ocean color, water-leaving radiances, biogeochemical model, radiative transfer model

OPEN ACCESS

Edited by:
Shubha Sathyendranth, Plymouth Marine Laboratory, UK

Reviewed by:
Tim Moore, University of New Hampshire, USA
Hajo Krasemann, Helmholtz-Zentrum Geesthacht Centre for Materials and Coastal Research (HZ), Germany

****Correspondence:***
Watson W. Gregg
watson.gregg@nasa.gov

Specialty section:
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

Received: *13 December 2016*
Accepted: *17 February 2017*
Published: *06 March 2017*

Citation:
Gregg WW and Rousseaux CS (2017) Simulating PACE Global Ocean Radiances. Front. Mar. Sci. 4:60. doi: 10.3389/fmars.2017.00060

INTRODUCTION

The now 19-year time series of routine global ocean color observations from space has led to advancements in the science of ocean biology beyond expectations. From chlorophyll interannual variability to inherent optical properties to physical-biological coupling, the time series has been an invaluable resource for scientists in a broad range of ocean and atmosphere-related fields. As is often the case in science, the proliferation of information from these moderate resolution missions has raised as many questions as it has answered. Coupled with improvements in detector technology, the time is now right for advancement of ocean biogeochemical science from space using higher spectral resolution missions.

Higher spectral resolution can potentially improve detection of optical constituents in the oceans that have important effects on biology, biogeochemistry, and light transmission. One major objective is the determination of phytoplankton groups from space. Research to detect phytoplankton groups from space has been going on for some time using the fleet of moderate spectral resolution sensors (e.g., Kamykowski et al., 2002; Alvain et al., 2005; Aiken et al., 2007; Bracher et al., 2009; Brewin et al., 2010, 2011; Kostadinov et al., 2010; Masotti et al., 2010; Hirata et al., 2011). Methods to identify size classes have also been pursued (e.g., Loisel et al., 2006; Brewin et al., 2011) but these only loosely relate to phytoplankton functionality/taxonomy. Several phytoplankton discrimination methods resolve dominant groups only (Sathyendranath et al., 2004; Alvain et al., 2005, 2008; Hirata et al., 2008; Raitsos et al., 2008). Hirata et al. (2011) provides taxonomic classifications, with relative and even absolute abundances quantified. Using satellite ocean chlorophyll concentrations rather than radiances, this empirical methodology essentially assumes that abundance reflects taxonomy, which is valid in many instances but not always (Rousseaux et al., 2013).

Moderate resolution ocean color sensors containing only a few discrete spectral bands, such as the global missions flown to date, do not contain sufficient spectral information to enable unequivocal phytoplankton functional/taxonomic discrimination. Many phytoplankton species/groups have subtle, but distinct spectral signatures. Use of hyper-spectral remote retrievals with many bands spanning the visible and ultraviolet spectrum holds potential for resolving these spectral distinctions (e.g., Bracher et al., 2009; Sadeghi et al., 2012; Palacios et al., 2015; Neukermans et al., 2016).

To close this knowledge gap, NASA has proposed the PACE mission, a global hyper-spectral sensor to test the ability to retrieve phytoplankton population distributions, as well as other important ocean constituents with optical signatures. The mission, proposed for launch in the early 2020's, can potentially demonstrate the feasibility and capability of hyper-spectral observations from space and enable scientists to observe and quantify these important ocean biological features. PACE is intended to follow future planned hyperspectral missions PRISM (Meini et al., 2015) and EnMAP (Foerster et al., 2015) with extended spectral range into the ultraviolet, faster observational repeat times, and emphasis on global ocean observational capability.

Since there is no global observational precedent, many mission development activities, design tradeoff assessments, operational strategies, and other issues, are speculative. Here we develop a dynamic global model at extreme hyper-spectral resolution (1 nm) to provide a platform to approximate realistic ocean conditions and help with resolving at least some of these issues and understand if such a simulation can assist in resolving many of the issues that inevitably arise in the design and testing of a new mission. The objective of this effort is to quantitatively assess the skill of a global model using a forward radiance representation to simulate global ocean water-leaving radiances. The skill is evaluated spectrally with explicit error characterization.

METHODS

Global Ocean Physical-Biogeochemical Model Configuration

The underlying biogeochemical constituents are simulated by the NOBM which is coupled to a global ocean circulation model, Poseidon (Schopf and Loughe, 1995). It spans the domain from -84° to 72° latitude in increments of 1.25° longitude by $2/3^\circ$ latitude, including only open ocean areas, where bottom depth >200 m. NOBM incorporates global coupled physical-biological processes, including four phytoplankton groups (diatoms, chlorophytes, cyanobacteria, and coccolithophores), which span much of the functionality of the global oceans, four nutrients (nitrate, ammonium, silicate, and dissolved iron), three detrital components (particulate organic carbon, silicate, and iron), and two carbon components (dissolved organic and inorganic carbon). It is a three-dimensional representation of coupled circulation/ biogeochemical processes in the global oceans (Gregg et al., 2003; Gregg and Casey, 2007).

Optically-active constituents have been added to NOBM to improve realism and complexity of the ocean simulation and better represent the ocean optical variability that will be observed by PACE. We have added particulate inorganic carbon (PIC) and chromophoric dissolved organic carbon (CDOC) as prognostic state variables. PIC is produced by coccolithophores as detached coccoliths and is lost via sinking and dissolution. PIC is produced as a fraction (25%) of the coccolithophore

Abbreviations: a_{CDM}, absorption coefficient of Chromophoric Dissolved and particulate organic Matter; a_{CDOC}, absorption coefficient of CDOC; BIOSOPE, Biogeochemistry and Optics South Pacific Experiment; CDOC, Chromophoric Dissolved Organic Carbon; CZCS, Coastal Zone Color Scanner; DOC, Dissolved Organic Carbon; EnMAP, Environmental MAPping and Analysis Program; GMAO, Global Modeling and Assimilation Office; MAP, Modeling, Analysis and Prediction; MERRA, Modern-Era Retrospective Analysis for Research and Applications; MODIS, MOderate Resolution Imaging Spectroradiometer; NIR, Near InfraRed; NOBM, NASA Ocean Biogeochemical Model; OASIM, Ocean-Atmosphere Spectral Irradiance Model; PACE, Plankton, Aerosol, Cloud and ocean Ecosystems ; PIC, Particulate Inorganic Carbon ; PRISMA, PRecursore IperSpettrale della Missione Applicativa; PSU, Practical Salinity Units; SeaWiFS, Sea-viewing Wide Field-of-view Sensor; S-NPP, Suomi National Polar-orbiting Partnership.

growth rate (Gregg and Casey, 2007) minus respiration. The PIC sinking rate is represented here as an exponential function of concentration, assuming that large concentrations of PIC are associated with larger coccolith size.

$$w_s(\mathrm{PIC}) = a_o \exp(a_1 * \mathrm{PIC}) \quad (1)$$

where w_s is the PIC sinking rate (m d^{-1}), PIC is in units of μgC l^{-1}, $a_0 = 0.1$ m d^{-1} and $a_1 = 2.0$ l μgC^{-1} (Gregg and Rousseaux, 2016). Dissolution follows Buitenhuis et al. (2001), except that no dissolution is allowed for depths shallower than the calcium carbonate compensation depth, which we define as 3500 m.

Chromophoric dissolved organic carbon (CDOC) represents the biogeochemical constituent necessary for the simulation of absorption by $a_{CDOC}(\lambda)$, the absorption coefficient, which is an optical quantity. CDOC is formed and destroyed the same as DOC, using Aumont et al. (2002) with an assumed DOC:CDOC production/loss ratio of 0.5. It is additionally destroyed by the absorption of spectral irradiance. We follow the methodology of Gregg and Rousseaux (2016) for photo-destruction (photolysis) of CDOC per unit irradiance quanta, with a different quantum yield $\varphi CDOC$ of 3.0E-6 (μM μmol photons absorbed m^{-3}) for results in reasonable agreement with MODIS-Aqua data (Maritorena et al., 2010).

Ocean-Atmosphere Spectral Irradiance Model

NOBM is coupled to OASIM (Gregg and Carder, 1990; Gregg, 2002; Gregg and Casey, 2009) to simulate the propagation of downward spectral irradiance in the oceans and the upwelling irradiance/radiance. The irradiance pathways for OASIM are shown in **Figure 1**. The atmosphere and ocean portions of the downwelling and upwelling irradiance are implemented at 25-nm spectral resolution. Higher spectral resolution is impractical for global models that integrate at 30 min time steps in our case. Upwelling radiance is produced at 1 nm resolution, however. Biases and uncertainties in the atmospheric component of OASIM have been characterized for clear sky high spectral resolution (1 nm; Gregg and Carder, 1990) and under mixed cloudy and clear skies for integrated spectral resolution (Gregg and Casey, 2009). We elaborate here on the ocean optical calculations.

Optical Properties of Ocean Constituents

The coupled NOBM-OASIM model includes optically active constituents, including seawater, phytoplankton, detritus, PIC, and CDOC each with unique spectral characteristics (**Figure 2**). All are prognostic state variables, with independent sources and sinks. The optical properties of each constituent are taken from various efforts in the peer reviewed literature.

Water

The spectral absorption and scattering properties of seawater was reported by Smith and Baker (1981) for the 200–800 nm spectral domain. Pope and Fry (1997) revised this for the range 380–720 nm, but this was for pure water. Morel et al. (2007) derived new data for absorption and scattering for the spectral range 300–500 nm using information in the clearest ocean waters of the South Pacific (although absorption values >420 nm were taken from Pope and Fry, 1997). Finally, Lee et al. (2015) reported absorption coefficients in the range 350–550 nm derived using remote sensing reflectance algorithms for the same clear ocean water data used by Morel et al. (2007). Mason et al. (2016) used laboratory observations to obtain new absorption coefficients for the spectral range 250–550 nm. Like Pope and Fry (1997), their results were specific to pure water.

Water absorption data used here are from Smith and Baker (1981) for 200–300 nm and 730–800 nm, Morel et al., 2007) for 300–350 nm, Lee et al. (2015) for 350–550 nm, Pope and Fry (1997) for 550–720 nm, Circio and Petty (1951) for 800 nm–2.5 μm, and Maul (1985) for 2.5–4 μm. Water scattering is from the method of Zhang et al. (2009), which accounts for temperature and salinity dependence. The backscattering-to-total scattering ratio $\tilde{b}_{bw}$ for water is 0.5.

Phytoplankton

Phytoplankton optical properties are obtained from various sources. Chlorophyll-specific absorption coefficients $a^*_p(\lambda)$ are derived by taking reported spectra and normalizing to the absorption at 440 nm [$a^*_p(440)$]. Normalized specific absorption spectra $[a^*_p(\lambda)]_N$ are computed for each of the four phytoplankton groups: diatom and chlorophyte $[a^*_p(\lambda)]_N$ are taken from Sathyendranath et al. (1987), cyanobacteria from Bricaud et al. (1988), and coccolithophores from Morel and Bricaud (1981). Then the specific spectral $a^*_p(\lambda)$ values are derived using mean values at 440 nm. Diatom $a^*_p(440)$ represents the mean of 5 observations containing 4 different spp., chlorophytes 6 observations from 4 spp., cyanobacteria 5 observations from 3 spp., and coccolithophores 3 observations of 1 spp.

Phytoplankton specific scattering coefficients $b^*_p(\lambda)$ are obtained from measurements at 590 nm and extended to the entire spectrum from specific attenuation coefficients (Bricaud et al., 1988). Diatom and chlorophyte specific scattering coefficients at 590 nm, $b^*_p(590)$ and $b*_p(590)$, are the mean of 5 observations and 6 observations, respectively, from Morel (1987), Bricaud and Morel (1986), and Bricaud et al. (1988). Cyanobacteria $b^*_p(590)$ is the mean of 8 observations from Morel (1987), Bricaud and Morel (1986), Bricaud et al. (1988), and Ahn et al. (1992). Coccolithophore $b^*_p(590)$ is derived from the mean of 3 observations from Bricaud and Morel (1986), Bricaud et al. (1988), and Ahn et al. (1992).

We assume no spectral dependence in the backscattering-to-total scattering ratio $\tilde{b}_{bp}$. Ahn et al. (1992) suggested a spectral dependence for cyanobacteria but generally none for the other groups. Reported values for $\tilde{b}_{bp}$ are 0.002 for diatoms (Morel, 1988), 0.00071 for chlorophytes, 0.0032 for cyanobacteria (Ahn et al., 1992), and 0.00071 for coccolithophores (Morel, 1988). Some of these values have come under question based on non-sphericity of many natural phytoplankton populations (Vaillancourt et al., 2004; Whitmire et al., 2010). Based on these results, we increased $\tilde{b}_{bp}$ for chlorophytes and coccolithophores by a factor of 10, but kept them as reported for diatoms and cyanobacteria.

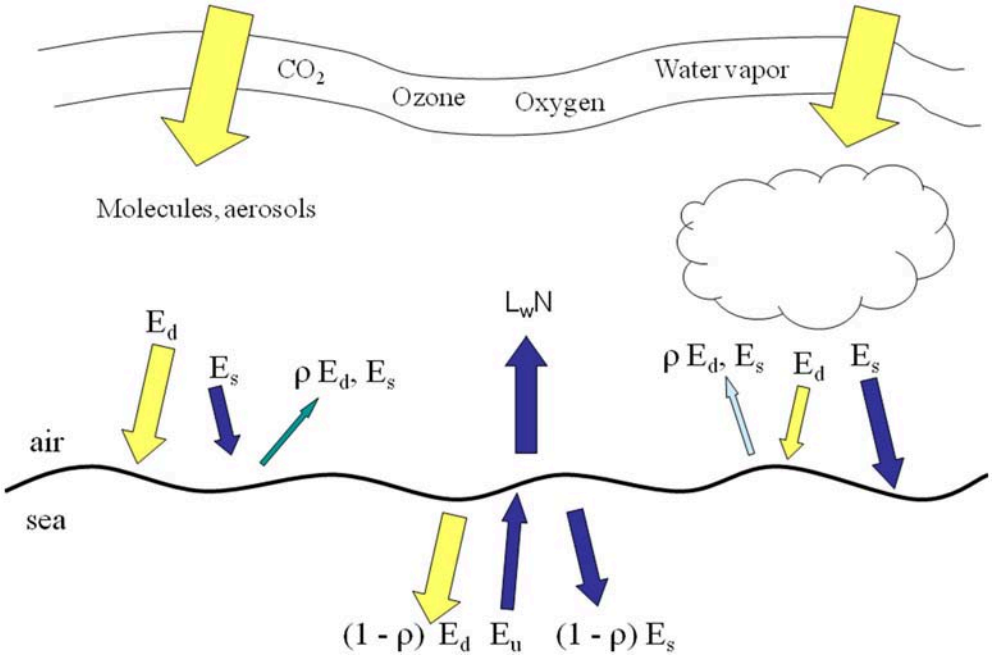


FIGURE 1 | irradiance pathways in OASIM. E_d is direct downwelling irradiance, E_s is diffuse downwelling. ρ surface reflectance, E_u is diffuse upwelling irradiance, and L_wN is normalized water-leaving radiance. All irradiances and radiances are spectrally resolved at 25 nm for E_d, E_s, and E_u. and 1 nm for L_wN.

Detritus

Detritus both absorbs and scatters light (**Figure 2**). Absorption is typically considered an exponential function of wavelength (Roesler et al., 1989; Gallegos et al., 2011).

$$a_d(\lambda) = Da^*_d \exp[-S_d(\lambda - 440)] \quad (2)$$

where $a_d(\lambda)$ is the absorption coefficient of detritus (m^{-1}), D is the concentration of detritus $\mu g\ C\ m^{-3}$, $S_d = 0.013\ nm^{-1}$ (Gallegos et al., 2011) and a^*_d is the mass-specific absorption coefficient of detritus, which is set to 8.0E-5 $m^2\ mg^{-1}$ for small detritus as typically found in oceanic waters (Gallegos et al., 2011). Only organic carbon detritus in the model is used for detrital optics.

Detritus scattering is also taken from Gallegos et al. (2011).

$$b_d(\lambda) = D\, b^*_d\, (550/\lambda)^{0.5} \quad (3)$$

where b_d is the total scattering coefficient, and $b*_d$ is the mass-specific scattering coefficient, which is set as 0.00115 $m^2\ mg^{-1}$, and the backscattering-to-total scattering ratio $\tilde{b}_{bd}$ is 0.005.

PIC

PIC optical properties have been evaluated by Gordon et al. (2009). We adopt this formulation for our simulation. PIC scatters irradiance but does not absorb

$$b_{PIC}(\lambda) = PIC\, b^*_{PIC}\,(\lambda) \quad (4)$$

where PIC is the concentration of PIC ($mgC\ m^{-3}$) and $b^*_{PIC}(\lambda)$ is PIC-specific spectral scattering coefficient from Gordon et al. (2009) in units of $m^2\ mgC^{-1}$. The backscattering-to-total scattering ratio $\tilde{b}_{bpic}$ is from Balch et al. (1996), using their lower bound of 0.01.

CDOC

As a dissolved component, CDOC only absorbs and does not scatter. Its spectral absorption is similar to detritus but with a different slope

$$a_{CDOC}(\lambda) = a^*_{cdoc} \exp[-S_{cdoc}(\lambda - 443)] \quad (5)$$

where a^*_{cdoc} is the mass-specific absorption coefficient of CDOC ($m^2\ mg^{-1}$), $S_{cdoc} = 0.014\ nm^{-1}$ (Bricaud et al., 1981, 2010). S is in the low end range of observations in surface waters of the Equatorial Atlantic (Andrew et al., 2013) but only slightly lower than those observed in the Mediterranean Sea (Organelli et al., 2014). There are few reports of the mass-specific absorption coefficient of CDOC a^*_{cdoc}. We have found three observations in the literature (Carder et al., 1989; Yacobi et al., 2003; and Tzortziou et al., 2007). The more recent two are in agreement at $2.98 \times 10^{-4}\ m^2\ mg^{-1}$ in 4 rivers in Georgia, USA (Yacobi et al., 2003) and $2.78 \times 10^{-4}\ m^2\ mg^{-1}$ as the mean of 4 stations in the Rhode River, Maryland, USA (Tzortziou et al., 2007). Carder et al. (1989) reported a mean over about nearly an order of magnitude lower in the Gulf of Mexico ($4.74 \times 10^{-5}\ m^2\ mg^{-1}$). We choose Yacobi et al. (2003) for our simulation.

Upwelling Spectral Radiance

OASIM uses 25-nm spectral resolution in the 350–7700 nm range in the coupled model for downwelling and upwelling irradiance

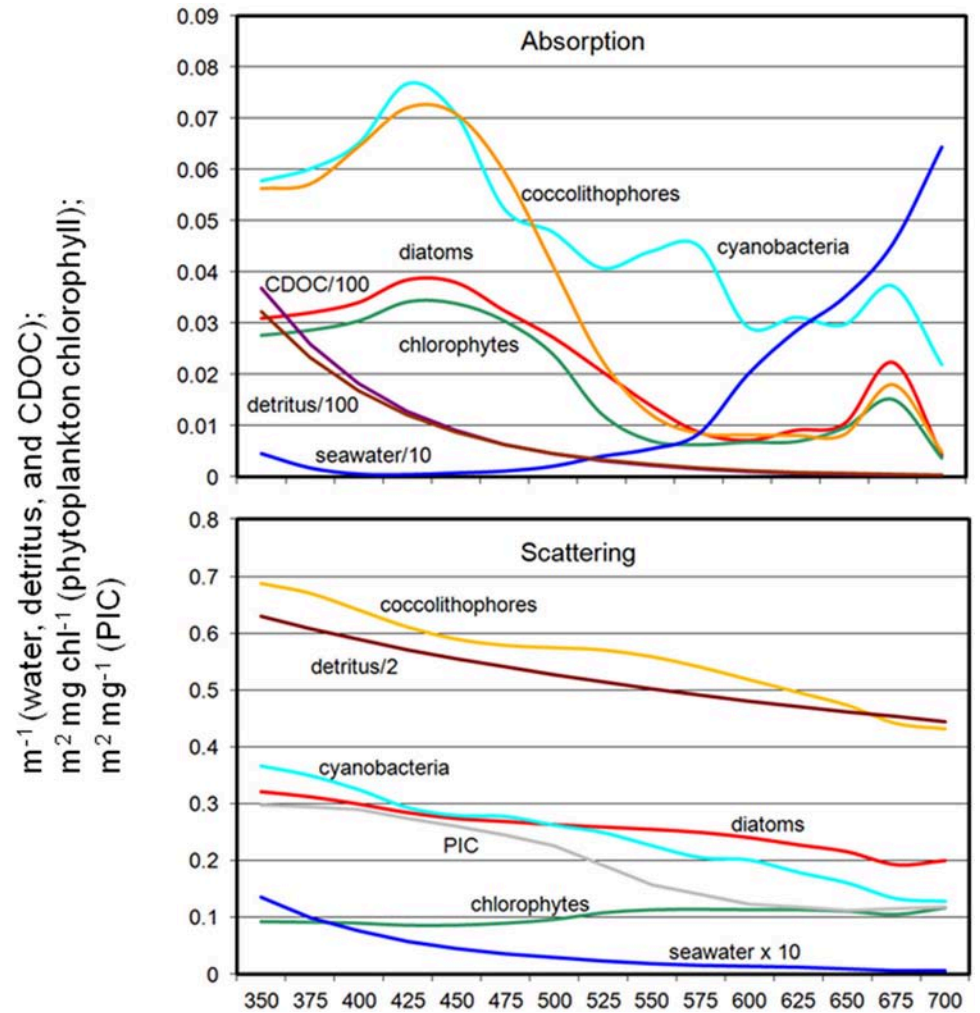


FIGURE 2 | Spectral absorption and scattering coefficients of water, phytoplankton, detritus, PIC, and CDOC in OASIM.

needed for phytoplankton growth and CDOC destruction. For enhanced realism of the PACE simulation of upwelling radiance we increase the spectral resolution to 1 nm. Since all of the optical properties data are available at 5 nm resolution or less, it is reasonable to simply interpolate the 5 nm data. The computation of upwelling spectral radiance $L_wN(\lambda)$ is derived from the coupled expressions of downwelling and upwelling irradiance by Aas (1987) as modified by Ackleson et al. (1994).

$$\frac{dE_d(\lambda)}{dz} = -C_d(\lambda)E_d(\lambda) \quad (6)$$

$$\frac{dE_s(\lambda)}{dz} = -C_s(\lambda)E_s(\lambda) + B_u(\lambda)E_u(\lambda) + F_d(\lambda)E_d(\lambda) \quad (7)$$

$$\frac{dE_u(\lambda)}{dz} = -C_u(\lambda)E_u(\lambda) - B_s(\lambda)E_s(\lambda) - B_d(\lambda)E_d(\lambda) \quad (8)$$

where $E_d(\lambda)$ is the spectral downwelling direct irradiance at the bottom of a model layer, $E_s(\lambda)$ is the downwelling diffuse irradiance, and $E_u(\lambda)$ is the upwelling diffuse irradiance. The attenuation terms C_x (where x is an indicator for the irradiance pathway d for direct downwelling, s for diffuse downwelling, and u for diffuse upwelling), backscattering terms B_x, and forward scattering F_x differ for each of the irradiance pathways because of different shape factors (Aas, 1987; Ackleson et al., 1994) and mean cosines.

$$C_d(\lambda) = [a(\lambda) + b(\lambda)]/\underline{\mu}_d \quad (9)$$

$$C_s(\lambda) = [a(\lambda) + r_s b_b(\lambda)]/\underline{\mu}_s \quad (10)$$

$$C_u(\lambda) = [a(\lambda) + r_u b_b(\lambda)]/\underline{\mu}_u \quad (11)$$

$$B_d(\lambda) = b_b(\lambda)/\underline{\mu}_d \quad (12)$$

$$B_s(\lambda) = r_s b_b(\lambda)/\underline{\mu}_s \quad (13)$$

$$B_u(\lambda) = r_u b_b(\lambda)/\underline{\mu}_u \quad (14)$$

$$F_d(\lambda) = (1 - b_b')b(\lambda)/\underline{\mu}_d \quad (15)$$

where a is the absorption coefficient, b is the total scattering coefficient, b_b is the backscattering coefficient, b_b' is the ratio

of backscattering to total scattering, and $\underline{\mu}$ is the mean cosine (constant for diffuse irradiance, but varies with solar zenith angle for direct irradiance). The shape factors are indicated by the r_x terms, and are specified as in Ackleson et al. (1994). Equation 5 can be solved a priori, which can then be used as a boundary condition, greatly simplifying the solution of the coupled Equations 6, 7.

Equation 8 can be simplified for normalized upwelling radiance since by its definition the surface downwelling irradiance does not include attenuation effects of the atmosphere and the solar zenith angle is assumed to be 0° with overhead sun (Gordon, 1997). Substituting the mean extraterrestrial irradiance (Thuillier et al., 2004) for downwelling irradiance, we can obtain upwelling normalized water-leaving radiance solving the Aas (1987) expressions and correcting for surface reflectance.

$$L_wN(\lambda) = F_o(\lambda, 0^-)(1-\rho)/(n^2Q) \quad (16)$$

where F_o is the mean extraterrestrial irradiance (mW cm^{-2} μm^{-1}) just below the ocean surface (0^-) derived using Aas (1987), ρ is the surface reflectance (0.021), n is the index of refraction (1.341) and Q is the radiance:irradiance distribution function (= π for normalized surface irradiance).

Using 1 nm spectral resolution L_wN not only supports testing PACE sensor and mission concepts, it also simplifies comparison with MODIS-Aqua L_wN by virtue of avoiding band mismatches. The pathways of optical constituents to optical properties to upwelling normalized water-leaving radiances as represented by the NOBM-OASIM global coupled physical-biogeochemical-optical model is depicted in **Figure 3**.

Data Assimilation

Global total chlorophyll from MODIS is assimilated into NOBM using the method described in Gregg (2008). Additionally, global PIC from MODIS (Balch et al., 2005) is assimilated, using the same methodology except that the data are not log-transformed before assimilation. CDOC is assimilated, however, it requires a transformation before the process is executed. There is no available satellite data for CDOC, but a satellite product called a_{CDM} is available (Garver and Siegel, 1997; Maritorena and Siegel, 2005; Maritorena et al., 2010). We use the products from MODIS-Aqua in this effort. This product represents the absorption of both CDOM and detritus (hence the usage of CDM to minimize confusion about its nature). Siegel et al. (2002) estimated the detrital contribution as 12%. We assume this is globally constant and apply a correction of 0.88 to the $a_{CDM}(443)$ data fields prior to assimilation. We recognize this is a potential error, but it is difficult to separate the two in a reflectance inversion methodology because the spectral slopes of absorption are quite similar. The satellite $a_{CDM}(443)$ is assimilated with model $a_{CDOC}(443)$, which is then easily converted to CDOC using the mass-specific absorption coefficient of CDOC (Yacobi et al., 2003).

Upwelling radiances are not assimilated. They are computed using the distributions of optical constituents in the model, their optical properties (**Figure 2**), and Equation 16 at 1 nm spectral resolution.

Model Setup

The model is integrated for 35 years from an initial state using climatological atmospheric forcing, with the new variables PIC and CDOC initialized to 0 concentrations. The model is then run forward in time from 2003 through 2007 using transient atmospheric forcing from MERRA (Rienecker et al., 2011) and assimilating MODIS-Aqua total chlorophyll, PIC, and CDOC.

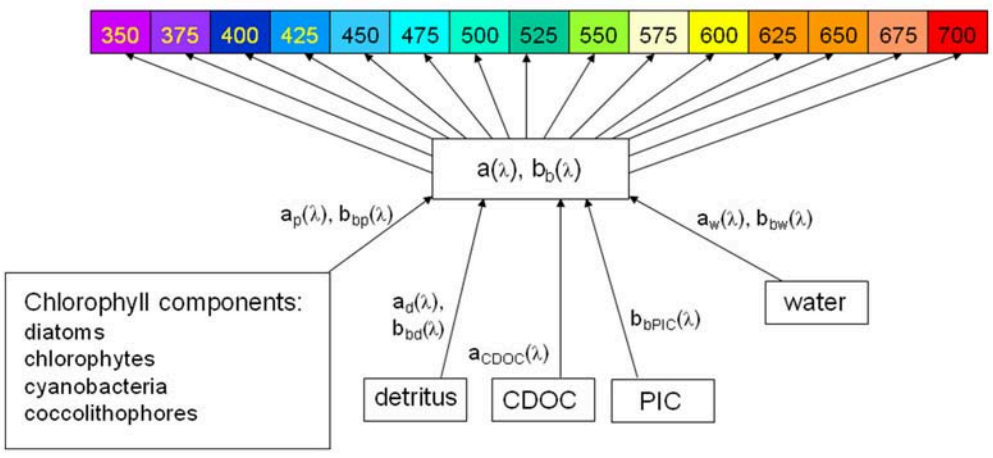


FIGURE 3 | OASIM spectral upwelling radiance and dependencies in the ocean. Shown are the visible bands. The spectral resolution for upwelling radiance is 1 nm. Inherent optical properties are derived from spectral characteristics of water, phytoplankton groups, detritus. PIC, and CDOC.

Statistical Comparison

The optical constituents of the NOBM-OASIM assimilation model are compared to in situ and/or satellite (MODIS) monthly data where and when available. Phytoplankton groups are compared to in situ data while total chlorophyll, PIC, and a_{CDOC} are compared to satellite estimates. The statistics are aggregated over the 12 basins of the global oceans, mean differences (biases) computed, and then correlations computed over the basins. This provides an estimate of large scale correlations and is very stringent considering the low number of observations. The major ocean basins are divided into 3 main regions, high latitudes (poleward of ± 40° latitude): North Atlantic and Pacific and Southern Ocean, mid-latitudes (between ± 40° and ±10° latitude): North Central Atlantic and Pacific, South Atlantic, Pacific and Indian, and North Indian, and tropical basins (between ± 10° latitude): Equatorial Atlantic, Pacific, and Indian. Comparison of assimilated model results with the data used for assimilation is typically insufficient for assessing assimilation performance (Gregg et al., 2009). However, in this case the objective is to simulate dynamic global water-leaving radiances to support a proposed mission, not to assess the assimilation methodology. Here, knowledge of the biases and uncertainties in the underlying ocean optical constituents derived from the assimilation model is best achieved using the satellite data inputs for assimilation. Normalized water-leaving radiance using OASIM and the computed optical constituent distributions are compared to MODIS at the available MODIS bands, 412, 443, 488, 531, 547, and 667 nm. Using 1 nm upwelling radiances at the center of MODIS bands, we can evaluate the simulated bias and uncertainty with MODIS data and avoid model/data band misalignment. These statistics are not aggregated by basin.

RESULTS

We evaluate ocean optical constituents, specifically phytoplankton, total chlorophyll, PIC, and a_{CDOC}, the latter three of which are provided as data sets from MODIS-Aqua. Water

TABLE 1 | Comparison of simulated optical constituents in NOBM-OASIM with data (*in situ* or satellite).

Optical constituent	Difference	Correlation	N (ocean basins)
Diatoms	17.0% (*in situ*)	0.890 $P < 0.05$	11
Chlorophytes	−16.2% (*in situ*)	−0.318 NS	10
Cyanobacteria	−2.4% (*in situ*)	0.732 $P < 0.05$	11
Coccolithophores	5.3% (*in situ*)	0.716 $P < 0.05$	10
Total Chlorophyll	−35.9% (satellite)	0.869 $P < 0.05$	12
PIC	−28.5% (satellite)	0.868 $P < 0.05$	12
a_{CDOC}	−24.6% (satellite)	0.890 $P < 0.05$	12
Detritus	NA	NA	NA

NS indicates not significant at 95% confidence. NA indicates data not available for comparison. The satellite comparison uses MODIS-Aqua and model data used are co-located and coincident with monthly mean MODIS data.

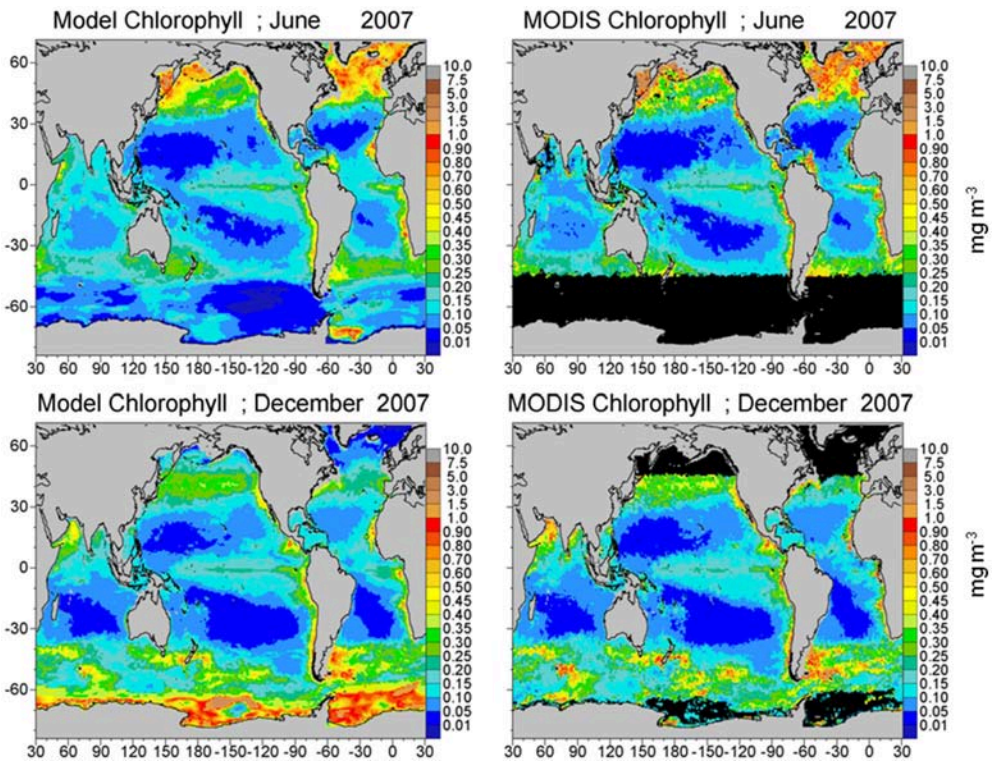


FIGURE 4 | Model assimilated total chlorophyll for June and December 2007 compared to MODIS-Aqua chlorophyll.

is a constant background and we are not aware of global data on detritus. We evaluate water-leaving radiances by comparing model upwelling radiances at MODIS-Aqua wavelengths with those MODIS-Aqua radiance data.

Global Ocean Optical Constituents

Total chlorophyll from the assimilated NOBM-OASIM model is within −35.9% of satellite data (model low), with a correlation across basins of 0.869 ($P < 0.05$; **Figure 4**; **Table 1**). The model is low because of uncorrected a_{CDM} in the satellite data, especially near coasts and river mouths, which artificially drives up the estimates of chlorophyll.

Phytoplankton group relative abundances are positively correlated with in situ data for diatoms, cyanobacteria, and coccolithophores ($P < 0.05$) but chlorophytes are not correlated (**Table 1**). All four groups have relative abundance biases < ±20% compared to in situ data, with diatoms the largest at 17%.

Assimilated PIC is correlated with satellite estimates ($P < 0.05$) and concentrations are within −28.5% (**Figure 5**; **Table 1**). Simulated PIC is overestimated and more widespread in the Southern Ocean in December, but otherwise exhibits similar variability as indicated by the correlation coefficient ($r = 0.868$). It is unable to capture the localized extreme high concentrations in June in the northern high latitudes, which leads to model underestimates globally. Model comparison of a_{CDOC} (443 nm) is within −24.6% of satellite estimates of a_{CDM} (443 nm) (**Table 1**), which represents the combined absorption of dissolved matter and particulate matter (detritus). A basin correlation coefficient of 0.890 ($P < 0.05$) is obtained (**Table 1**). Maps of global distributions for June and December 2007 illustrate the comparison between model and data (**Figure 6**). Although river discharge is not included in the model, high a_{CDOC} is produced at major river mouths (e.g., Amazon, Orinoco, Congo) via the assimilation of a_{CDM} (see **Figure 6**).

Global Normalized Water-Leaving Radiances

The mean of the global median difference of model normalized water-leaving radiances with MODIS-Aqua radiances for all 6 bands for the period 2003–2007 is −0.074 mW cm^{-2} μm^{-1} sr^{-1} (−10.4%) with a mean semi-interquartile range of 0.077 and a significant correlation of 0.706 ($P < 0.05$). There is a positive and significant correlation with all the simulated radiances with satellite data (**Figure 7**). The largest relative difference (−30%) and lowest correlation ($r = 0.48$) occurs in the longest MODIS band, 667 nm (**Figure 7**). Band 1 (412 nm) has the largest absolute difference (−0.19 mW cm^{-2} μm^{-1} sr^{-1}; **Figure 7**), but only the third largest relative difference with a mean of −12.5%, and it has a high correlation of 0.946. All simulated radiances are low relative to data (**Figure 7**). Correlations of the longer visible wavelengths, 531, 547, and 667 nm are much lower than those of the shorter wavelengths.

Global maps of water-leaving radiances illustrate the spatial agreement and discrepancies between the model and satellite data

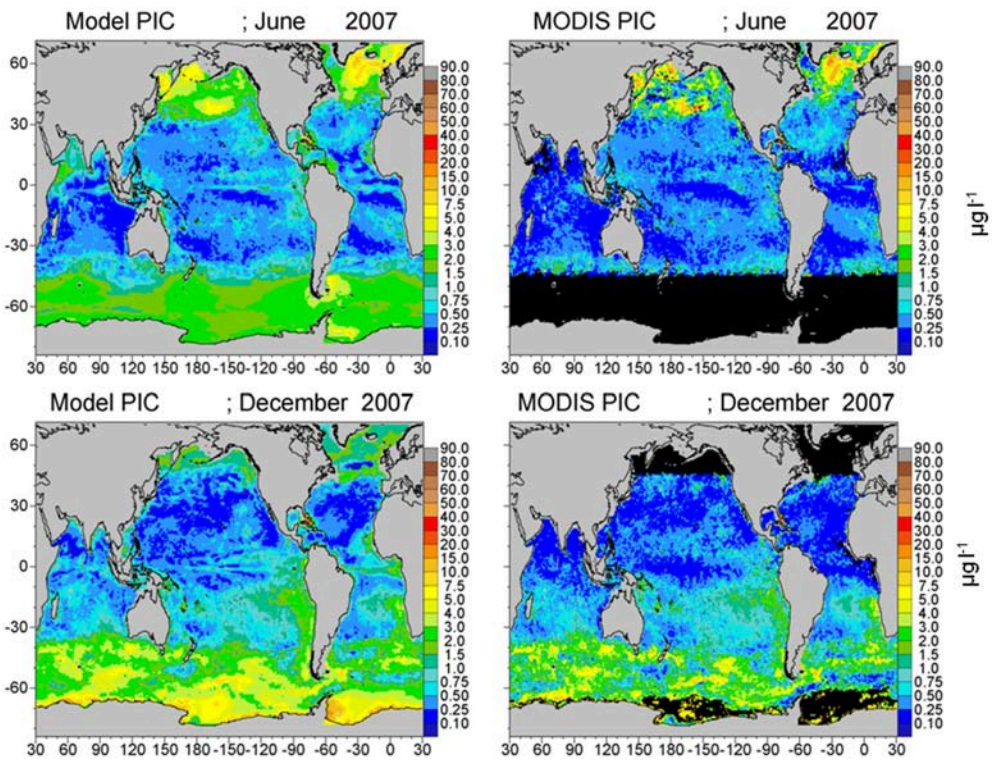


FIGURE 5 | Model assimilated PIC for June and December2007 compared to MODIS-Aqua PIC.

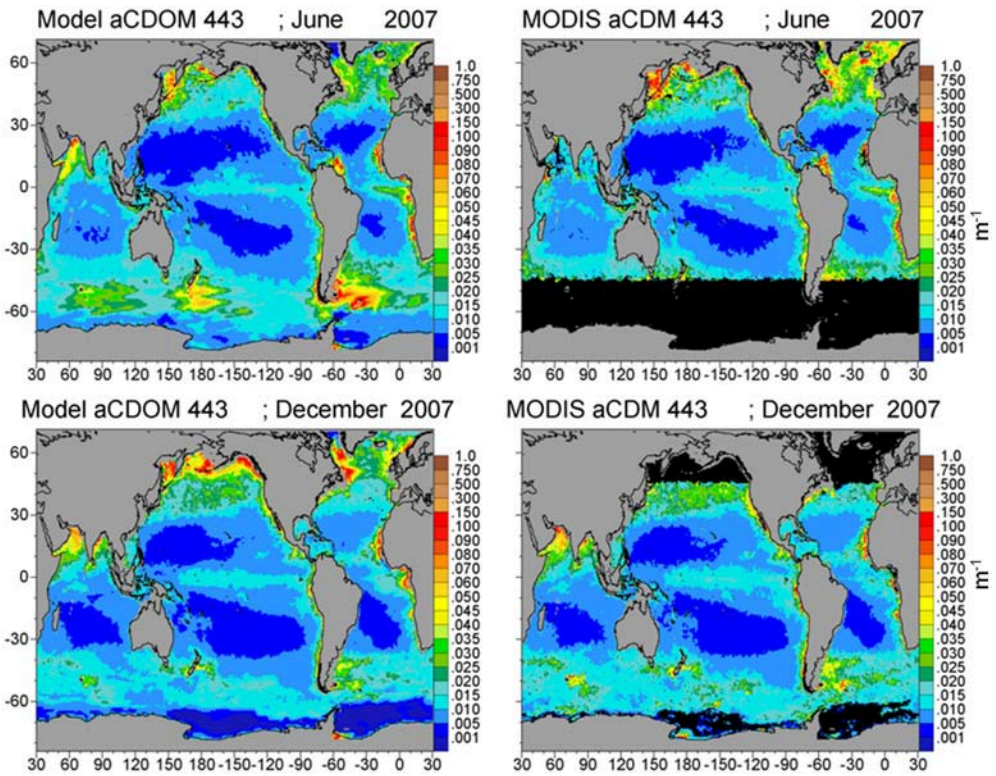


FIGURE 6 | Model assimilated a_{CDOC} 443 nm for June and December 2007 compared to MODIS-Aqua a_{CDM} 443 nm.

(**Figures 8**–**10**). The spatial distributions reflect the biases and correlations shown in **Figure 7**. Low biases in model radiances are apparent for all bands, but the locations differ. Low model radiances are most apparent for the shorter wavelengths (412 and 443 nm) in the central gyres (**Figure 8**). Mid-range bands (531 and 547 nm) show low model biases in the northern high latitudes (**Figures 9**, **10**). The longest MODIS band (667 nm) does not exhibit a model bias as shown in **Figure 7**, but the bias is below the spectral resolution of the figure.

Maps of normalized water-leaving radiances at various wavelengths from the 1 nm hyper-spectral resolution capability are shown in **Figures 11**, **12**. The radiance wavelengths are broken into the two figures to capture variability over the widely-ranging radiance values shown. The second set of radiance maps (**Figure 12**) uses a different scale for radiance values. Otherwise, spatial variability in these radiances is not visible.

Two locations in the North Pacific Ocean are selected to show hyperspectral variability in different oceanic environments (**Figure 13**). One is a low-chlorophyll central gyre location which is characterized by low chlorophyll, PIC and CDOC, southwest of Hawaii. The other is in the high latitude North Pacific just south of the Aleutian Islands, where high chlorophyll, PIC and CDOC prevail. Hyperspectral 1 nm normalized water-leaving radiances show considerable differences in magnitude and local spectral slopes, suggesting the potential for discrimination of ocean constituents from PACE.

DISCUSSION

We have described a comprehensive model of optical constituents and their influences on hyper-spectral upwelling radiance in the global oceans. The model contains a representation of major optical constituents, namely, water, total chlorophyll, four major phytoplankton taxonomic/functional groups, organic detritus, PIC, and CDOC. All except water are prognostic variables in the model with individual sources and sinks, and with full dynamical capability arising from advection and diffusion processes in the global oceans.

Normalized water-leaving radiances from the global distributions of optical constituents have been quantitatively compared to MODIS-Aqua radiances for the 6 wavelengths available at 412, 443, 488, 531, 547, and 667 nm. These 6 discrete wavelengths provide only a partial basis for estimating the potential of a global dynamical model to represent the hyper-spectral capability of the next generation PACE mission. Thus, the error estimation is incomplete, and relevance to PACE and its ability to simulate future global hyper-spectral radiances is unconfirmed. However, the comparison of the model with the 6 MODIS bands suggests a level of skill sufficient to support some analysis of mission capability and design, and the level of caution necessary to proceed in these activities is quantified here.

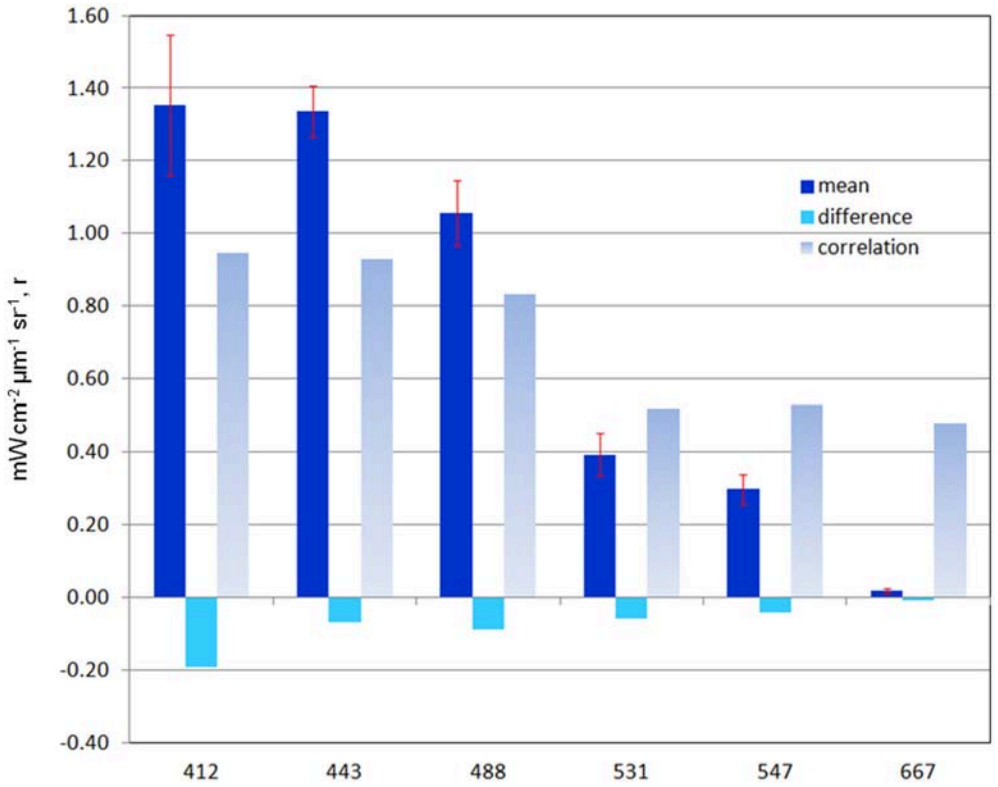


FIGURE 7 | Global statistics on model normalized water-leaving radiances $L_wN(\lambda)$ compared to MODIS-Aqua data for 2003–2011. Mean radiance and difference is mW cm^{-2} μm^{-1} sr^{-1}. Correlation is *r*-value. All correlations are significant ($P < 0.05$: $N > 3.7 \times 10^6$. Error bars represent semi-interquartile range.

Global Ocean Optical Constituents

The global ocean biology model is optically comprehensive, but it is not complete. There are optical constituents in the oceans that are not included in the model. Some can be important, sometimes globally but most often regionally. For example, bacteria and virus scattering is not present in the model. Bacteria scattering is considered an important component of the scattering from the living part of the particulate pool, possibly dominating the phytoplankton (Balch et al., 2002; Stramski et al., 2004). However, the scattering contributions from the living components are estimated to be small relative to detritus (Stramski et al., 2004). We assume here that bacteria covary with detritus. Virus scattering is disputed. Balch et al. (2002) suggest it may be important while Stramski et al. (2004) consider it negligible.

Minerals/suspended sediments are not included. These are most important near river mouths at times of high discharge, but they also occur from particulate deposition from the atmosphere, such as desert dust (Wozniak Stramski, 2004) or organic carbon from biomass burning. Absorption by mycosporine-like amino acids (Moisan and Mitchell, 2001) is not included in the model. This is most important in the ultraviolet spectrum, and casts suspicion on the simulated representations of water-leaving radiances in this spectral region by the model. PACE is nominally expected to detect as low as 350 nm (PACE Mission Science Definition Team Report, 2012), but there may be interest in expanding that range if it is technically and economically feasible. The most recent configuration concept is to expand the detection limit to 320 nm. Inclusion of the effects of mycosporine-like amino acids should be included in future improvements of the biological global model.

Finally, four phytoplankton groups cannot possibly represent the range and complexity of the phytoplankton taxa living in the oceans. Unfortunately, detailed knowledge of the optical, physical, and physiological properties of the world's ocean phytoplankton, which is required to parameterize our coupled optical, physical, and biological model, is not available. We recognize our four groups as a shortcoming, but they do capture a substantial range of functionality. Diatoms represent the fast growing, fast sinking component particularly important in the carbon and silicon cycles. Cyanobacteria represent the functional opposite, as a slow growing, nearly floating, very small phytoplankton that occupy the nutrient-desolate vast ocean gyres, and additionally have a limited nitrogen-fixing capability (Rousseaux et al., 2013). Coccolithophores represent a unique category of calcium-producing phytoplankton, which scatter light out of the oceans effectively and play a role in the carbon cycle by affecting alkalinity in addition to photosynthesis and respiration processes. Finally, chlorophytes represent (or at least are intended to represent) intermediate phytoplankton with characteristics between diatoms and cyanobacteria. It

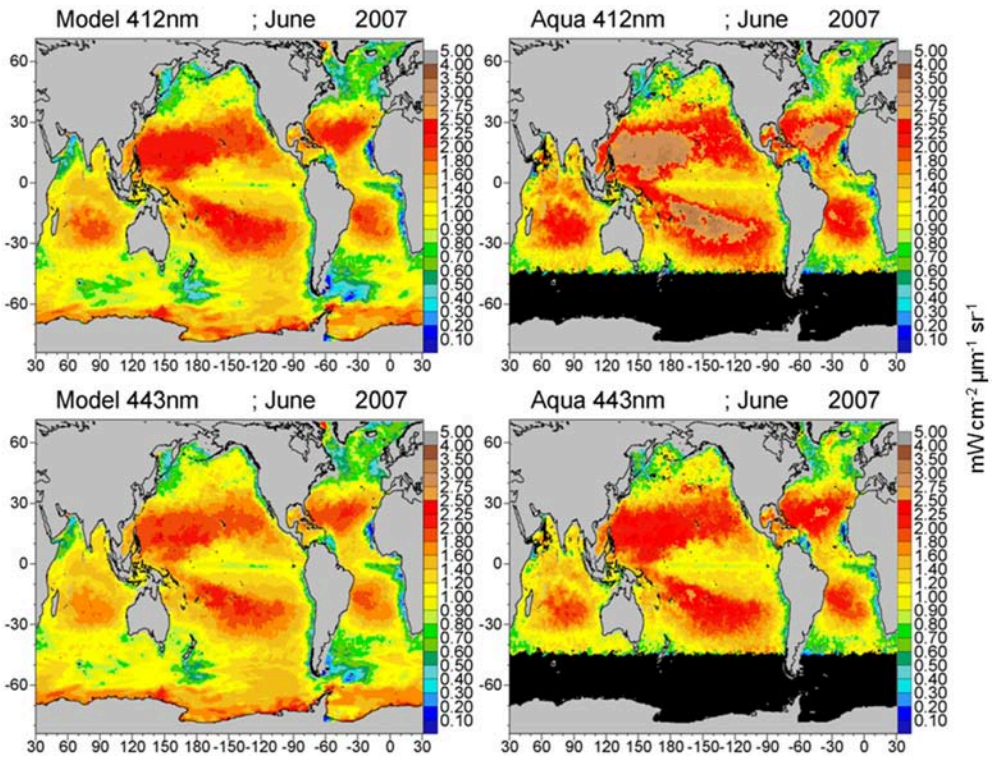


FIGURE 8 | Model normalized water-leaving radiances $L_wN(\lambda)$ for 412 and 443 nm compared to MODIS-Aqua radiances.

is this intermediate category that is most under-represented here and is where much of the diversity of the global ocean arises.

The fact that chlorophytes are not significantly correlated with in situ data in the model is particularly important because they are the only group in the model representative of the diverse phytoplankton component between the functional extremes of diatoms and cyanobacteria, save for the unique coccolithophore class. This is a deficiency in the model as it pertains to PACE and we acknowledge that their lack of correlation with data is important. However, in the model we assume chlorophytes represent a very wide range of phytoplankton, often reported to as nanoplankton. Since in situ data sets rarely specifically identify chlorophytes, we compare our model chlorophytes to in situ data reports of nanoplankton, non-diatoms or non-pico-prokarytotes, representing this middle ground between diatoms and cyanobacteria. We note that most of the lack of correlation with in situ data occurs in the high latitudes, where chlorophytes are not common, but other types on nanoplankton are sometimes abundant. The abundance of these reported nanoplankton in the high latitudes, coupled with the near-absence of chlorophytes in the model, is the cause of the lack of correlation. The model representation of chlorophyte abundance corresponds much more closely with reported observations of nanoplankton in the lower latitudes, suggesting that simulation of PACE radiances in these basins is likely to be more realistic.

Using Data Assimilation to Improve the Representation of Global Optical Constituents

The assimilation of chlorophyll has been demonstrated to improve the representation of distributions regionally and globally (Hu et al., 2012; Fontana et al., 2013; Gregg and Rousseaux, 2014). Assimilation of PIC and a_{CDM} has not been attempted globally, to our knowledge. Our purpose in assimilating PIC and a_{CDM} is not novelty but fidelity. The optical properties of PIC have been established (Balch et al., 1996; Gordon et al., 2009) and one can find models of production and dissolution in the literature (Buitenhuis et al., 2001; Gangsto et al., 2011; Barrett et al., 2014). Our parameterization of sinking processes is a matter of trial and error using global satellite fields of PIC from MODIS-Aqua. Assimilation of a_{CDM} is a larger challenge. Although assimilation of optical properties, in particular the diffuse attenuation coefficient, has shown value (Ciavatta et al., 2014), the assimilation of a_{CDM} is more problematic because there a few examples of its use in coupled physical-biogeochemical models (e.g., Buitenhuis et al., 2001; Xiu and Chai, 2014; Dutkiewicz et al., 2015) We approach the problem in a bottom-up fashion, adding a dynamical tracer to the biogeochemical model suite, i.e., CDOC, which has the optical properties of $a_{CDOC}(\lambda)$. The characterization of the biological production and loss terms for CDOC is more or less straightforward, as it can be related to those from the optically inert DOC (e.g., Aumont et al., 2002). Loss of CDOC via the

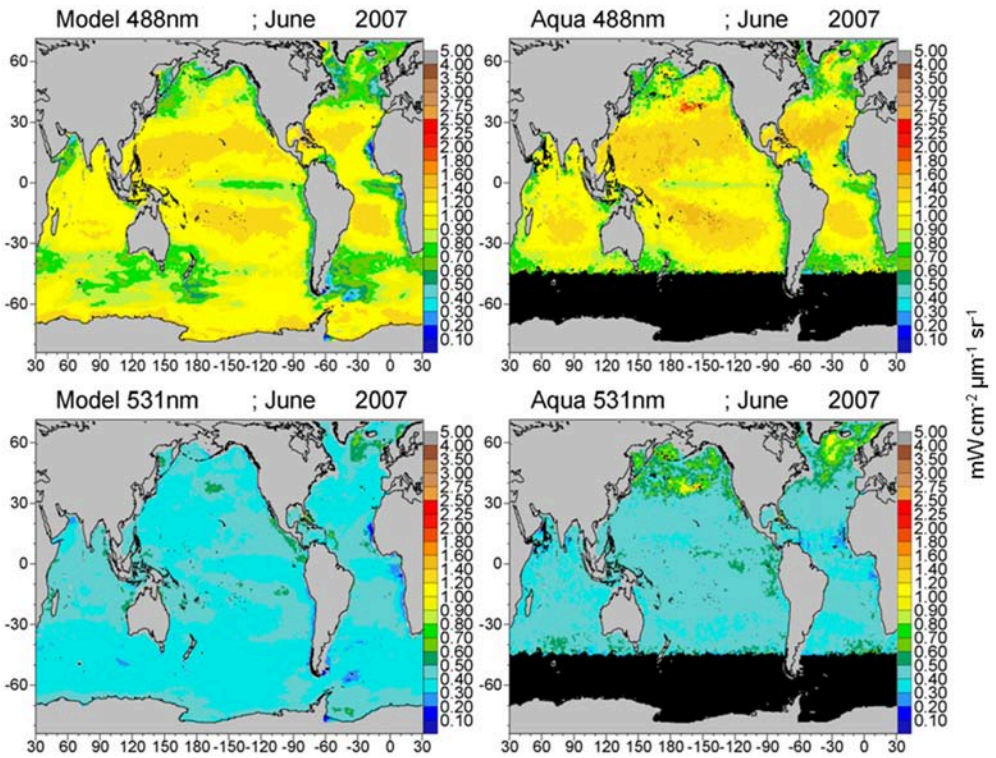


FIGURE 9 | Model normalized water-leaving radiances $L_wN(\lambda)$ for 488 and 531 nm compared to MODIS-Aqua radiances.

absorption of spectral irradiance is more difficult. Although the absorption characteristics are well-established, how that relates to CDOC concentration and subsequent destruction is difficult to quantify. There is regional information on defining a quantum yield for CDOC photolysis, φ_{cdoc} (e.g., Reader and Miller, 2012, 2014), but we require a global spectrally integrated solution. We consider our parameterization of φ_{cdoc} to be tenuous, but we take consolation that the assimilation guides us to a reasonable result in the end, and even rectifies the absence of river input in the model, which is a major source of CDOC to the oceans. For the present purpose of providing a model to assist in the early stages of development of a future mission, we believe our approach has support as an initial step. The statistical comparison of CDOC distributions with satellite data supports this approach as well (**Table 1**; **Figure 6**).

Global Water-Leaving Radiances

The comparison of model water-leaving radiances with MODIS-Aqua at the 6 MODIS bands suggests some skill in the simulation: the mean of the global median difference is -0.077 ± 0.079 mW cm^{-2} µm^{-1} sr^{-1} (-10.4%). A statistically significant correlation with all the simulated radiances with satellite data is found (**Figure 7**), although some of the correlation coefficients are low. We emphasize that the radiances are not assimilated. We emphasize that the radiances are not assimilated. Rather, they are the result of the distribution of optical constituents in the coupled model.

The longer visible wavelengths, 531, 547, and 667 nm have lower correlations with satellite data than the shorter ones. There is much less spatial variability in the longer wavelengths (**Figures 9**, **10**). Ocean color sensors have much larger uncertainty in these wavelengths (Mélin et al., 2016) which contributes to the decrease in correlation of these radiances here.

The model is always low relative to the MODIS normalized water-leaving radiances. The low model radiances occur in different regions for the different bands. For the shortest MODIS wavelengths, 412 and 443 nm, largest biases occur in the ocean gyres (**Figure 8**), where ocean biological optical constituents are at their lowest magnitudes. The 412 nm band has a larger model-data discrepancy than the 443 nm band (**Figure 7**). For the mid-range bands 531 and 547 nm, the model-data discrepancies occur in the northern high latitudes.

The model low bias for $L_wN(412)$ and $L_wN(443)$ in the central gyres suggests either missing scattering in the model or overestimated absorption. These regions are biologically the most barren regions in the global ocean, where the main optical constituent is water. The southeast Pacific gyre has been the subject of an intensive field campaign (BIOSOPE), and several investigators have relied upon this data set to revise the understanding of the optical properties of seawater (Morel et al., 2007; Lee et al., 2015), CDOM and particulate detrital absorption (Bricaud et al., 2010), and total particulate backscattering (Twardowski et al., 2007). The Lee et al. (2015) seawater absorption revision reduced the absorption coefficients,

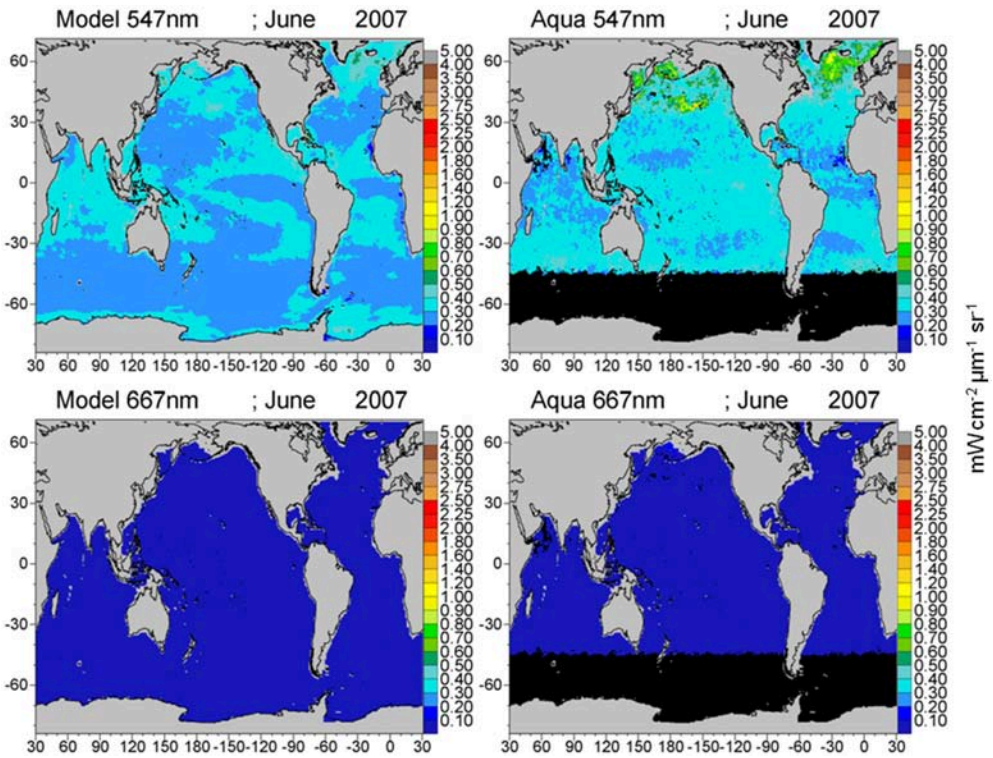


FIGURE 10 | Model normalized water-leaving radiances $L_wN(\lambda)$ for 547 and 667 nm compared to MODIS-Aqua radiances.

thus producing more scattering, which has helped in our model here, since the revision is used in our calculations. Residual underestimation of scattering and/or overestimation of absorption still prevails in the simulation.

It is possible that the exclusion of mineral scattering in the model is important in the central gyres. However, this argument would be more persuasive for the North Central Pacific and North Central Atlantic gyres than the South Pacific gyre, since there are few atmospheric depositions to this region. One cannot neglect the possibility of radiative model error as well. Perhaps the use of empirical constants in a remote sensing reflectance algorithm, such as Lee et al. (2002) or Gordon et al. (1988), would improve radiances. However, this would sever connections in the radiative modeling system, which uses an analytical model for simulation of both irradiance transmittance in the ocean and the irradiance and radiance re-emerging to and above the surface.

Finally, the spectral slope of detrital absorption $S_d(\lambda)$ used here, 0.013 nm^{-1}, which was derived from assessment of small particulates in the Chesapeake Bay (Gallegos et al., 2011), is higher than that derived from the southeast Pacific by Bricaud et al. (2010), 0.0094 nm^{-1}. This could lead to the higher absorption and subsequent lower backscatter, especially in the shorter wavelengths, as we observe here. How much will depend upon the concentration of detritus in this region and the other central gyres.

The model also exhibits low radiances compared to MODIS for the 531 and 547 nm bands (**Figures 9**, **10**), except these are mostly located in the northern high latitudes. These discrepancies appear to be related to the distributions of PIC (**Figure 5**). Model PIC distributions here largely correspond with satellite distributions, although local maxima in the southern central North Pacific and the Greenland Sea are subdued in the model (**Figure 5**). These two locations are responsible for the largest disagreements. However, additional local maxima in satellite PIC occur in the northern Bering Sea and western Sea of Okhotsk (**Figure 5**), that are not accompanied by high water-leaving radiances in the MODIS 531 and 547 nm bands (**Figures 10**, **11**). High chlorophyll (**Figure 4**) and a_{CDOM} (**Figure 6**) in the model and MODIS likely suppress the scattering of PIC in the northern Bering Sea and Sea of Okhotsk. But the lack of representation of the high scattering by PIC in the south-central North Pacific and Greenland Sea results from the spatially smoother PIC distributions in the model compared to MODIS (**Figure 5**). Overall widespread higher radiance dispersed throughout the northern basins in likely due to inadequate PIC scattering in the model, considering the correspondence between model and satellite PIC distributions. Excessive absorption by other constituents in the model can contribute to the differences in radiances between model and data here. Such high absorption would likely be due to phytoplankton (particularly diatoms, which are predominant in the North Pacific), or coccolithophores which are prevalent in the North Atlantic.

Global maps of selected normalized water-leaving radiances other than those coincident with MODIS-Aqua show

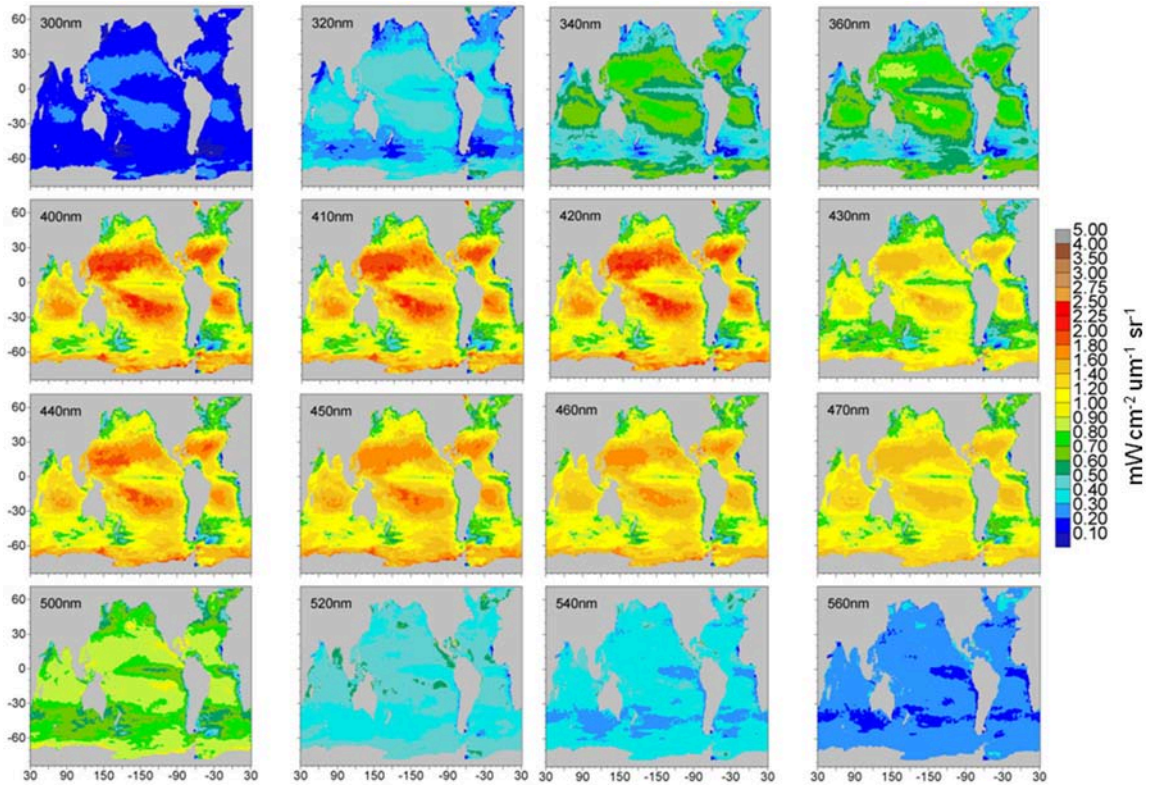


FIGURE 11 | Model normalized water-leaving radiances for selected wavelengths in the ultraviolet and visible region.

considerable spectral and spatial variability (**Figures 11**, **12**). The figures are divided into two groups because the spectral range is so large that different scales must be utilized. **Figure 10** shows radiances from two ultraviolet-b bands (300 and 320 nm), to an ultraviolet-a band (340 nm), and 13 through mid-range visible (360–560 nm). There is a steady increase in radiance intensity as we progress from shorter to longer wavelengths until about 400–410 nm, then a slow decline to 560 nm. An exception to this trend is the radiance at 430 nm, which shows a sharp decline relative to its neighbors at 420 and 440 nm (**Figure 11**). This is due to a local minimum in the extraterrestrial irradiance that is employed at 1 nm resolution (Thuillier et al., 2004). These local minima and maxima occur occasionally in the radiance spectrum and represent a potential issue when choosing band locations for PACE. There can be very large swings in signal strength in short wavelength segments.

The second selection of radiance wavelengths, at extreme ultraviolet-b along with the long end of visible and 3 near infrared wavelengths (**Figure 12**), shows increasing intensity from 250 through 270 nm, and another from 600 to 630 nm, before reversing from 650 to 720 nm. There is very little normalized water-leaving irradiance at 720 nm and spatial variability will require another scale change to be visible. There is another anomaly, this time a local maximum, at 270 nm, again due to the high spectral variability in the extraterrestrial irradiance. This set of radiances, with the possible exceptions of the shorter 600 nm bands, suggests that ocean signal detection from a satellite will be challenging. The longer 600 nm wavelengths are conventionally used for atmospheric correction since there is so little ocean contribution to the normalized water-leaving radiance (e.g., Gordon, 1997) while NIR bands (e.g., Wang et al., 2016) have shown additional promise for the rare conditions when the ocean does contribute here.

Potential Uses for Pace Mission Design and Analysis

The hyper-spectral 1 nm resolution ocean model presented here suggests skill for simulating global normalized water-leaving radiances, as shown by the comparison with the moderate resolution bands for MODIS-Aqua. Quantitative error characterization shows the limits of usefulness in the MODIS bands and the potential for simulating radiances outside the current satellite observational capability. This suggests at least some usefulness for pre-launch PACE design and analysis activities, guided by due caution of the limits of the simulation.

Representation of remotely-sensed normalized water-leaving radiances may be approached using airborne (e.g., Airborne Visible/Infrared Imaging Spectrometer, Portable Remote Imaging SpectroMeter), or in situ data, or coastal spaceborne

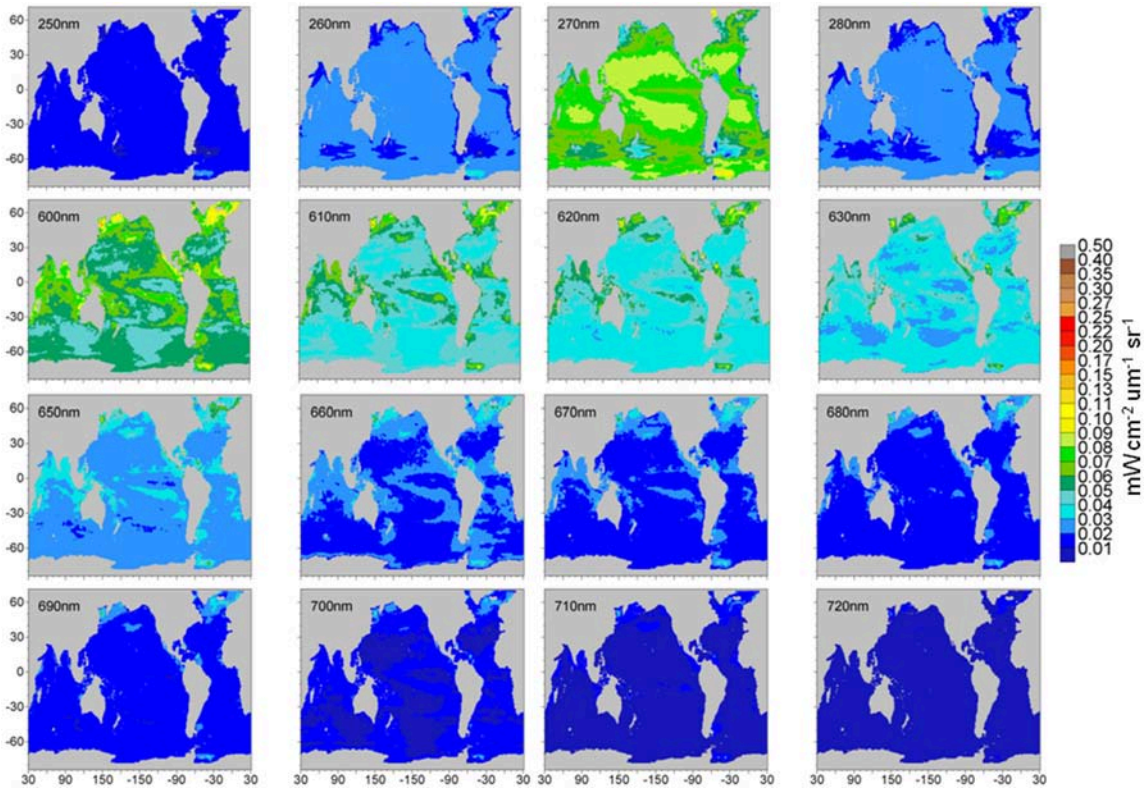


FIGURE 12 | Model normalized water-leaving radiances for selected wavelengths in the ultraviolet, long visible, and near-infrared region. Note scale change.

imagers, such as the Hyperspectral Imager for the Coastal Ocean. However, the global observing simulation capability of the present assimilated model can contribute in other important ways that airborne, in situ and limited spaceborne data cannot.

The most important attribute that separates PACE from previous ocean color missions is its global hyper-spectral resolution capability. The global simulation described here at 1 nm can help clarify questions about band selection, specifically choice of bands, band widths, number of bands and their center location. Variability over orbital tracks encountering a range of solar and satellite angles complicates band selection decisions in ways that in situ and most airborne activities cannot resolve. The global seasonal nature of the simulation assists in understanding potential signal strength issues over the diverse regions and seasons encountered in a global mission. It is possible to sample the simulated 1 nm bands in various scenarios to observe and optimize their location and widths, subject to the viewing constraints of an orbiting platform. Optical effects, such as spectral response function can be included in the analysis. As mission design and construction proceeds, issues can arise and tradeoffs must be assessed. These often include signal-to-noise ratios, detector saturation effects, gain selection and operation (if applicable), stray light, and bright target recovery. The existence of the simulation described here can provide numerical answers from an orbital perspective, even if approximate, as these issues emerge. The limitations of the model are quantitatively characterized here and can be factored into the decisions on how to proceed. A much more modest simulation, using only a single global map of ocean color data derived from the entire CZCS mission (Gregg et al., 1997), proved helpful in designing and managing the SeaWiFS mission, which, like PACE, had no global observational precedent.

The second most important feature of this simulation is to provide a platform for algorithm development activities. Although the phytoplankton differentiation in the model is necessarily simplified, it can be used in coarse algorithm activities. At worst, algorithms that cannot differentiate among the simple phytoplankton assemblage in the simulation would likely have difficulties in actual ocean observations, where the phytoplankton diversity is enormous.

The simulation can also assist in studies of data collection strategies on orbit. Seasonal variability in phytoplankton/PIC/CDOC distributions is explicitly incorporated in the simulation to include a full representation of optical combinations as seen to date with current missions. If coupled with a similarly comprehensive and hyper-spectral atmospheric simulation, and an orbital viewing platform, the combined models can be used to

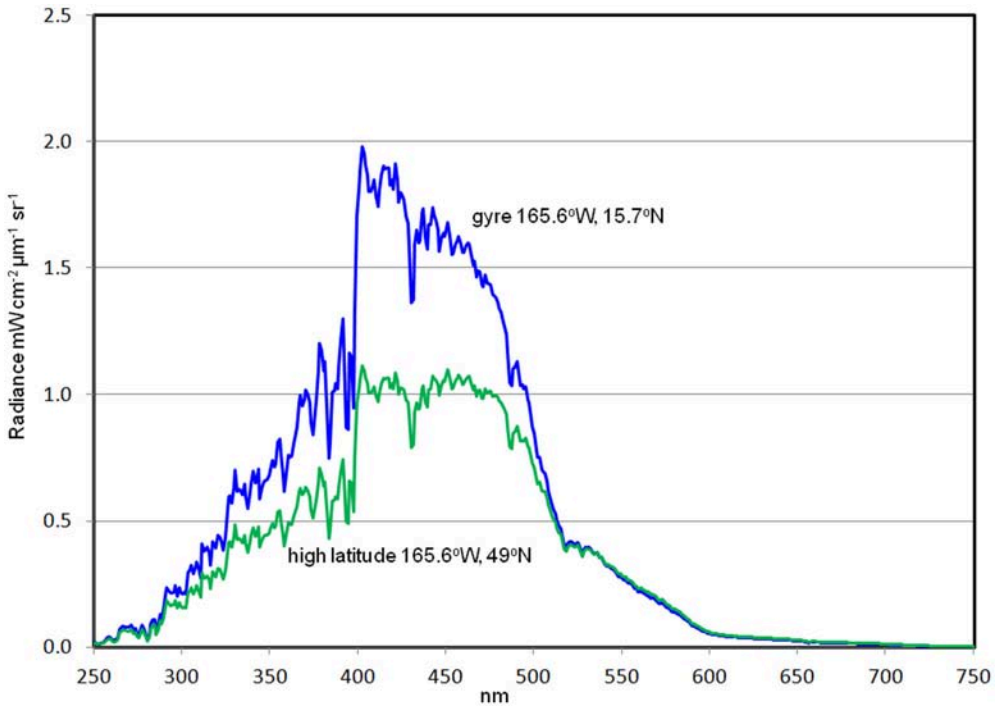


FIGURE 13 | Normalized water-leaving radiances from two locations in the Pacific Ocean: a gyre location (low chlorophyll) and a high latitude location (high chlorophyll).

explore signal retrieval at the sensor and help maximize the ability to meet the challenging goals of this ambitious mission.

AUTHOR CONTRIBUTIONS

WG was responsible for writing and organizing the manuscript. CR was responsible for deriving hyper-spectral data and assisting in the writing and reviewing of the manuscript.

ACKNOWLEDGMENTS

We thank the NASA/MERRA Project, the MODIS Ocean Color Processing Team, and the algorithm developers for PIC and a_{CDM} for the data sets and public availability. We also thank the members of the PACE Science Team for optics parameters and data and two reviewers. This work was supported by NASA PACE, S-NPP, CMS, and MAP Programs. Hyperspectral model radiances are available at the GMAO web site https://gmao.gsfc.nasa.gov/research/oceanbiology/data.php.

REFERENCES

Aas, E. (1987). Two-stream irradiance model for deep waters. *Appl. Opt.* 26, 2095–2101.

Ackleson, S. G., Balch, W. M., and Holligan, P. M. (1994). Response of water-leaving radiance to particulate calcite and chlorophyll *a* concentrations: a model for Gulf of Maine coccolithophore blooms. *J. Geophys. Res.* 99, 7483–7499.

Ahn, Y.-H., Bricaud, A., and Morel, A. (1992). Light backscattering efficiency and related properties of some phytoplankters. *Deep Sea Res.* 39, 1835–1855.

Aiken, J., Fishwick, J. R., Lavender, S., Barlow, R., Moore, G. F., Sessions, H., et al. (2007). Validation of MERIS reflectance and chlorophyll during the BENCAL cruise October 2002: preliminary validation of new demonstration products for phytoplankton functional types and photosynthetic parameters. *Int. J. Remote Sens.* 28, 497–516. doi: 10.1080/01431160600821036

Alvain, S., Moulin, C., Dandonneau, Y., and Bréon, F. M. (2005). Remote sensing of phytoplankton groups in case 1 waters from global SeaWiFS imagery. *Deep Sea Res. I Oceanogr. Res. Pap.* 52, 1989–2004. doi: 10.1016/j.dsr.2005.06.015

Alvain, S., Moulin, C., Dandonneau, Y., and Loisel, H. (2008). Seasonal distribution and succession of dominant phytoplankton groups in the global ocean: a satellite view. *Global Biogeochem. Cycles* 22:GB3001. doi: 10.1029/2007GB003154

Andrew, A. A., DelVecchio, R., Subramaniam, A., and Blough, N. V. (2013). Chromophoric dissolved organic matter (CDOM) in the equatorial atlantic ocean: optical properties and their relation to CDOM structure and source. *Mar. Chem.* 148, 33–43. doi: 10.1016/j.marchem.2012.11.001

Aumont, O., Belviso, S., and Monfray, P. (2002). Dimethylsulfoniopropionate (dmsp) and dimethylsulfide (dms) sea surface distributions simulated from a global three-dimensional ocean carbon cycle model. *J. Geophys. Res.* 107:3029. doi: 10.1029/1999JC000111

Balch, W. M., Gordon, H. R., Bowler, B. C., Drapeau, D. T., and Booth, E. S. (2005). Calcium carbonate measurements in the surface global ocean based on Moderate-Resolution Imaging Spectroradiometer data. *J. Geophys. Res.* 110:C07001. doi: 10.1029/2004JC002560

Balch, W. M., Kilpatrick, K. A., and Trees, C. C. (1996). The 1991 coccolithophore bloom in the central North Atlantic. 1. Optical properties and factors affecting their distribution. *Limnol. Oceanogr.* 41, 1669–1683.

Balch, W. M., Vaughn, J. M., Novotny, J. F., Drapeau, D. T., Goes, J. I., Booth, E., et al. (2002). Fundamental changes in light scattering associated with infection of marine bacteria by bacteriophage. *Limnol. Oceanogr.* 47, 1554–1561. doi: 10.4319/lo.2002.47.5.1554

Barrett, P. M., Resing, J. A., Buck, N. J., Feely, R. A., Bullister, J. L., Buck, C. S., et al. (2014). Calcium carbonate dissolution in the upper 1000 m of the eastern North Atlantic. *Global Biogeochem. Cycles* 28, 386–397. doi: 10.1002/2013GB004619

Bracher, A., Vountas, M., Dinter, T., Burrows, J. P., Rottgers, R., and Peeken, I. (2009). Quantitative observation of cyanobacteria and diatoms from space using PhytoDOAS on SCIAMACHY data. *Biogeosciences* 6, 751–764. doi: 10.5194/bg-6-751-2009

Brewin, R. J. W., Hardman-Mountford, N. J., Lavender, S. J., Raitsos, D. E., Hirata, T., Uitz, J., et al. (2011). An intercomparison of bio-optical techniques for detecting dominant phytoplankton size class from satellite remote sensing. *Remote Sens. Environ.* 115, 325–339. doi: 10.1016/j.rse.2010.09.004

Brewin, R. J. W., Sathyendranath, S., Hirata, T., Lavender, S. J., Barciela, R. M., and Hardman-Mountford, N. J. (2010). A three-component model of phytoplankton size class for the Atlantic Ocean. *Ecol. Modell.* 221, 1472–1483. doi: 10.1016/j.ecolmodel.2010.02.014

Bricaud, A., Babin, M., Claustre, H., Ras, J., and Tieche, F. (2010). Light absorption properties and absorption budget of Southeast Pacific waters. *J. Geophys. Res.* 115:C08009. doi: 10.1029/2009JC005517

Bricaud, A., Bedhomme, A.-L., and Morel, A. (1988). Optical properties of diverse phytoplanktonic species: experimental results and theoretical interpretation. *J. Plankton Res.* 10, 851–873.

Bricaud, A., and Morel, A. (1986). Light attenuation and scattering by phytoplanktonic cells: a theoretical modeling. *Appl. Opt.* 25, 571–580.

Bricaud, A., Morel, A., and Prieur, L. (1981). Absorption by dissolved organic matter of the sea (yellow substance) in the UV and visible domains. *Limnol. Oceanogr.* 26, 43–53.

Buitenhuis, E. T., van der Wal, P., and de Baar,. H. J. W. (2001). Blooms of Emiliana huxleyi are sinks of atmospheric carbon dioxide: a field and mesocosm study derived simulation. *Global Biogeochem. Cycles* 15, 577–587. doi: 10.1029/2000GB001292

Carder, K. L., Steward, R. G., Harvey, G. R., and Ortner, P. B. (1989). Marine humic and fulvic acids: their effects on remote sensing of ocean chlorophyll. *Limnol. Oceanogr.* 34, 68–81.

Ciavatta, S., Torres, R., Martinez-Vicente, V., Smyth, T., Dall'Olmo, G., Polimene, L., et al. (2014). Assimilation of remotely-sensed optical properties to improve marine biogeochemistry modeling. *Prog. Oceanogr.* 127, 74–95. doi: 10.1016/j.pocean.2014.06.002

Circio, J. A., and Petty, C. C. (1951). The near infrared absorption spectrum of liquid water. *J. Opt. Soc. Am.* 41, 302–308.

Dutkiewicz, S., Hickman, A. E., Jahn, O., Gregg, W. W., Mouw, C. B., and Follows, M. J. (2015). Capturing optically important constituents and properties in a marine biogeochemical and ecosystem model. *Biogeosciences* 12, 4447–4481. doi: 10.5194/bg-12-4447-2015

Foerster, S., Carrère, V., Rast, M., and Staenz, K. (2015). Preface: the Environmental Mapping and Analysis Program (EnMAP) Mission: preparing for Its scientific exploitation remote. *Sensing* 8:957. doi: 10.3390/rs8110957

Fontana, C., Brasseur, P., and Brankart, J.-M. (2013). Toward a multivariate reanalysis of the North Atlantic Ocean biogeochemistry during 1998–2006 based on the assimilation of SeaWiFS chlorophyll data. *Ocean Sci.* 9, 37–56. doi: 10.5194/os-9-37-2013

Gallegos, C. L., Werdell, P. J., and McClain, C. (2011). Long-term changes in light scattering in Chesapeake Bay inferred from Secchi depth, light attenuation, and remote sensing measurements. *J. Geophys. Res.* 116:C00H08, doi: 10.1029/2011JC007160

Gangsto, R., Joos, F., and Gehlen, M. (2011). Sensitivity of pelagic calcification to ocean acidification. *Biogeosciences* 8, 433–458. doi: 10.5194/bg-8-433-2011

Garver, S. A., and Siegel, D. A. (1997). Inherent optical property inversion of ocean color spectra and its biogeochemical interpretation. I. Time series from the Sargasso Sea. *J. Geophys. Res.* 102, 18607–18625.

Gordon, H. R. (1997). Atmospheric correction of ocean color imagery in the Earth Observing System era. *J. Geophys. Res.* 102, 17081–17106.

Gordon, H. R., Brown, O. B., Evans, R. H., Brown, J. W., Smith, R. C., Baker, K. S., et al. (1988). A semianalytic radiance model of ocean color. *J. Geophys. Res.* 93, 10909–11092.

Gordon, H. R., Smyth, T. J., Balch, W. M., Boynton, G. C., and Tarran, G. A. (2009). Light scattering by coccoliths detached from *Emiliania huxleyi*. *Appl. Opt.* 48, 6059–6073. doi: 10.1364/AO.48.006059

Gregg, W. W. (2002). "A coupled ocean-atmosphere radiative model for global ocean biogeochemical models," in *NASA Global Modeling and Assimilation Series*, Vol. 22, ed M. Suarez (Greenbelt, MD: NASA Technical Memorandum 2002-104606), 33.

Gregg, W. W. (2008). Assimilation of SeaWiFS ocean chlorophyll data into a three-dimensional global ocean model. *J. Mar. Syst.* 69, 205–225. doi: 10.1016/j.jmarsys.2006.02.015

Gregg, W. W., and Carder, K. L. (1990). A simple spectral solar irradiance model for cloudless maritime atmospheres. *Limnol. Oceanogr.* 35, 1657–1675.

Gregg, W. W., and Casey, N. W. (2007). Modeling coccolithophores in the global oceans. *Deep Sea Res. II Top. Stud. Oceanogr.* 54, 447–477. doi: 10.1016/j.dsr2.2006.12.007

Gregg, W. W., and Casey, N. W. (2009). Skill assessment of a spectral ocean-atmosphere radiative model. *J. Mar. Syst.* 76, 49–63. doi: 10.1016/j.jmarsys.2008.05.007

Gregg, W. W., and Rousseaux, C. S. (2014). Decadal trends in global pelagic ocean chlorophyll: a new assessment integrating multiple satellites, *in situ* data, and models. *J. Geophys. Res. Oceans* 119, 5921–5933. doi: 10.1002/2014JC010158

Gregg, W. W., and Rousseaux, C. S. (2016). Directional and spectral irradiance in ocean models: effects on simulated global phytoplankton, nutrients, and primary production. *Front. Mar. Sci.* 3:240. doi: 10.3389/fmars.2016.00240

Gregg, W. W., Friedrichs, M. A. M., Robinson, A. R., Rose, K. A., Schlitzer, R., Thompson, K. R., et al. (2009). Skill assessment in ocean biological data assimilation. *J. Mar. Syst.* 76, 16–33. doi: 10.1016/j.jmarsys.2008.05.006

Gregg, W. W., Ginoux, P., Schopf, P. S., and Casey, N. W. (2003). Phytoplankton and iron: validation of a global three-dimensional ocean biogeochemical model. *Deep Sea Res. II* 50, 3143–3169. doi: 10.1016/j.dsr2.2003.07.013

Gregg, W. W., Patt, F. S., and Woodward, R. H. (1997). Development of a simulated data set for the SeaWiFS mission. *IEEE Trans. Geosci. Remote Sens.* 35, 421–435.

Hirata, T., Aiken, J., Hardman-Mountford, N., Smyth, T. J., and Barlow, R. G. (2008). An absorption model to determine phytoplankton size classes from satellite ocean colour. *Remote Sens. Environ.* 112, 3153–3159. doi: 10.1016/j.rse.2008.03.011

Hirata, T., Hardman-Mountford, N. J., Brewin, R. J. W., Aiken, J., Barlow, R., Suzuki, K., et al. (2011). Synoptic relationships between surface Chlorophyll-a and diagnostic pigments specific to phytoplankton functional types. *Biogeosciences* 8, 311–327. doi: 10.5194/bg-8-311-2011

Hu, J., Fennel, K., Mattern, J. P., and Wilkin, J. (2012). Data assimilation with a local Ensemble Kalman Filter applied to a three-dimensional biological model of the Middle Atlantic Bight. *J. Mar. Syst.* 94, 145–156. doi: 10.1016/j.jmarsys.2011.11.016

Kamykowski, D., Zentara, S. J., Morrison, J. M., and Switzer, A. C. (2002). Dynamic global patterns of nitrate, phosphate, silicate, and iron availability and phytoplankton community composition from remote sensing data. *Global Biogeochem. Cycles* 16, 1077. doi: 10.1029/2001GB001640

Kostadinov, T. S., Siegel, D. A., and Maritorena, S. (2010). Global variability of phytoplankton functional types from space: assessment via the particle size distribution. *Biogeosci. Discus.* 7, 4295–4340. doi: 10.5194/bgd-7-4295-2010

Lee, Z. P., Carder, K. L., and Arnone, R. (2002). Deriving inherent optical properties from water color: a multi-band quasianalytical algorithm for optically deep waters. *Appl. Opt.* 41, 5755–5772. doi: 10.1364/AO.41.005755

Lee, Z., Wei, J., Voss, K., Lewis, M., Bricaud, A., and Huot, Y. (2015). Hyperspectral absorption coefficient of "pure" seawater in the range of 350–550 nm inverted from remote sensing reflectance. *Appl. Opt.* 54, 546–558. doi: 10.1364/AO.54.000546

Loisel, H., Nicolas, J. M., Sciandra, A., Stramski, D., and Poteau, A. (2006). Spectral dependency of optical backscattering by marine particles from satellite remote sensing of the global ocean. *J. Geophys. Res.* 111:C09024. doi: 10.1029/2005JC003367

Maritorena, S., Hembise, O., Fanton d'Andon, A., Mangin, A., and Siegel, D. A. (2010). Merged satellite ocean color data products using a bio-optical Model: Characteristics, benefits and issues. *Remote Sens. Environ.* 114, 1791–1804. doi: 10.1016/j.rse.2010.04.002

Maritorena, S., and Siegel, D. A. (2005). Consistent merging of satellite ocean color data using a semi-analytical model. *Remote Sens. Environ.* 94, 429–440. doi: 10.1016/j.rse.2004.08.014

Mason, J. D., Cone, M. T., and Fry, E. S. (2016). Ultraviolet (250–550 nm) absorption spectrum of pure water. *Appl. Opt.* 55, 7163–7172. doi: 10.1364/AO.55.007163

Masotti, I., Moulin, C., Alvain, S., Bopp, L., Tagliabue, A., and Antoine, D. (2010). Large-scale shifts in phytoplankton groups in the Equatorial Pacific during ENSO cycles. *Biogeosci. Discus.* 7, 2523–2548. doi: 10.5194/bgd-7-2523-2010

Maul, G. A. (1985). *Introduction to Satellite Oceanography*. Boston, MA: Martinus Nijhoff, 606.

Meini, M., Fossati, E., Giunti, L., Molina, M., Formaro, R., Longo, F., et al. (2015). "The PRISMA mission hyperspectral payload," in *IAC-15-B1.3.7, 66th International Astronautical Congress*, (Jerusalem).

Mélin, F., Sclep, G., Jackson, T., and Sathyendranath, S. (2016). Uncertainty estimates of remote sensing reflectance derived from comparison of ocean color satellite data sets. *Remote Sens. Environ.* 177, 107–124. doi: 10.1016/j.rse.2016.02.014

Moisan, T. A., and Mitchell, B. G. (2001). UV absorption by mycosporine-like amino acids in *Phaeocystis antarctica* Karsten induced by photosynthetically available radiation. *Mar. Biol.* 138, 217–227. doi: 10.1007/s002270000424

Morel, A. (1987). Chlorophyll-specific scattering coefficient of phytoplankton. A simplified theoretical approach. *Deep Sea Res.* 34, 1093–1105.

Morel, A. (1988). Optical modeling of the upper ocean in relation to its biogenous mattercontent (Case I waters). *J. Geophys. Res.* 93, 10749–10768.

Morel, A., and Bricaud, A. (1981). Theoretical results concerning light absorption in a discrete medium, and application to specific absorption of phytoplankton. *Deep Sea Res.* 28, 1375–1393.

Morel, A., Gentili, B., Claustre, H., Babin, M., Bricaud, A., Ras, J., et al. (2007). Optical properties of the "clearest" natural waters. *Limnol. Oceanogr.* 52, 217–229. doi: 10.4319/lo.2007.52.1.0217

Neukermans, G., Reynolds, R. A., and Stramski, D. (2016). Optical classification and characterization of marine particle assemblages within the western Arctic Ocean. *Limnol. Oceanogr.* 61, 1472–1494. doi: 10.1002/lno.10316

Organelli, E., Bricaud, A., Antoine, D., and Matsuoka, A. (2014). Seasonal dynamics of light absorption by chromophoric dissolved organic matter (CDOM) in the NW Mediterranean Sea (BOUSSOLE Site). *Deep Sea Res. I Oceanogr. Res. Pap.* 191, 72–85. doi: 10.1016/j.dsr.2014.05.003

PACE Mission Science Definition Team Report (2012). Available online at: http://decadal.gsfc.nasa.gov/pace-resources.html.

Palacios, S. L., Kudela, R. M., Guild, L. S., Negrey, K. H., Torres-Perez, J., and Broughton, J. (2015). Remote sensing of phytoplankton functional types in the coastal oceanfrom the HyspIRI Preparatory Flight Campaign. *Remote Sens. Environ.* 167, 269–280. doi: 10.1016/j.rse.2015.05.014

Pope, R. M., and Fry, E. S. (1997). Absorption spectrum (380-700 nm) of pure water. II. Integrating cavity measurements. *Appl. Opt.* 46, 8710–8723.

Raitsos, D. E., Lavender, S. J., Maravelias, C. D., Haralabous, J., Richardson, A. J., and Reid, P. C. (2008). Identifying four phytoplankton functional types from space: an ecological approach. *Limnol. Oceanogr.* 53, 605–613. doi: 10.4319/lo.2008.53.2.0605

Reader, H. E., and Miller, W. L. (2012). Variability of carbon monoxide and carbon dioxide apparent quantum yield spectra in three coastal estuaries of the South Atlantic Bight. *Biogeosciences* 9, 4279–4294. doi: 10.5194/bg-9-4279-2012

Reader, H. E., and Miller, W. L. (2014). The efficiency and spectral photon dose dependence of photochemically induced changes to the bioavailability of dissolved organic carbon. *Limnol. Oceanogr.* 59, 182–194. doi: 10.4319/lo.2014.59.1.0182

Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., and Liu, E. (2011). MERRA-NASA's modern-era retrospective analysis for research and applications. *J. Clim.* 24, 3624–3648. doi: 10.1175/JCLI-D-11-00015.1

Roesler, C. S., Perry, M. J., and Carder, K. L. (1989). Modeling in situ phytoplankton absorption from total absorption spectra in productive inland marine waters. *Limnol. Oceanogr.* 34, 1510–1523.

Rousseaux, C. S., Hirata, T., and Gregg, W. W. (2013). Satellite views of global phytoplankton community distributions using an empirical algorithm and a numerical model. *Biogeosci. Discus.* 10, 1083–1109. doi: 10.5194/bgd-10-1083-2013

Sadeghi, A. T., Dinter, T., Vountas, M., Taylor, B. B., Altenburg-Soppa, M., IPeeken, I., et al. (2012). Improvement to the PhytoDOAS method for identification of coccolithophores using hyper-spectral satellite data. *Ocean Sci.* 8, 1055–1070. doi: 10.5194/os-8-1055-2012

Sathyendranath, S., Lazzara, L., and Prieur, L. (1987). Variations in the spectral values of specific absorption of phytoplankton. *Limnol. Oceanogr.* 32, 403–415.

Sathyendranath, S., Watts, L., Devred, E., Platt, T., Caverhill, C., and Maass, H. (2004). Discrimination of diatoms from other phytoplankton using ocean-colour data. *Mar. Ecol. Prog. Ser.* 272, 59–68. doi: 10.3354/meps272059

Schopf, P. S., and Loughe, A. (1995). A reduced gravity isopycnal ocean model: hindcasts of El Nino. *Mon. Weather Rev.* 123, 2839–2863.

Siegel, D. A., Maritorena, S., Nelson, N. B., Hansell, D. A., and Lorenzi-Kayser, M. (2002). Global distribution and dynamics of colored dissolved and detrital organic materials. *J. Geophys. Res.* 107:3228. doi: 10.1029/2001JC000965

Smith, R. C., and Baker, K. S., (1981). Optical properties of the clearest natural waters (200-800 nm). *Appl. Opt.* 20, 177–184.

Stramski, D., Boss, E., Bogucki, D., and Voss, K. J. (2004). The role of seawater constituents in light backscattering in the ocean. *Prog. Oceanogr.* 61, 27–56. doi: 10.1016/j.pocean.2004.07.001

Thuillier, G., Floyd, L., Woods, T. N., Cebula, R., Hilsenrath, E., Hers'e, M., et al. (2004). Solar irradiance reference spectra for two solar active levels. *Adv. Space Res.* 3, 256–261. doi: 10.1016/j.asr.2002.12.004

Twardowski, M. S., Claustre, H., Freeman, S. A., Stramski, D., and Huot, Y. (2007). Optical backscattering properties of the "clearest" natural waters. *Biogeosciences* 4, 1041–1058. doi: 10.5194/bg-4-1041-2007

Tzortziou, M., Osburn, C. L., and Neale, P. J. (2007). Photobleaching of dissolved organic material from a tidal marsh-estuarine system of the chesapeake bay. *Photochem. Photobiol.* 83, 782–792. doi: 10.1111/j.1751-1097.2007.00142.x

Vaillancourt, R. D., Brown, C. W., Guillard, R. R. L., and Balch (2004). Light backscattering properties of marine phytoplankton: relationships to cell size, chemical composition, and taxonomy. *J. Plankton Res.* 26, 191–212. doi: 10.1093/plankt/fbh012

Wang, M. H., Shi, W., Jiang, L. D., and Voss, K. (2016). NIR- and SWIR-based on-orbit vicarious calibrations for satellite ocean color sensors. *Opt. Express* 24, 20437–20453. doi: 10.1364/oe.24.020437

Whitmire, A. L., Pegau, W. S., Karp-Boss, L., Boss, E., and Cowles, T. J. (2010). Spectral backscattering properties of marine phytoplankton cultures. *Opt. Express* 18, 15073–15093. doi: 10.1364/oe.18.015073

Wozniak and Stramski, D. (2004). Modeling the optical properties of mineral particles suspended in seawater and their influence on ocean reflectance and chlorophyll estimation from remote sensing algorithms. *Appl. Opt.* 43, 3489–3503. doi: 10.1364/AO.43.003489

Xiu, P., and Chai, F. (2014). Connections between physical, optical and biogeochemical processes in the Pacific Ocean. *Prog. Oceanogr.* 122, 30–53. doi: 10.1016/j.pocean.2013.11.008

Yacobi, Y. Z., Alberts, J. J., Takács, M., and Michelle McElvaine, M. (2003). Absorption spectroscopy of chromophoric dissolved organic carbon in Georgia (USA) rivers: the impact of molecular size distribution. *J. Limnol.* 62, 41–46. doi: 10.4081/jlimnol.2003.41

Zhang, X., Hu, L., and He, M.-X. (2009). Scattering by pure seawater: Effect of salinity. *Opt. Express* 17, 5698–5710. doi: 10.1364/OE17.005698

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Gregg and Rousseaux. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Marine Science

ORIGINAL RESEARCH
published: 04 April 2017
doi: 10.3389/fmars.2017.00090

Light Absorption by Phytoplankton in the Upper Mixed Layer of the Black Sea: Seasonality and Parametrization

Tanya Churilova[1], Vyacheslav Suslin[2], Olga Krivenko[1], Tatiana Efimova[1], Nataliia Moiseeva[1], Vladimir Mukhanov[1] and Liliya Smirnova[1]*

[1] A.O. Kovalevsky Institute of Marine Biological Research, Sevastopol, Russia, [2] Department of Ocean Dynamics, Marine Hydrophysical Institute, Sevastopol, Russia

Standard NASA ocean color algorithm OC4 was developed on the basis of ocean optical data and while appropriate for Case 1 oceanic waters could not be adequately applied for the Black Sea waters due to its different bio-optical properties. OC4 algorithm is shown to overestimate chlorophyll concentration (Chl-*a*) in summer and underestimate Chl-*a* during early spring phytoplankton blooms in the Black Sea. For correct conversion of satellite data to Chl-*a*, primary production and other indicators regional algorithms should be developed taking into account bio-optical properties of the Black Sea waters. Light absorption by phytoplankton pigments– $a_{ph}(\lambda)$ have been measured in open sea and shelf Black Sea waters in different seasons since 1998. It was shown that the first optical depth was located within the upper mixed layer (UML) for most of the year with the exception of the spring when seasonal stratification was developing. As a result spectral features of water leaving radiance were determined by optical properties of the UML. Significant seasonal differences in Chl-*a* specific light absorption coefficients of phytoplankton within UML have been revealed. These differences were caused by adaptive changes of composition and intracellular pigment concentration due to variable environment conditions–mainly light intensity. Empirical relationships between $a_{ph}(\lambda)$ and Chl-*a* were derived by least squares fitting to power functions for different seasons. Incorporation of these results will refine the regional ocean color models and provide improved and seasonally adjusted estimates of chlorophyll a concentration, downwelling radiance and primary production in the Black Sea based on satellite data.

Keywords: phytoplankton, light absorption, parameterization, chlorophyll *a* concentration, upper mixed layer, the Black Sea

OPEN ACCESS

Edited by:
Shubha Sathyendranth, Plymouth Marine Laboratory, UK

Reviewed by:
Bob Brewin, Plymouth Marine Laboratory, UK
Toru Hirawake, Hokkaido University, Japan

****Correspondence:***
Tanya Churilova
tanya.churilova@gmail.com

Specialty section:
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

Received: *29 December 2016*
Accepted: *15 March 2017*
Published: *04 April 2017*

Citation:
Churilova T, Suslin V, Krivenko O, Efimova T, Moiseeva N, Mukhanov V and Smirnova L (2017) Light Absorption by Phytoplankton in the Upper Mixed Layer of the Black Sea: Seasonality and Parametrization. Front. Mar. Sci. 4:90. doi: 10.3389/fmars.2017.00090

INTRODUCTION

Visible spectral radiometric data are used widely to assess water productivity (Saba et al., 2011) and to study effect of climate change on ocean productivity (Behrenfeld et al., 2006). Optical scanners of Sea-viewing Wide Field-of-view Sensor (SeaWiFS), MEdium Resolution Imaging Spectrometer (MERIS), Moderate Resolution Imaging Spectroradiometer aboard the Terra and Aqua satellites (MODIS-Aqua/Terra) measure water leaving radiance at several spectral bands (R_{RS}) (Feldman and McClain, 2013). The spectral distribution of R_{RS} is influenced by particulate scattering and absorbance of solar radiance by all in-water optically active components: phytoplankton, non-algal particles (NAP), colored dissolved organic matter (CDOM) and pure water

(Kirk, 1994). Light absorption by particles ($a_p(\lambda)$), phytoplankton ($a_{ph}(\lambda)$), NAP ($a_{NAP}(\lambda)$) and CDOM ($a_{CDOM}(\lambda)$) have been studied in different regions of the global ocean since the 80-s (Hoepffner and Sathyendranath, 1992; Bricaud et al., 1995, 1998; Cleveland, 1995; Babin et al., 2003) to develop algorithms for assessment of water productivity based on remote sensing. Inherent optical properties (IOPs) vary throughout the world ocean. Due to high variability in light absorption and scattering by optically active components, the world ocean needs to be subdivided into various provinces based on regional IOPs, and their features could be used to improve remote-sensing algorithms for each province (Hoepffner and Sathyendranath, 1992; Lutz et al., 1996; Suzuki et al., 1998). Originally the standard NASA algorithm could be applied if there was a high correlation between $a_{ph}(\lambda)$ and absorption by colored dissolved and suspended organic matter ($a_{CDM}(\lambda)$) (Morel and Prieur, 1977). Although NASA standard algorithms are continually being updated (O'Reilly et al., 2000), the latest versions (OC4 for SeaWiFS, and OC3M for MODIS-Aqua /Terra) do not provide an adequate assessment of chlorophyll *a* concentration (Chl-*a*) in the Black Sea waters (Suslin and Churilova, 2016) which belong to the Case 2 (Suslin et al., 2007). Berthon et al. (2008) underlined that important uncertainties for the retrieval of marine products like Chl-*a* still persisted in areas (including the Black Sea) where relatively high CDOM absorption and optically active water constituents CDOM and NAP do not co-vary in a predictable manner with Chl-*a*.

For correct conversion of optical scanner signals into water productivity indices regional algorithms need to be developed taking into account bio-optical properties of the Black Sea. The assessment of the Chl-*a* needs to derive $a_{ph}(\lambda)$ from total light absorption by all optically active components and then estimate Chl-*a* based on relationship between $a_{ph}(\lambda)$ and Chl-*a*. This relationship is also required for development of regional algorithms of downwelling radiance and primary production by spectral approach. Early versions of the regional bio-optical algorithms (Suslin et al., 2008; Churilova et al., 2009; Churilova and Suslin, 2010) were based on limited amount of bio-optical data available. The bio-optical properties of the Black Sea (namely $a_{ph}(\lambda)$, $a_{NAP}(\lambda)$ and $a_{CDOM}(\lambda)$) have been studied since 1995 in open and coastal waters of the Black Sea (Churilova, 2001; Churilova and Berseneva, 2004; Churilova et al., 2004; Chami et al., 2005; Berthon et al., 2008; Dmitriev et al., 2009). Variability in $a_{ph}(\lambda)$ spectral distributions and coefficient values in coastal (Churilova and Berseneva, 2004; Chami et al., 2005; Dmitriev et al., 2009) and open waters (Churilova et al., 2004; Berthon et al., 2008) have been demonstrated but seasonal variability in Chl-*a* specific phytoplankton light absorption coefficients remains not known in details. The bio-optical data measured in the deep waters of the Black Sea from 2011 to 2015 will be examined in this study.

The aim of the current research is to analyze seasonal variability of relationship between phytoplankton light absorption coefficients and chlorophyll *a* concentrations in upper mixed layer (UML) of the Black Sea and derive season-specific modeling parameters.

MATERIALS AND METHODS

Sampling

Bio-optical measurements were carried out during 7 cruises of RV "Professor Vodyanitsky" in different seasons during 2011–2015 in the deep-water areas (deeper 100 m isobath) of Black Sea (**Table 1**, **Figure 1**). Water samples were collected at 5–7 depths within euphotic zone with 10 liter Niskin bottles of CTD/rosette system MARK-III (Neil Brown Ocean Sensors, Inc) or SBE-911plus (Sea Bird Electronics). Sampling depths were chosen based on water transparency by Secchi disc depth (Zs), as well as temperature and salinity profiles measured by CTD system. The euphotic zone (Z_{eu}), determined as penetration depth for 1% of photosynthetically available radiance (PAR), was calculated based on the light (I) attenuation with depth (z) (Kirk, 1994):

$$I(z) = I(0) \times e^{(-K_d \times z)}, \quad (1)$$

where K_d–light attenuation coefficient on average for both euphotic layer and for visible light waves (400–700 nm). Z_{eu} was calculated based on the Equation (1):

$$Z_{eu} = \frac{4.6}{K_d} \quad (2)$$

Values of K_d were estimated based on the relationship between K_d and Z_s obtained for the Black Sea (Vedernikov, 1989). Average light intensity in the UML (PAR_{UML}) was estimated in accordance with (Babin et al., 1996):

$$PAR_{UML} = PAR(0) \times \frac{\left(1 - e^{\left(-4.6 \times \frac{Z_{uml}}{Z_{eu}}\right)}\right)}{\left(4.6 \times \frac{Z_{uml}}{Z_{eu}}\right)}, \quad (3)$$

where PAR(0)–PAR at the sea surface, data from Suslin et al. (2015), Z_{uml}–UML depth was determined using temperature difference criterion (0.5°C) and mean water temperature at 0–3 m as reference level. Optical depth (ζ) of Z_{uml} was assessed using K_d calculated based on Z_s (Vedernikov, 1989).

Pigment Analysis

For chlorophyll and phaeopigment concentration analysis 1–2 L water samples were gently vacuum filtered (<25 kPa)

TABLE 1 | Information about scientific cruises of RV Professor Vodyanitsky (PV) in the Black Sea.

Cruise	Year	Date	Investigation area of the Black Sea
PV 69	2011	2–11 August	Deep western
PV 70	2011	19–27 August	Deep western
PV 77	2014	3–7 September	Deep eastern
PV 78	2014	28 November–9 December	Deep eastern
PV 79	2015	25–30 September	Deep eastern
PV 81	2015	3–10 November	Deep western
PV 82	2015	5–9 December	Deep eastern

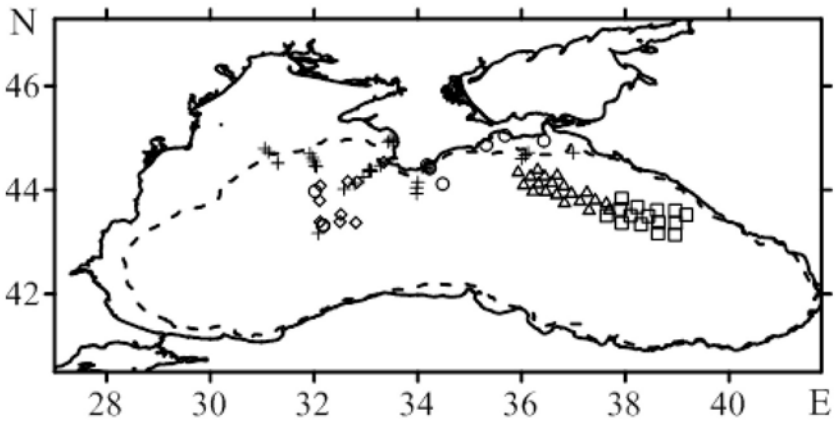


FIGURE 1 | Map of bio-optical stations in the scientific cruises on RV Professor Vodyanitskiy (PV) in the Black Sea: № 69 (crosses), 70 (circles), 77 (square), 78 (square, triangles), 79 (triangles), 81 (rhombs), 82 (triangles). The dotted line defines the 100 m isobath.

onto 25 mm diameter Whatman GF/F glass fiber filters. Filters were wrapped in an aluminum foil and stored in a liquid nitrogen until analysis on a laboratory. Filters were placed in 5 ml of 90% acetone in a 10 ml glass centrifuge tube, then were treated with vibration for 20 s using a vibration mixer (FALK Falc instruments, Italy), extracted at 5°C or below, for at least 10 h and then centrifuged. The above procedure was repeated using an additional 5 ml of 90% acetone for a more complete extraction of phytoplankton pigments. The second extraction of the pigments contributed 15% on average to total concentration values. The extracts were then analyzed for pigment content by spectrophotometric method (Lorenzen, 1967; Jeffrey and Humphrey, 1975) using spectrophotometer Lambda 35 (Perkin Elmer). Proportion of non-photosynthetic pigments in the total phytoplankton pigment content (NPP) was determined in accordance with relationship between environmental light condition (PAR) and NPP proposed by Babin et al. (1996).

Phytoplankton Light Absorption

Optical densities of particulate matter were determined by the filter pad technique ("wet filter technique") (Yentsch, 1962; Mitchell and Kiefer, 1988). $a_{ph}(\lambda)$ was determined by the difference between $a_p(\lambda)$ and $a_{NAP}(\lambda)$:

$$a_{ph}(\lambda) = a_p(\lambda) - a_{NAP}(\lambda) \quad (4)$$

Values of $a_{ph}(\lambda)$ were obtained from optical densities after correction for differential scattering (setting the mean absorption between 740 and 750 nm to zero) and for the path length amplification factor, converting decimal to natural logarithms, taking into account the volume filtered and the filter area of filtration, and subtracting $a_{NAP}(\lambda)$ (Churilova, 2001). The sample optical densities were measured from 350 to 750 nm on Perkin Elmer Lambda 35 spectrophotometer equipped with an integrating sphere. $a_{NAP}(\lambda)$ values were experimentally determined using the chemical (bleaching by NaClO solution) method (Tassan and Ferrari, 1995). The path length amplification factor (beta-correction) was estimated applying the quadratic equation described by Mitchell (1990). To get Chl-*a* specific light absorption coefficients of phytoplankton ($a^*_{ph}(\lambda)$) the values of $a_{ph}(\lambda)$ (m^{-1}) were divided by the sum of chlorophyll *a* and phaeopigments concentrations (Chl-*a*) (mg m^{-3}). Relationships between $a_{ph}(\lambda)$ and Chl-*a* were derived by least squares fitting to power functions for visible spectral domain 400–700 nm with 1 nm resolution.

Phytoplankton

Identification of phytoplankton species (micro- and nano-size fractions), counting of cells and cell size measurements were performed with transmission microscope Ergaval (Carl Zeiss Jena) using Naumann chamber. Water samples (2–5 L) were concentrated by inverse filtration method using nuclepore filters with 1 μm pore diameter. The concentrated samples were fixed with a solution (4% final concentration) of 25 g paraformaldehyde dissolved in 100 ml of hot (80°C) 25% glutaraldehyde, clarified with few drops of 1 N NaOH solution. Cells were sized and its volumes were assessed using geometrical figures (sphere, ellipsoid or cylinder) corresponding to the cell shapes. Phytoplankton analysis was conducted only at selected stations in August 2011, September 2014 and December 2014, 2015.

In August 2011, September 2014 and December 2014, 2015 flow cytometric analysis was performed by flow cytometer Cytomics FC 500 (Beckman Coulter, USA) equipped with a single-phase argon laser (488 nm) (Marie et al., 1999; Schapira et al., 2010). For all detected particles phycoerythrin fluorescence emission (575 nm), and chlorophyll fluorescence emission (675 nm) were measured. The samples were fixed with formaldehyde (final concentration 2%) immediately after the sampling, then the samples were frozen in liquid nitrogen (−80°C) and stored at −20°C until analysis in the laboratory. The cytometer measurements were calibrated by the addition of a known concentration of the Fluorospheres Flow-CheckTM (Beckman

Coulter). Cytometric data were analyzed using CXP software (Beckman Coulter).

RESULTS

Chlorophyll *a* Concentration

Chl-*a* in surface layer of the deep water regions of the Black Sea were low in a summer. In the deep western part of the sea in August 2011 (**Figure 1**) Chl-*a* in the UML were in a range 0.15–0.30 mg·m^{-3}. At this time, seasonal thermocline was well developed with maximum of temperature gradient 3.3 ± 1.1°C m^{-1} at the 12 ± 2.3 m depth, where optical depth (ζ) was 1.5 ± 0.42 (**Table 2**). In August water transparency was high. Values of Z_s and K_d were 16 ± 2.1 m and 0.12 ± 0.013 m^{-1}, correspondingly. Z_{eu} values were 37 ± 4.0 m. Vertical Chl-*a* profiles were characterized by rather homogeneous Chl-*a* distribution within UML and deep Chl-*a* maximum (DCM) located near the bottom of the euphotic zone with Chl-*a* values 5–10 times higher than in the UML (**Figure 2**). In the surface layer of the deep eastern part of the Black Sea in September 2014 and 2015 (**Figure 1**) Chl-*a* values (0.21–0.35 mg·m^{-3}) were very similar to these measured in summer. In September maximum temperature gradient (4.3 ± 1.2°C m^{-1}) and its location (9.5 ± 2.7 m with ζ = 1.1 ± 0.40) were similar to those observed in the summer. Vertical Chl-*a* distribution was similar to that observed in August (**Figure 2**), but with less variability in Chl-*a*: Chl-*a* concentration in the DCM was 3 times higher than in the UML in comparison with 5–10 times differences in August. Water transparency in the August and September was comparable (Z_s = 16 ± 1.4 m; K_d = 0.12 ± 0.0073 m^{-1}; Z_{eu} = 38 ± 2.6 m).

In the western deep part of the Black Sea in the late autumn (November 2015) the seasonal thermocline was substantially destructed. It resulted in an enlargement of the UML (28 ± 3.4 m), which become ∼3 times deeper than in the summer (**Figure 2**). Chl-*a* in the surface layer in November 2015 varied from 0.54 to 1.4 mg m^{-3} with less transparency (Z_s = 13 ± 1.0 m; K_d = 0.15 ± 0.009 m^{-1}) (**Table 2**). Maximum temperature gradient (1.5 ± 0.38°C m^{-1}) was located at the optical depth of 4.2 ± 0.72. Consequently, Z_{eu} (31 ± 2.0 m) was close to UML depth. Chl-*a* was distributed homogeneously within UML and decreased sharply in thermocline (**Figure 2**). Thus, vertical pigment distribution in November contrasted with that in summer, when seasonal thermocline divided euphotic zone into two quasi isolated layers with different environments. In fact, in late autumn phytoplankton was present in UML only. In December 2014 and 2015 in surface layer of eastern deep water part of the Black Sea (**Figure 1**) Chl-*a* varied from 1.0 to 2.0 mg m^{-3} (1.3 ± 0.25 mg m^{-3}). UML was 32 ± 7.0 m. Vertical distribution of Chl-*a* was homogeneous within UML similar to Chl-*a* profiles observed in western waters in November 2015. In December Z_s and K_d were equal to November 2015 data (12 ± 2.5 m and 0.15 ± 0.023 m^{-1}, correspondingly). Maximum temperature gradient was 0.93 ± 0.45°C m^{-1} and located at optical depths of 4.9 ± 1.2 (**Table 2**). Consequently, in December euphotic zone (30 ± 4.9 m) occurred within UML as it was observed in November. In both December and November phytoplankton was present within UML only.

Phytoplankton

In August 2011 in UML of western deep waters of the Black Sea phytoplankton was dominated by dinoflagellates. Wet biomass of phytoplankton was 540 ± 310 mg·m^{-3} on average. Assuming intracellular organic carbon (C) content at 10% of wet biomass, C to Chl-*a* ratio (C/Chl-*a*) was 145 ± 76 mg mg^{-1}. Biomass of photosynthetic picoplankton was 2.7 ± 0.46 mg·m^{-3} on average. The contribution of picoplankton to total phytoplankton biomass was <1%. In September 2014 phytoplankton biomass in UML was assessed at selected stations. Wet biomass was ∼450 mg m^{-3}. The phytoplankton was dominated mainly (50–70%) by dinoflagellates *Gymnodinium spp* (*Gymnodinium fungiforme* and *Gymnodinium paululum*). C/Chl-*a* ratio was ∼110 mg mg^{-1}. Photosynthetic picoplankton biomass was equal 1.7 ± 1.0 mg·m^{-3} on average and its contribution to total phytoplankton biomass was < 1%. In December 2014 and 2015 wet phytoplankton biomass in UML varied from 190 to 430 mg m^{-3}. In 2014 *Proboscia alata* dominated (by biomass) in phytoplankton community. In 2015 phytoplankton was represented mainly by large diatoms *Pseudosolenia calcar-avis* with cell volume 19000–83000 μ m^3. C/Chl-*a* ratio was ∼25–40 mg mg^{-1}. In December 2014 and 2015 biomass of photosynthetic picoplankton was 11.0 ± 4.9 and 13.0 ± 4.4 mg·m^{-3} on average correspondingly, and picoplankton contribution to total phytoplankton biomass was ∼5%.

Phytoplankton Light Absorption

Phytoplankton light absorption spectra measured in UML are presented on **Figure 3**. To examine the relationship between $a_{ph}(\lambda)$ and Chl-*a* in the UML results were grouped into 2 datasets: (1) summer dataset that included results from August 2011, September 2014, 2015; (2) winter dataset with results from November 2014, December 2014 and 2015 (**Table 1**). September 2014, 2015 was considered part to the summer season, due to persistence of strong seasonal stratification with typical "summer" type of vertical distribution of pigments. November 2015 was considered part of winter, because of water column structure similarity to that in December 2014 and 2015. In November depths of UML and euphotic zone were close and all phytoplankton was present within UML as it was in December.

In the $a_{ph}(\lambda)$ spectra two main peaks were observed: in blue (near 440 nm) and red (near 678 nm) spectrum domains (**Figure 3**). The seasonal differences in the phytoplankton light absorption were manifested in both spectral shapes and values of chlorophyll *a* specific coefficients. In the summer $a^*_{ph}(\lambda)$ were relatively high in the blue spectrum domain. Ratio between blue and red peaks (R) was 3.4 (± 0.61) on average in summer, which was significantly higher than in winter (2.2 ± 0.45) (**Figure 3**). In both winter and summer R values decreased if Chl-*a* increased. The variations of $a_{ph}(\lambda)$ as a function of Chl-*a* are shown in **Figure 4** at two wavelengths (∼440 and 678 nm) corresponding to the blue and red peaks of the spectra. To describe the relationship between $a_{ph}(\lambda)$ and Chl-*a* a power function was used (**Figure 4**):

$$a_{ph}(\lambda) = A(\lambda) \times (\text{Chl-}a)^{B(\lambda)}, \quad (5)$$

TABLE 2 | Hydrophysical characteristics: maximum temperature gradient (ΔT) and depth (Z_{tc})/optical depth (ζ) of its location; Secchi disc depth visibility (Z_s); euphotic zone (Z_{eu}); diffuse attenuation coefficient for downwelling irradiance over the Z_{eu} (K_d); photosynthetically availably radiance incident on the Black Sea surface [PAR(0)]and averaged over upper mixed layer [PAR(UML)].

	Z_{tc}, m	ζ	ΔT, °C m^{-1}	PAR (UML), E m^{-2} d^{-1}	Z_s, m	K_d, m^{-1}	Z_{eu}, m	PAR(0), E m^{-2} d^{-1}
AUGUST (2011)								
mean	12	1.5	3.3	27	16	0.12	37	52
SD	2.3	0.42	1.1	8.1	2.1	0.013	4.0	1.2
min	8.0	0.84	2.2	31	12	0.10	30	46
max	13	2.0	5.7	24	20	0.15	46	56
SEPTEMBER (2014, 2015)								
mean	9.5	1.1	4.3	22.5	16	0.12	38	38
SD	2.7	0.40	1.2	2.9	1.4	0.0073	2.6	4.1
min	5	0.55	2.9	21	12	0.11	30	28
max	14	2.1	7.8	31	17	0.15	45	46
NOVEMBER (2015)								
mean	28	4.2	1.5	3.9	13	0.15	31	17
SD	3.4	0.72	0.38	1.2	1.0	0.0090	2.0	2.5
min	24	3.12	1.0	2.0	11	0.13	28	9
max	35	5.6	2.1	8.6	14	0.16	34	22
DECEMBER (2014, 2015)								
mean	32	4.9	0.93	2.4	12	0.15	30	12
SD	7.0	1.2	0.45	0.8	2.5	0.023	4.9	1.7
min	27	2.9	0.24	2.5	7	0.11	20	7.8
max	55	12.7	2.0	1.2	19	0.23	43	15

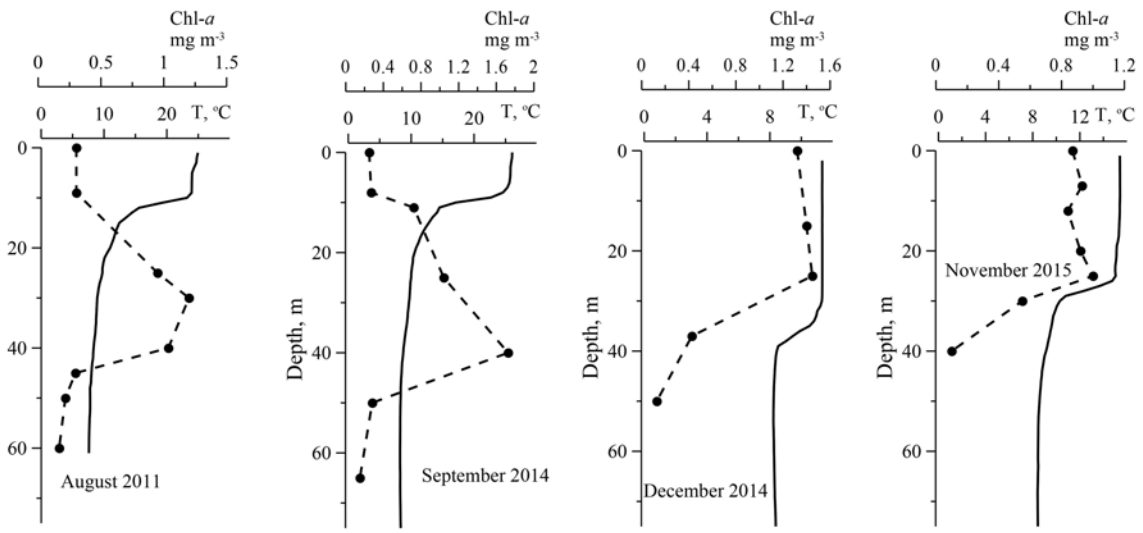


FIGURE 2 | Examples of vertical profiles of temperature (T, line) and the sum of chlorophyll *a* and phaeopigments concentrations (Chl-*a*, circles) in the deep water of the Black Sea in different time.

where A(λ)–spectral coefficient, which is equal to $a^*_{ph}(\lambda)$ in case when Chl-*a* equal to 1 mg m^{-3}.

For two data sets following fit equations were obtained (**Figure 4**):

(1) In summer:

$$a_{ph}(440) = 0.076 \times (\text{Chl-}a)^{0.84} (r^2 = 0.66) \quad (6)$$

$$a_{ph}(678) = 0.024 \times (\text{Chl-}a)^{0.95} (r^2 = 0.63) \quad (7)$$

(2) In winter:

$$a_{ph}(440) = 0.045 \times (\text{Chl-}a)^{0.81} (r^2 = 0.78) \quad (8)$$

$$a_{ph}(678) = 0.021 \times (\text{Chl-}a)^{0.95} (r^2 = 0.88) \quad (9)$$

To infer $a_{ph}(\lambda)$ spectral distribution from Chl-*a* relationship between these parameters needs to be determined for entire visible spectrum (400–700 nm). Based on two empirical data sets the $a_{ph}(\lambda)$ vs Chl-*a* dependencies were parameterized

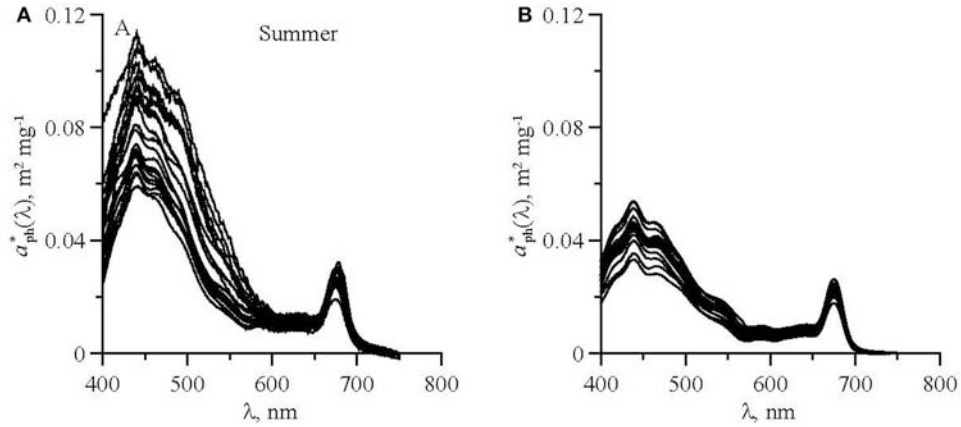


FIGURE 3 | Chl-*a* specific phytoplankton light absorption coefficients $a^*_{ph}(\lambda)$ in upper mixed layer of the Black Sea in summer (A) and winter **(B)**.

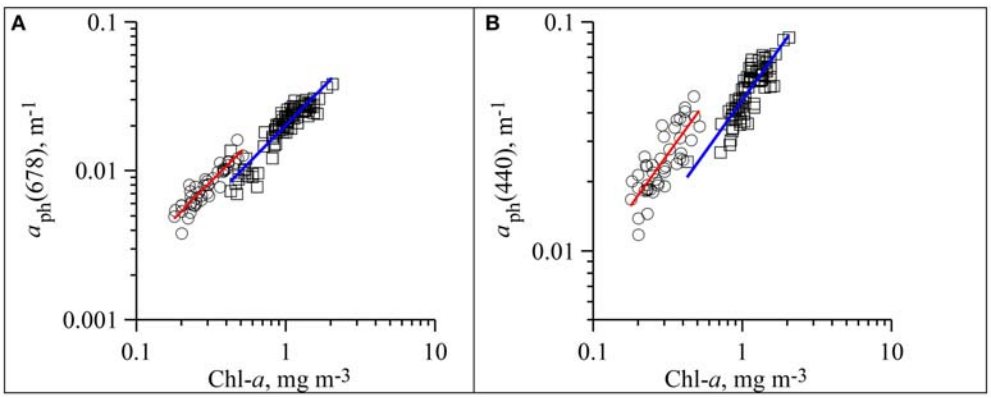


FIGURE 4 | Dependence of phytoplankton light absorption coefficients at red (a_{ph}(678)) (A) and blue (a_{ph}(440)) **(B)** peaks of spectra on the sum of chlorophyll a and phaeopigments concentrations (Chl-*a*) in upper mixed layer of the Black Sea in summer (circles, red line of fitting) and winter (squares, blue line of fitting).

using Equation (5) for summer and winter. The results of the parameterization performed from 400 to 700 nm with 1 nm spectral resolution are presented in **Figure 5** and in **Tables 3, 4**. It is evident that $a^*_{ph}(\lambda)$ values are higher in summer than those in winter (**Figure 3**). This seasonal difference is more pronounced in the blue spectrum domain. For summer phytoplankton the value of A(λ) coefficient at 440 nm is about twice higher than that for winter.

Photosynthetically available radiance incident on the Black Sea surface [PAR(0)] varied seasonally (Suslin et al., 2015). In August and September PAR(0) was on average 52 ± 1.2 and 38 ± 2.6 E m^{-2} d^{-1} correspondingly (**Table 2**). In November and December PAR(0) was 17 ± 2.5 and 12 ± 1.7 E m^{-2} d^{-1}, which were about 3 times lower than those in warm months. PAR in UML depends not only on PAR(0) but also on water transparency and ratio between Z_{UML} and Z_{eu}. In winter, waters were less transparent than in summer. Moreover, UML was comparable with Z_{eu} in winter while in summer Z_{UML} was located between first and second optical depths. As results PAR in UML differed more (∼10 times) between summer and winter in comparison with seasonal dynamics of PAR(0) (**Table 2**). PAR in UML was equal in August and September 27 ± 8.1 and 23 ± 2.9 E m^{-2} d^{-1}, correspondingly (**Table 2**). In November and December PAR in UML was equal 3.9 ± 1.2 and 2.4 ± 0.8 E m^{-2} d^{-1}, correspondingly (**Table 2**).

DISCUSSION

The first optical depth which determines water leaving radiance spectral patterns (Gordon and McGlunev, 1975) detectable by remote scanners is located within the UML in the deep open waters of the Black Sea. It should be noted that K_d averaged over the euphotic zone was used in our assessment. However, domination of CDOM in total light absorption in the Black Sea results in sharply decreasing K_d values with depth. In the subsurface layer K_d values were estimated to be ∼1.6 times higher than mean K_d for euphotic zone (Churilova et al., 2009). Therefore, in our assessment Z_{uml} was

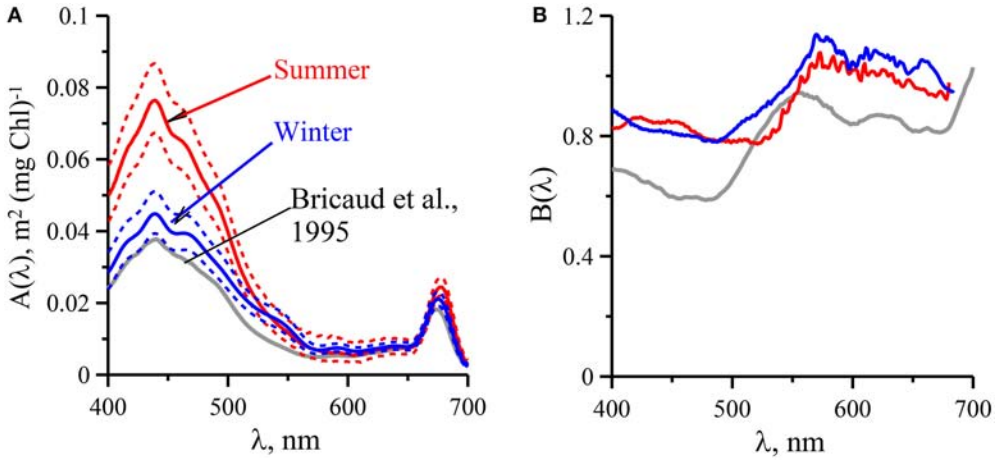


FIGURE 5 | Spectral values of the constants *A*(λ) (A) and *B*(λ) **(B)** obtained when fitting the variations of phytoplankton light absorption ($a_{ph}(\lambda)$) vs. the sum of chlorophyll *a* and phaeopigments concentration (Chl-*a*) to power laws of the form $a_{ph}(\lambda) = A(\lambda)(\text{Chl-}a)^{B(\lambda)}$: for upper mixed layer of the Black Sea in summer (red lines) and winter (blue lines) with comparison with data (gray lines) follow in Bricaud et al. (1995).

underestimated by using average K_d. In fact, it gives more reasonable conclusion about location of the first optical depth within UML. Consequently, bio-optical properties of the UML in the Black Sea determine remotely sensed optical signals and could be used for development and refining of the regional models of productivity indicators.

Analysis of the link between phytoplankton light absorption coefficients and chlorophyll *a* concentration revealed seasonal differences in UML (**Figures 3**, **4** and **Tables 3, 4**) which were related to difference of $a^*_{ph}(\lambda)$ values between summer and winter. The difference was more pronounced in the blue spectrum domain (**Figure 3**). In summer values of parameterization coefficient A(λ) relevant to red and blue peaks were on average 15 and 70% higher than those for winter (**Figure 5**), respectively. Seasonal differences in normalized (on Chl-*a*) phytoplankton light absorbance capacity were related to strong changes in environmental conditions in UML, mainly due to the averaged PAR within UML (**Figure 6**). Observed seasonal dynamics of (UML) and euphotic zone in the deep open waters of the Black Sea are consistent with intraannual changes of these parameters (Z_{uml} and Z_{eu}) in the Black Sea outlined earlier (Ivanov and Belokopytov, 2011; Agirbas et al., 2014).

In winter [PAR(0)] decreased but the ratio between Z_{UML} and Z_{eu} increased in comparison with summer. As the result in winter average light field within UML decreased in almost 10-fold in comparison with PAR_{UML} in summer. Seasonal changes of environmental conditions in UML caused ∼5–7 fold variability in C/Chl-*a* ratio between winter and summer. Observed C/Chl-*a* variability agrees with a change of intracellular concentration of chlorophyll *a* (MacIntyre et al., 2002; Behrenfeld et al., 2005) due to physiological acclimation of algae cultures and phytoplankton to light intensity decreased in the same range as PAR varied in the UML of the Black Sea. Intracellular pigment concentration defines degree of pigment packaging, which in turn effects on $a^*_{ph}(\lambda)$ (Morel and Bricaud, 1981; Bidigare et al., 1990; Hoepffner and Sathyendranath, 1991; Kirk, 1994; Fujiki and Taguchi, 2002). In the current research it was shown that $a^*_{ph}(\lambda)$ and C/Chl-*a* were significantly less in winter than values of those parameters in summer, which were relevant to "pigment packaging" effected on $a^*_{ph}(\lambda)$ at red peak (∼678 nm) where light quanta are absorbed by chlorophyll *a* and phaeopigments only (Jeffrey et al., 1997). At shorter wavelengths (in blue spectrum domain) seasonal variation in $a^*_{ph}(\lambda)$ was more pronounced than at red wavelengths (**Figure 3**). In the blue part of the spectrum besides chlorophyll *a*, other accessory pigments absorb light quanta (Bidigare et al., 1990; Jeffrey et al., 1997) which lead to "smoothing" of spectra due to accessory pigment "packaging" if Chl-*a* specific absorption coefficients are considered. Ratio of accessory pigment-to-Chl-*a* changes due to photoacclimation of algae (MacIntyre et al., 2002; Grant and Louda, 2010), which is related mainly to photoprotective (i.e., non-photosynthetic) pigments (NPP). In review of photoacclimation of different microalgae taxons it was shown that ∼order increase of light intensity resulted in ∼3–4 times increase of photoprotective xanthophyll to Chl-*a* ratio on average (Figure 9 in MacIntyre et al., 2002). Investigations of phytoplankton accessory pigments variability have demonstrated that photoprotective pigments tend to be greater in the surface low Chl-*a* waters at latitudes where radiance incident on the sea surface is relatively high (Stuart et al., 1998; Barlow et al., 2004; Sathyendranath et al., 2005). Variability of R ($a_{ph}(440)$ /$a_{ph}(678)$) was shown to be correlated with NPP to Chl-*a* ratio (Lutz et al., 2003). Altogether the increase of $a^*_{ph}(440)$ is related to low Chl-*a* waters with lower intracellular concentrations of pigments, and a greater proportion of photoprotective pigments occurred in stratified, high light, nutrient-limited regions (Bricaud et al., 1995; Cleveland, 1995; Aguirre-Hernandez et al., 2004). In general these results are in a good agreement with the Black Sea observations. Although in current research pigment composition

TABLE 3 | Spectral values of the constant obtained when fitting the variations of $a_{ph}(\lambda)$ vs. the (chlorophyll *a* + phaeopigment) concentration (Chl-*a*) to power laws of the form.

λ	A(λ)	SD	B(λ)	λ	A	SD	B(λ)
400	0.050	0.00779	0.82	544	0.013	0.00332	0.88
402	0.051	0.00774	0.83	546	0.013	0.0033	0.88
404	0.052	0.00777	0.83	548	0.012	0.00329	0.91
406	0.054	0.00799	0.83	550	0.012	0.00315	0.92
408	0.056	0.00805	0.83	552	0.011	0.00319	0.94
410	0.058	0.00813	0.84	554	0.010	0.00319	0.95
412	0.060	0.00845	0.84	556	0.010	0.00296	0.97
414	0.061	0.0086	0.85	558	0.009	0.00302	1.00
416	0.063	0.0087	0.86	560	0.009	0.00323	0.98
418	0.064	0.00896	0.86	562	0.008	0.00363	0.99
420	0.065	0.00901	0.86	564	0.008	0.00323	1.03
422	0.066	0.00911	0.86	566	0.008	0.00311	1.04
424	0.067	0.00949	0.86	568	0.007	0.00358	1.03
426	0.069	0.00966	0.86	570	0.007	0.00343	1.04
428	0.070	0.00974	0.86	572	0.007	0.00334	1.07
430	0.071	0.00989	0.85	574	0.007	0.00319	1.06
432	0.073	0.00991	0.85	576	0.007	0.00331	1.03
434	0.075	0.00983	0.84	578	0.006	0.00354	1.03
436	0.076	0.00977	0.84	580	0.006	0.00316	1.03
438	0.076	0.0097	0.85	582	0.006	0.00333	1.05
440	0.076	0.00946	0.84	584	0.006	0.00325	1.06
442	0.076	0.00941	0.85	586	0.006	0.00375	1.03
444	0.074	0.00934	0.84	588	0.006	0.00319	1.01
446	0.072	0.00914	0.85	590	0.006	0.00308	1.04
448	0.071	0.00898	0.85	592	0.006	0.00335	1.05
450	0.069	0.00876	0.84	594	0.006	0.00323	1.05
452	0.068	0.00888	0.85	596	0.006	0.00307	1.01
454	0.067	0.00873	0.85	598	0.006	0.0029	1.01
456	0.066	0.00871	0.85	600	0.006	0.00274	1.02
458	0.066	0.00875	0.84	602	0.006	0.00291	1.02
460	0.065	0.00865	0.84	604	0.005	0.00286	1.02
462	0.065	0.00856	0.83	606	0.005	0.00258	0.99
464	0.064	0.00842	0.83	608	0.006	0	1.01
466	0.063	0.00849	0.82	610	0.006	0.00289	1.04
468	0.062	0.00846	0.82	612	0.006	0.00281	1.00
470	0.061	0.00804	0.81	614	0.006	0.00261	1.01
472	0.059	0.00796	0.81	616	0.006	0.00257	1.02
474	0.058	0.00779	0.80	618	0.006	0.00237	1.01
476	0.056	0.00765	0.79	620	0.007	0.00246	1.00
478	0.055	0.00759	0.79	622	0.007	0.00253	1.00
480	0.054	0.00738	0.79	624	0.007	0.00251	1.02
482	0.053	0.00731	0.79	626	0.007	0.00217	1.02
484	0.051	0.00719	0.79	628	0.007	0.00223	1.01
486	0.051	0.00708	0.78	630	0.007	0.00248	1.00
488	0.049	0.00692	0.79	632	0.007	0.00227	1.01
490	0.048	0.00684	0.78	634	0.007	0.00212	0.99
492	0.047	0.00664	0.79	636	0.008	0.00214	0.98
494	0.046	0.00652	0.78	638	0.008	0.00216	0.98
496	0.044	0.00631	0.78	640	0.008	0.00214	0.97
498	0.042	0.00615	0.78	642	0.008	0.00187	0.96

(Continued)

TABLE 3 | Continued

λ	A(λ)	SD	B(λ)	λ	A	SD	B(λ)
500	0.040	0.00599	0.78	644	0.008	0.00206	0.98
502	0.038	0.00583	0.78	646	0.008	0.00209	0.99
504	0.036	0.00564	0.78	648	0.008	0.0021	0.98
506	0.034	0.00539	0.79	650	0.008	0.00189	0.97
508	0.032	0.00519	0.78	652	0.008	0.00204	0.96
510	0.030	0.00503	0.79	654	0.008	0.00217	0.96
512	0.028	0.00498	0.79	656	0.008	0.00181	0.96
514	0.027	0.00453	0.78	658	0.009	0.0017	0.95
516	0.025	0.00434	0.78	660	0.010	0.00178	0.95
518	0.024	0.00414	0.78	662	0.012	0.00162	0.94
520	0.023	0.00416	0.78	664	0.013	0.00184	0.94
522	0.022	0.00391	0.78	666	0.016	0.00213	0.94
524	0.021	0.00376	0.78	668	0.018	0.00225	0.93
526	0.020	0.00393	0.79	670	0.020	0.00244	0.94
528	0.019	0.00372	0.80	672	0.022	0.00249	0.95
530	0.018	0.00386	0.79	674	0.024	0.00253	0.95
532	0.017	0.00365	0.80	676	0.024	0.00251	0.93
534	0.017	0.00362	0.82	678	0.024	0.00262	0.92
536	0.016	0.00343	0.81	680	0.024	0.00262	0.98
538	0.015	0.00349	0.82	682	0.022	0.00245	0.98
540	0.014	0.00343	0.84	684	0.020	0.00234	0.98
542	0.014	0.00337	0.89	686	0.017	0.00224	0.98
				688	0.014	0.00195	0.98
				690	0.011	0.00162	0.98
				692	0.009	0.00132	0.98
				694	0.007	0.00112	0.98
				696	0.005	0.00102	0.98
				698	0.004	0.00105	0.98
				700	0.003	0.0011	0.98

$a_{ph}(\lambda) = A(\lambda)\ (Chl\text{-}a)^{B(\lambda)}$ and determination coefficients on the log-transformed data r^2 (summer).

was not analyzed, but rough assessment of NPP (share of photoprotective pigments in total weight all pigments) based on dependence of NPP on light intensity (Babin et al., 1996) showed that NPP in UML was ∼5 times higher in summer in comparison with NPP in winter (**Figure 6**).

Seasonal phytoplankton succession observed is typical for the deep-water ecosystem of the Black Sea (Georgieva, 1993; Berseneva et al., 2004; Mikaelyan et al., 2005). In general biomass of the phytoplankton consists of Bacillariaphyceae, Dinophyceae, and Prymnesiophyceae (presented mainly by coccolithophores). Two-weekly monitoring at fixed stations in the western deep-water part of the Black Sea showed a change in phytoplankton species composition within an year (Berseneva et al., 2004): in general diatoms were dominating in winter and in yearly spring "blooms," dinoflagellates and coccolithophores were prevailing in the community in summer. Coccolithophores "bloom" in May-June.

Shift in species dominating in phytoplankton community is attributed with changes in size and shape of the cells. Cells size effects on pigment package within the cells which results in decreasing of $a^*_{ph}(\lambda)$ due to self-shading of pigments within large cells (Morel and Bricaud, 1981; Sosik and Mitchell, 1994; Fujiki and Taguchi, 2002). In different ocean regions variability in the $a^*_{ph}(\lambda)$ was related with change in phytoplankton species composition and cell size (Bricaud et al., 1995; Cleveland, 1995; Millan-Nunez et al., 2004). Package effect caused by cell size is detected by decreasing of $a^*_{ph}(678)$ (Fujiki and Taguchi, 2002) because at shorter wavelengths $a^*_{ph}(\lambda)$ is affected by accessory pigments as well. Values of $a^*_{ph}(678)$ decreased in winter by ∼15% compared with summer due to both C/Chl-*a* and phytoplankton variability (**Figures 3**, **5**) although the large diatoms (*Pseudosolenia calcar-avis*) were dominated in phytoplankton community.

The cells of *Pseudosolenia calcar-avis* have cylindrical shape unlike dinoflagellates, cells of which are closer to the ellipsoid. Volume of *Pseudosolenia calcar-avis* cell exceeds ∼2 orders of magnitude the volume of the dinoflagellates (*Gymnodinium spp*) cells. However, in the case of the cylindrically shaped cells the large volume is not critical for the cell's capacity to absorb

TABLE 4 | Spectral values of the constant obtained when fitting the variations of $a_{ph}(\lambda)$ vs. the (chlorophyll *a* + phaeopigment) concentration (Chl-*a*) to power laws of the form.

λ	A(λ)	SD	B(λ)	λ	A	SD	B(λ)
400	0.0285	0.00521	0.89	544	0.0147	0.00266	0.96
402	0.0294	0.00513	0.88	546	0.0143	0.00264	0.97
404	0.0306	0.00508	0.88	548	0.0137	0.00261	0.98
406	0.0318	0.00508	0.87	550	0.0131	0.00255	0.99
408	0.0331	0.00510	0.87	552	0.0125	0.00247	1.00
410	0.0343	0.00512	0.86	554	0.0118	0.00235	1.02
412	0.0354	0.00518	0.86	556	0.0111	0.00222	1.02
414	0.0363	0.00521	0.85	558	0.0104	0.00206	1.04
416	0.0371	0.00526	0.84	560	0.0097	0.00191	1.05
418	0.0376	0.00528	0.84	562	0.0090	0.00178	1.06
420	0.0381	0.00529	0.83	564	0.0085	0.00165	1.06
422	0.0386	0.00535	0.83	566	0.0080	0.00152	1.09
424	0.0391	0.00539	0.83	568	0.0076	0.00141	1.12
426	0.0397	0.00543	0.82	570	0.0073	0.00130	1.14
428	0.0405	0.00550	0.82	572	0.0070	0.00124	1.13
430	0.0415	0.00559	0.82	574	0.0069	0.00119	1.12
432	0.0426	0.00569	0.82	576	0.0069	0.00115	1.12
434	0.0436	0.00578	0.82	578	0.0069	0.00112	1.13
436	0.0444	0.00583	0.82	580	0.0070	0.00110	1.12
438	0.0448	0.00586	0.82	582	0.0071	0.00109	1.12
440	0.0448	0.00584	0.82	584	0.0072	0.00110	1.11
442	0.0444	0.00578	0.82	586	0.0073	0.00110	1.08
444	0.0436	0.00565	0.82	588	0.0074	0.00111	1.08
446	0.0426	0.00554	0.81	590	0.0075	0.00110	1.07
448	0.0416	0.00541	0.81	592	0.0075	0.00110	1.08
450	0.0407	0.00530	0.81	594	0.0074	0.00110	1.07
452	0.0400	0.00521	0.81	596	0.0073	0.00109	1.03
454	0.0396	0.00514	0.80	598	0.0071	0.00107	1.02
456	0.0394	0.00508	0.80	600	0.0069	0.00106	1.05
458	0.0393	0.00503	0.81	602	0.0068	0.00104	1.03
460	0.0394	0.00501	0.80	604	0.0066	0.00102	1.04
462	0.0394	0.00496	0.80	606	0.0065	0.00102	1.07
464	0.0394	0.00492	0.80	608	0.0065	0.00101	1.06
466	0.0393	0.00485	0.80	610	0.0066	0.00101	1.07
468	0.0390	0.00479	0.80	612	0.0066	0.00100	1.09
470	0.0386	0.00470	0.80	614	0.0067	0.00099	1.05
472	0.0381	0.00459	0.80	616	0.0069	0.00099	1.07
474	0.0374	0.00448	0.79	618	0.0070	0.00099	1.08
476	0.0366	0.00437	0.78	620	0.0071	0.00099	1.08
478	0.0358	0.00425	0.78	622	0.0072	0.00100	1.07
480	0.0350	0.00416	0.79	624	0.0073	0.00100	1.07
482	0.0341	0.00406	0.78	626	0.0074	0.00100	1.07
484	0.0334	0.00399	0.78	628	0.0075	0.00000	1.06
486	0.0326	0.00392	0.78	630	0.0076	0.00101	1.06
488	0.0319	0.00388	0.78	632	0.0077	0.00102	1.06
490	0.0311	0.00385	0.79	634	0.0079	0.00103	1.06
492	0.0304	0.00380	0.79	636	0.0080	0.00104	1.03
494	0.0296	0.00377	0.80	638	0.0080	0.00104	1.04
496	0.0288	0.00371	0.80	640	0.0080	0.00104	1.05
498	0.0279	0.00364	0.81	642	0.0080	0.00107	1.04

(Continued)

TABLE 4 | Continued

λ	A(λ)	SD	B(λ)	λ	A	SD	B(λ)
500	0.0270	0.00355	0.81	644	0.0080	0.00107	1.02
502	0.0261	0.00344	0.82	646	0.0079	0.00108	1.01
504	0.0252	0.00333	0.83	648	0.0078	0.00109	1.01
506	0.0243	0.00321	0.84	650	0.0078	0.00107	1.02
508	0.0235	0.00310	0.85	652	0.0079	0.00107	1.03
510	0.0227	0.00300	0.85	654	0.0082	0.00107	1.03
512	0.0220	0.00290	0.86	656	0.0087	0.00109	1.05
514	0.0213	0.00280	0.86	658	0.0095	0.00115	1.05
516	0.0206	0.00273	0.87	660	0.0106	0.00126	1.05
518	0.0200	0.00267	0.87	662	0.0121	0.00142	1.05
520	0.0194	0.00262	0.87	664	0.0137	0.00160	1.04
522	0.0189	0.00258	0.88	666	0.0156	0.00174	1.03
524	0.0184	0.00254	0.89	668	0.0175	0.00188	1.02
526	0.0180	0.00252	0.90	670	0.0191	0.00196	1.01
528	0.0176	0.00251	0.90	672	0.0203	0.00202	0.99
530	0.0172	0.00251	0.92	674	0.0210	0.00202	0.98
532	0.0169	0.00252	0.91	676	0.0211	0.00200	0.96
534	0.0165	0.00254	0.92	678	0.0206	0.00196	0.96
536	0.0162	0.00257	0.93	680	0.0194	0.00188	0.95
538	0.0159	0.00260	0.94	682	0.0178	0.00180	0.95
540	0.0155	0.00263	0.94	684	0.0156	0.00171	0.95
542	0.0152	0.00266	0.95	686	0.0132	0.00160	0.95
				688	0.0108	0.00147	0.95
				690	0.0086	0.00131	0.95
				692	0.0066	0.00114	0.95
				694	0.0051	0.00097	0.95
				696	0.0038	0.00083	0.95
				698	0.0029	0.00070	0.95
				700	0.0022	0.00062	0.95

$a_{ph}(\lambda) = A(\lambda)\ (Chl\text{-}a)^{B(\lambda)}$ and determination coefficients on the log-transformed data r^2 (winter).

light. The light absorption capacity of cylindrically shaped cells is determined by the diameter of their section (Kirk, 1976). Therefore, despite the difference in the cell volume optically significant size of cylindrical diatoms (10–30 μm) was similar to that of dinoflagellates (10–40 μm). Consequently, in this case the effect of the size (volume) of the cells on the degree of pigment packaging is not as significant as in the case of large spherical cells (Morel and Bricaud, 1981). It explains weak (∼15%) seasonal difference in $a^*_{ph}(678)$ observed in the Black Sea (**Figure 5**).

CONCLUSIONS

Seasonal differences in chlorophyll-*a* specific phytoplankton light absorption coefficients are caused by annual dynamics in environmental conditions in the (UML) and adaptive response of algae cells/population (via variation of pigment composition and concentration in the cell) and of phytoplankton community (via shift in phytoplankton species composition with attributed changes in size and shape of cell). Consequently, parameterization of the relationship between phytoplankton light absorption coefficients and chlorophyll *a* concentration proceeded for different seasons (summer and winter) will allow to refine the regional algorithm of Chl-*a* assessment based on remote sensing (Suslin and Churilova, 2016). Because in the Black Sea light absorption by dissolved organic matter there is relatively high and not correlated with phytoplankton absorption or chlorophyll *a* concentration regional Chl-*a* algorithm requires splitting of light absorption into $a_{ph}(\lambda)$ and $a_{CDM}(\lambda)$ (Suslin and Churilova, 2016) and then Chl-*a* is retrieved from the $a_{ph}(\lambda)$ at 490 nm. Relationships between Chl-*a* and $a_{ph}(\lambda)$ obtained for the summer and winter conditions in the Black Sea differ by coefficients A(λ) in power equation, but coefficients B(λ) are practically the same (**Figure 5**). Consequently, values of A(λ) coefficient define the seasonal difference in retrieval of Chl-*a* based on $a_{ph}(\lambda)$. Values of A(λ) at 490 nm are equal 0.048 and 0.031 $m^2\ mg^{-1}$ correspondently for summer and winter conditions in UML of the Black Sea (**Tables 3, 4**). For instance, using the summer relationship between Chl-*a* and summer a_{ph} (490) values one can get Chl-*a* equal to 0.2–0.3 mg m^{-3}, but using the winter link between these parameters or link

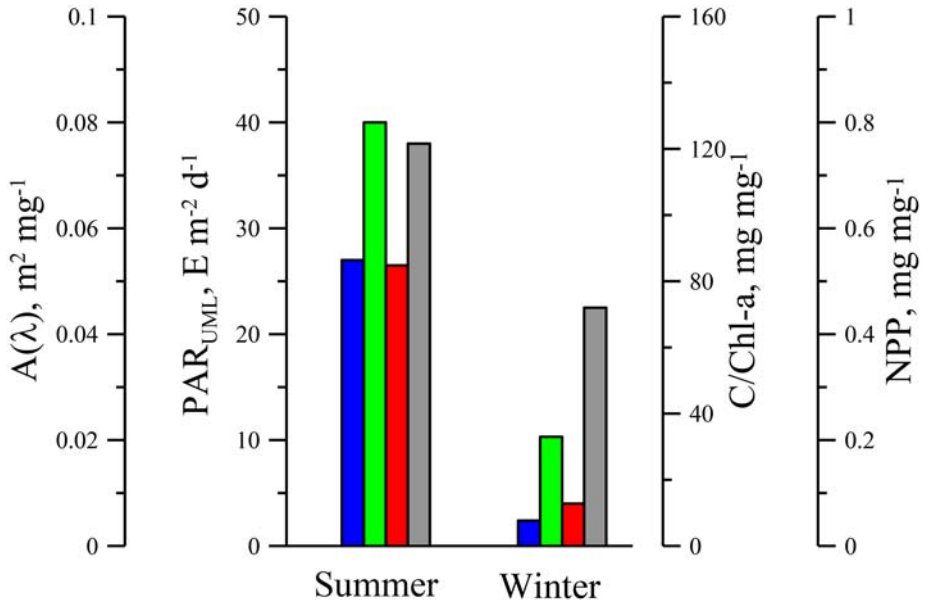


FIGURE 6 | Photosynthetically available radiance in upper mixed layer (PAR_{UML}, blue bars), carbon to chlorophyll *a* ratio (C/Chl-*a*, green bars), share of non-photosynthetic photoprotective pigments (NPP, red bars) in total pigments and coefficient [A(440), gray bars] in phytoplankton light absorption parameterization fulfilled with power equation $a_{ph}(\lambda) = A(\lambda)(\text{Chl-}a)^{B(\lambda)}$ for upper mixed layer in the Black Sea in summer and winter.

obtained for different regions of the world ocean (Bricaud et al., 1995) one gets Chl-*a* equal to 0.36–0.53 or to 0.34–0.55 mg m^{-3}. Consequently, the retrieved Chl-*a* values become almost twice lower if one takes into account the Black Sea summer conditions and relevant relationship between $a_{ph}(\lambda)$ and Chl-*a*. Undoubtedly accuracy of splitting of light absorption into $a_{ph}(\lambda)$ and $a_{CDM}(\lambda)$ also affect the accuracy of Chl-*a* assessment (Suslin and Churilova, 2016).

Moreover, seasonal difference in links between $a_{ph}(\lambda)$ and Chl-*a* could provide more correct assessment of downwelling radiance and primary production in the Black Sea using spectral approaches (Churilova et al., 2016). However, it should be noted that application of the obtained $a_{ph}(\lambda)$ parametrization is limited by the rather narrow range of Chl-*a*, which was measured in the deep waters. The relatively narrow range of Chl-*a* caused the high values of B(λ) coefficients in comparison of those obtained based on numerical data measured in different regions of World Ocean with Chl-*a* covering the range 0.02–25 mg m^{-3} (**Figure 5**; Bricaud et al., 1995).

The parameterization obtained based on bio-optical data measured in deep waters is unlikely to be correct for coastal waters. Coastal waters may differ from deep waters in nutrient availability, transparency and turbulence. These different environmental conditions would results in change of intracellular pigment concentration and phytoplankton species composition which in turn effect on $a^*_{ph}(\lambda)$. In this regard since 2014 bio-optical properties have been investigated in the Crimean coastal waters in different seasons. These new data will be merged with summer results measured before (Churilova and Berseneva, 2004; Dmitriev et al., 2009) and then analyzed to determine the seasonality in $a_{ph}(\lambda)$ parameterization.

AUTHOR CONTRIBUTIONS

TC, idea, management and writing. VS, mathematical proceeding and interpretation data. OK, contribution to the discussion. TE, chlorophyll and absorption measurements, calculations. NM, chlorophyll and absorption measurements, calculations. VM, phytoplankton data analysis. LS, phytoplankton species identification.

FUNDING

The work presented in this paper was carried out in the framework of RF state task according to plan of scientific research of the AO. Kovalevsky Institute of Marine Biological Research (theme # 0828-2014-0016) and of the Marine Hydrophysical Institute (theme # 0827-2014-0011). The work was partially supported by RFBR (project 17-05-00113).

ACKNOWLEDGMENTS

The authors are very thankful to ESA and PML for invitation us to attend the workshop "Color and Light in the Ocean (CLEO)" held at ESRIN, Frascati, Italy on 6-8 September, 2016. The authors would like to thank the reviewers for their helpful and constructive comments that greatly contributed to improving the paper. The authors thank very much ESA for covering of the article processing charges.

REFERENCES

Agirbas, E., Feyzioglu, A. M., and Kopuz, U. (2014). Seasonal changes of phytoplankton chlorophyll *a*, primary production and their relation in the continental shelf area of the South Eastern Black Sea. *Turk. J. Fish. Aquat. Sci.* 14, 713–726. doi: 10.4194/1303-2712-v14_3_14

Aguirre-Hernandez, E., Gaxiola-Castro, G., Najera-Martınez, S., Baumgartner, T., Kahru, M., and Mitchell, B. G. (2004). Phytoplankton absorption, photosynthetic parameters, and primary production off Baja California: summer and autumn 1998. *Deep-Sea Res. II.* 51, 799–816. doi: 10.1016/j.dsr2.2004.05.015

Babin, M., Morel, A., Claustre, H., Bricaud, A., Kolber, Z., and Falkowski, P. (1996). Nitrogen- and iraadiance- dependent variations of maximum quantum yield of carbon fixation in eutrophic, mesotrophic and oligotrophic marine systems. *Deep-Sea Res. I.* 43, 1241 – 1272. doi: 10.1016/0967-0637(96)00058-1

Babin, M., Stramski, D., Ferrari, G. M., Claustre, H., Bricaud, A., Obolensky, G., et al. (2003). Variations in the light absorption coefficients of phytoplankton, non-algal particles, and dissolved organic matter in coastal waters around Europe. *J. Geophys. Res.* 108, 3211. doi: 10.1029/2001jc000882

Barlow, R. G., Aiken, J., Moore, G. F., Holligan, P. M., and Lavender, S. (2004). Pigment adaptations in surface phytoplankton along the eastern boundary of the Atlantic Ocean. *Mar. Ecol. Prog. Ser.* 281, 13–26. doi: 10.3354/meps281013

Behrenfeld, M. J., Boss, E., Siegel, D. A., and Shea, D. M. (2005). Carbon-based ocean productivity and phytoplankton physiology from space. *Global Biogeochem. Cycles* 19:GB1006. doi: 10.1029/2004GB002299

Behrenfeld, M. J., O'Malley, R. T., Siegel, D. A., McClain, C. R., Sarmiento, J. L., Feldman, G. C., et al. (2006). Climate-driven trends in contemporary ocean productivity. *Nature* 444, 752–755. doi: 10.1038/nature05317

Berseneva, G., Churilova, T., and Georgieva, L. (2004). Seasonal variability in pigment concentration, taxonomic structure of phytoplankton community in western part of the Black Sea. *Oceanology* 44, 211–219.

Berthon, J.-F., Mélin, F., and Zibordi, G. (2008). "Ocean Color remote sensing of the optically complex european seas," in *Remote Sensing of the European Seas,* eds V. Barale and M. Gade (Dordrecht: Springer Science-Business Media), 35–52.

Bidigare, R. R., Ondrusek, M. E., Morrow, J. H., and Kiefer, D. A. (1990). *In vivo* absorption properties of algal pigments. *Ocean Opt. X.* 1302, 290–302. doi: 10.1117/12.21451

Bricaud, A., Babin, M., Morel, A., and Claustre, H. (1995). Variability in the chlorophyll-specific absorption coefficients of natural phytoplankton: analysis and parameterization. *J. Geophys. Res.* 100, 13321–13332. doi: 10.1029/95JC00463

Bricaud, A., Morel, A., Babin, M., and K., Allali, H. (1998). Claustre Variations of light absorption by suspended particles with chlorophyll *a* concentration in oceanic (case 1) waters: analysis and implications for bio-optical models. *J. Geophys. Res.* 103, 31 033–31 044.

Chami, M., Shybanov, E. B., Churilova, T. Y., Khomenko, G. A., Lee, M. E. G., Martynov, O. V., et al. (2005). Optical properties of the particles in the Crimea coastal waters (Black Sea). *J. Geophys. Res.* 110, 1–17. doi: 10.1029/2005JC003008

Churilova, T. (2001). Light absorption by phytoplankton and detritus in the Black Sea in spring. *Oceanology* 41, 687–695.

Churilova, T., and Berseneva, G. (2004). Absorption of light by phytoplankton, detritus, and dissolved organic substances in the coastal region of the Black Sea (July–August 2002). *Phys. Oceanogr.* 14, 221–233. doi: 10.1007/s11110-005-0016-3

Churilova, T., Berseneva, G., and Georgieva, L. (2004). Variability in bio-optical characteristics of phytoplankton in the Black Sea. *Oceanology* 44, 192–204.

Churilova, T., and Suslin, V. (2010). "Parameterization of light absorption by all in-water optically active components in the Black Sea: Impact for underwater irradiance and primary production modeling," in *Coastal to Global Operational Oceanography: Achievements and Challenges, Proceedings of the Fifth International Conference on EuroGOOS,* (Exeter, UK).

Churilova, T., Suslin, V., Krivenko, O., Efimova, T., and Moiseeva, N. (2016). Spectral approach to assessment of phytoplankton photosynthesis rate in the black sea based on satellite. Methodological aspects of the regional model development. *J. Siberian Federal Univ.* 4, 367–384.

Churilova, T., Suslin, V., and Sosik, H. M. (2009). A spectral model of underwater irradiance in the Black Sea. *Phys. Oceanogr.* 19, 366–378. doi: 10.1007/s11110-010-9060-8

Cleveland, J. S. (1995). Regional models for phytoplankton absorption as a function of chlorophyll *a* concentration. *J. Geophys.* 100, 333–344.

Dmitriev, E. V., Khomenko, G., Chami, M., Sokolov, A. A., Churilova, T. Y., and Korotaev, G. K. (2009). Parameterization of light absorption by components of seawater in optically complex coastal waters of the Crimea Peninsula (Black Sea). *Appl. Opt.* 48, 1249–1261. doi: 10.1364/AO.48.001249

Feldman, G. C., McClain, C. R. (2013). *Ocean Color Web SeaWiFS Reprocessing 2010.0, MODIS-Terra Reprocessing 2013.0, MODIS-Aqua Reprocessing 2013.1, MERIS Reprocessing 2012.0, NASA Goddard Space Flight Center*. eds N. Kuring and S. W. Bailey. Available online at: http://oceancolor.gsfc.nasa.gov/cms/

Fujiki, T., and Taguchi, S. (2002). Variability in chlorophyll *a* specific absorption coefficient in marine phytoplankton as a function of cell size and irradiance. *J. Plankton Res.* 24, 859–874. doi: 10.1093/plankt/24.9.859

Georgieva, L. V., (1993). "Species composition and dynamics of phytoplankton," in *Plankton of the Black Sea,* ed V. E. Zaika (Kiev: Naukova Dumka), 31–55.

Gordon, H. R., and McGlunev, W. R. (1975). Estimation of the depth of sunlight penetration in the sea for remote sensing. *Appl. Opt.* 14, 413–416. doi: 10.1364/AO.14.000413

Grant, C. S., and Louda, J. W. (2010). Microalgal pigment ratios in relation to light intensity: implications for chemotaxonomy. *Aquat. Biol.* 11, 27–138, doi: 10.3354/ab00298

Hoepffner, N., and Sathyendranath, S. (1991). Effect of pigment composition on absorption properties of phytoplankton. *Mar. Ecol. Prog. Ser.* 73, 11–23. doi: 10.3354/meps073011

Hoepffner, N., and Sathyendranath, S. (1992). Bio-optical characteristics of coastal waters: absorption spectra of phytoplankton and pigment distribution in the western North Atlantic. *Limnol. Oceanogr.* 37, 1660–1679. doi: 10.4319/lo.1992.37.8.1660

Ivanov, V. A., and Belokopytov, V. N. (2011). *Oceanography of the Black Sea. National Academy of Science of Ukraine.* (Sevastopol: ECOSI-Gidrofizika), 211.

Jeffrey, S. W., and Humphrey, G. F. (1975). New spectrophotometric equations for determining chiorophylls a, b, c1 and c2 in higher plants, algae and phytoplankton. *Biochem. Physiol. Pflanzen.* 167, 191–197. doi: 10.1016/S0015-3796(17)30778-3

Jeffrey, S. W., Mantoura, R. F. C., and Wright, S. W. (eds). (1997). *Phytoplankton Pigments in Oceanography: Guidelines to Modern Methods.* Paris: UNESCO publishing.

Kirk, J. T. O. (1976). A Theoretical analysis of the contribution of algae cells to the attenuation of light within natural waters. III. Cylindrical and spheroidal cells. *New Phytol.* 77, 341–358. doi: 10.1111/j.1469-8137.1976.tb01524.x

Kirk, J. T. O. (1994). *Light and Photosynthesis in Aquatic Ecosystems*. New York, NY: Cambridge University Press.

Lorenzen, C. J. (1967). Determination of chlorophyll and pheopigments: spectrophotometric equations. *Limnol. Oceanogr.* 12, 343–346. doi: 10.4319/lo.1967.12.2.0343

Lutz, V. A., Sathyendranath, S., and Head, E. J. H. (1996). Absorption coefficient of phytoplankton: regional variations in the North Atlantic. *Mar. Ecol. Prog. Ser.* 135, 197–213. doi: 10.3354/meps135197

Lutz, V. A., Sathyendranath, S., Head, E. J. H., and Li, W. K. W. (2003). Variability in pigment composition and optical characteristics of phytoplankton in the Labrador Sea and the Central North Atlantic. *Mar. Ecol. Prog. Ser.* 260, 1–18. doi: 10.3354/meps260001

MacIntyre, H. L., Kana, T. M., Anning, J., and Geider, R. (2002). Photoacclimation of photosynthesis irradiance response curves and photosynthetic pigments in microalgae and cyanobacteria. *J. Phycol.* 38, 17–38. doi: 10.1046/j.1529-8817.2002.00094.x

Marie, D., Partensky, F., Vaulot, D., and Brussaard, C. P. D. (1999). "Enumeration of phytoplankton, bacteria and viruses in marine samples," in *Current Protocols in Cytometry Supplement 10,* eds J. P. Robingson, Z. Darzynkiewicz, P. N. Dean, A. Orfao, P. Rabinovitch, C. C. Stewart et al. (New York, NY: John Wiley and Sons Inc.), 11.11.1–11.11.15.

Mikaelyan, A. S., Pautova, L. A., Pogosyan, S. I., and Sukhanova, I. N. (2005). Summer bloom of Coccolothophorids in the Northeastern Black Sea. *Oceanology* 45, S127–S138.

Millan-Nunez, E., Sieracki, M. E., Millan-Nunez, R., Lara-Lara, J. R., and Gaxiola-Castro, G. C. C. (2004). Trees Specific absorption coefficient and phytoplankton biomass in the southern region of the California Current. *Deep-Sea Res. II.* 51, 817–826. doi: 10.1016/j.dsr2.2004.05.023

Mitchell, B. G. (1990). "Algorithms for determining the absorption coefficient of aquatic particulates using the quantitative filter technique (QFT)," in *Ocean Optics X*. ed R. Spinrad (Washington, DC: SPIE Bellingham). 137–148.

Mitchell, B. G., and Kiefer, D. A. (1988). Chlorophyll *a* specific absorption and fluorescence excitation spectra for light limited phytoplankton. *Deep-Sea Res.* 35, 639–663. doi: 10.1016/0198-0149(88)90024-6

Morel, A., and Bricaud, A. (1981). Theoretical results concerning light absorption in a discrete medium, and application to specific absorption of phytoplankton. *Deep-Sea Res.* 28, 1375–1393. doi: 10.1016/0198-0149(81)90039-X

Morel, A., and Prieur, L. (1977). Analysis of variations in ocean color. *Limnol. Oceanogr.* 22, 709–722. doi: 10.4319/lo.1977.22.4.0709

O'Reilly, J. E., Maritorena, S., O'Brien, M. C., Siegel, D. A., Toole, D., Menzies, D., et al. (2000). "SeaWiFS Postlaunch Calibration and Validation Analyses, Part 3," in *NASA Technical Memorandum. 2000-206892,* Vol. 11. eds S. B. Hooker and E. R. Firestone (Greenbelt, MD: NASA Goddard Space Flight Center), 49.

Saba, V. S., Friedrichs, M. A. M., Antoine, D., Armstrong, R. A., and Asanuma, I. M. J. (2011). An evaluation of ocean color model estimates of marine primary productivity in coastal and pelagic regions across the globe. *Biogeosciences.* 8, 489–503. doi: 10.5194/bg-8-489-2011

Sathyendranath, S., Stuart, V., Platt, T., Bouman, H., Ulloa, O., and Maass, H. (2005). Remote sensing of ocean color: towards algorithms for retrieval of pigment composition. *Indian J. Mar. Sci.* 34, 333–340.

Schapira, M., Buscot, M. J., Pollet, T., Leterme, S. C., and Seuront, L. (2010). Distribution of picophytoplankton communities from brackish to hypersaline waters in a South Australian coastal lagoon. *Saline Syst.* 6, 2–16. doi: 10.1186/1746-1448-6-2

Sosik, H. M., and Mitchell, B. G. (1994). Effects of temperature on growth, light absorption and quantum yield in *Dunaliella tertiolecta* (Chlorophyceae). *J. Phycol.* 30, 833–840. doi: 10.1111/j.0022-3646.1994.00833.x

Stuart, V., Sathyendranath, S., Platt, T., Maass, H., and Irwin, B. D. (1998). Pigments and species composition of natural phytoplankton populations: effect on the absorption spectra. *J. Plankton Res.* 20, 187–217. doi: 10.1093/plankt/20.2.187

Suslin, V., and Churilova, T. (2016). The Black Sea regional algorithm of separation of light absorption by phytoplankton and colored detrital matter using ocean color scanner's bands from 480-560 nm. *Int. J. Remote Sens.* 37, 4380–4400. doi: 10.1080/01431161.2016.1211350

Suslin, V. V., Churilova, T. Y., and Sosik, H. M. (2008). The SeaWiFS algorithm of chlorophyll *a* in the Black Sea. *Mor. Ekol. Zh.* 7, 24–42.

Suslin, V. V., Korolev, S. N., Kucheryaviy, A. A., Churilova, T., and Ya Krivenko, O. V. (2015). "Photosynthetically available radiation on surface of the Black Sea based on ocean color data," in *Proceedings SPIE 9680, 21st International Symposium Atmospheric and Ocean Optics: Atmospheric Physics,* (Bellingham, WA).

Suslin, V. V., Suetin, V. S., Korolev, S. N., and Kucheryaviy, A. A. (2007). "Use of SeaWiFS data to estimate water optical properties of the Black Sea," in *Proceedings SPIE 6615, Current Research on Remote Sensing, Laser Probing, and Imagery in Natural Waters,* (Bellingham, WA).

Suzuki, K., Kishino, M., Sasaoka, K., Saitoh, S. I., and Saino, T. (1998). Chlorophyll-specific absorption coefficients and pigments of phytoplankton off sanriku, Northwestern North Pacific. *J. Oceanogr.* 54, 517–526. doi: 10.1007/BF02742453

Tassan, S., and Ferrari, G. (1995). An alternative approach to absorption measurements of aquatic particles retained on filters. *Limnol. Oceanogr.* 40, 1358–1368. doi: 10.4319/lo.1995.40.8.1358

Vedernikov, V. I. (1989). "Primary production and chlorophyll in the Black Sea in summer and fall," in *Struktura i Produktsionnye Kharakteristiki Planktonnykh Soobshchestv: Sbornikh Nauchnykh Rabot,* eds M. E. Vonogradov and M. V. Flint (Moscow: Nauka), 65–83.

Yentsch, C. S. (1962). Measurement of visible light absorption by particulate matter in the ocean. *Limnol. Oceanogr.* 7, 207–217. doi: 10.4319/lo.1962.7.2.0207

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer BB and handling Editor declared their shared affiliation, and the handling Editor states that the process nevertheless met the standards of a fair and objective review.

Copyright © 2017 Churilova, Suslin, Krivenko, Efimova, Moiseeva, Mukhanov and Smirnova. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Marine Science

ORIGINAL RESEARCH
published: 13 April 2017
doi: 10.3389/fmars.2017.00104

Uncertainty in Ocean-Color Estimates of Chlorophyll for Phytoplankton Groups

Robert J. W. Brewin[1,2], Stefano Ciavatta[1,2], Shubha Sathyendranath[1,2], Thomas Jackson[1], Gavin Tilstone[1], Kieran Curran[1], Ruth L. Airs[1], Denise Cummings[1], Vanda Brotas[3], Emanuele Organelli[1], Giorgio Dall'Olmo[1,2] and Dionysios E. Raitsos[1,2]*

[1] Plymouth Marine Laboratory, Plymouth, UK, [2] National Centre of Earth Observation, Plymouth Marine Laboratory, Plymouth, UK, [3] Faculdade de Ciências, Marine and Environmental Sciences Centre, Universidade de Lisboa, Lisboa, Portugal

OPEN ACCESS

Edited by:
Chris Bowler, École Normale Supérieure, France

Reviewed by:
Ramaiah Nagappa, National Institute of Oceanography, India
Salvatore Marullo, National Agency For New Technologies, Energy and Sustainable Economic Development, Italy

***Correspondence:**
Robert J. W. Brewin
robr@pml.ac.uk

Specialty section:
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

Received: *13 January 2017*
Accepted: *27 March 2017*
Published: *13 April 2017*

Citation:
Brewin RJW, Ciavatta S, Sathyendranath S, Jackson T, Tilstone G, Curran K, Airs RL, Cummings D, Brotas V, Organelli E, Dall'Olmo G and Raitsos DE (2017) Uncertainty in Ocean-Color Estimates of Chlorophyll for Phytoplankton Groups. Front. Mar. Sci. 4:104. doi: 10.3389/fmars.2017.00104

Over the past decade, techniques have been presented to derive the community structure of phytoplankton at synoptic scales using satellite ocean-color data. There is a growing demand from the ecosystem modeling community to use these products for model evaluation and data assimilation. Yet, from the perspective of an ecosystem modeler these products are of limited use unless: (i) the phytoplankton products provided by the remote-sensing community match those required by the ecosystem modelers; and (ii) information on per-pixel uncertainty is provided to evaluate data quality. Using a large dataset collected in the North Atlantic, we re-tune a method to estimate the chlorophyll concentration of three phytoplankton groups, partitioned according to size [pico- ($<2\,\mu$m), nano- (2–20 μm) and micro-phytoplankton ($>20\,\mu$m)]. The method is modified to account for the influence of sea surface temperature, also available from satellite data, on model parameters and on the partitioning of microphytoplankton into diatoms and dinoflagellates, such that the phytoplankton groups provided match those simulated in a state of the art marine ecosystem model (the European Regional Seas Ecosystem Model, ERSEM). The method is validated using another dataset, independent of the data used to parameterize the method, of more than 800 satellite and *in situ* match-ups. Using fuzzy-logic techniques for deriving per-pixel uncertainty, developed within the ESA Ocean Colour Climate Change Initiative (OC-CCI), the match-up dataset is used to derive the root mean square error and the bias between *in situ* and satellite estimates of the chlorophyll for each phytoplankton group, for 14 different optical water types (OWT). These values are then used with satellite estimates of OWTs to map uncertainty in chlorophyll on a per pixel basis for each phytoplankton group. It is envisaged these satellite products will be useful for those working on the validation of, and assimilation of data into, marine ecosystem models that simulate different phytoplankton groups.

Keywords: phytoplankton, size, function, chlorophyll, ocean-color, uncertainty

1. INTRODUCTION

The size structure and taxonomic composition of phytoplankton influence many processes in phytoplankton biology, marine biogeochemistry and marine ecology (Chisholm, 1992; Raven, 1998; Le Quéré et al., 2005; Marañón, 2009, 2015; Finkel et al., 2010). Photosynthesis, growth, light absorption, nutrient uptake, carbon export, and the transfer of energy through the marine food chain, are all influenced by phytoplankton community structure (Platt and Denman, 1976, 1977, 1978; Morel and Bricaud, 1981; Prieur and Sathyendranath, 1981; Probyn, 1985; Geider et al., 1986; Legendre and LeFevre, 1991; Maloney and Field, 1991; Chisholm, 1992; Sunda and Huntsman, 1997; Raven, 1998; Laws et al., 2000; Ciotti et al., 2002; Bricaud et al., 2004; Devred et al., 2006; Guidi et al., 2009; Briggs et al., 2011). In the face of considerable challenges (Shimoda and Arhonditsis, 2016), growing emphasis has been placed on the representation of biogeochemistry in ecosystem models by explicitly incorporating different phytoplankton groups as state variables, often partitioned according to their size or taxonomic composition (Aumont et al., 2003; Blackford et al., 2004; Le Quéré et al., 2005; Kishi et al., 2007; Marinov et al., 2010; Ward et al., 2012; Butenschön et al., 2016). With this aspiration comes a demand for observations on phytoplankton groups (e.g., for model validation and data assimilation) that is not being met with current *in situ* observations that are sparse in time and space. To address the issue of data availability, the past decade has seen many attempts to estimate phytoplankton groups using satellite remote-sensing (IOCCG, 2014), which is capable of viewing the ocean with high temporal and spatial coverage.

Current techniques to estimate phytoplankton groups using satellite data can be partitioned into three categories: spectral, abundance and ecological approaches (Nair et al., 2008; Brewin et al., 2011b; IOCCG, 2014). Spectral-based approaches seek to use the optical signatures of the phytoplankton groups directly for their detection from space. Abundance-based approaches invoke relationships between the phytoplankton groups and some index of phytoplankton abundance or biomass (e.g., chlorophyll concentration) that can be retrieved from satellites. Ecological-based approaches use ocean-color together with additional environmental data (e.g., sea surface temperature (SST), irradiance, wind) that can also be retrieved from satellite to identify ecological niches where particular phytoplankton communities may be found. Spectral-based approaches are more direct as they target known optical signatures, whereas abundance-based and ecological-based approaches are indirect, in that they use satellite remote-sensing as a means to extrapolate known relationships between the phytoplankton groups and a property that can by derived accurately from space (e.g., chlorophyll concentration, SST). Though it would appear more sensible to use a direct approach, issues with spectral-based techniques can arise when the signal-to-noise ratio in the ocean-color data is too low to detect the targeted signature (Garver et al., 1994; Wang et al., 2005), when the phytoplankton group being targeted has a similar optical signature to other groups, when the spectral signatures are not known sufficiently well, or when the spectral resolution is not adequate for detecting the target signature. In such cases, an indirect method (e.g., ecological or abundance based) would be more suitable. Future ocean-color missions will help address some of these issues through improved accuracy and spectral resolution. For instance, the recently launched Ocean and Land Color Instrument (OLCI) on-board ESA's Sentinel-3a satellite offers more spectral wavebands than its predecessor (MERIS), and NASA's planned Pre-Aerosol Clouds and ocean Ecosystem (PACE) mission will aim to provide hyperspectral ocean-color data, improving the potential for phytoplankton group retrievals. For further details on all of these methods, the reader is referred to the works of Nair et al. (2008), Brewin et al. (2011b), De Moraes Rudorff and Kampel (2012), IOCCG (2014), and Mouw et al. (2017). Recently, efforts have been made to combine abundance and ecological-based approaches, for instance, Brewin et al. (2015) and Ward (2015) modified the relationship between the chlorophyll concentration of the phytoplankton groups and total chlorophyll (abundance-based) according to the environmental (ecological-based) conditions (e.g., temperature or light availability).

Phytoplankton group-specific satellite products are now being used for the validation of (Ward et al., 2012; Hirata et al., 2013; Hashioka et al., 2013; Rousseaux et al., 2013; Vogt et al., 2013; Holt et al., 2014; de Mora et al., 2016; Laufkötter et al., 2016), or assimilation of data into (Xiao and Friedrichs, 2014), ecosystem models. However, there are two challenges that modelers face when undertaking such analyses (Bracher et al., 2017). Firstly, there is often a mismatch between phytoplankton products provided by the remote-sensing community and those required by the ecosystem modelers. These difficulties arise in cases where a phytoplankton group adopted by the ecosystem modeler has similar optical properties to other phytoplankton groups, meaning they may not be detected directly using spectral-based methods, or the phytoplankton group does not co-vary in a predictable manner with variables amenable from remote-sensing, limiting abundance-based and ecological-based methods and rendering the use of satellite products difficult. Greater dialog between ecosystem modelers and the remote-sensing community is required to bridge this mismatch where feasible.

The second challenge is associating a level of uncertainty to the satellite phytoplankton group products, ideally on a per-pixel basis (per grid cell of the model). This is an essential prerequisite for both ecosystem model validation and data assimilation. If the uncertainties in the satellite products are too high they may not be useful for validation and may have little impact on a data assimilation scheme, since the target for data assimilation is to modify model simulations such that they agree with the observations within their uncertainties (e.g., Gregg et al., 2009; Ford et al., 2012; Ciavatta et al., 2014, 2016). Whereas many approaches have been proposed to derive satellite phytoplankton group products (IOCCG, 2014), few provide estimates of per-pixel uncertainty.

There are two methods commonly used to estimate uncertainty in ocean-color products: error propagation, or model-based uncertainties, and comparison of satellite estimates with *in situ* data (validation). Error propagation typically involves propagation of errors from input to output products, knowing the uncertainties in the input and model parameters.

These techniques have been used for estimating uncertainties in chlorophyll concentration and inherent optical properties (Maritorena et al., 2010; Lee et al., 2011; Werdell et al., 2013a), and for some satellite phytoplankton group products (Kostadinov et al., 2009, 2016; Roy et al., 2013; Brewin et al., 2017). In addition to estimating per-pixel uncertainty, these techniques can be very useful for understanding the sensitivity of model parameters and model inputs on the output products (Roy et al., 2013; Kostadinov et al., 2016; Brewin et al., 2017).

In a user consultation of ocean-color products, conducted as part of the ESA Ocean Colour Climate Change Initiative (OC-CCI), there seemed to be a preference from ecosystem modelers for estimates of uncertainties based on comparison with *in situ* data, rather than model-based uncertainties (Sathyendranath, 2011). For most techniques, satellite phytoplankton group products have been validated with *in situ* data (see Table 3 of Mouw et al., 2017). However, this information is typically provided as a single statistic (e.g., root mean square error), which can be difficult to convert to a per-pixel error, considering uncertainties are likely to vary with the environmental conditions and the magnitude of the product. Furthermore, the distribution of data used in validation datasets may not be an adequate representation of the spatial and temporal variability in the region under study.

To overcome these issues, Moore et al. (2001, 2009, 2012) proposed the use of an optical classification of pixels, together with fuzzy-logic statistics, to estimate per-pixel errors in satellite ocean-color products based on comparison with *in situ* data. In this approach, satellite and *in situ* match-ups are segregated into dominant optical water types (ranging from oligotrophic to turbid waters), then error statistics are computed for each dominant optical water-type. An ocean-color spectrum (at a given pixel) is then compared with all the optical water type spectra to determine its fuzzy membership. The fuzzy membership is then used to compute the error by weighting the errors in each dominant optical water type according to the fuzzy membership. This approach can, to a certain degree, overcome issues with the distribution of data used in the validation, and account for uncertainties varying with the conditions and the magnitude of the product. It has been adopted in the ESA OC-CCI project and is used to provide per-pixel errors (root mean square error and bias) for all OC-CCI products, including: chlorophyll, diffuse attenuation coefficient, and the inherent optical properties of oceanic waters. However, this approach has not been applied to satellite phytoplankton group products.

The Copernicus Marine Environment Monitoring Service (CMEMS) project "Toward Operational Size-class Chlorophyll Assimilation (TOSCA)" seeks to address these issues by: (i) providing remotely-sensed products on phytoplankton groups that map onto those simulated by the European Regional Seas Ecosystem model (ERSEM; Butenschön et al., 2016), which is the ecosystem model adopted in this project; and (ii) provide uncertainty estimates for the remotely-sensed products on a per-pixel basis, based on *in situ* match-ups (the preferred choice for ecosystem modelers; Sathyendranath, 2011). In this paper, we re-tuned an abundance-based method (Brewin et al., 2010, 2015) to estimate the chlorophyll concentration of three phytoplankton groups, partitioned according to size, from satellite data in the North Atlantic. The abundance-based method was modified to account for the influence of SST (i.e., combining the method with an ecological-approach), and partition microphytoplankton into diatoms and dinoflagellates, so that the phytoplankton groups provided by the satellite approach match those simulated by ERSEM. Using an optical classification of pixels with fuzzy-logic statistics (Moore et al., 2001, 2009, 2012; Jackson and Sathyendranath, 2015), we present a method for deriving per-pixel uncertainty for each phytoplankton group based on a validation dataset of satellite and *in situ* match-ups, which is independent of the data used to parameterize the method.

2. METHODS

2.1. Study Area: The North Atlantic

The chosen study site was the North Atlantic (**Figure 1**), spanning 46° W to 13° E and 20° N to 66° N, and categorized by the CMEMS Ocean Colour Thematic Assembley Centre (OCTAC) as the Atlantic (ATL) region. This region encompasses a range of bio-optical conditions from clear, deep open-ocean waters to shallower optically-complex shelf seas. We chose this site because of two factors: (i) it is a region that has been extensively sampled over the past few decades, resulting in a relatively large number of *in situ* observations on phytoplankton groups when compared with other regions of the ocean; and (ii) it has been subject to many studies on marine ecosystem modeling (e.g., Holt et al., 2014). The North Atlantic is also home

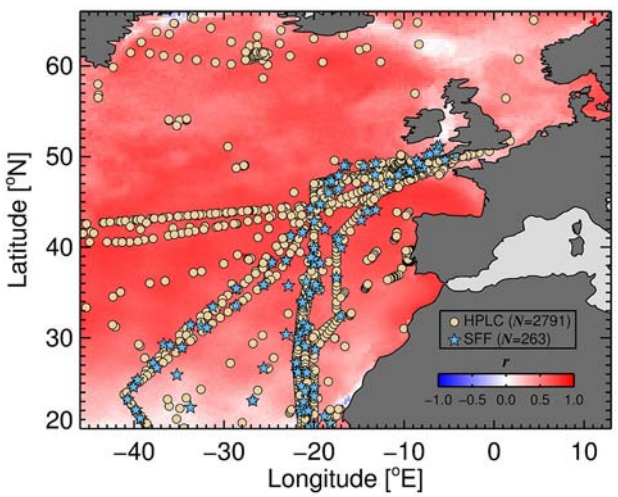


FIGURE 1 | Locations of High Performance Liquid Chromatography (HPLC) and size-fractionated filtration (SFF) *in situ* data (<20 m depth) used in this study (CMEMS OCTAC ATL region). Background color show pixel-by-pixel correlation coefficients (*r*) of monthly Sea Surface Temperature (ESA SST products) and monthly average light in the mixed-layer between 2000 and 2010 [computed using Equation 11 of Brewin et al. (2015) with a monthly climatology of mixed-layer depth (de Boyer Montégut et al., 2004), monthly photosynthetic available radiation products from NASA SeaWiFS (http://oceancolor.gsfc.nasa.gov/), and K_d estimated from Morel et al. (2007) using OC-CCI monthly chlorophyll products].

to one of the largest spring phytoplankton blooms on the planet (Ducklow and Harris, 1993) and is known as a major region for the biological drawdown of seawater CO_2 (Takahashi et al., 2002, 2009) and primary production (Tilstone et al., 2014).

2.2. Statistical Tests

To compare the *in situ* and satellite chlorophyll concentrations, we used the root mean square error (Ψ) and bias (δ), consistent with the statistical tests adopted in the ESA OC-CCI project and used to provide per-pixel errors. The Ψ and δ values were computed according to

$$\Psi = \left[\frac{1}{N} \sum_{i=1}^{N} \left(X_i^E - X_i^M \right)^2 \right]^{1/2}, \quad (1)$$

and

$$\delta = \frac{1}{N} \sum_{i=1}^{N} \left(X_i^E - X_i^M \right), \quad (2)$$

where X is the variable (chlorophyll concentration) and N is the number of samples. The superscript E denotes the estimated variable (e.g., satellite estimate) and M the measured variable (e.g., *in situ*). Note that the unbiased root mean square error (Δ) can be computed from Ψ and δ according to $\Delta = (\Psi^2 - \delta^2)^{1/2}$. In addition we also used the Pearson linear correlation coefficient (r), to see how well estimated variables and measured variables are correlated. All statistical tests were performed in $\log_{10}$ space, considering that the chlorophyll concentration is approximately log-normally distributed (Campbell, 1995). Definitions for all symbols used in the paper are provided in **Table 1**.

2.3. Data

2.3.1. High Performance Liquid Chromatography (HPLC) Pigment Data

A total of 2,791 samples collected in the North Atlantic region and analyzed by High Performance Liquid Chromatography (HPLC) were used in this study (**Figure 1**), spanning 1995–2014. This dataset comprised of samples from: the Atlantic Meridional Transect (AMT) cruises 1-23 (Gibb et al., 2000; Barlow et al., 2002; Aiken et al., 2009; Brewin et al., 2010; Airs and Martinez-Vicente, 2014a,b,c; Brewin et al., 2015); the GeP&CO program (Dandonneau et al., 2004); the North Atlantic bloom experiment (Werdell et al., 2003; Westberry et al., 2010); the eastern Atlantic Ocean (Brotas et al., 2013); the North Atlantic, collected by the Bedford Institute of Oceanography (Sathyendranath et al., 2001; Devred et al., 2006); the Western Channel Observatory in the English Channel (Station L4 and E1; Smyth et al., 2010); a series of UK NERC-funded research cruises (D261, D262, D264, D325, JC011, JC037, and JCR656) in the North Atlantic and North Sea (Tilstone et al., 2015); and from the NASA bio-Optical Marine Algorithm Dataset (NOMAD Version 2.0 ALPHA, Werdell and Bailey, 2005), following the removal of any AMT data so as to avoid duplication. Details of HPLC methods used can be found in the aforementioned references.

Only samples collected within the top 20 m of the water column (or within the 1st optical depth as in the case of the NASA NOMAD dataset) were used [i.e., within the surface mixed-layer depth (rarely <20 m; de Boyer Montégut et al., 2004)]. To control the quality of the pigment data, we used only HPLC data for which the total chlorophyll concentration was greater than 0.001 mg m^{-3} (Uitz et al., 2006), and the difference between the total chlorophyll concentration and the total accessory pigments was less than 30% of the total pigment concentration (Trees et al., 2000; Aiken et al., 2009; Brewin et al., 2015).

2.3.1.1. Size-fractionated chlorophyll estimates from HPLC

The fractions of total chlorophyll for the three phytoplankton size classes (F_p, F_n, and F_m, for pico-, nano-, and microplankton, respectively) were estimated following the methods of Brewin et al. (2015), adapted from Vidussi et al. (2001), Uitz et al. (2006), Brewin et al. (2010), and Devred et al. (2011). Note, whenever we refer to microplankton, nanoplankton and picoplankton, we are referring to phytoplankton. First, the total chlorophyll concentration (C) was estimated from the weighted sum of the seven diagnostic pigments, hereafter denoted C_w, according to

$$C_w = \sum_{i=1}^{7} W_i P_i, \quad (3)$$

where, the weights are denoted [**W**], and the diagnostic pigments [**P**] = {fucoxanthin; peridinin; 19′-hexanoyloxyfucoxanthin; 19′-butanoyloxyfucoxanthin; alloxanthin; total chlorophyll-b; zeaxanthin}. We computed the weights [**W**] using multi-linear regression on the 2,791 samples. Retrieved values for the weights compare reasonably to values derived globally (**Table 2**), and total chlorophyll (C) and total chlorophyll estimated from Equation (3) (C_w) were in good agreement (r = 0.99, Ψ = 0.10). Having derived C_w, the fractions of chlorophyll in each size class relative to the total chlorophyll concentration were estimated.

Following Brewin et al. (2015), the fraction of picoplankton chlorophyll concentration (F_p) was computed according to

$$F_p = \begin{cases} \frac{(-12.5C+1)W_3P_3}{C_w} + \frac{\sum_{i=6}^{7} W_iP_i}{C_w} & \text{if } C \le 0.08 \text{ mg m}^{-3} \\ \frac{\sum_{i=6}^{7} W_iP_i}{C_w} & \text{if } C > 0.08 \text{ mg m}^{-3}. \end{cases} \quad (4)$$

The fraction of nanoplankton chlorophyll concentration (F_n) was estimated by first apportioning part of the fucoxanthin pigment (P_1) to the nanoplankton pool, as conducted by Devred et al. (2011), such that

$$P_{1,n} = 10^{\{q_1 \log_{10}(P_3) + q_2 \log_{10}(P_4)\}}, \quad (5)$$

where P_3 and P_4 refer to 19′-hexanoyloxyfucoxanthin and 19′-butanoyloxyfucoxanthin. This is to account for the fact that fucoxanthin is a precursor to 19′-hexanoyloxyfucoxanthin and 19′-butanoyloxyfucoxanthin (Devred et al., 2011). We recomputed these coefficients (q_1 and q_2) using the 2,791 HPLC samples, and arrived at values of $q_1 = 0.14$ and $q_2 = 1.35$. For any sample where $P_{1,n}$ was higher than P_1, then $P_{1,n}$ was set to equal P_1. Following Brewin et al. (2015), the fraction of

TABLE 1 | Symbols and definitions.

Symbol	Definition	Units
C	Total chlorophyll concentration	$mg\,m^{-3}$
C_w	Total chlorophyll concentration estimated from the seven diagnostic pigments (Equation 3)	$mg\,m^{-3}$
C_p	Chlorophyll concentration for picophytoplankton (cells $< 2\,\mu$m)	$mg\,m^{-3}$
$C_{p,n}$	Chlorophyll concentration for combined nano-picophytoplankton (cells $< 20\,\mu$m)	$mg\,m^{-3}$
C_n	Chlorophyll concentration for nanophytoplankton (cells $2 - 20\,\mu$m)	$mg\,m^{-3}$
C_m	Chlorophyll concentration for microphytoplankton (cells $> 20\,\mu$m)	$mg\,m^{-3}$
C_{diat}	Chlorophyll concentration for diatoms	$mg\,m^{-3}$
C_{dino}	Chlorophyll concentration for dinoflagellates	$mg\,m^{-3}$
$C^m_{p,n}$	Asymptotic maximum value of $C_{p,n}$ (cells $<20\,\mu$m)	$mg\,m^{-3}$
C^m_p	Asymptotic maximum value of C_p (cells $<2\,\mu$m)	$mg\,m^{-3}$
C_i^{SST}	Chlorophyll concentration for group i (where $i = p, n, m$, *diat* and *dino*) estimated using the SST dependent paramaterizations (Equations 10–16)	$mg\,m^{-3}$
$D_{p,n}$	Fraction of total chlorophyll in combined nano-picophytoplankton (cells $<20\,\mu$m) as total chlorophyll tends to zero	Dimensionless
D_p	Fraction of total chlorophyll in picophytoplankton (cells $<2\,\mu$m) as total chlorophyll tends to zero	Dimensionless
F_p	Fraction of total chlorophyll for picophytoplankton (cells $<2\,\mu$m)	Dimensionless
$F_{p,n}$	Fraction of total chlorophyll for combined nano-picophytoplankton (cells $< 20\,\mu$m)	Dimensionless
F_n	Fraction of total chlorophyll for nanophytoplankton (cells $2 - 20\,\mu$m)	Dimensionless
F_m	Fraction of total chlorophyll for microphytoplankton (cells $>20\,\mu$m)	Dimensionless
F_{diat}	Fraction of total chlorophyll for diatoms	Dimensionless
F_{dino}	Fraction of total chlorophyll for dinoflagellates	Dimensionless
G_1	Parameter of Equation (12) controlling lower and/or upper bound in $C^m_{p,n}$	$mg\,m^{-3}$
G_2	Parameter of Equation (12) controlling slope of change in $C^m_{p,n}$ with SST	$^\circ C^{-1}$
G_3	Parameter of Equation (12) controlling the SST mid-point of G_2	$^\circ C$
G_4	Parameter of Equation (12) controlling lower and/or upper bound in $C^m_{p,n}$	$mg\,m^{-3}$
H_1	Parameter of Equation (13) controlling lower and/or upper bound in C^m_p	$mg\,m^{-3}$
H_2	Parameter of Equation (13) controlling slope of change in C^m_p with SST	$^\circ C^{-1}$
H_3	Parameter of Equation (13) controlling the SST mid-point of H_2	$^\circ C$
H_4	Parameter of Equation (13) controlling lower and/or upper bound in C^m_p	$mg\,m^{-3}$
J_1	Parameter of Equation (14) controlling lower and/or upper bound in $D_{p,n}$	Dimensionless
J_2	Parameter of Equation (14) controlling slope of change in $D_{p,n}$ with SST	$^\circ C^{-1}$
J_3	Parameter of Equation (14) controlling the SST mid-point of J_2	$^\circ C$
J_4	Parameter of Equation (14) controlling lower and/or upper bound in $D_{p,n}$	Dimensionless
K_1	Parameter of Equation (15) controlling lower and/or upper bound in D_p	Dimensionless
K_2	Parameter of Equation (15) controlling slope of change in D_p with SST	$^\circ C^{-1}$
K_3	Parameter of Equation (15) controlling the SST mid-point of K_2	$^\circ C$
K_4	Parameter of Equation (15) controlling lower and/or upper bound in D_p	Dimensionless
P_i	Diagnostic pigments (where $i = 1$ to 7) for: fucoxanthin (1), peridinin (2), 19′-hexanoyloxyfucoxanthin (3), 19′-butanoyloxyfucoxanthin (4), alloxanthin (5), total chlorophyll-b (6), and zeaxanthin (7)	$mg\,m^{-3}$
$P_{1,n}$	Diagnostic pigment fucoxanthin in nanophytoplankton	$mg\,m^{-3}$
$q_{1\rightarrow2}$	Empirical coefficients used to compute $P_{1,n}$ from P_3 and P_4 (Equation 5)	Dimensionless
r	Pearson correlation coefficient	Dimensionless
SST	Sea surface temperature	$^\circ C$
T_i	Membership for each Optical Water Type (OWT)	Dimensionless
W_i	Weights in Equation (3) (where $i = 1$ to 7) for: fucoxanthin (1), peridinin (2), 19′-hexanoyloxyfucoxanthin (3), 19′-butanoyloxyfucoxanthin (4), alloxanthin (5), total chlorophyll-b (6), and zeaxanthin (7)	Dimensionless
α	Parameter of Equation (16) controlling slope of change in C_{dino}/C_m with SST	$^\circ C^{-1}$
β	Parameter of Equation (16) controlling the SST mid-point of α	$^\circ C$
δ	Bias between $\log_{10}$-transformed concentrations from estimated and measured data	Dimensionless
Δ	Unbiased root mean square error between $\log_{10}$-transformed concentrations from estimated and measured data	Dimensionless
Ψ	Root mean square error between $\log_{10}$-transformed concentrations from estimated and measured data	Dimensionless

TABLE 2 | Key taxonomic groups of phytoplankton, their typical size class, their category in the ERSEM model and their diagnostic pigment.

Key taxonomic groups	Typical size class[&]	ERSEM Group[#]	Pigment	[**W**]		
				This study (N. Atlantic)[$]	Brewin et al. (2015) (Global)[$]	Uitz et al. (2006) (Global)
Diatoms	Micro[a]	Diatoms	Fucoxanthin[e] (P_1)	1.65 (±0.01)	1.51 (±0.01)	1.41
Dinoflagellates	Micro	Dinoflagellates[d]	Peridinin (P_2)	1.04 (±0.03)	1.35 (±0.02)	1.41
Prymnesiophytes	Nano[b]	Nano	19′-hexanoyloxyfucoxanthin[2] (P_3)	0.78 (±0.01)	0.95 (±0.01)	1.27
Pelagophytes	Nano	Nano	19′-butanoyloxyfucoxanthin (P_4)	1.19 (±0.03)	0.85 (±0.02)	0.35
Cryptophytes	Nano	Nano	Alloxanthin (P_5)	3.14 (±0.04)	2.71 (±0.05)	0.60
Chlorophytes, Prochlorophytes	Pico[c]	Pico	Total Chlorophyll-b* (P_6)	1.38 (±0.02)	1.27 (±0.01)	1.01
Cyanobacteria, Prochlorophytes	Pico	Pico	Zeaxanthin (P_7)	1.02 (±0.01)	0.93 (±0.00)	0.86

*The table also shows a comparison of the weights ([**W**]) computed for Equation (3) using the 2791 HPLC data samples collected in this study, with weights derived from two other studies of the global ocean.*

[*] Total Chlorophyll-b refers to the sum of Chlorophyll-b and divinyl chlorophyll-b.

[&] Micro refers to cell cells >20 μm, Nano cells 2–20μm and Pico cells <2 μm in size.

[$] Bracketed values refer to the standards deviations for each coefficient.

[#] Phytoplankton state variables in ERSEM model.

[a] Diatoms can be found in the nano size class.

[b] Prymnesiophytes and 19′-hexanoyloxyfucoxanthin pigment can be found in the pico size class.

[c] Some chlorophytes can be found in the nanoplankton size class (Latasa et al., 2004).

[d] Also named microplankton in ERSEM.

[e] Fucoxanthin can be found in the nano size class.

nanoplankton chlorophyll concentration (F_n) was then estimated according to

$$F_n = \begin{cases} \frac{12.5CW_3P_3}{C_w} + \frac{\sum_{i=4}^{5} W_iP_i + W_1P_{1,n}}{C_w} & \text{if } C \leq 0.08 \text{ mg m}^{-3} \\ \frac{\sum_{i=3}^{5} W_iP_i + W_1P_{1,n}}{C_w} & \text{if } C > 0.08 \text{ mg m}^{-3}. \end{cases} \quad (6)$$

Finally, following Devred et al. (2011) and Brewin et al. (2015), the fraction of microplankton chlorophyll concentration (F_m) was estimated as

$$F_m = \frac{\sum_{i=1}^{2} W_iP_i - W_1P_{1,n}}{C_w}. \quad (7)$$

Note that F_m can also be computed by simply subtracting F_n and F_p from one. The fractions of chlorophyll in each size class were then multiplied by the corresponding HPLC-derived total chlorophyll concentration (C) to derive the size-specific chlorophyll concentrations for each sample (C_p, $C_{p,n}$, C_n, and C_p, where the subscripts "p" refers to pico-, "n" nano- and "m" microphytoplankton, and the subscript "p, n" refers to combined pico and nanophytoplankton).

2.3.1.2. Partitioning the fraction of microphytoplankton chlorophyll into fractions of diatoms and dinoflagellates

The fraction of microphytoplankton chlorophyll concentration (F_m) is estimated from two diagnostic pigments, fucoxanthin in microphytoplankton ($P_{1,m}$) and peridinin (P_2). It is generally assumed that fucoxanthin in microphytoplankton is the primary pigment for diatoms (Stauber and Jeffrey, 1988) and peridinin for dinoflagellates, as the majority of photosynthetic dinoflagellates contain a chloroplast with peridinin as the major carotenoid (see Table 1 and Zapata et al., 2012). Following Hirata et al. (2011), this assumption was used to partition microphytoplankton chlorophyll into the concentrations of the two groups.

The fraction of microplankton diatoms to total chlorophyll (F_{diat}) and the fraction of microplankton dinoflagellates to total chlorophyll (F_{dino}) were computed as

$$F_{diat} = \frac{W_1P_1 - W_1P_{1,n}}{C_w}, \quad (8)$$

and

$$F_{dino} = \frac{W_2P_2}{C_w}, \quad (9)$$

respectively. The chlorophyll concentrations for diatoms and dinoflagellates (C_{diat} and C_{dino}) were then obtained by multiplying the fractions by the corresponding HPLC-derived total chlorophyll concentration (C).

2.3.2. Size-Fractionated Filtration (SFF) Data

A total of 263 size-fractionated fluorometric chlorophyll (SFF) measurements collected previously in the North Atlantic region were also used in this study (**Figure 1**), spanning 1996–2015. This comprised of samples from: the Atlantic Meridional Transect cruises 2–23 (see Marañón et al., 2001; Serret et al., 2001; Robinson et al., 2002; Brewin et al., 2014a,b; Tilstone et al., 2017, for details); the Western Channel Observatory in the English Channel (Station L4 and E1; see Barnes et al., 2014, for details); and the NERC shelf seas biogeochemistry programme.

In all cases, ∼200–300 ml samples were sequentially filtered through 20, 2, and 0.2 μm polycarbonate filters. Following filtration, pigments were extracted by storing the filters in 90% acetone at −20°C for between 10 and 24 h. Samples were then analyzed using a Turner Design Fluorometer, pre- and post-calibrated using pure chlorophyll-a in 90% acetone as a standard. The total chlorophyll concentration was taken as the sum of the size fractions for each sample. The concentration of chlorophyll passing through the 2 μm filter was designated C_p (picoplankton

chlorophyll), chlorophyll retained on the 20 μm filter designated C_m (microplankton chlorophyll) and the chlorophyll retained on the 2 μm filter, having passed through the 20 μm filter, designated C_n (nanoplankton chlorophyll).

2.4. Merging of *in situ* Datasets

Systematic biases in size-fractionated chlorophyll estimated from HPLC pigments and from SFF have been observed in the Atlantic Ocean (Brewin et al., 2014a), with implications for models that estimate size-fractionated chlorophyll as a function of total chlorophyll (Brewin et al., 2014b) and models that estimate size-fractionated primary production (Brewin et al., 2017). Therefore, care needs to be taken when combining these two datasets. **Figure 2** shows a comparison of 31 concurrent and co-located data points of total chlorophyll (**Figure 2A**), picoplankton chlorophyll (**Figure 2B**), nanoplankton chlorophyll (**Figure 2C**) and microplankton chlorophyll (**Figure 2D**), from the HPLC and SFF dataset used here.

Despite there being biases in size-fractionated chlorophyll consistent with those observed by Brewin et al. (2014a) (**Figure 2**), these biases are notably smaller (e.g., for picoplankton chlorophyll $\delta = -0.07$ compared with $\delta = -0.27$ in see their Figure 3 Brewin et al. (2014a), and for nanoplankton $\delta = 0.15$ compared with $\delta = 0.22$), suggesting for surface waters in the North Atlantic, there is reasonable agreement between the two methods, at least for the datasets used here. Given the good agreement in **Figure 2**, the two datasets were combined into a single dataset, providing 3,054 measurements of size-fractionated chlorophyll (2,791 HPLC and 263 SFF). **Figure 3** shows a schematic diagram of how the datasets were combined and subsequently used for model parameterization and validation.

For each sample, SST data were extracted by matching each *in situ* sample in time (daily temporal match-up) and space (closest latitude and longitude) with daily, 1/4° resolution Optimal Interpolation Sea Surface Temperature (OISST) data (Version 2.0; Reynolds et al., 2007) acquired from the NOAA website (http://www.esrl.noaa.gov/psd/data/gridded/data.noaa.oisst.v2.highres.html).

2.5. Partitioning Into Parameterization and Validation Datasets

The merged dataset was matched to daily, level 3 (4 km sinusoidal projected) satellite chlorophyll and optical water type (OWT) data, from version 3.0 of the Ocean Colour Climate Change Initiative (OC-CCI, a merged MERIS, MODIS-Aqua, SeaWiFS and VIIRS product available at http://www.oceancolour.org/), between 1997 and 2015. Each *in situ* sample was matched with a single satellite pixel in time (daily match-up) and space (closest pixel with a distance <4 km away). Of the 3,054 samples, there

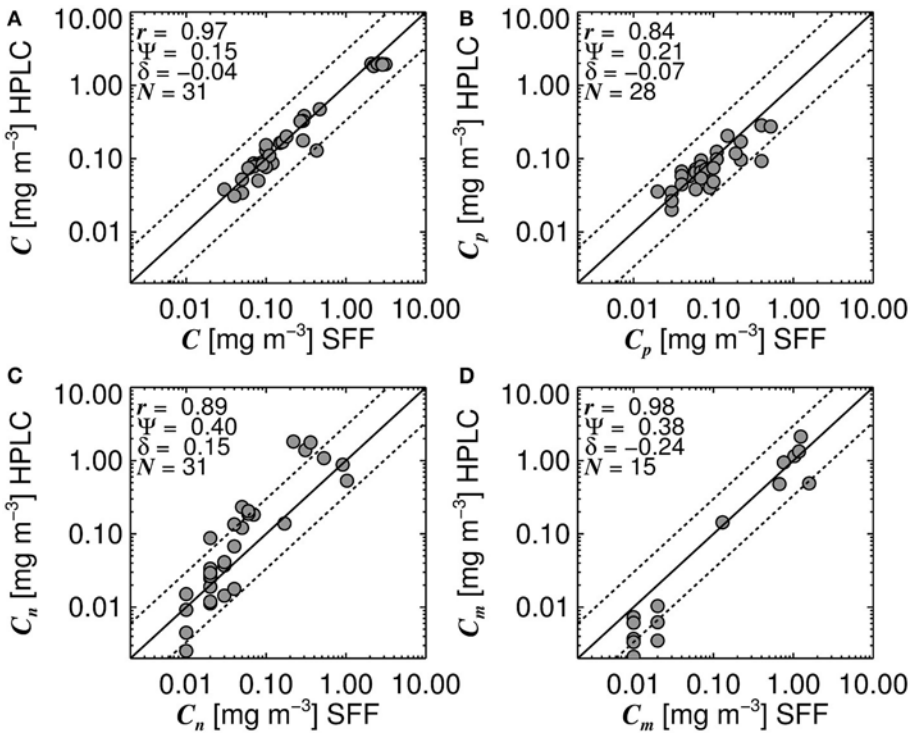


FIGURE 2 | Concurrent and co-located size-fractionated chlorophyll estimated from High Performance Liquid Chromatography (HPLC) and size-fractionated filtration (SFF) for surface waters in the North Atlantic region. (A) shows a comparison of total chlorophyll (*C*), **(B)** picoplankton chlorophyll (C_P), **(C)** nanoplankton chlorophyll (C_n), and **(D)** microplankton chlorophyll (C_m). Black line represents the 1:1 line and dotted lines represent the 1:1 line ±30% $\log_{10}$ chlorophyll. *N* refers to the number of samples used to compute statistics, *r* refers to the Pearson linear correlation coefficient, Ψ the root mean square error (Equation 1) and δ the bias (Equation 2).

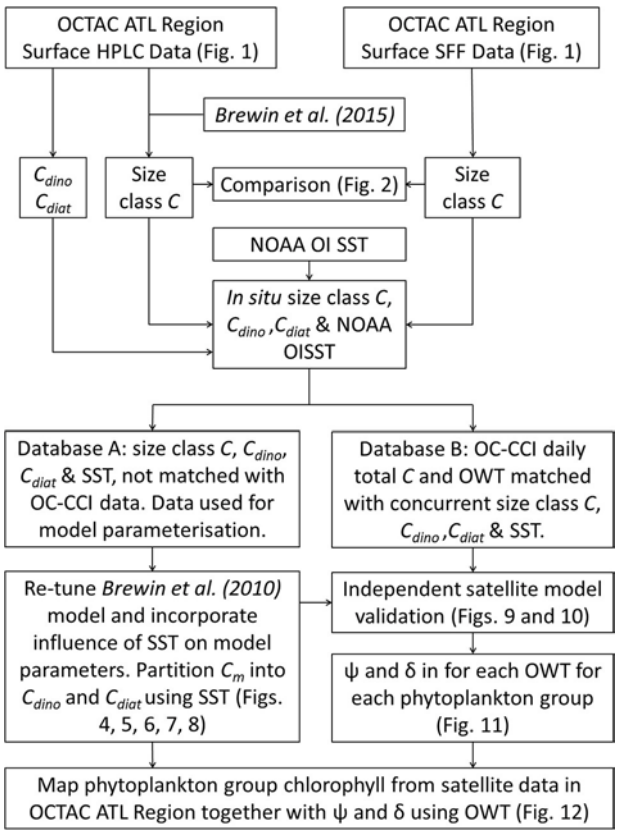


FIGURE 3 | A flow chart of the processing techniques. Data collected in the OCTAC ATL region [both High Performance Liquid Chromatography (HPLC) and size-fractionated filtration (SFF)] were partitioned into two databases [parameterization (Database A) and satellite validation (Database B)], and used to re-tune, adapt and validate the model of Brewin et al. (2010), compute the root mean square error (Ψ) and bias (δ) for each optical water type (OWT), and map phytoplankton group products and associated errors using ocean-color data.

were 815 corresponding satellite chlorophyll and optical water type (OWT) data. These 815 measurements were set aside and used for independent validation of the satellite model and for characterizing per-pixel error, leaving 2,239 measurements that were used for model development (parameterization). **Figure 3** shows a schematic diagram of how the data were partitioned into the parameterization and validation dataset.

The OWT data provided in version 3.0 of the OC-CCI dataset contains the per-pixel membership of 14 different optical classes, ranging from oligotrophic (e.g., OWT 1) to very turbid (OWT 14) waters. Building on the work of Moore et al. (2001, 2009, 2012), this new set of optical classes were constructed for use with OC-CCI remote sensing reflectance (R_{rs}) spectra (Jackson and Sathyendranath, 2015). These classes were trained using R_{rs} spectra from satellite data, rather than using a database of *in situ* observations, as conducted in Moore et al. (2009), and the number of optical water classes were increased to 14, to better cover the range of R_{rs} spectra observed in the global oceans, particularly the oligotrophic gyres. For further details of the training and production of the 14 OWT the reader is referred to Jackson and Sathyendranath (2015).

2.6. Satellite Model of Phytoplankton Groups

2.6.1. Three-Component Model of Brewin et al. (2010)

As a starting point, we used the three-component model of Brewin et al. (2010) to estimate the chlorophyll concentrations in three phytoplankton size classes [pico- (<2 μm), nano- (2–20 μm), and micro-phytoplankton (>20 μm)] as a function of total chlorophyll in the study region (**Figure 1**). This approach has been successfully tuned to the global ocean (Brewin et al., 2015; Ward, 2015) as well as different oceanic regions, including: the Atlantic Ocean (North and South; Brewin et al., 2010, 2014b; Tilstone et al., 2014); the North East Atlantic (Brotas et al., 2013); the Indian Ocean (Brewin et al., 2012b); the Western Iberian coastline (Brito et al., 2015); the Mediterranean Sea (Sammartino et al., 2015); and the South China Sea (Lin et al., 2014). Estimating size-fractionated chlorophyll from satellite data (using satellite total chlorophyll as input to the three-component model) has been tested extensively with *in situ* data in different oceanic regions (Brewin et al., 2010, 2012b; Lin et al., 2014; Brewin et al., 2015).

The three-component model is based on two exponential functions (Sathyendranath et al., 2001), where the chlorophyll concentration of picoplankton (C_p, cells <2 μm) and combined pico- and nanoplankton ($C_{p,n}$, cells <20 μm) are obtained from

$$C_{p,n} = C^m_{p,n}[1 - \exp(-\frac{D_{p,n}}{C^m_{p,n}}C)], \quad (10)$$

and

$$C_p = C^m_p[1 - \exp(-\frac{D_p}{C^m_p}C)]. \quad (11)$$

The parameters $D_{p,n}$ and D_p determine the fraction of total chlorophyll in the two size classes (<20 μm and <2 μm, respectively) as total chlorophyll tends to zero, and $C^m_{p,n}$ and C^m_p are the asymptotic maximum values for the two size classes (<20 μm and <2 μm respectively). The chlorophyll concentration of nano-phytoplankton (C_n) and micro-phytoplankton (C_m) are simply calculated as $C_n = C_{p,n} - C_p$ and $C_m = C - C_{p,n}$.

A single set of model parameters was first derived by fitting (Equations 10 and 11) using a standard, nonlinear least-squared fitting procedure (Levenberg-Marquardt, IDL Routine MPFITFUN, Moré, 1978; Markwardt, 2008) with relative weighting (Brewin et al., 2011a). The parameters $D_{p,n}$ and D_p were constrained to be less than or equal to one, since size-fractionated chlorophyll cannot exceed total chlorophyll. We used the method of bootstrapping (Efron, 1979; Brewin et al., 2015) to compute a parameter distribution, and from the resulting parameter distribution, median values and 95% confidence intervals were computed (see **Table 3**). The parameters $D_{p,n}$ and D_p were found to be significantly different from the global parameters derived in Brewin et al. (2015) (see **Table 3**). The model was found to capture the trends in the fractions (F_p, F_n, $F_{p,n}$, and F_m) and absolute concentrations

TABLE 3 | Parameter values for Equations 10 and 11 compared with global parameters derived in Brewin et al. (2015).

Study	Parameters for equations 10 and 11				Location	$N^{\#}$
	$C_{p,n}^{m}$ *	C_{p}^{m} *	$D_{p,n}$	D_p		
Brewin et al. (2015)$^{\$}$	0.77 (0.72↔0.84)	0.13 (0.12↔0.14)	0.94 (0.93↔0.95)	0.80 (0.78↔0.82)	Global	5841
This study$^{\$}$	0.82 (0.76↔0.88)	0.13 (0.12↔0.13)	0.87 (0.86↔0.89)	0.73 (0.71↔0.76)	N Atlantic	2239
This study$^{\$}$ (<15°C)	1.83 (1.47↔2.44)	0.31 (0.24↔0.47)	0.60 (0.58↔0.63)	0.26 (0.23↔0.30)	N Atlantic	1017
This study$^{\$}$ (≥15°C)	0.86 (0.79↔0.96)	0.13 (0.12↔0.14)	0.93 (0.91↔0.94)	0.74 (0.72↔0.77)	N Atlantic	1222

$^{\$}$*Model parameters are computed as the median of the bootstrap parameter distribution and bracket parameter values refer to the 2.5 and 97.5% confidence intervals on the distribution.*
$^{\#}$*N = Number of samples used for model parameterization*
**Denotes units in mg m*$^{-3}$.

(C_p, C_n, $C_{p,n}$, and C_m) of the size classes as a function of total chlorophyll for the North Atlantic parameterization dataset (**Figure 4**).

2.6.2. Modification of Three-Component Model Using SST

Brewin et al. (2015) and Ward (2015) have investigated the influence of light availability and SST respectively on the parameterization of the three-component model. In the North Atlantic, seasonal variations in SST and the average light in the mixed-layer are highly correlated (**Figure 1**). Therefore, considering: (i) that there is, regionally, a covariation of SST with the average light in the mixed-layer (**Figure 1**); (ii) that three inputs are required to compute the average light in the mixed-layer (photosynthetically-active radiation, diffuse attenuation and mixed-layer depth), one of which is not amenable from remote-sensing (mixed-layer depth); and (iii) that the maturity (operational use) and accuracy of SST retrievals is very high (Merchant et al., 2014), we chose to investigate the influence of SST on model parameters in the study area, similar to the study of Ward (2015) for a global dataset.

Figure 4 illustrates the general inverse correlation between SST and total chlorophyll ($r = -0.67$ for SST and $\log_{10}(C)$), highlighting that higher fractions of smaller cells (lower fractions of large cells) are typically associated with higher SST. To investigate if SST has any influence on the parameters of the three-component model, we partitioned the parameterization data into lower temperature waters ($< 15^{\circ}$C) and higher temperature waters ($\geq 15^{\circ}$C), and fitted the model separately to the two datasets of a roughly equal number (>1,000, see **Table 3**). We observed significantly different model parameters for high and low temperature waters (see **Table 3** and **Figure 4**), suggesting a relationship between SST and model parameters. We then sorted the dataset according to SST, and conducted a running fit of the three-component model (Equations 10 and 11) as a function of SST with a bin size of 600 samples [chosen to ensure each fit had reasonable representation of observations over the entire trophic range (low to high chlorophyll)]. We used the method of bootstrapping (100 iterations) and derived median values and 95% confidence intervals on each parameter distribution (**Figure 5**).

Significant relationships between all model parameters ($C_{p,n}^{m}$, C_{p}^{m}, $D_{p,n}$, and D_p) and SST were observed (**Figure 5**). The relationship between SST and model parameters could be represented using a logistic function, such that $C_{p,n}^{m}$ and C_{p}^{m} may be expressed as

$$C_{p,n}^{m} = 1 - \{\frac{G_1}{1 + \exp[-G_2(\mathrm{SST} - G_3)]} + G_4\}, \quad (12)$$

and

$$C_{p}^{m} = 1 - \{\frac{H_1}{1 + \exp[-H_2(\mathrm{SST} - H_3)]} + H_4\}, \quad (13)$$

where G_1 and G_4 control the upper and lower bounds of $C_{p,n}^{m}$, G_2 represents the slope of change in $C_{p,n}^{m}$ with SST, and G_3 is the SST mid-point of the slope between $C_{p,n}^{m}$ and SST. For C_{p}^{m}, H_i, where $i = 1$–4, is analogous to G_i for $C_{p,n}^{m}$. The parameter $D_{p,n}$ and D_p were expressed as

$$D_{p,n} = \frac{J_1}{1 + \exp[-J_2(\mathrm{SST} - J_3)]} + J_4, \quad (14)$$

and

$$D_{p} = \frac{K_1}{1 + \exp[-K_2(\mathrm{SST} - K_3)]} + K_4, \quad (15)$$

where J_1 and J_4 control the upper and lower bounds of $D_{p,n}$, J_2 represents the slope of change in $D_{p,n}$ with SST, and J_3 is the SST mid-point of the slope between $D_{p,n}$ and SST. For D_p, K_i is analogous to J_i for $D_{p,n}$. The parameters for Equations (12)–(15) were fitted using a nonlinear least-squared fitting procedure (Levenberg-Marquardt) with bootstrapping, and parameter values are provided in **Table 4**. The equations are seen to capture the relationships between parameters and SST accurately (**Figure 5** and **Table 4**).

Figure 6 shows simulations of size-fractionated chlorophyll as a function of total chlorophyll for different SST, when incorporating (Equations 12–15) into the three-component model (Equations 10 and 11). In general, the performance for all size classes improved when using the SST-dependent parameterization, when compared with that using a single set of parameters (**Figure 7**), with a significant improvement in the correlation coefficient for C_p (Z-test, $p < 0.05$). Whereas modeled $C_{p,n}$, C_n, and C_p reach static asymptotes at high concentrations when using a single set of parameters (see **Figure 7**, top-row, horizontal purple dashed lines), the SST-dependent parameterization does not, and captures the variability in the size-fractionated chlorophyll at these higher concentrations.

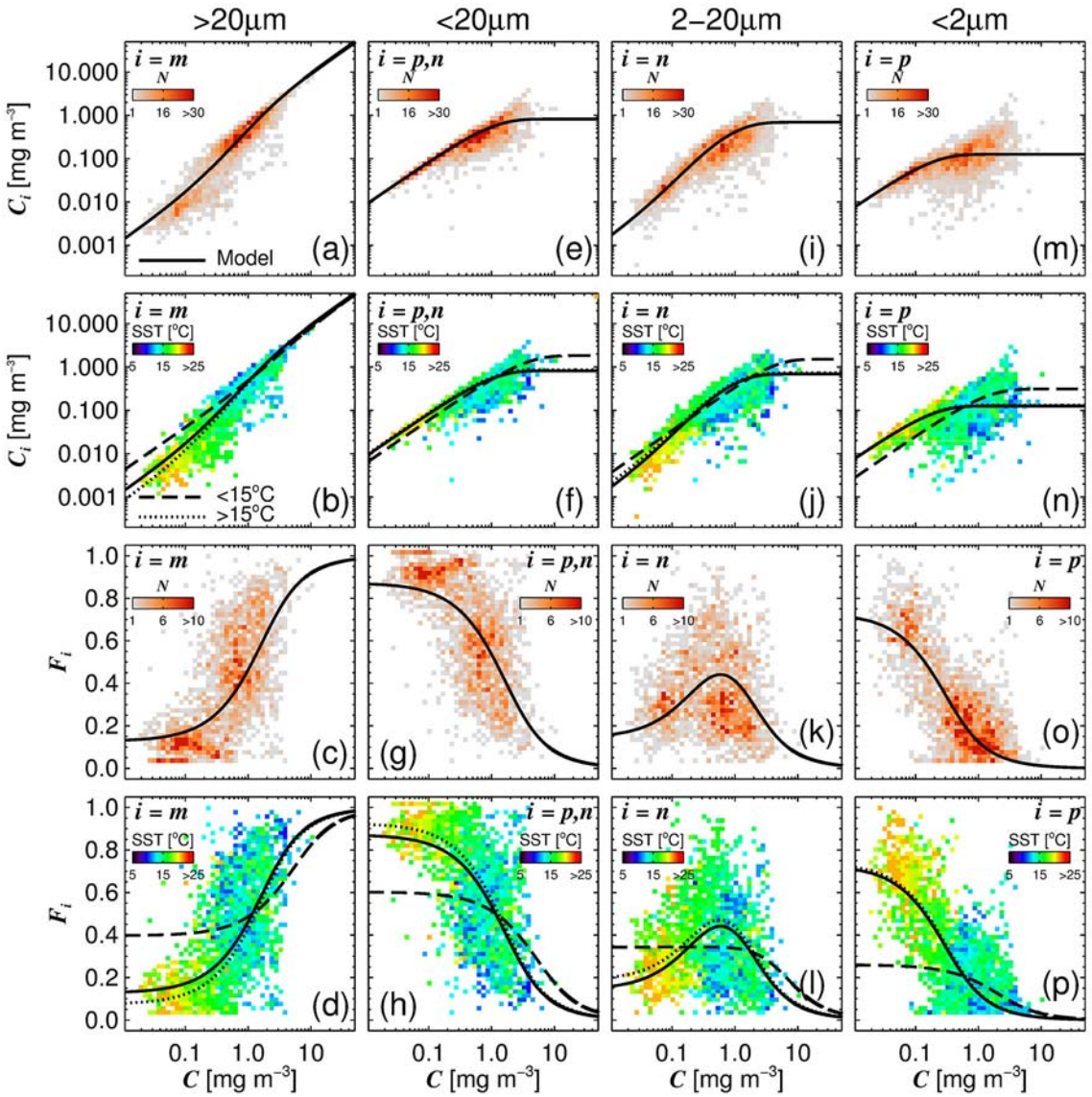


FIGURE 4 | The absolute chlorophyll concentrations [C_m (a,b), $C_{p,n}$ (e,f), C_n (i,j), and C_p (m,n)] and fractions [F_m (c,d), $F_{p,n}$ (g,h), F_n (k,l) and F_p (o,p)] in the parameterization dataset plotted as a function of total chlorophyll concentration (C), with the re-tuned (Brewin et al., 2010) model (parameters from Table 3), overlain. The top row **(a,e,i,m)** and middle-bottom row **(c,g,k,o)** show bivariate histogram plots with the shading indicating the number of observations (N). The bottom row **(d,h,l,p)** and middle-top row **(b,f,j,n)** show the same bivariate plots but the shading represents the median sea surface temperature (SST) of the data points that lie within the bins.

2.6.3. Partitioning of Microphytoplankton Chlorophyll Into Diatoms and Dinoflagellates

Considering diatoms are known to dominate the microphytoplankton community in the North Atlantic during the initiation of the spring bloom when SST is still relatively low and nutrient concentrations high (Ducklow and Harris, 1993; Sieracki et al., 1993; Savidge et al., 1995), and that dinoflagellates typically increase in late summer and early autumn (McQuatters-Gollop et al., 2007; Widdicombe et al., 2010) when SST is generally at its highest in the North Atlantic, we investigated the use of SST to partition microplankton chlorophyll (C_m) into diatoms (C_{diat}) and dinoflagellates (C_{dino}). **Figure 8A** shows a significant relationship between ratio of C_{dino} to C_m and SST ($r = 0.28$, $p < 0.001$), with the ratio increasing with increasing SST. We modeled this relationship by fitting a logistic function to the data (**Figure 8A**), such that

$$\frac{C_{dino}}{C_m} = \frac{1}{1 + \exp[-\alpha(\mathrm{SST} - \beta)]}, \tag{16}$$

where $\alpha = 0.10$ (0.08↔0.13) and $\beta = 32.5$ (29.7↔36.1). **Figures 8B,C** show model estimates of C_{dino} (obtained by

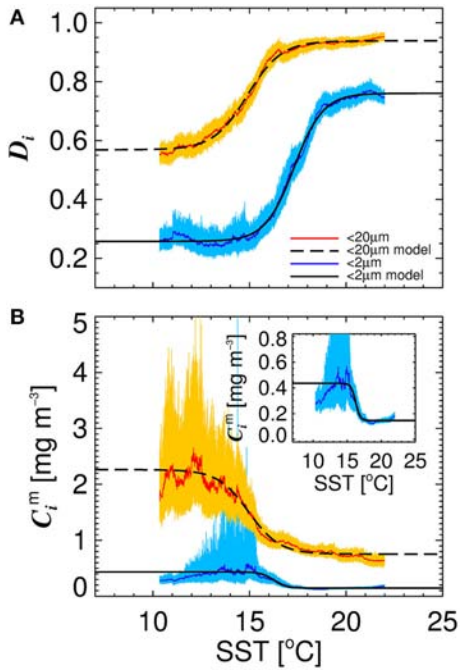


FIGURE 5 | The relationship between sea surface temperature (SST) and the parameters of the re-tuned (Brewin et al., 2010) model for the North Atlantic dataset. (A) Shows the relationship between $D_{p,n}$ and SST, and D_p and SST. **(B)** Shows the relationship between $C^m_{p,n}$ and SST, and C^m_p and SST. Solid color lines show median values on the bootstrap parameter distribution and lighter shades represent 95% confidence intervals. Black solid and dashed lines represent logistic models fitted between the parameters and SST shown in Equations (12–15), with the parameters provided in **Table 4**.

multiplying the modeled ratio (Equation 16) by C_m) plotted against measured C_{dino}, and estimates of C_{diat} (obtained as $C_m(1-(C_{dino}/C_m))$) against measured C_{diat}. In general, there is good agreement between the estimates and measurements, with higher correlations and lower root mean square errors for C_{diat} compared with C_{dino} (**Figures 8B,C**). Combining estimates of C_m using the three component model (Equations 10–15) with estimates of the ratio of C_{dino} to C_m (Equation 16), C_{dino} and C_{diat} can be estimated as a function of total chlorophyll (C) and SST.

2.7. Validation of the Satellite Model, Estimates of Per-Pixel Uncertainty and Application to Satellite Data

The satellite match-up dataset (not used for model parameterization) was used to validate the model by using satellite-derived total chlorophyll (OC-CCI) and SST (NOAA OISST) as inputs to Equations (10–15) and comparing the results with independent *in situ* chlorophyll concentrations for each phytoplankton group. In addition, the satellite match-ups were partitioned into 14 OWT by selecting the highest OWT membership for each sample. The root mean square error (Ψ, Equation 1) and bias (δ, Equation 2) in the satellite estimates were computed separately for each OWT and for each phytoplankton group.

We applied the model to a relatively cloud-free 8-day chlorophyll (OC-CCI) and SST (NOAA OISST) composite for the data between 17th and 24th June 2008, to illustrate its application to a satellite image. Uncertainties (Ψ and δ) in each pixel of the study area were computed by weighing the uncertainties in each OWT by their membership. For instance, Ψ at a given pixel for a hypothetical phytoplankton group would be computed as

$$\Psi = \frac{\sum_{i=1}^{14} \Psi_i T_i}{\sum_{i=1}^{14} T_i}, \tag{17}$$

where i represents each OWT and T represents the membership of each OWT.

3. RESULTS AND DISCUSSION

3.1. Satellite Validation

Considering the agreement between total satellite and *in situ* chlorophyll in the validation dataset ($r = 0.86$, $\Psi = 0.29$, $\delta = -0.01$), the satellite estimates of size-fractionated chlorophyll compare well with the independent *in situ* data (**Figure 9**, $r = 0.49$ to 0.86, and $\Psi = 0.30$ to 0.45), in agreement with previous studies (Brewin et al., 2010, 2012b; Lin et al., 2014; Brewin et al., 2015). Although the SST-dependent parameterization (C_i^{SST}) has a similar statistical performance compared with that obtained when using a single set of parameters, the SST-dependent parameterization is not constrained by static asymptotes for $C_{p,n}$, C_n, and C_p (**Figure 9** top-row, horizontal purple dashed lines) and captures better the variability in the size-fractionated chlorophyll at these higher concentrations. Correlation coefficients for picoplankton chlorophyll are higher for the SST-dependent parameterization (C_p^{SST}) when compared with the single set of parameters (C_p) in both the parameterization (**Figure 7**) and validation (**Figure 9**) datasets. This finding is consistent with results from Pan et al. (2013) who highlighted the benefits of including SST when estimating zeaxanthin (diagnostic pigment for picoplankton) from satellite data.

Satellite estimates of diatom and dinoflagellate chlorophyll also compare reasonably well with the independent *in situ* data (**Figure 10**). Satellite estimates of diatom chlorophyll have higher correlation coefficient (r) and lower error (Ψ) when compared with dinoflagellate chlorophyll estimates, suggesting better performance for this phytoplankton group. High errors in satellite estimates of dinoflagellate chlorophyll reflect how challenging it is to retrieve this phytoplankton group from space (Raitsos et al., 2008; Shang et al., 2014), though it is encouraging to observe significant correlations between the satellite and *in situ* dinoflagellate chlorophyll concentrations ($r > 0.64$, $p < 0.001$) in the validation dataset, especially when considering the lower range of chlorophyll variability in dinoflagellates (**Figure 10**).

3.2. Changes in Performance With Optical Water Types (OWT)

For each of the 14 OWT and for each of the phytoplankton groups, the root mean square error (Ψ), bias (δ) and number of observations (N) for match-ups in the validation dataset are

TABLE 4 | Parameter values for Equations (12) and (15).

Model parameter	Equation	Parameters for equations 12 and 15				$r^{\#}$	$p^{\&}$
$C^m_{p,n}$*	12	$G_1 = -1.51$ (−1.57↔−1.43)	$G_2 = -1.25$ (−1.41↔ −1.25)	G_3 =14.95 (14.87↔15.05)	$G_4 = 0.25$ (0.23↔0.26)	0.98	<0.001
C^m_p*	13	$H_1 = 0.29$ (0.28↔0.30)	$H_2 = 3.05$ (2.87↔3.26)	H_3 =16.24 (16.19↔16.29)	$H_4 = 0.56$ (0.55↔0.57)	0.91	<0.001
$D_{p,n}$	14	$J_1 = 0.370$ (0.367↔0.373)	$J_2 = 1.13$ (1.10↔1.16)	J_3 =14.89 (14.87↔14.91)	$J_4 = 0.569$ (0.566↔0.571)	1.00	<0.001
D_p	15	$K_1 = 0.503$ (0.501↔0.505)	$K_2 = 1.33$ (1.31↔1.37)	K_3 =17.31 (17.28↔17.32)	$K_4 = 0.258$ (0.256↔0.259)	1.00	<0.001

$ *Model parameters are computed as the median of the bootstrap parameter distribution and bracket parameter values refer to the 2.5 and 97.5% confidence intervals on the distribution.*
\# *Correlation coefficients (r) were computed using the median parameter values reported.*
& *p refers to the significance of each correlation (<0.001 is highly significant), computed using the correlation coefficient (r) and the number of samples (N), based on the probability that the correlation could have been produced by random data.*
**Denotes units in mg m^{-3}.*

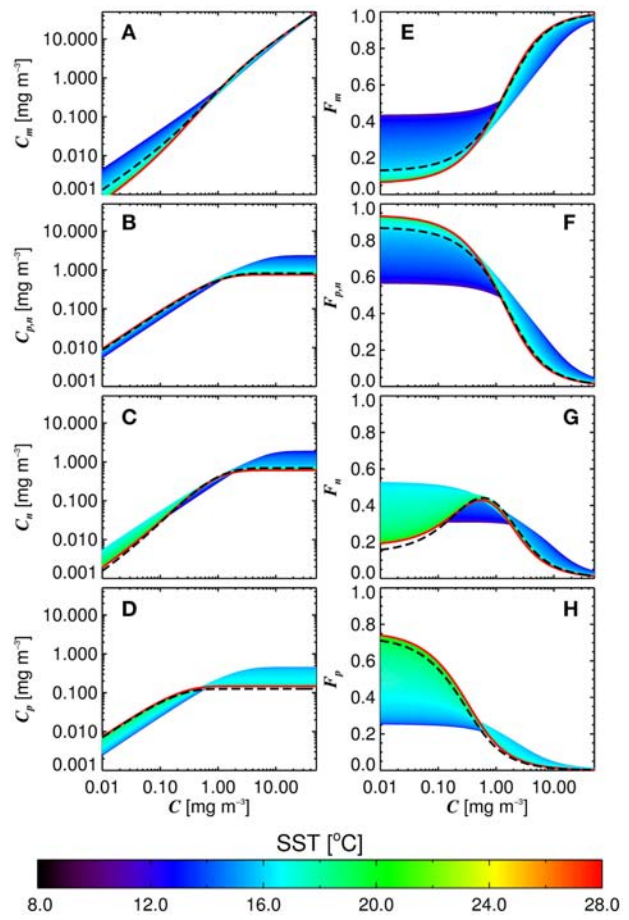


FIGURE 6 | Size-fractionated chlorophyll (A–D) and the fractions of total chlorophyll in each size class (E–H) plotted as a function of the total chlorophyll using the re-tuned (Brewin et al., 2010) model, and varying the parameters according to the sea surface temperature (SST) (Equations 12–15). Dashed black lines refer to the re-tuned model using a single set of parameters (**Table 3**).

provided in **Table 5**. The Ψ, δ, and N are also plotted in **Figure 11** for satellite estimates of total chlorophyll and chlorophyll for the four phytoplankton groups using the SST-dependent parameterization (Equations 12 to 15). The Ψ values in each OWT for total chlorophyll are consistent with those provided in version 3.0 of the OC-CCI dataset, based on a much larger global match-up dataset (∼14,500) (**Figure 11A**). The Ψ values for total chlorophyll increase from lower OWTs (characteristic of oligotrophic open-ocean waters) to higher OWTs (characteristic of more optically complex turbid coastal waters). A result that is also consistent with the original work of Moore et al. (2009), see their Table 2, and the theoretical limitations of using empirical ocean-color chlorophyll algorithms in optically-complex waters (IOCCG, 2000). Biases (δ) in total chlorophyll are generally quite low (**Figure 11B**), consistent with version 3.0 of the OC-CCI dataset, though do not always have the same sign, and are much higher for OWT14, probably due to very few match-ups ($N = 4$) in this class (**Figure 11B**).

Consistent with satellite estimates of total chlorophyll, there is a tendency for Ψ to increase from lower to higher OWTs for all the phytoplankton groups (**Table 5**, **Figures 11C,E,G,I**), particularly for smaller cells (pico and nano-plankton) and for dinoflagellates. This is likely due to: i) the satellite estimates of total chlorophyll, which are used as input to the phytoplankton group model, having larger errors at higher OWTs (**Figure 11A**); and ii) possible deviations in the relationships between the phytoplankton groups and total chlorophyll in optically complex waters, when compared with typical open-ocean conditions. With the exception of diatoms, there is a slight tendency for the models to overestimate chlorophyll for the phytoplankton groups at higher OWTs (e.g., 8–14), as indexed by a positive bias (**Table 5**, **Figures 11F,H,J**).

3.3. Application of the Model to a Satellite Image

Figure 12 illustrates the application of the phytoplankton group model (SST-dependent parameterization; Equations 10–15) to satellite chlorophyll (OC-CCI) and SST (NOAA OISST) composites for the period 17th to 24th June 2008. Satellite products used as inputs to the model – chlorophyll (**Figure 12A**), OWT membership (plotted by dominance (highest membership) in **Figure 12B**) and SST (**Figure 12C**)—highlight the different biogeochemical areas in the region, with oligotrophic waters to the south (high SST, low total chlorophyll, low OWT), more productive waters to the north (lower SST, higher chlorophyll and OWT), and very productive coastal waters (variable SST, high chlorophyll and OWT). **Figures 12D,G,J,M**, show estimates of chlorophyll for the four phytoplankton groups,

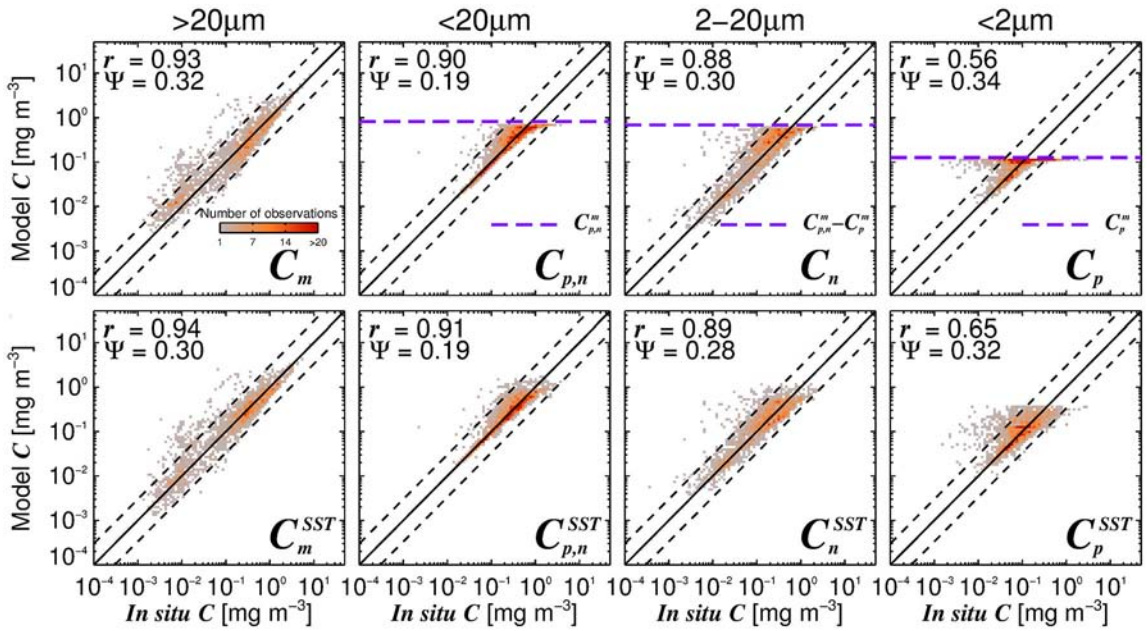


FIGURE 7 | The modeled size-fractionated chlorophyll plotted against *in situ* size-fractionated chlorophyll in the parameterization dataset, for the re-tuned (Brewin et al., 2010) model with a single set of parameters (top-row, Table 3) and using the SST-dependent parameterization (bottom row, Equations 12–15, Table 4). The superscript SST denotes the modeled chlorophyll concentrations using the SST-dependent parameterization. The correlation coefficient (*r*) and root-mean-square-error (Ψ) are also shown. Statistical tests were computed using the parameter values reported in **Tables 3**, **4**. The top panels also show the maximum attainable concentrations for the different size classes as purple dashed horizontal lines when using a single set of parameters (**Table 3**).

diatoms (C_{diat}), dinoflagellates (C_{dino}), nanoplankton (C_n) and picoplankton (C_p), respectively. Picoplankton (C_p) are the dominant group in the warm oligotrophic waters, nanoplankton (C_n) in intermediate (mesotrophic waters), and diatoms (C_{diat}) in the northern productive waters and coastal regions (eutrophic waters). Dinoflagellates rarely dominate (i.e., rarely have the highest chlorophyll of the four groups), but typically have higher concentrations in coastal regions.

In addition to the concentrations, per-pixel uncertainties (Ψ and δ) are plotted for each phytoplankton group (**Figure 12**), through application of Equation (17) on a per-pixel basis, using per-pixel OWT membership provided by the OC-CCI products and statistics from **Table 5**. In general, lower Ψ is observed in the oligotrophic waters to the south of the region, with Ψ increasing toward more productive waters. Dinoflagellates have the highest Ψ in these productive waters, reflecting higher uncertainty in deriving the concentrations of this phytoplankton group (see also **Figure 10**). Lower Ψ are seen for nano- and picoplankton, when compared with the larger size classes. Diatoms display a less variable Ψ throughout the entire region, when compared with the other three phytoplankton groups.

Biases (δ) are close to zero for all phytoplankton groups in the warm oligotrophic waters (**Figure 12**), with positive biases seen for dinoflagellates, nanoplankton and picoplankton in the more productive waters, implying a slight overestimation in chlorophyll by the satellite model in these waters. These biases can be caused by two reasons: (i) biases in model input (total chlorophyll); and (ii) biases in model parameters used for partitioning total chlorophyll into the phytoplankton groups. There were no major biases (with the exception of OWT14) in total chlorophyll (model input) in the validation dataset (**Figure 11B**). Nonetheless, it is likely that the use of alternative input chlorophyll algorithms (e.g., a semi-analytical algorithm) will impact these biases. The positive biases seen for dinoflagellates, nanoplankton and picoplankton in the more productive waters are likely caused by biases in model parameters at higher OWTs. In the future, with a larger database, modifications to model parameters according to OWT could be feasible, and would likely reduce observed biases.

As well as varying within the region as illustrated in **Figure 12**, temporal variations in chlorophyll concentration and associated per-pixel errors can be captured by application of the model to satellite data over the course of the seasons.

3.4. Potential Caveats In the Approach

3.4.1. *In situ* Estimates of Phytoplankton Group Chlorophyll

The performance of a model is tightly related to the quality of data used to tune it. We used estimates of phytoplankton group chlorophyll principally from HPLC. Whereas recent refinements in the use of HPLC to infer size-fractionated chlorophyll (Uitz et al., 2006; Brewin et al., 2010; Devred et al., 2011; Brewin et al., 2015) were used, diagnostic pigments determined by HPLC can be found in a variety of phytoplankton taxa and size classes, such that its use as a single *in situ* method may not always be dependable (Nair et al., 2008). Therefore, we combined data on size-fractionated chlorophyll estimated from HPLC with those from SFF, which encouragingly, were found to

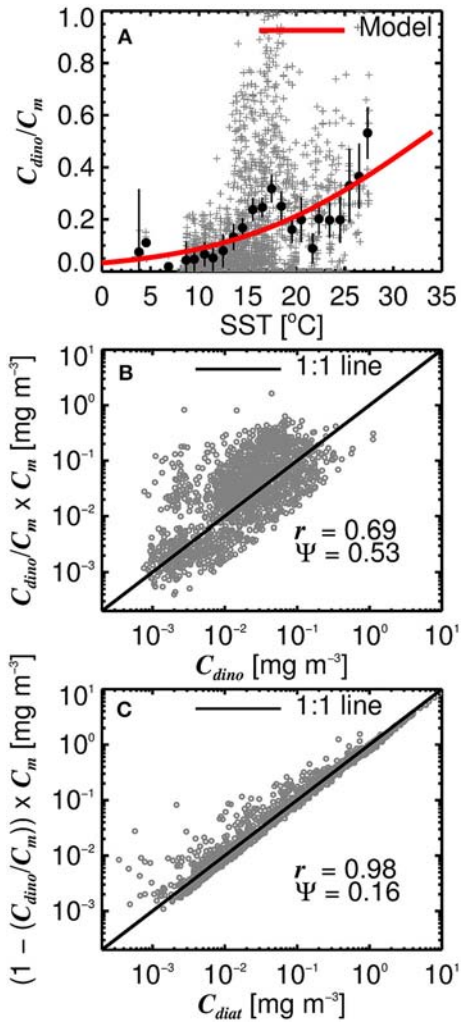


FIGURE 8 | (A) The ratio of dinoflagellate chlorophyll (C_{dino}) to microplankton chlorophyll (C_m) plotted as a function of sea surface temperature (SST). Gray points show raw values, black dots are binned averages with 95% confidence intervals on the averages, and red line show the fitted model (Equation 16). **(B)** Shows the modeled ratio (Equation 16) multiplied by microplankton chlorophyll (C_m) to estimate C_{dino}, plotted against measured C_{dino}. **(C)** Shows one minus the modeled ratio (Equation 16) multiplied by microplankton chlorophyll (C_m) to estimate C_{diat}, plotted against measured C_{diat}. r is the correlation coefficient and Ψ the root-mean-square-error.

be in reasonable agreement with each other for surface waters in the North Atlantic region (**Figure 2**). Yet, biases between the two techniques have been observed in Atlantic waters (Brewin et al., 2014a). Uncertainties in the SFF technique can arise from filter clogging, inaccurate pore sizes and cell breakage. The partitioning of microplankton chlorophyll into diatoms and dinoflagellates was based on the assumption that fucoxanthin in microphytoplankton can be attributed to diatoms and peridinin to dinoflagellates (Equations 8 and 9). Yet, there can also be fucoxanthin-containing dinoflagellates (e.g., *Kryptoperidinium foliaceum*) in Atlantic waters (Kempton et al., 2002), though there occurrence is generally not well known. Greater efforts to combine other sources of *in situ* data (e.g., flow cytometry, video imagery, optical measurements and microscopy) should help improve, and quantify uncertainty in, estimates of phytoplankton group chlorophyll *in situ* and ultimately, the parameterization of satellite models.

3.4.2. The Satellite Phytoplankton Group Model

The conceptual framework of the Brewin et al. (2010) model has been supported by data from: phytoplankton spectral absorption measurements (Brewin et al., 2011a); spectral particle backscattering measurements (Brewin et al., 2012a); chlorophyll estimated by size-fractionated filtration (Raimbault et al., 1988; Chisholm, 1992; Riegman et al., 1993; Gin et al., 2000; Marañón et al., 2012; Brewin et al., 2014a; Ward, 2015); flow cytometry and microscopy (Brotas et al., 2013). The model has also been found to reproduce inter-annual variations in size structure consistent with theories on coupling between physical-chemical processes and ecosystem structure (Brewin et al., 2012b), and found to reproduce the typical normalized-biomass size-spectrum of phytoplankton (Brewin et al., 2014b). The model has captured relationships between size structure and total chlorophyll in a variety of contrasting regions (e.g., Lin et al., 2014; Brito et al., 2015; Sammartino et al., 2015).

Yet, as with any abundance-based method, the model does not directly detect the phytoplankton groups: it simply infers the concentrations of chlorophyll in each group based on relationships, developed using data collected in the past, with properties that can by derived accurately from space (e.g., chlorophyll concentration and sea surface temperature). The model is not expected to capture blooms that deviate from the general trends observed in the parameterization dataset (**Figures 4**,**5**). For this reason, such techniques may not be appropriate for certain applications. For instance, under a climate-change scenario, there is the possibility that the relationships between properties (e.g., total chlorophyll and group-specific chlorophyll) may change, which may not be detected using an abundance-based approach (Sathyendranath et al., submitted). For such applications, spectral-based methods are likely to be preferable.

Two versions of the re-tuned Brewin et al. (2010) model were carried forward in this study: one using a fixed set of parameters (**Table 3**); and the other where the parameters were tied with SST (**Table 4**). The Brewin et al. (2010) model with a fixed parameter set has an advantage that only four parameters are required to compute the size fractions (**Table 3**), compared with 16 that are used in the SST-dependent model (**Table 4**). A larger dataset is required to tune the SST-dependent model for regional applications, when compared with the model with a fixed parameter set. Furthermore, when considering all samples together, only a slight improvement in model performance (Ψ and δ) was achieved when using the SST-dependent model (**Figures 7**, **9**). Yet, the SST-dependent model captured variations in model parameters, such as the asymptotic maximum values for small cells ($C^m_{p,n}$ and C^m_p), that are known to vary with changes in bottom-up (e.g., nutrients and light) and top-down (grazing) processes (Riegman et al., 1993; Brewin et al., 2014b). The

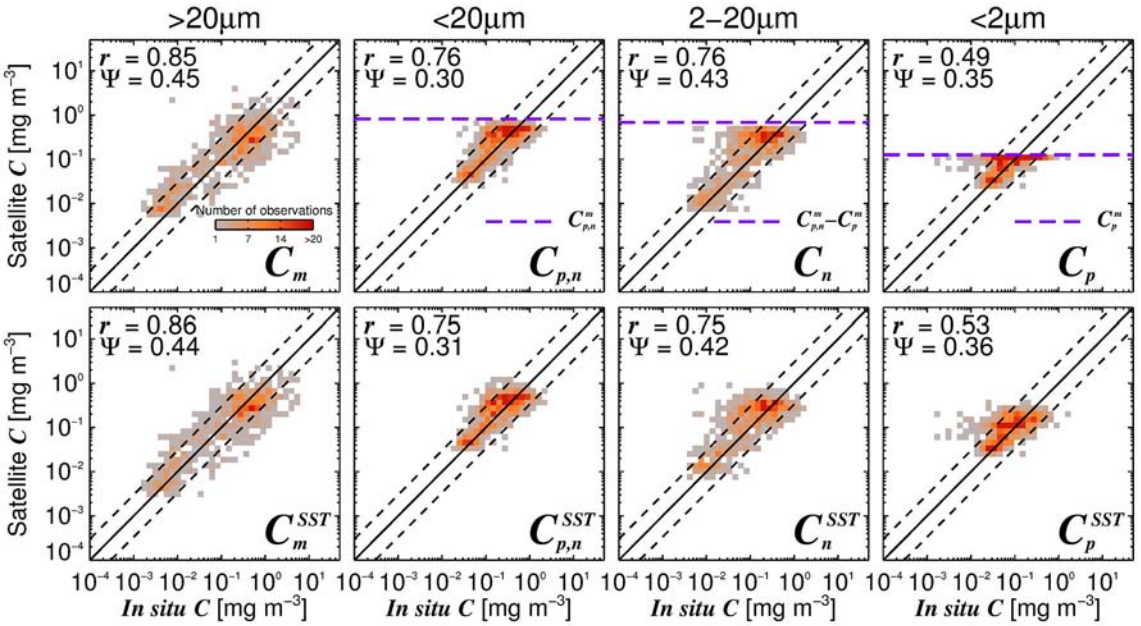


FIGURE 9 | Satellite estimates of size-fractionated chlorophyll plotted against independent *in situ* size-fractionated chlorophyll in the validation dataset, for the re-tuned (Brewin et al., 2010) model with a single set of parameters (top-row, Table 3) and using the SST-dependent parameterization (bottom row, Equations 12–15). The superscript SST denotes the modeled chlorophyll concentrations using the SST-dependent parameterization. The correlation coefficient (*r*) and root-mean-square-error (Ψ) are also shown. Statistical tests were computed using the parameter values reported in **Tables 3**, **4**. The top panels also show the maximum attainable concentrations for the different size classes as purple dashed horizontal lines when using a single set of parameters (**Table 3**).

fixed parameter model simply failed to capture these variations, resulting in unrealistic static asymptotes (**Figures 7**, **9** top-row, horizontal purple dashed lines).

Variations in the relationships of size structure with total chlorophyll and with SST were generally consistent with those proposed by Ward (2015), with the fractions of larger cells (e.g., microplankton) generally increasing with decreasing SST, for concentrations of total chlorophyll less than 1 mg m^{-3}, and the fractions of small cells (picoplankton) increasing (**Figure 6**). Yet, in the Ward (2015) study these variations were typically observed at lower temperature (<5°C) than those shown in this study (<17°C). Results are also relatively consistent for small cells (picoplankton) with those proposed by Brewin et al. (2015), when using average light in the mixed-layer, rather than SST, to vary model parameters, though differ for microplankton (see Figures 4, 5 of Brewin et al., 2015). Differences between studies are possibly due to the regional-tuning of the model when compared with the global studies of Ward (2015) and Brewin et al. (2015). There are also differences in the two approaches: whereas Ward (2015) introduces an additional term to the three-component model to account for temperature dependence, here we have let the model parameters change in response to SST variation.

Motivated by the need to provide satellite products of phytoplankton groups that match those as defined in ecosystem models, particularly ERSEM (**Table 2**), we proposed a partitioning of microplankton chlorophyll (C_m) into diatoms (C_{diat}) and dinoflagellates (C_{dino}), by modeling the ratio of C_{dino} to C_m as a function of SST (**Figure 8A**). This differs to that proposed by Hirata et al. (2011) which is based solely on total chlorophyll. We observed a significant relationship between C_{dino}/C_m and SST that was consistent with known seasonal variations of the two phytoplankton groups in the region (McQuatters-Gollop et al., 2007; Widdicombe et al., 2010). Yet, there still are significant variations surrounding this relationship (**Figure 8A**), and C_{dino} was found to have the highest errors in the satellite model (**Figures 11**, **12**). The approach may fail to capture blooms of microplankton chlorophyll (C_m) entirely dominated by dinoflagellates (**Figure 8A**), that can occur in the region (Widdicombe et al., 2010). Future improvements in C_{dino} satellite estimates may be possible by incorporating spectral information (Shang et al., 2014) or other environmental data (Raitsos et al., 2008). Such improvements may significantly aid ecosystem models considering the difficulties in modeling this group due to their motility and complex trophic behavior (Ciavatta et al., 2011).

3.4.3. Per-Pixel Uncertainties

In-line with methods used in the OC-CCI project (Jackson and Sathyendranath, 2015), our satellite estimates of the chlorophyll concentration of each phytoplankton group come with per-pixel uncertainty (**Figure 12**), an essential requirement for use in many applications, such as ecosystem model validation, data assimilation and quantifying evidence of trends in a time-series. Yet, estimates of uncertainty we provide are based on the assumption that the *in situ* data is the truth. As discussed in the previous section, *in situ* measurements of phytoplankton group chlorophyll also have their uncertainties, which are difficult to quantify (Brewin et al., 2014a). In addition, the estimates of

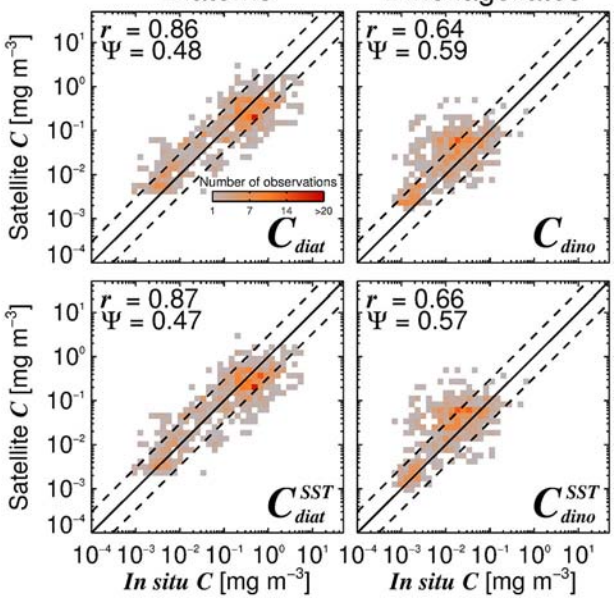


FIGURE 10 | Satellite estimates of diatom (C_{diat}) and dinoflagellate (C_{dino}) chlorophyll plotted against independent *in situ* estimates of C_{diat} and C_{dino} in the validation dataset, using the Brewin et al. (2010) model with a single set of parameters (top-row, Table 3) together with estimates of C_{dino}/C_m (Equation 16), and using the SST-dependent parameterization (bottom row, Equations 12–15) together with estimates of C_{dino}/C_m (Equation 16). The superscript SST denotes the modeled microplankton chlorophyll using the SST-dependent parameterization (Equations 12–15) multiplied by estimates of C_{dino}/C_m (Equation 16). The correlation coefficient (r) and root-mean-square-error (Ψ) are also shown.

uncertainty are based on comparisons of co-incident discrete *in situ* point measurements, representing volumes of sea water of the order of 5 litres or less, with 4 km satellite pixels representing a signal from $\sim 16 \times 10^{10}$ litres of water, assuming a 10 m optical depth. Additional uncertainties can occur because of vast differences in the temporal scales associated with the two types of measurements. In the future, such uncertainties may be reduced with the aid of new *in situ* methods capable of continuously measuring the optical and biogeochemical properties of the water (Dall'Olmo et al., 2012; Boss et al., 2013; Chase et al., 2013; Werdell et al., 2013b; Brewin et al., 2016).

By computing uncertainty statistics for each OWT, we can overcome issues with the distribution of data used in the validation. For instance, in our validation dataset, the majority of samples came from three OWTs (10, 11, and 12, see **Figure 11**), yet in the satellite image (**Figure 12B**), the majority of the region is dominated by OWTs less than 10. If one were to consider a single value of any statistical metric (as provided in **Figures 9**, **10**) as representative of the uncertainty in the entire satellite data, it would not be well representative of the majority of the region. Yet, as the number of samples in each OWT vary, so does our confidence in the error statistics for each OWT. Some OWTs (e.g., 1, 2, and 14) have very few observations (**Table 5**), and consequently we have low confidence in the uncertainty estimates for these OWTs.

TABLE 5 | Root mean square error (Ψ) and bias (δ) for 14 OC-CCI optical water types (OWT) for the four phytoplankton groups, using the two approaches (SST-dependent with superscript SST, and single set of parameters) to estimate phytoplankton group chlorophyll from satellite data.

OWT	Picoplankton						Nanoplankton						Diatoms						Dinoflagellates					
	C_p^{SST}			C_p			C_n^{SST}			C_n			C_{diat}^{SST}			C_{diat}			C_{dino}^{SST}			C_{dino}		
	Ψ	δ	N	Ψ	δ	N	Ψ	δ	N	Ψ	δ	N	Ψ	δ	N	Ψ	δ	N	Ψ	δ	N	Ψ	δ	N
1	0.13	−0.03	13	0.14	−0.06	13	0.37	−0.11	13	0.39	−0.16	13	0.28	−0.04	9	0.34	0.19	9	0.18	−0.11	6	0.19	0.12	6
2	0.28	−0.13	9	0.29	−0.15	9	0.40	0.01	9	0.41	−0.04	9	0.50	−0.30	6	0.41	−0.07	6	0.15	−0.07	4	0.20	0.16	4
3	0.16	0.04	67	0.16	0.02	67	0.28	0.09	58	0.27	0.05	58	0.37	0.05	50	0.45	0.26	50	0.17	0.03	40	0.29	0.24	40
4	0.19	0.06	42	0.19	0.04	42	0.30	0.11	39	0.29	0.06	39	0.41	−0.10	33	0.42	0.08	33	0.28	0.12	22	0.39	0.30	22
5	0.22	0.08	36	0.22	0.07	36	0.26	0.09	33	0.26	0.03	33	0.40	0.03	26	0.43	0.18	26	0.30	0.26	14	0.44	0.42	14
6	0.20	0.08	41	0.23	0.11	41	0.32	0.13	41	0.32	0.06	41	0.54	0.23	39	0.59	0.32	39	0.32	0.05	27	0.36	0.14	27
7	0.49	0.22	41	0.50	0.23	41	0.35	0.11	38	0.34	0.09	38	0.60	0.05	36	0.62	0.11	36	0.47	−0.05	27	0.47	0.01	27
8	0.38	0.17	36	0.38	0.20	36	0.39	−0.07	36	0.37	−0.07	36	0.49	0.15	34	0.51	0.14	34	0.33	0.03	31	0.35	0.02	31
9	0.39	0.20	79	0.36	0.19	79	0.40	0.12	67	0.40	0.14	67	0.56	−0.17	64	0.57	−0.15	64	0.42	0.13	49	0.43	0.14	49
10	0.41	0.15	111	0.38	0.06	111	0.38	0.11	108	0.38	0.14	108	0.52	−0.22	99	0.53	−0.21	99	0.50	0.32	91	0.51	0.33	91
11	0.35	0.13	158	0.33	−0.04	158	0.48	0.20	152	0.48	0.22	152	0.42	−0.19	145	0.42	−0.17	145	0.70	0.45	139	0.72	0.47	139
12	0.42	0.08	147	0.42	−0.10	147	0.50	0.21	139	0.54	0.27	139	0.38	−0.05	132	0.38	−0.05	132	0.67	0.39	129	0.68	0.39	129
13	0.58	0.21	21	0.55	0.08	21	0.63	0.18	21	0.61	0.22	21	0.55	0.03	21	0.54	0.04	21	0.83	0.07	17	0.82	0.10	17
14	0.44	0.41	4	0.15	−0.04	4	0.70	0.68	4	0.62	0.60	4	0.79	0.79	4	0.88	0.87	4	1.44	1.37	4	1.52	1.45	4

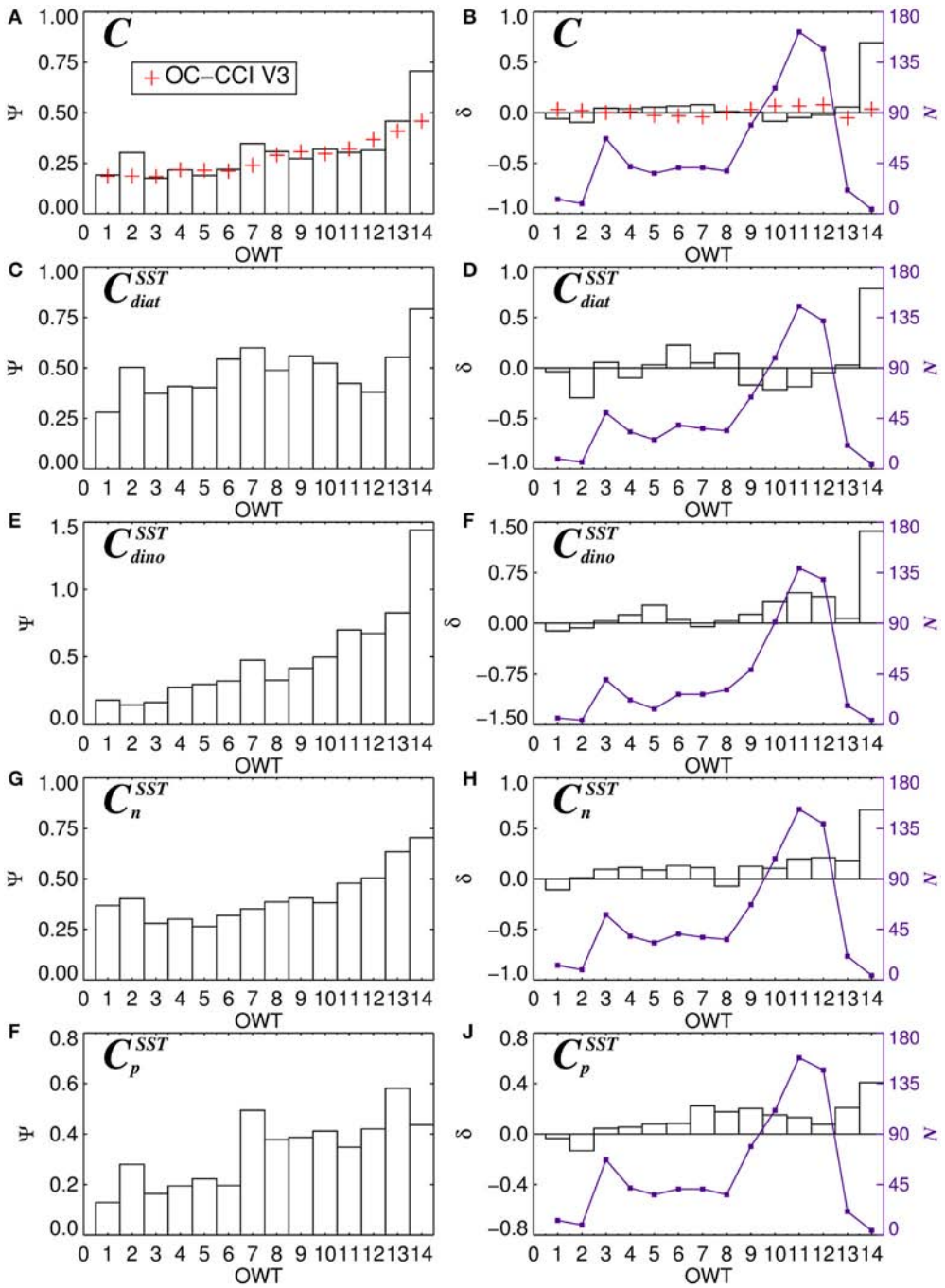


FIGURE 11 | The average root-mean-square-error (Ψ) and bias (Ψ) for 14 dominant OC-CCI Optical Water Types (OWT) for total chlorophyll (A,B), diatom chlorophyll (C,D), dinoflagellate chlorophyll (E,F), nanoplankton chlorophyll (G,H), and picoplankton chlorophyll (I,J). *N* (violet lines and squares) shows the number of observations of each dominant OWT. Plots **(C–J)** are for the SST-dependent parameterization (Equations 12–15) together with estimates of C_{dino}/C_m (Equation 16).

4. SUMMARY

We re-tuned an abundance-based model (Brewin et al., 2010, 2015) for estimating the chlorophyll concentration of three phytoplankton size classes as a function of total chlorophyll (available from satellite data) in the North Atlantic region using a large dataset of size-fractionated chlorophyll measurements. The model was modified to account for the influence of sea surface temperature (SST, also available from satellite data) on model parameters, and on the partitioning of chlorophyll in

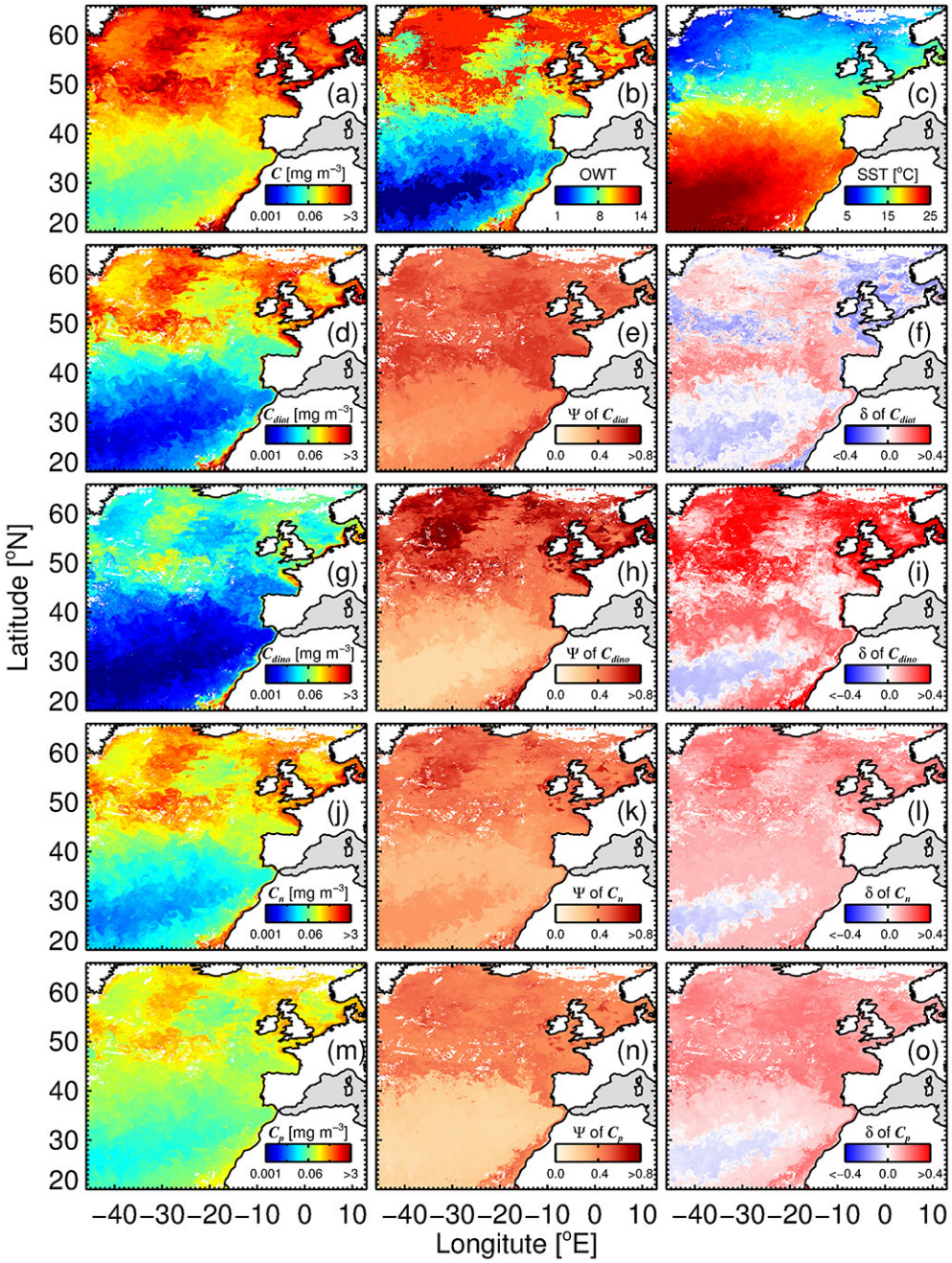


FIGURE 12 | Satellite estimates of phytoplankton group chlorophyll and per-pixel errors for an 8 day (relatively clear sky) composite (17th to 24th June 2008) of OC-CCI chlorophyll (a), (dominant) optical water type (b) and SST (NOAA OISST) data (c). Example shown is using the SST-dependent parameterization (Equations 12–15) together with estimates of C_{dino}/C_m (Equation 16): **(d)** Diatom chlorophyll (C_{diat}); **(e)** per-pixel root-mean-square-error (Ψ) of C_{diat}; **(f)** per-pixel bias (δ) of C_{diat}; **(g)** dinoflagellate chlorophyll (C_{dino}); **(h)** Ψ of C_{dino}; **(i)** δ of C_{dino}; **(J)** nanoplankton chlorophyll (C_n); **(k)** Ψ of C_n; **(l)** δ of C_n; **(m)** picoplankton chlorophyll (C_p); **(n)** Ψ of C_p; and **(o)** δ of C_p.

large phytoplankton (microphytoplankton) into diatoms and dinoflagellates, so that the phytoplankton groups provided matched those used in a marine ecosystem model (ERSEM). Results indicate that in the North Atlantic: (i) the relationship between size-fractionated chlorophyll and total chlorophyll changes with the environmental conditions (SST); and (ii) the ratio of dinoflagellate chlorophyll to microplankton chlorophyll increases with SST.

Application of the method to satellite estimates of total chlorophyll and SST was validated using an independent dataset

of satellite and *in situ* match-ups. This dataset was used with information on the optical water type, based on fuzzy-logic statistics developed within the ESA OC-CCI project, to derive uncertainties in 14 different optical water types, which were then used to map uncertainties in chlorophyll on a per-pixel basis for each phytoplankton group in a satellite image. These satellite products will be useful for those evaluating the performance of the ERSEM model and assimilating chlorophyll for each phytoplankton group into ERSEM in research and operational applications. Such an approach could be extended to other ecosystem models that simulate phytoplankton functional groups in the oceans.

AUTHOR CONTRIBUTIONS

RB synthesized the data, re-tuned and further-developed the algorithm, organized, prepared and wrote the first version of the manuscript, and prepared all figures and tables. SC, SS, TJ, EO, GD, and DR contributed to the intellectual development of the algorithms, and GT, KC, RA, DC, and VB collected and processed parts of the datasets used in the paper. All authors contributed to the final version of the manuscript.

FUNDING

This work has been carried out as part of the Copernicus Marine Environment Monitoring Service (CMEMS) project "Toward Operational Size-class Chlorophyll Assimilation (TOSCA)." CMEMS is implemented by MERCATOR OCEAN in the framework of a delegation agreement with the European Union. This work was also supported by the UK National Centre for Earth Observation (NCEO). Additional support from the Ocean Colour Component of the Climate Change Initiative of the European Space Agency (ESA) is gratefully acknowledged. Data collection by GT was supported by NERC-UK ECOMAR (grant no: NE/C513018/1). We thank ESA for covering publication costs.

ACKNOWLEDGMENTS

The authors would like to acknowledge all scientists and crew involved in the collection of the *in situ* data used in this manuscript, without which this work would not have been feasible. We owe a debit of gratitude to all those involved in data collection. AMT data were funded through the UK Natural Environment Research Council, through the UK marine research institutes' strategic research programme Oceans 2025 awarded to PML and the National Oceanography Centre. The authors would like to thank European Space Agency (ESA) for CCI data used, NOAA for the OISST products, and NASA for MODIS-Aqua and SeaWiFS products used. This is a contribution to MARE - UID/MAR/04292/2013, the Ocean Colour Climate Change Initiative of ESA and contribution number 311 of the AMT programme.

REFERENCES

Aiken, J., Pradhan, Y., Barlow, R., Lavender, S., Poulton, A., Holligan, P., et al. (2009). Phytoplankton pigments and functional types in the Atlantic Ocean: A decadal assessment, 1995-2005. *Deep Sea Res. I* 56, 899–917. doi: 10.1016/j.dsr2.2008.09.017

Airs, R., and Martinez-Vicente, V. (2014a). *AMT18 (JR20081003) HPLC Pigment Measurements from CTD Bottle Samples*. Liverpool: British Oceanographic Data Centre - Natural Environment Research Council.

Airs, R., and Martinez-Vicente, V. (2014b). *AMT19 (JR20081003) HPLC Pigment Measurements from CTD Bottle Samples.* Liverpool: British Oceanographic Data Centre - Natural Environment Research Council.

Airs, R., and Martinez-Vicente, V. (2014c). *AMT20 (JR20081003) HPLC Pigment Measurements from CTD Bottle Samples.* Liverpool: British Oceanographic Data Centre - Natural Environment Research Council.

Aumont, O., Maier-Reimer, E., Blain, S., and Monfray, P. (2003). An ecosystem model of the global ocean including Fe, Si, P colimitations. *Global Biogeochem. Cycles* 17:1060. doi: 10.1029/2001GB001745

Barlow, R. G., Aiken, J., Holligan, P. M., Cummings, D. G., Mariotena, S., and Hooker, S. (2002). Phytoplankton pigment and absorption characteristics along meridional transects in the Atlantic Ocean. *Deep Sea Res. I* 49, 637–660. doi: 10.1016/S0967-0637(01)00081-4

Barnes, M. K., Tilstone, G. H., Smyth, T. J., Suggett, D. J., Astoreca, R., Lancelot, C., et al. (2014). Absorption-based algorithm of primary production for total and size-fractionated phytoplankton in coastal waters. *Mar. Ecol. Prog. Ser.* 504, 73–89. doi: 10.3354/meps10751

Blackford, J. C., Allen, J. I., and Gilbert, F. J. (2004). Ecosystem dynamics at six contrasting sites: a generic modelling study. *J. Mar. Syst.* 52, 191–215. doi: 10.1016/j.jmarsys.2004.02.004

Boss, E., Picheral, M. P., Leeuw, T., Chase, A., Karsenti, E., Gorsky, G., et al. (2013). The characteristics of particulate absorption, scattering and attenuation coefficients in the surface ocean; Contribution of the Tara Oceans expedition. *Methods Oceanogr.* 7, 52–62. doi: 10.1016/j.mio.2013.11.002

Bracher, A., Bouman, H., Brewin, R. J. W., Bricaud, A., Brotas, V., Ciotti, A. M., et al. (2017). Obtaining phytoplankton diversity from ocean color: a scientific roadmap for future development. *Front. Mar. Sci.* 4:55. doi: 10.3389/fmars.2017.00055

Brewin, R. J. W., Dall'Olmo, G., Pardo, S., van Dongen-Vogel, V., and Boss, E. S. (2016). Underway spectrophotometry along the Atlantic Meridional Transect reveals high performance in satellite chlorophyll retrievals. *Remote Sens. Environ.* 183, 82–97. doi: 10.1016/j.rse.2016.05.005

Brewin, R. J. W., Dall'Olmo, G., Sathyendranath, S., and Hardman-Mountford, N. J. (2012a). Particle backscattering as a function of chlorophyll and phytoplankton size structure in the open-ocean. *Opt. Exp.* 20, 17632–17652. doi: 10.1364/OE.20.017632

Brewin, R. J. W., Devred, E., Sathyendranath, S., Hardman-Mountford, N. J., and Lavender, S. J. (2011a). Model of phytoplankton absorption based on three size classes. *Appl. Opt.* 50, 4535–4549. doi: 10.1364/AO.50.004535

Brewin, R. J. W., Hardman-Mountford, N. J., Lavender, S., Raitsos, D. E., Hirata, T., Uitz, J., et al. (2011b). An intercomparison of bio-optical techniques for detecting dominant phytoplankton size class from satellite remote sensing. *Remote Sens. Environ.* 115, 325–339. doi: 10.1016/j.rse.2010.09.004

Brewin, R. J. W., Hirata, T., Hardman-Mountford, N. J., Lavender, S., Sathyendranath, S., and Barlow, R. (2012b). The influence of the Indian Ocean Dipole on interannual variations in phytoplankton size structure as revealed by Earth Observation. *Deep Sea Res. II* 77–80, 117–127. doi: 10.1016/j.dsr2.2012.04.009

Brewin, R. J. W., Sathyendranath, S., Hirata, T., Lavender, S. J., Barciela, R., and Hardman-Mountford, N. J. (2010). A three-component model of phytoplankton size class for the Atlantic Ocean. *Ecol. Model.* 221, 1472–1483. doi: 10.1016/j.ecolmodel.2010.02.014

Brewin, R. J. W., Sathyendranath, S., Jackson, T., Barlow, R., Brotas, V., Airs, R., et al. (2015). Influence of light in the mixed layer on the parameters of a three-component model of phytoplankton size structure. *Remote Sens. Environ.* 168, 437–450. doi: 10.1016/j.rse.2015.07.004

Brewin, R. J. W., Sathyendranath, S., Lange, P. K., and Tilstone, G. (2014a). Comparison of two methods to derive the size-structure of natural populations of phytoplankton. *Deep Sea Res. I* 85, 72–79. doi: 10.1016/j.dsr.2013.11.007

Brewin, R. J. W., Sathyendranath, S., Tilstone, G., Lange, P. K., and Platt, T. (2014b). A multicomponent model of phytoplankton size structure. *J. Geophys. Res.* 119, 3478–3496. doi: 10.1002/2014JC009859

Brewin, R. J. W., Tilstone, G., Jackson, T., Cain, T., Miller, P., Lange, P. K., et al. (2017). Modelling size-fractionated primary production in the Atlantic Ocean from remote sensing. *Prog. Oceanogr.* doi: 10.1016/j.pocean.2017.02.002

Bricaud, A., Claustre, H., Ras, J., and Oubelkheir, K. (2004). Natural variability of phytoplanktonic absorption in oceanic waters: influence of the size structure of algal populations. *J. Geophys. Res.* 109:C11010. doi: 10.1029/2004JC002419

Briggs, N., Perry, M. J. P., Cetinić, I., Lee, C., D'Asaro, E., Gray, A. M., et al. (2011). High-resolution observations of aggregate flux during a sub-polar North Atlantic spring bloom. *Deep Sea Res. I* 58, 1031–1039. doi: 10.1016/j.dsr.2011.07.007

Brito, A. C., Sà, C., Brotas, V., Brewin, R. J. W., Silva, T., Vitorino, J., et al. (2015). Effect of phytoplankton size classes on bio-optical properties of phytoplankton in the Western Iberian coast: application of models. *Remote Sens. Environ.* 156, 537–550. doi: 10.1016/j.rse.2014.10.020

Brotas, V., Brewin, R. J. W., Sá, C., Brito, A. C., Silva, A., Mendes, C. R., et al. (2013). Deriving phytoplankton size classes from satellite data: validation along a trophic gradient in the eastern Atlantic Ocean. *Remote Sens. Environ.* 134, 66–77. doi: 10.1016/j.rse.2013.02.013

Butenschön, M., Clark, J., Aldridge, J. N., Allen, J. I., Artioli, Y., Blackford, J., et al. (2016). ERSEM 15.06: a generic model for marine biogeochemistry and the ecosystem dynamics of the lower trophic levels. *Geosci. Model Dev.* 9, 1293–1339. doi: 10.5194/gmd-9-1293-2016

Campbell, J. W. (1995). The lognormal distribution as a model for bio-optical variability in the sea. *J. Geophys. Res.* 100, 13237–13254. doi: 10.1029/95JC00458

Chase, A., Boss, E., Zaneveld, J. R. V., Bricaud, A., Claustre, H., Ras, J., et al. (2013). Decomposition of *in situ* particulate absorption spectra. *Methods Oceanogr.* 7, 110–124. doi: 10.1016/j.mio.2014.02.002

Chisholm, S. W. (1992). "Phytoplankton size," in *Primary Productivity and Biogeochemical Cycles in the Sea*, eds P. G. Falkowski and A. D. Woodhead (New York, NY: Springer), 213–237.

Ciavatta, S., Kay, S., Saux-Picart, S., Butenschön, M., and Allen, J. I. (2016). Decadal reanalysis of biogeochemical indicators and fluxes in the North West European shelf-sea ecosystem. *J. Geophys. Res.* 121, 1824–1845. doi: 10.1002/2015jc011496

Ciavatta, S., Torres, R., Martinez-Vicente, V., Smyth, T., Dall'Olmo, G., Polimene, L., et al. (2014). Assimilation of remotely-sensed optical properties to improve marine biogeochemistry modelling. *Prog. Oceanogr.* 127, 74–95. doi: 10.1016/j.pocean.2014.06.002

Ciavatta, S., Torres, R., Saux-Picart, S., and Allen, J. I. (2011). Can ocean color assimilation improve biogeochemical hindcasts in shelf seas? *J. Geophys. Res.* 116:C12. doi: 10.1029/2011JC007219

Ciotti, A. M., Lewis, M. R., and Cullen, J. J. (2002). Assessment of the relationships between dominant cell size in natural phytoplankton communities and the spectral shape of the absorption coefficient. *Limnol. Oceanogr.* 47, 404–417. doi: 10.4319/lo.2002.47.2.0404

Dall'Olmo, G., Boss, E., Behrenfeld, M., and Westberry, T. K. (2012). Particulate optical scattering coefficients along an Atlantic Meridional Transect. *Opt. Exp.* 20, 21532–21551. doi: 10.1364/OE.20.021532

Dandonneau, Y., Deschamps, P. Y., Nicolas, J. M., Loisel, H., Blanchot, J., Montel, Y., et al. (2004). Seasonal and interannual variability of ocean colour and composition of phytoplankton communities in the North Atlantic, equatorial Pacific and South Pacific. *Deep Sea Res. II* 51, 303–318. doi: 10.1016/j.dsr2.2003.07.018

de Boyer Montégut, C., Madec, G., Fisher, A. S., Lazar, A., and Iudicone, D. (2004). Mixed layer depth over the global ocean: an examination of profile data and a profile-based climatology. *J. Geophys. Res.* 109:C12003. doi: 10.1029/2004JC002378

de Mora, L., Butenschön, M., and Allen, J. I. (2016). The assessment of a global marine ecosystem model on the basis of emergent properties and ecosystem function: a case study with ERSEM. *Geosci. Model Dev.* 9, 59–76. doi: 10.5194/gmd-9-59-2016

De Moraes Rudorff, N., and Kampel, M. (2012). Orbital remote sensing of phytoplankton functional types: a new review. *Int. J. Remote Sens.* 33, 1967–1990. doi: 10.1080/01431161.2011.601343

Devred, E., Sathyendranath, S., Stuart, V., Maas, H., Ulloa, O., and Platt, T. (2006). A two-component model of phytoplankton absorption in the open ocean: theory and applications. *J. Geophys. Res.* 111:C03011. doi: 10.1029/2005JC002880

Devred, E., Sathyendranath, S., Stuart, V., and Platt, T. (2011). A three component classification of phytoplankton absorption spectra: applications to ocean-colour data. *Remote Sens. Environ.* 115, 2255–2266. doi: 10.1016/j.rse.2011.04.025

Ducklow, H. W., and Harris, R. P. (1993). Introduction to the JGOFS North Atlantic bloom experiment. *Deep Sea Res. Part II Top. Stud. Oceanogr.* 40, 1–8. doi: 10.1016/0967-0645(93)90003-6

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Ann. Stat.* 7, 1–26. doi: 10.1214/aos/1176344552

Finkel, Z. V., Beardall, J., Flynn, K., Quigg, A., Rees, T. A. V., and Raven, J. A. (2010). Phytoplankton in a changing world: cell size and elemental stoichiometry. *J. Plank. Res.* 32, 119–137. doi: 10.1093/plankt/fbp098

Ford, D. A., Edwards, K. P., Lea, D., Barciela, R. M., Martin, M. J., and Demaria, J. (2012). Assimilating GlobColour ocean colour data into a pre-operational physical-biogeochemical model. *Ocean Sci.* 8, 751–771. doi: 10.5194/os-8-751-2012

Garver, S. A., Siegel, D. A., and Mitchell, B. G. (1994). Variability in near-surface particulate absorption spectra: what can a satellite ocean color imager see? *Limnol. Oceanogr.* 39, 1349–1367. doi: 10.4319/lo.1994.39.6.1349

Geider, R. J., Platt, T., and Raven, J. A. (1986). Size dependence of growth and photosynthesis in diatoms: a synthesis. *Mar. Ecol. Prog. Ser.* 30, 93–104. doi: 10.3354/meps030093

Gibb, S. W., Barlow, R. G., Cummings, D. G., Rees, N. W., Trees, C. C., Holligan, P. M., et al. (2000). Surface phytoplankton pigment distributions in the Atlantic Ocean: an assessment of basin scale variability between 50°N and 50°S. *Prog. Oceanogr.* 45, 339–368. doi: 10.1016/S0079-6611(00)00007-0

Gin, K. Y.-H., Lin, X., and Zhang, S. (2000). Dynamics and size structure of phytoplankton in the coastal waters of Singapore. *J. Plank. Res.* 22, 1465–1484. doi: 10.1093/plankt/22.8.1465

Gregg, W. W., Friedrichs, M. A. M., Robinson, A. R., Rose, K. A., Schlitzer, R., Thompson, K. R., et al. (2009). Skill assessment in ocean biological data assimilation. *J. Mar. Syst.* 76, 16–33. doi: 10.1016/j.jmarsys.2008.05.006

Guidi, L., Stemmann, L., Jackson, G. A., Ibanez, F., Claustre, H., Legendre, L., et al. (2009). Effects of phytoplankton community on production, size and export of large aggregates: a world-ocean analysis. *Limnol. Oceanogr.* 54, 1951–1963. doi: 10.4319/lo.2009.54.6.1951

Hashioka, T., Vogt, M., Yamanaka, Y., Le Quéré, C., Aita, M. N., Alvain, S., et al. (2013). Phytoplankton competition during the spring bloom in four plankton functional type models. *Biogeosciences* 10, 6833–6850. doi: 10.5194/bg-10-6833-2013

Hirata, T., Hardman-Mountford, N. J., Brewin, R. J. W., Aiken, J., Barlow, R., Suzuki, K., et al. (2011). Synoptic relationships between surface chlorophyll-a and diagnostic pigments specific to phytoplankton functional types. *Biogeosciences* 8, 311–327. doi: 10.5194/bg-8-311-2011

Hirata, T., Saux Picart, S., Hashioka, T., Aita-Noguchi, M., Sumata, H., Shigemitsu, M., et al. (2013). A comparison between phytoplankton community structures derived from a global 3D ecosystem model and satellite observation. *J. Mar. Syst.* 109–101, 129–137. doi: 10.1016/j.jmarsys.2012.01.009

Holt, J., Allen, J. I., Anderson, T. R., Brewin, R. J. W., Butenschön, M., Harle, J., et al. (2014). Challenges in integrative approaches to modelling the marine ecosystems of the North Atlantic: physics to fish and coasts to ocean. *Prog. Oceanogr.* 129(Part B), 285–313. doi: 10.1016/j.pocean.2014.04.024

IOCCG (2000). *Remote Sensing of Ocean Colour in Coastal, and Other Optically Complex Waters*. Technical report, ed S. Sathyendranath, Reports of the

International Ocean-Colour Coordinating Group, No. 3, IOCCG, Dartmouth, NS.

IOCCG (2014). *Phytoplankton Functional Types from Space*. Technical Report, ed S. Sathyendranath, Reports of the International Ocean-Colour Coordinating Group, No. 15, IOCCG, Dartmouth, NS.

Jackson, T., and Sathyendranath, S. (2015). *Ocean Colour Climate Change Initiative (OC-CCI) Phase 2- Comprehensive Error Characterization Report 2.0.*

Kempton, J., Wolny, J., Tengs, T., Rizzo, P., Morris, R., Tunnell, J., et al. (2002). *Kryptoperidinium foliaceum* blooms in South Carolina: a multi-analytical approach to identification. *Harmful Algae* 1, 383–392. doi: 10.1016/S1568-9883(02)00051-3

Kishi, M. J., Kashiwai, M., Ware, D. M., Megrey, B. A., Eslinger, D. L., Werner, F. E., et al. (2007). NEMURO–a lower trophic level model for the North Pacific marine ecosystem. *Ecol. Model.* 202, 12–25. doi: 10.1016/j.ecolmodel.2006.08.021

Kostadinov, T. S., Milutinović, S., Marinov, I., and Cabré, A. (2016). Carbon-based phytoplankton size classes retrieved via ocean color estimates of the particle size distribution. *Ocean Sci.* 12, 561–575. doi: 10.5194/os-12-561-2016

Kostadinov, T. S., Siegel, D. A., and Maritorena, S. (2009). Retrieval of the particle size distribution from satellite ocean color observations. *J. Geophys. Res.* 114:C09015. doi: 10.1029/2009JC005303

Latasa, M., Scharek, R., Gall, F. L., and Guillou, L. (2004). Pigment suites and taxonomic groups in prasinophyceae. *J. Phycol.* 40, 1149–1155. doi: 10.1111/j.1529-8817.2004.03136.x

Laufkötter, C., Vogt, M., Gruber, N., Aumont, O., Bopp, L., Doney, S., et al. (2016). Projected decreases in future marine export production: the role of the carbon flux through the upper ocean ecosystem. *Biogeosciences* 13, 4023–4047. doi: 10.5194/bg-13-4023-2016

Laws, E. A., Falkowski, P. G., Smith, W. O. Jr., Ducklow, H., and McCarth, J. J. (2000). Temperature effects on export production in the open ocean. *Global Biogeochem. Cycles* 14, 1231–1246. doi: 10.1029/1999GB001229

Le Quéré, C., Harrison, S. P., Prentice, C. I., Buitenhuis, E. T., Aumont, O., Bopp, L., et al. (2005). Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. *Global Change Biol.* 11, 2016–2040. doi: 10.1111/j.1365-2486.2005.1004.x

Lee, Z., Arnone, R., Hu, C., Werdell, P. J., and Lubac, B. (2011). Uncertainties of optical parameters and their propagations in an analytical ocean color inversion algorithm. *Appl. Optics* 49, 369–381. doi: 10.1364/AO.49.000369

Legendre, L., and LeFevre, J. (1991). "From individual plankton cells to pelagic marine ecosystems and to global biogeochemical cycles," in *Particle Analysis in Oceanography*, ed S. Demers (Berlin: Springer), 261–300.

Lin, J., Cao, W., Wang, G., and Hu, S. (2014). Satellite-observed variability of phytoplankton size classes associated with a cold eddy in the South China Sea. *Mar. Pollut. Bull.* 83, 190–197. doi: 10.1016/j.marpolbul.2014.03.052

Maloney, C. L., and Field, J. G. (1991). The size-based dynamics of plankton food webs. I. A simulation model of carbon and nitrogen flows. *J. Plank. Res.* 13, 1003–1038. doi: 10.1093/plankt/13.5.1003

Marañón, E. (2009). "Phytoplankton size structure," in *Encyclopedia of Ocean Sciences*, eds J. H. Steele, K. Turekian, and S. A. Thorpe (Oxford: Academic Press). doi: 10.1016/b978-012374473-9.00661-5

Marañón, E. (2015). Cell size as a key determinant of phytoplankton metabolism and community structure. *Ann. Rev. Mar. Sci.* 7, 241–264. doi: 10.1146/annurev-marine-010814-015955

Marañón, E., Cermeño, P., Latasa, M., and Tadonléké, R. D. (2012). Temperature, resources, and phytoplankton size structure in the ocean. *Limnol. Oceanogr.* 57, 1266–1278. doi: 10.4319/lo.2012.57.5.1266

Marañón, E., Holligan, P. M., Barciela, R., González, N., Mouriño, B., Pazó, M. J., et al. (2001). Patterns of phytoplankton size structure and productivity in contrasting open-ocean environments. *Mar. Ecol. Prog. Ser.* 216, 43–56. doi: 10.3354/meps216043

Marinov, I., Doney, S. C., and Lima, I. D. (2010). Response of ocean phytoplankton community structure to climate change over the 21st century: partitioning the effects of nutrients, temperature and light. *Biogeosciences* 7, 3941–3959. doi: 10.5194/bg-7-3941-2010

Maritorena, S., Fanton d'Andon, O. H., Mangin, A., and Siegel, D. A. (2010). Merged satellite ocean color data products using a bio-optical model: characteristics, benefits and issues. *Remote Sens. Environ.* 114, 1791–1804. doi: 10.1016/j.rse.2010.04.002

Markwardt, C. B. (2008). "Non-linear least squares fitting in IDL with MPFIT," in *Proceedings of the Astronomical Data Analysis Software and Systems XVIII, ASP Conference Series, Quebec, Canada*, vol. 411, eds D. Bohlender, P. Dowler, and D. Duran (San Francisco, CA: Astronomical Society of the Pacific).

McQuatters-Gollop, A., Raitsos, D. E., Edwards, M., and Attrill, M. J. (2007). Spatial patterns of diatom and dinoflagellate seasonal cycles in the NE Atlantic Ocean. *Mar. Ecol. Prog. Ser.* 339, 301–306. doi: 10.3354/meps339301

Merchant, C. J., Embury, O., Roberts-Jones, J., Fiedler, E., Bulgin, C. E., Corlett, G. K., et al. (2014). Sea surface temperature datasets for climate applications from Phase 1 of the European Space Agency Climate Change initiative (SST CCI). *Geosci. Data J.* 1, 179–191. doi: 10.1002/gdj3.20

Moore, T., Campbell, J. W., and Feng, H. (2001). A fuzzy logic classification scheme for selecting and blending satellite ocean color algorithms. *IEEE Trans. Geosci. Remote Sens.* 39, 1764–1776. doi: 10.1109/36.942555

Moore, T. S., Campbell, J. W., and Dowell, M. D. (2009). A class-based approach to characterizing and mapping the uncertainty of the MODIS ocean chlorophyll product. *Remote Sens. Environ.* 113, 2424–2430. doi: 10.1016/j.rse.2009.07.016

Moore, T. S., Dowell, M. D., and Franz, B. A. (2012). Detection of coccolithophore blooms in ocean color satellite imagery: a generalized approach for use with multiple sensors. *Remote Sens. Environ.* 117, 249–263. doi: 10.1016/j.rse.2011.10.001

Moré, J. (1978). *The Levenberg-Marquardt Algorithm: Implementation and Theory. In: Numerical Analysis*. Berlin: Springer-Verlag.

Morel, A., and Bricaud, A. (1981). Theoretical results concerning light absorption in a discrete medium, and application to specific absorption of phytoplankton. *Deep Sea Res.* 28, 1375–1393. doi: 10.1016/0198-0149(81)90039-X

Morel, A., Huot, Y., Gentili, B., Werdell, P. J., Hooker, S. B., and Franz, B. A. (2007). Examining the consistency of products derived from various ocean color sensors in open ocean (case 1) waters in the perspective of a multi-sensor approach. *Remote Sens. Environ.* 111, 69–88. doi: 10.1016/j.rse.2007.03.012

Mouw, C. B., Hardman-Mountford, N. J., Alvain, S., Bracher, A., Brewin, R. J. W., Bricaud, A., et al. (2017). A consumer's guide to satellite remote sensing of multiple phytoplankton groups in the global ocean. *Front. Mar. Sci.* 4:41. doi: 10.3389/fmars.2017.00041

Nair, A., Sathyendranath, S., Platt, T., Morales, J., Stuart, V., Forget, M.-H., et al. (2008). Remote sensing of phytoplankton functional types. *Remote Sens. Environ.* 112, 3366–3375. doi: 10.1016/j.rse.2008.01.021

Pan, X., Wong, G. T., Ho, T. Y., Shiah, F. K., and Liu, H. (2013). Remote sensing of picophytoplankton distribution in the northern South China Sea. *Remote Sens. Environ.* 128, 162–175. doi: 10.1016/j.rse.2012.10.014

Platt, T., and Denman, K. L. (1976). The relationship between photosynthesis and light for natural assemblages of coastal marine phytoplankton. *J. Phycol.* 12, 421–430. doi: 10.1111/j.1529-8817.1976.tb02866.x

Platt, T., and Denman, K. L. (1977). Organisation in the pelagic ecosystem. *HelgoaẼnder Wissenschaftliche Meeresuntersuchungen* 30, 575–581. doi: 10.1007/BF02207862

Platt, T., and Denman, K. L. (1978). The structure of pelagic marine ecosystems. *Rapp. P.-v. Réun. Cons. perm. int. Explor. Mer.* 173, 60–65.

Prieur, L., and Sathyendranath, S. (1981). An optical classification of coastal and oceanic waters based on the specific spectral absorption curves of phytoplankton pigments, dissolved organic matter and other particulate materials. *Limnol. Oceanogr.* 26, 617–689. doi: 10.4319/lo.1981.26.4.0671

Probyn, T. A. (1985). Nitrogen uptake by size-fractionated phytoplankton populations in the southern Benguela upwelling system. *Mar. Ecol. Prog. Ser.* 22, 249–258. doi: 10.3354/meps022249

Raimbault, P., Rodier, M., and Taupier-Letage, I. (1988). Size fraction of phytoplankton in the Ligurian Sea and the Algerian Basin (Mediterranean Sea): size distribution versus total concentration. *Mar. Microb. Food Webs* 3, 1–7.

Raitsos, D. E., Lavender, S. J., Maravelias, C. D., Haralambous, J., Richardson, A. J., and Reid, P. C. (2008). Identifying four phytoplankton functional types from space: an ecological approach. *Limnol. Oceanogr.* 53, 605–613. doi: 10.4319/lo.2008.53.2.0605

Raven, J. A. (1998). Small is beautiful: the picophytoplankton. *Funct. Ecol.* 12, 503–513. doi: 10.1046/j.1365-2435.1998.00233.x

Reynolds, R. W., Smith, T. M., Liu, C., Chelton, D. B., Casey, K. S., and Schlax, M. G. (2007). Daily high-resolution-blended analysis for sea surface temperature. *J. Climate* 20, 5473–5496. doi: 10.1175/2007JCLI1824.1

Riegman, R., Kuipers, B. R., Noordeloos, A. A. M., and Witte, H. J. (1993). Size-differential control of phytoplankton and the structure of plankton communities. *Netherlands J. Sea Res.* 31, 225–265. doi: 10.1016/0077-7579(93)90026-O

Robinson, C., Serret, P., Tilstone, G., Teira, E., Zubkov, M. V., Rees, A. P., et al. (2002). Plankton respiration in the Eastern Atlantic Ocean. *Deep Sea Res. I* 49, 787–813. doi: 10.1016/S0967-0637(01)00083-8

Rousseaux, C. S., Hirata, T., and Gregg, W. W. (2013). Satellite views of global phytoplankton community distributions using an empirical algorithm and a numerical model. *Biogeosci. Discuss.* 10, 1083–1109. doi: 10.5194/bgd-10-1083-2013

Roy, S., Sathyendranath, S., Bouman, H., and Platt, T. (2013). The global distribution of phytoplankton size spectrum and size classes from their light-absorption spectra derived from satellite data. *Remote Sens. Environ.* 139, 185–197. doi: 10.1016/j.rse.2013.08.004

Sammartino, M., Di Cicco, A., Marullo, S., and Santoleri, R. (2015). Spatio-temporal variability of micro-, nano- and pico-phytoplankton in the Mediterranean Sea from satellite ocean colour data of SeaWiFS. *Ocean Sci.* 11, 759–778. doi: 10.5194/os-11-759-2015

Sathyendranath, S. (2011). Ocean Colour Climate Change Initiative (OC-CCI) - Phase One User Requirements Document.

Sathyendranath, S., Stuart, V., Cota, G., Maas, H., and Platt, T. (2001). Remote sensing of phytoplankton pigments: a comparison of empirical and theoretical approaches. *Int. J. Remote Sens.* 22, 249–273. doi: 10.1080/014311601449925

Savidge, G., Boyd, P., Pomroy, A., Harbour, D., and Joint, I. (1995). Phytoplankton production and biomass estimates in the north-east Atlantic Ocean. *Deep Sea Res. Part I* 42, 599–617. doi: 10.1016/0967-0637(95)00016-Y

Serret, P., Robinson, C., Fernández, E., Teira, E., and Tilstone, G. (2001). Latitudinal variation of the balance between plankton photosynthesis and respiration in the eastern Atlantic Ocean. *Limnol. Oceanogr.* 46, 1642–1652. doi: 10.4319/lo.2001.46.7.1642

Shang, S., Wu, J., Huang, B., Lin, G., Lee, Z., Liu, J., et al. (2014). A new approach to discriminate dinoflagellate from diatom blooms from space in the East China Sea. *J. Geophys. Res.* 119, 4653–4668. doi: 10.1002/2014JC009876

Shimoda, Y., and Arhonditsis, G. B. (2016). Phytoplankton functional type modelling: running before we can walk? A critical evaluation of the current state of knowledge. *Ecol. Model.* 320, 29–43. doi: 10.1016/j.ecolmodel.2015.08.029

Sieracki, M. E., Verity, P. G., and Stoecker, D. K. (1993). Plankton community response to sequential silicate and nitrate depletion during the 1989 North Atlantic spring bloom. *Deep Sea Res. Part II Top. Stud. Oceanogr.* 40, 213–225. doi: 10.1016/0967-0645(93)90014-E

Smyth, T. J., Fishwick, J., Al-Moosawi, L., Cummings, D. G., Harris, R., Kitidis, V., et al. (2010). A broad spatio-temporal view of the Western English Channel observatory. *J. Plank. Res.* 32, 585–601. doi: 10.1093/plankt/fbp128

Stauber, J. L., and Jeffrey, S. W. (1988). Photosynthetic pigments in fifty-one species of marine diatoms. *J. Phycol.* 24, 158–172. doi: 10.1111/j.1529-8817.1988.tb04230.x

Sunda, W. G., and Huntsman, S. A. (1997). Interrelated influence of iron, light and cell size on marine phytoplankton growth. *Nature* 390, 389–392. doi: 10.1038/37093

Takahashi, T., Sutherland, S. C., Sweeney, C., Poisson, A., Metzl, N., Tilbrook, B., et al. (2002). Global sea–air CO_2 flux based on climatological surface ocean pCO_2, and seasonal biological and temperature effects. *Deep Sea Res. Part II Top. Stud. Oceanogr.* 49, 1601–1622. doi: 10.1016/S0967-0645(02)00003-6

Takahashi, T., Sutherland, S. C., Wanninkhof, R., Sweeney, C., Feely, R. A., Chipman, D. W., et al. (2009). Climatological mean and decadal change in surface ocean pCO_2, and net sea–air CO_2 flux over the global oceans. *Deep Sea Res. Part II Top. Stud. Oceanogr.* 56, 554–577. doi: 10.1016/j.dsr2.2008.12.009

Tilstone, G. H., Lange, P. K., Misra, A., Brewin, R. J. W., and Cain, T. (2017). Micro-phytoplankton photosynthesis, primary production and potential export production in the Atlantic Ocean. *Prog. Oceanogr.* doi: 10.1016/j.pocean.2017.01.006

Tilstone, G., Miller, P., Brewin, R. J. W., and Priede, I. G. (2014). Enhancement of primary production in the North Atlantic outside of the spring bloom, identified by remote sensing of ocean colour and temperature. *Remote Sens. Environ.* 146, 77–86. doi: 10.1016/j.rse.2013.04.021

Tilstone, G. H., Taylor, B. H., Blondeau-Patissier, D., Powell, T., Groom, S. B., Rees, A. P., et al. (2015). Comparison of new and primary production models using SeaWiFS data in contrasting hydrographic zones of the northern North Atlantic. *Remote Sens. Environ.* 156, 473–489. doi: 10.1016/j.rse.2014.10.013

Trees, C. C., Clark, D. K., Bidigare, R. R., Ondrusek, M. E., and Mueller, J. L. (2000). Accessory pigments versus chlorophyll a concentrations within the euphotic zone: a ubiquitous relationship. *Limnol. Oceanogr.* 45, 1130–1143. doi: 10.4319/lo.2000.45.5.1130

Uitz, J., Claustre, H., Morel, A., and Hooker, S. B. (2006). Vertical distribution of phytoplankton communities in open ocean: an assessment based on surface chlorophyll. *J. Geophys. Res.* 111:C08005. doi: 10.1029/2005JC003207

Vidussi, F., Claustre, H., Manca, B. B., Luchetta, A., and Marty, J. C. (2001). Phytoplankton pigment distribution in relation to upper thermocline circulation in the eastern Mediterranean Sea during winter. *J. Geophys. Res.* 106, 19939–19956. doi: 10.1029/1999JC000308

Vogt, M., Hashioka, T., Payne, M., Buitenhuis, E. T., Le Quéré, C., Alvain, S., et al. (2013). The distribution, dominance patterns and ecological niches of plankton functional groups in Dynamic Green Ocean Models and satellite estimates. *Biogeosci. Discuss.* 10, 17193–17247. doi: 10.5194/bgd-10-17193-2013

Wang, P., Boss, E., and Roesler, C. (2005). Uncertainties of inherent optical properties obtained from semianalytical inversions of ocean color. *Appl. Opt.* 44, 4074–4085. doi: 10.1364/AO.44.004074

Ward, B. A. (2015). Temperature-correlated changes in phytoplankton community structure are restricted to polar waters. *PLoS ONE* 10:e0135581. doi: 10.1371/journal.pone.0135581

Ward, B. A., Dutkiewicz, S., Jahn, O., and Follows, M. J. (2012). A size-structured food-web model for the global ocean. *Limnol. Oceanogr.* 57, 1877–1891. doi: 10.4319/lo.2012.57.6.1877

Werdell, P. J., and Bailey, S. W. (2005). An improved *in-situ* bio-optical data set for ocean colour algorithm development and satellite data production validation. *Remote Sens. Environ.* 98, 122–140. doi: 10.1016/j.rse.2005.07.001

Werdell, P. J., Bailey, S. W., Fargion, G. S., Pietras, C., Knobelspiesse, K. D., Feldman, G. C., et al. (2003). Unique data repository facilitates ocean color satellite validation. *EOS Trans. Am. Geophys. Union* 84, 377–387. doi: 10.1029/2003EO380001

Werdell, P. J., Franz, B. A., Bailey, S. W., Feldman, G. C., Boss, E., Brando, V. E., et al. (2013a). Generalized ocean color inversion model for retrieving marine inherent optical properties. *Appl. Opt.* 52, 2019–2037. doi: 10.1364/AO.52.002019

Werdell, P. J., Proctor, C. W., Boss, E., Leeuw, T., and Ouhssain, M. (2013b). Underway sampling of marine inherent optical properties on the Tara Oceans expedition as a novel resource for ocean color satellite data product validation. *Methods Oceanogr.* 7, 40–51. doi: 10.1016/j.mio.2013.09.001

Westberry, T. K., Dall'Olmo, G., Behrenfeld, M., and Moutin, T. (2010). Coherence of particulate beam attenuation and backscattering coefficients in diverse open ocean environments. *Opt. Exp.* 18, 15419–15425. doi: 10.1364/OE.18.015419

Widdicombe, C., Eloire, D., Harbour, D., Harris, R., and Somerfield, P. (2010). Long-term phytoplankton community dynamics in the Western English Channel. *J. Plankt. Res.* 32, 643–655. doi: 10.1093/plankt/fbp127

Xiao, Y., and Friedrichs, M. A. M. (2014). The assimilation of satellite-derived data into a one-dimensional lower trophic level marine ecosystem model. *J. Geophys. Res.* 119, 2691–2712. doi: 10.1093/plankt/fbp127

Zapata, M., Fraga, S., Rodríguez, F., and Garrido, J. L. (2012). Pigment-based chloroplast types in dinoflagellates. *Mar. Ecol. Prog. Ser.* 465, 33–52. doi: 10.3354/meps09879

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Brewin, Ciavatta, Sathyendranath, Jackson, Tilstone, Curran, Airs, Cummings, Brotas, Organelli, Dall'Olmo and Raitsos. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Marine Science

ORIGINAL RESEARCH
published: 08 May 2017
doi: 10.3389/fmars.2017.00133

Impact of El Niño Variability on Oceanic Phytoplankton

Marie-Fanny Racault[1,2]*, *Shubha Sathyendranath*[1,2], *Robert J. W. Brewin*[1,2], *Dionysios E. Raitsos*[1,2], *Thomas Jackson*[1] and *Trevor Platt*[1]

[1] Plymouth Marine Laboratory, Plymouth, UK, [2] Plymouth Marine Laboratory, National Centre for Earth Observation, Plymouth, UK

Oceanic phytoplankton respond rapidly to a complex spectrum of climate-driven perturbations, confounding attempts to isolate the principal causes of observed changes. A dominant mode of variability in the Earth-climate system is that generated by the El Niño phenomenon. Marked variations are observed in the centroid of anomalous warming in the Equatorial Pacific under El Niño, associated with quite different alterations in environmental and biological properties. Here, using observational and reanalysis datasets, we differentiate the regional physical forcing mechanisms, and compile a global atlas of associated impacts on oceanic phytoplankton caused by two extreme types of El Niño. We find robust evidence that during Eastern Pacific (EP) and Central Pacific (CP) types of El Niño, impacts on phytoplankton can be felt everywhere, but tend to be greatest in the tropics and subtropics, encompassing up to 67% of the total affected areas, with the remaining 33% being areas located in high-latitudes. Our analysis also highlights considerable and sometimes opposing regional effects. During EP El Niño, we estimate decreases of −56 TgC/y in the tropical eastern Pacific Ocean, and −82 TgC/y in the western Indian Ocean, and increase of +13 TgC/y in eastern Indian Ocean, whereas during CP El Niño, we estimate decreases −68 TgC/y in the tropical western Pacific Ocean and −10 TgC/y in the central Atlantic Ocean. We advocate that analysis of the dominant mechanisms forcing the biophysical under El Niño variability may provide a useful guide to improve our understanding of projected changes in the marine ecosystem in a warming climate and support development of adaptation and mitigation plans.

Keywords: El Niño variability, ENSO, climate, ocean-color, ESA climate change initiative, phytoplankton

OPEN ACCESS

Edited by:
Laura Lorenzoni,
University of South Florida, USA

Reviewed by:
Monique Messié,
Monterey Bay Aquarium Research Institute, USA
Peter Allan Thompson,
Commonwealth Scientific and Industrial Research Organisation (CSIRO), Australia

***Correspondence:**
Marie-Fanny Racault
mfrt@pml.ac.uk

Specialty section:
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

Received: *17 January 2017*
Accepted: *20 April 2017*
Published: *08 May 2017*

Citation:
Racault M-F, Sathyendranath S, Brewin RJW, Raitsos DE, Jackson T and Platt T (2017) Impact of El Niño Variability on Oceanic Phytoplankton. Front. Mar. Sci. 4:133. doi: 10.3389/fmars.2017.00133

INTRODUCTION

Phytoplankton, the microscopic vegetal cells living at the surface of the oceans, yield globally and annually some fifty billion tons of organic carbon through primary production (Longhurst et al., 1995), contributing to the oceanic uptake of ∼25% of the carbon dioxide (CO_2) emitted to the atmosphere every year (Le Quéré et al., 2015). The rates of primary production are not uniformly distributed across the ocean domain: the most highly productive oceanic regions are found at high-latitudes and in coastal upwelling systems. Oceanic primary producers are under the control of physical forcing on a broad spectrum of scales, and the forcing will be modified under climate change. In the latest assessment report (AR5), the Intergovernmental Panel on Climate Change (IPCC) has recognized "medium evidence" for the response of the highly productive oceanic regions to recent warming (especially since the 1970s) and "low confidence"

in the understanding of how equatorial upwelling systems might change in response to El Niño variability (Hoegh-Guldberg et al., 2014).

El Niño activity is characterized by anomalous warming of Sea Surface Temperature (SST) in the tropical Pacific, linked to a perturbation of atmospheric circulation patterns known as the Southern Oscillation. This ocean-atmosphere coupling, called the El Niño Southern Oscillation (ENSO), is a dominant mode of variability in the Earth-climate system with a typical frequency of 2–7 years (McPhaden et al., 2006; Cai et al., 2015). Each El Niño event is unique, exhibiting differences in surface and subsurface water temperature amplitude, duration, and spatial patterns (Capotondi et al., 2015). As an aid to classify El Niño events, the location of maximum anomalous SST warming observed during boreal winter has been used to delineate two extreme types of El Niño (Trenberth and Stepaniak, 2001; Larkin and Harrison, 2005; Ashok et al., 2007; Yu and Kao, 2007; Ashok and Yamagata, 2009; Kao and Yu, 2009; Kug et al., 2009; Lee and McPhaden, 2010; Takahashi et al., 2011; Cai et al., 2015; Capotondi et al., 2015). The Eastern-Pacific (EP) El Niño, also referred to as the "typical" or canonical El Niño, is characterized by maximum anomalous SST warming in the eastern tropical Pacific. In contrast, the Central-Pacific (CP) El Niño, variously referred to as El Niño Modoki (a Japanese word meaning pseudo), warm-pool El Niño, or dateline El Niño, is characterized by weak anomalous SST warming along the western coast of South America and maximum anomalous SST warming in the central tropical Pacific. The climatic perturbations generated by these two types of El Niño are induced through different atmospheric teleconnections: both have been associated with changes in temperature and rainfall patterns over the continental U.S. (Yu et al., 2012; Yu and Zou, 2013), in storm tracks in the Southern Hemisphere (Kao and Yu, 2009) and in cyclone trajectories in the North Atlantic (Kim et al., 2009). Since the beginning of the 1900s, two most extreme EP and CP El Niño events (in terms of amplitude of maximum SST anomalies in the Eastern and Central Pacific regions) have occurred within the last 20 years, in 1997/1998 and 2009/2010 respectively (Capotondi et al., 2015). The latest 2015/2016 El Niño has been reported with comparable magnitude of SST anomalies to the 1982/1983 and 1997/1998 events but with more limited intensity in the Eastern Pacific region (Paek et al., 2017; L'Heureux et al., in press).

Contrasting influence of the extreme El Niño events of 1997/1998 and 2009/2010 on oceanic phytoplankton has been characterized in the tropical Pacific domain (Gierach et al., 2012; Radenac et al., 2012). In this region, ENSO is recognized as the main driver of inter-annual phytoplankton variability. Tropical, as well as extra-tropical, influences of ENSO and ENSO Modoki have been demonstrated using statistical analyses based on a range of indices applied to ocean-color remote-sensing observations (Yoder and Kennelly, 2003; Behrenfeld et al., 2006; Chavez et al., 2011; Vantrepotte and Mélin, 2011; Messié and Chavez, 2012, 2013; Racault et al., 2012, 2017; Couto et al., 2013; Raitsos et al., 2015). One of the most widely-used environmental indices to characterize influence of ENSO on ocean biology is the Multivariate ENSO Index (MEI). This index is based on combined analysis of fields of sea level pressure, surface winds, SST, surface air temperature, and cloudiness for the entire Tropical Pacific domain (Wolter and Timlin, 1993). Due to this broad domain and multivariate statistical approach, the MEI encompasses the whole continuum of ENSO events (from most extreme SST anomalies located in the Central to Eastern Pacific regions), but as a result, the index does not allow us to separate effects of EP and CP El Niño variations. Separating EP and CP El Niño variations requires indices isolating the centroid of El Niño activity along the equatorial Pacific.

To date, the longitudinal position of the center of maximum SST anomalies has been delineated in the Niño1+2 (0°–10°S, 90°–80°W), Niño3 (5°–5°N, 150°W–90°W), Niño3.4 (5°N–5°S, 170°W–120°W) and Niño4 (5°–5°N, 160°E–150°W) regions (Ashok et al., 2007). Based on analyses of the SST anomalies variations in these regions, a range of El Niño indices have been constructed. In the present study, the EP and CP El Niño signals are characterized using the EP and CP index defined by Kao and Yu (2009). The EP index was calculated by first applying regression analysis of SST anomalies onto the Niño4 index (average SST anomalies over the Niño4 region) to remove the influence of the SST anomaly component associated with central Pacific warming, and then using Empirical Orthogonal Function (EOF) analysis to determine the spatial patterns and associated temporal index of EP events. Similarly, the CP index was calculated by applying regression analysis of SST anomalies onto the Niño1+2 index to remove the influence of the SST anomaly component associated with east Pacific warming, and then using EOF analysis to characterize pattern and index of CP events (Kao and Yu, 2009; Yu et al., 2012). Using partial correlation and EOF analyses, the characterization of the canonical El Niño and El Niño Modoki signals has also been achieved based on SST anomalies from the Niño3 region and by differentiating the influence of SST anomalies variations from a combination of regions in the tropical Pacific. The specific combination and definition of regions have been formulated as the Trans-Nino Index TNI (Trenberth and Stepaniak, 2001), the El Niño Modoki Index EMI (Ashok et al., 2007), and the Improved EMI (Li et al., 2010). A comprehensive review of El Niño indices definition is presented by Capotondi et al. (2015).

The main obstacles to distinguishing the ecosystem effects associated with El Niño variability may be summarized to arise from: (i) the challenges to construct continuous, synoptic-scale, long-term time-series of marine ecosystem state at high temporal and spatial resolutions for the global oceans (Sathyendranath and Krasemann, 2014); (ii) the diversity and complexity of the mechanisms driving the biophysical interactions in different oceanic sub-regions or provinces (Longhurst et al., 1995; Boyd et al., 2014); (iii) the difficulties to elucidate the roles of the local and remote-forcing mechanisms associated with different El Niño events (Cai et al., 2015; Capotondi et al., 2015); (iv) the issues of lag in the transmission of El Niños influences at higher latitudes and to other basins via different teleconnection mechanisms (Ashok et al., 2007; Couto et al., 2013); and finally (v) the broad ranges of meridional position, amplitude and evolution of SST anomalies

observed during different El Niño events, which militate against consensus in the choice of method to estimate indices of El Niño variability (Capotondi et al., 2015). Here, we propose an original approach that overcome some of these obstacles based on climate-quality ocean-color products and state-of-the-art reanalysis datasets, and allow us to establish an atlas of the impact of CP and EP types of El Niño on primary producers in the global oceans. Finally, we document the associated environmental changes and identify the dominant mechanisms driving the diverse biophysical interactions involved.

MATERIALS AND METHODS

The list of biological and physical datasets obtained for the analyses is summarized in **Table 1**. Based on datasets availability, two periods of study have been considered: (1) 1997–2012 (15 years) for biological and physical datasets, and (2) 1979–2014 (35 years) for physical datasets only.

Biological Datasets

Remotely-Sensed Chlorophyll Concentration and Associated Uncertainty Estimates

Chlorophyll is at the heart of primary production, it is the state variable used in photosynthesis-irradiance models to compute primary production; it has a distinct optical signature which makes it one of the easiest phytoplankton properties to measure, both by *in-situ* and satellite methods. Ocean-color sensors on satellites provide estimates of chlorophyll concentration at high spatial and temporal resolution and at global scale. Because they provide data consistently and frequently and over long periods of time, they are suitable for computations of certain ecological indicators and for studying long-term trends in the state of the marine ecosystem (Platt and Sathyendranath, 2008; Racault et al., 2014). However, ocean-color sensors do have a finite lifespan, and differences in instrument design and algorithms make it difficult to compare data from multiple sensors. When overlapping data are available from two or more sensors, such data can be used to establish inter-sensor bias and correct for it. Recently, under the European Space Agency (ESA) Climate Change Initiative (CCI), the ocean-color project (http://www.esa-oceancolour-cci.org) has produced new, improved products, merging observations from the Sea-viewing Wide Field-of-View Sensor (SeaWiFS, 1997–2010), the Moderate Resolution Imaging Spectroradiometer (MODIS, 2002-present) and the MEdium Resolution Imaging Spectrometer (MERIS, 2002–2012) to provide a 15-year (1997–2012 OC-CCIv2) global scale, climate-quality controlled, bias-corrected, and error-characterized data record of ocean-color (Sathyendranath and Krasemann, 2014). Furthermore, implementation of the coupled ocean-atmosphere POLYMER correction algorithm (MERIS period; Steinmetz et al., 2011) has increased significantly the coverage of chlorophyll observations (Sathyendranath and Krasemann, 2014; Racault et al., 2015).

The OC-CCI v2.0 Level 3 Mapped data of chlorophyll concentration, root-mean-square-difference (RMSD) and bias estimates of monthly log-transformed (base 10) Chl, were obtained at 4 km spatial resolution, and monthly temporal resolution from the ESA CCI Ocean Color website at http://www.esa-oceancolour-cci.org. To remain coherent with the resolution of the different datasets used in the analyses (**Table 1**), the chlorophyll concentration was mapped onto a $1^\circ \times 1^\circ$ regular grid by averaging all available data points within each new, larger pixel. Standard deviation of chlorophyll product (computed from bias and RMSD) was calculated by aggregating pixel values at the same spatial ($1^\circ \times 1^\circ$) resolution. Then, the standard deviation of the log10Chl was converted to its untransformed value. The standard error in the reported mean value at each pixel for the period 1997–2012 was computed, and values ranging between $\pm 1\%$ in the tropical gyres (associated with low chlorophyll concentration) and below $\pm 0.5\%$ at higher latitudes (associated with high chlorophyll concentration) were observed (Supplementary Figure 1). Uncertainties in the anomalies of relative (%) changes in chlorophyll in selected box areas were calculated using the standard methods for computing propagation of errors (Topping, 1972).

Remotely-Sensed Primary Production (PP)

Global observations of water-column PP were obtained from the Open Ocean Transboundary Water Assessment Programme (TWAP) using the algorithm of Platt and Sathyendranath (1988), with OC-CCI v2.0 Chlorophyll, SeaWiFS and MODIS spectrally-resolved light (i.e., PAR) as inputs. The model parameters (i.e., vertical structure of Chlorophyll and the photosynthesis-irradiance parameters) are assigned following the partitioning of the ocean into biogeographic provinces (Longhurst, 1998). The TWAP primary production estimates have been shown to compare consistently well with other global ocean primary production models (Longhurst et al., 1995; Antoine et al., 1996; Behrenfeld et al., 2005). The PP data have been obtained at 9 km spatial resolution, and monthly temporal resolution from https://www.oceancolour.org/thredds/catalog/TWAP-PProd/catalog.html. The PP data were regridded to $1^\circ \times 1^\circ$ spatial resolution by averaging all available data points within each new, larger pixel.

Physical Datasets

Remotely-Sensed Photosynthetically Active Radiation (PAR)

The Level 3 Mapped data of PAR, collected during the SeaWiFS and MODIS missions (Frouin et al., 2012), were obtained at 9 km spatial resolution and monthly resolution from the NASA website at http://oceancolor.gsfc.nasa.gov/cms/. The PAR data were regridded to $1^\circ \times 1^\circ$ spatial resolution by averaging all available data points within each new, larger pixel.

Remotely-Sensed Sea Surface Temperature (SST)

The sea surface temperature SST-CCI vexp1.2 Mapped gap-filled daily blend of the Advanced Very High Resolution Radiometers and the Along-Track Scanning Radiometers data (Merchant et al., 2014) were obtained at $1^\circ \times 1^\circ$ spatial resolution, and monthly temporal resolution from the ESA-CCI National Centre for Earth Observation (NCEO) portal at http://gws-access.ceda.ac.uk/public2/nceo_uor/sst/L3S/EXP1.2/.

TABLE 1 | Datasets obtained for the analysis.

Dataset	Version	Period	Time	Res.	Source
Chlorophyll Concentration	v2.0	1997–2013	Monthly	1° × 1°	Ocean Color-CCI ESA
Primary Production	v2.0	1997–2012	Monthly	1° × 1°	TWAP project
Sea Level Anomaly	v1.1	1993–2013	Monthly	0.25° × 0.25°	Sea Level-CCI ESA
Sea Surface Temperature	vEXP1.2	1991–2015	Monthly	1° × 1°	SST-CCI ESA
Sea Surface Temperature	ERA Interim	1979–2014	Monthly	0.75° × 0.75°	ECMWF
Wind	ERA Interim	1979–2014	Monthly	0.75° × 0.75°	ECMWF
Surface Air Temperature	Jan-15	1948–2014	Monthly	2.5° × 2.5°	NCEP/NCAR/NOAA
Precipitation	Jun-16	1979–2013	Monthly	2.5° × 2.5°	NCAR/NOAA/ESRL CMAP
Photosynthetically Active Radiation	R2010	1997–2012	Monthly	1° × 1°	NASA SeaWiFS & MODIS
Ocean subsurface temperature	v2p2p4	1950–2008	Monthly	0.5° × 0.5°	SODA
EP and CP El Niño indices	Jan-15	1948–2014	Monthly	–	(Kao and Yu, 2009; Yu et al., 2012)
Multivariate ENSO Index	Jan-13	1950–2013	Monthly	–	ESRL/NOAA

Information about the regridding procedure and selected period of study are provided in the Materials and Methods section in the manuscript. Res, Spatial Resolution.

Remotely-Sensed Sea Level (SL)

The sea level SL-CCI v1.1 Mapped gap-filled blend of the Topex/Poseidon, Jason-1/2 with the ERS-1/2 and Envisat missions data (Ablain et al., 2015) were obtained at 0.25° × 0.25° spatial resolution, and monthly temporal resolution from the ESA SL-CCI website at http://www.esa-sealevel-cci.org.

Reanalysis Products of SST and Wind

ERA Interim reanalysis of monthly 10 m U wind component, 10 m V wind component, 10 m wind speed (Dee et al., 2011) were obtained on 0.75° × 0.75° global grid-box from ECMWF at http://apps.ecmwf.int/datasets/data/interim_full_moda/.

Reanalysis Product of Surface Air Temperature (SAT)

NCEP/NCAR reanalysis of monthly surface air temperature (Kalnay et al., 1996) were obtained on 2.5° × 2.5° global grid-box from National Oceanic Atmospheric Administration/Office of Oceanic and Atmospheric Research/Earth System research Laboratory at http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.surface.html. This dataset was chosen to be consistent with the data used by Yu et al. (2012) in their analysis on El Niño impact on U.S. winter air surface temperature.

Reanalysis Product of Precipitation

CPC Merged Analysis of Precipitation (CMAP) interpolated data (Xie and Arkin, 1997) were obtained on 2.5° × 2.5° global grid-box at monthly temporal resolution from National Oceanic Atmospheric Administration/Office of Oceanic and Atmospheric Research/Earth System research Laboratory at http://www.esrl.noaa.gov/psd/data/gridded/data.cmap.html.

Mixed Layer Depth (MLD)

The mixed layer depth (MLD) was estimated as the shallowest depth at which ±0.2°C change is observed compared with the temperature at 10 m depth, based on the temperature criterion of de Boyer Montégut et al. (2004). The vertical profiles of temperature were downloaded from Simple Ocean Data Assimilation (SODA) model output v2p2p4 http://iridl.ldeo.columbia.edu/SOURCES/.CARTON-GIESE/.SODA/.v2p2p4/ at monthly temporal resolution and 0.25° × 0.4° × 40-level spatial and vertical resolutions (Carton and Giese, 2008). The data were regridded to 1° × 1° spatial resolution by averaging all available data points within each new, larger pixel.

Zonal Surface Currents

Annual average zonal surface currents data were obtained at 1° × 1° resolution for the global oceans from NOAA Ocean Surface Current Analyses Real-time (OSCAR) at http://www.esr.org/oscar_index.html.

Nitrate Concentration

Annual average surface nitrate concentration data were obtained at 1° × 1° resolution for the global oceans from the World Ocean Atlas Climatology (Boyer et al., 2013) at https://www.nodc.noaa.gov/OC5/woa13/.

Climate Impact Analysis

Climate Indices

Time-series of MEI based on principal component analysis of six atmosphere-ocean variable fields in the tropical Pacific basin i.e., SL, SST, SAT, U, and V wind components, and total cloudiness fraction of the sky (Wolter and Timlin, 1993) were obtained at http://www.esrl.noaa.gov/psd/enso/mei/table.html. Time-series of Eastern Pacific and Central Pacific El Niño indices based a combination of regression and empirical orthogonal function analyses applied to SST data in the tropical Pacific (Kao and Yu, 2009) were obtained at http://www.ess.uci.edu/~yu/2OSC/.

Statistical Analysis

The influences of EP and CP El Niño events may propagate across the world at different speed through different teleconnection mechanisms (Ashok et al., 2007). To avoid implementing impact analyses on monthly anomalies, which would involve different lag-coefficients, the influence of the different types of El Niño is characterized based on annual mean anomalies. Anomalies of physical and biological variables were computed first by removing the monthly mean climatology over the period 1997–2012. Then, annual mean anomalies were calculated by averaging

the monthly anomalies over the periods from June (of year *t*) to May (of year *t* + *1*) (i.e., spanning over two calendar years). This 12-month delineation period was chosen to follow the seasonality of ENSO activity, which generally peaks in the month of November to January (i.e., higher SST anomalies in the Equatorial Pacific, Niño 1–4 regions). Because chlorophyll concentrations can span three orders of magnitude, relative percent differences in chlorophyll were calculated such as:

$$C_{ra} = \left(Ct - \bar{C}\right) / \left(\left(C_t + \bar{C}\right)/2\right) * 100$$

where C_{ra} is relative chlorophyll anomalies in percent, C_t is annual chlorophyll concentration in mg.m^{-3} in year *t*, and $\bar{C}$ is the mean of annual chlorophyll concentrations over the 15 years study period. Note that the results of the EP and CP impacts (see method below) were not sensitive to the choice of normalization function.

The climate impact analysis to identify the oceanic regions that are most sensitive to El Niño variability is based on a statistical approach initially developed in a study of El Niño impact on U.S. winter air surface temperature using EP and CP indices (Yu et al., 2012). In the present study, the global and regional influences associated with each type of El Niño are extracted by separately regressing at each 1° × 1° grid point the EP and CP El Niño indices with: (a) annual mean anomalies of chlorophyll concentration time-series (as a key measure of phytoplankton population), and (b) annual mean anomalies of primary production (i.e., the rate of phytoplankton growth) time-series. To identify the mechanisms driving the regionally-different biological responses associated with each type of El Niño, we further applied the statistical analysis based on the EP and CP indices, to annual mean anomalies of SAT, SST, SL, wind, and precipitation.

The statistical significance of the regression coefficients was estimated according to Student *t*-test. The autocorrelation of the time-series was considered and the effective degrees of freedom, which enter the significance test was determined based on the method presented in Lin and Derome (1998).

Validation of EP and CP Impact on Interannual to Decadal Time-Scales

The analyses of impact on phytoplankton and primary production have been limited to 15 years by the availability of consistent climate-quality controlled satellite data (**Table 1**). To assess whether the results may be skewed due to the specific 1997/1998 EP event during the ocean-color satellite era, and to investigate the validity of the results over multi-decadal time scales, we have estimated the impact of EP and CP El Niño (considering autocorrelation, Lin and Derome, 1998) on sea and air surface temperatures, wind and precipitation during the period 1979–2014 (35-years) and compared the results with the analysis during the shorter period 1997–2012 (15-year). These two periods of 15 and 35 years respectively have been selected based on availability of physical and biological data products (**Table 1**). The period 1997–2012 includes one strong EP El Niño event in 1997–1998, and three CP events in 2002–2003, 2004–2005, and 2009–2010. The period 1979–2014 includes one additional strong EP El Niño event in 1982–1983, and three additional CP events in 1986–1987, 1991–1992, and 1994–1995 (**Figure 1**).

RESULTS AND DISCUSSION

EP and CP El Niño Impact on Oceanic Phytoplankton

The confidence level of the response patterns (**Figures 2A,B**) identified with the statistical analysis is assessed as very likely (i.e., within 90–100% probability range; IPCC Climate Change, 2013). During EP El Niño, at the global scale, the median chlorophyll impact is found to be −6.5%, mostly driven by a large decrease of −7.5% in the tropics (3,529 pixels), and limited increases in the Northern and Southern Hemispheres of +6.6 and +5.5% respectively (505 and 643 pixels respectively). During CP El Niño, global median chlorophyll impact is found to be −7.4%, which is the resultant of large decrease estimated for the tropics of −8.7% (1,948 pixels), and limited decrease estimated for the Northern Hemisphere of −4.9% (342 pixels) and increase for the Southern Hemisphere of +4.1% (500 pixels) (Supplementary Table 1 and Supplementary Figure 2). The impact values are estimated based on EP and CP indices equal to one (annual mean index values over the period 1997–2012 are presented in **Figure 1**). It is noteworthy that monthly impact values may be ∼3–4 times higher, particularly during the peak of El Niño events, such as in December 1997 when the EP El Niño index reached value of 3.9, and in December 2009 when the CP El Niño index reached value of 2.6 (Yu, 2016).

The regions identified as most sensitive to the EP and CP El Niño climatic perturbations compare well with the locations where significant trends in chlorophyll have been estimated previously using contemporary satellite records (Vantrepotte and Mélin, 2011; Gregg and Rousseaux, 2014; Hoegh-Guldberg et al., 2014). Furthermore, the biological and physical response patterns to each type of El Niño observed in the present work over the satellite record of 1997–2012 are consistent with contemporary case studies of specific EP and CP El Niño events in the Equatorial Pacific (Turk et al., 2011; Gierach et al., 2012; Radenac et al., 2012), the Indian Ocean (Webster et al., 1999), the continental U.S. (Yu et al., 2012; Yu and Zou, 2013), and the global oceans (Behrenfeld et al., 2001; Messié and Chavez, 2012, 2013). In addition, the response patterns observed for the physical variables have also been shown to persist over decadal timescales during the period 1979–2014 (see Section Physical Forcing Mechanisms Associated with El Niño Variability). Finally, when both EP and CP indices are equal to one, the sum of the observed impacts observed in response to EP and CP El Niño (Supplementary Figure 3) is shown to be approximately equal to the impact observed using the MEI. This is coherent as the MEI encompasses effects of both EP and CP El Niño variations.

Major Factors Influencing Phytoplankton Growth

To understand the specific influence of the climatic perturbations on phytoplankton, we must first identify the mechanisms

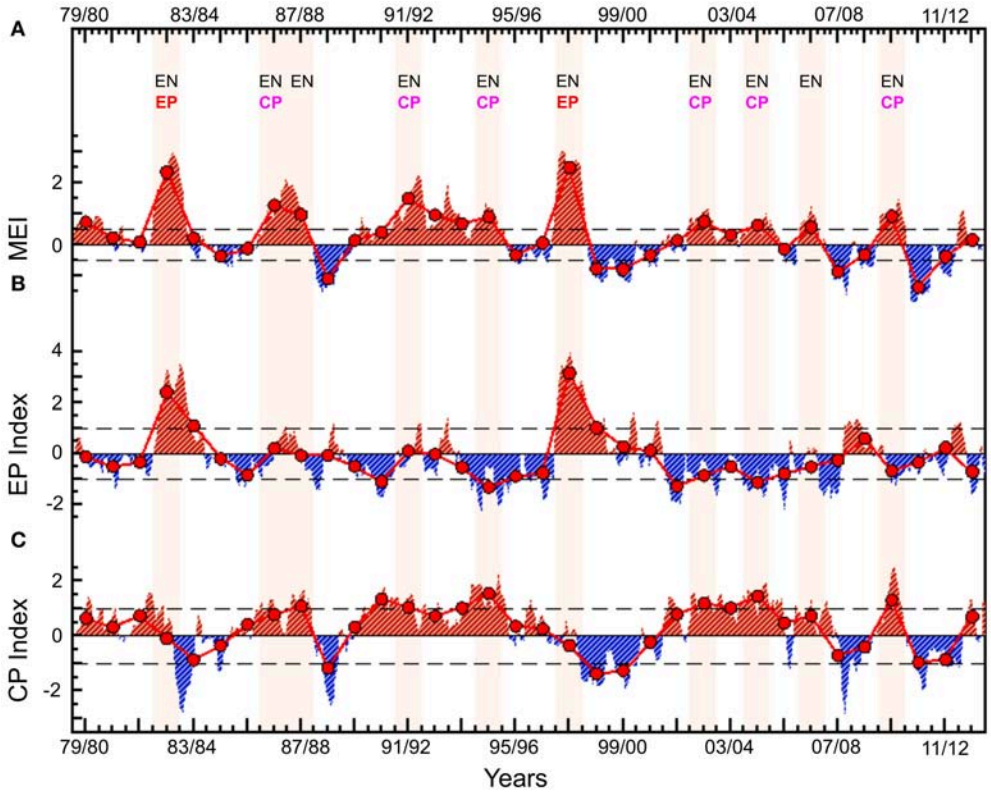


FIGURE 1 | Climate indices of El Niño events during the period 1979–2014. Monthly anomalies of **(A)** Multivariate ENSO Index; **(B)** Eastern Pacific El Niño Index; and **(C)** Central Pacific El Niño Index. Annual mean anomalies (thick red line and large dots) were calculated by averaging the monthly anomalies over the periods from June (of year *t*) to May (of year *t* + *1*). Classification of El Niño as EP and CP is based on Radenac et al. (2012) and Yu et al. (2012). EN, El Niño, EP, Eastern Pacific, CP, Central Pacific.

driving the biophysical interactions at the global and regional scales. Phytoplankton growth is light-limited at high-latitudes where annual mean nitrate concentration is high and monthly means of chlorophyll and PAR show positive correlation, and monthly means of chlorophyll and MLD show negative correlation (i.e., chlorophyll increases when MLD is shallower and light availability is higher; **Figures 3A,B**). In contrast, phytoplankton growth is nutrient-limited in the tropics and subtropics where light-availability is plentiful all-year-round, annual mean nitrate concentration is low (**Figure 3C**) and monthly means of chlorophyll and MLD show positive correlation (i.e., chlorophyll increases when MLD is deeper, and nutrient-rich deep waters are mixed with nutrient-poor surface waters, increasing nutrient availability for phytoplankton growth to occur). Further to nutrient supply from vertical mixing, the tropics display strong zonal surface currents (**Figure 3D**), which can increase horizontal advection of nutrient and, in turn, enhance phytoplankton growth. The latter mechanism can explain the weak correlation coefficients observed between monthly means of chlorophyll and MLD in some areas of the tropics and subtropics. Finally, the negative correlation shown between monthly means of chlorophyll and MLD in the eastern Equatorial Pacific is coherent with the observed high annual mean nitrate concentration (i.e., macronutrients are not limiting; **Figure 3C**), and previously reported iron limitation (i.e., limiting trace nutrient) occurring in the region (Gordon et al., 1997; Moore et al., 2013). In some specific areas of the North and Equatorial Pacific Ocean, and the Southern Ocean, known as High Nutrient-Low Chlorophyll (HNLC) regions, the low trace nutrients concentration (iron, manganese) present all year round, limit phytoplankton production.

The results presented in **Figure 3** are consistent with the different physical regimes and global climatological relationships previously demonstrated between open ocean satellite surface observations of chlorophyll and subsurface parameters of MLD, thermocline and nutricline depths at global scale (Wilson and Coles, 2005; Messié and Chavez, 2012; Brewin et al., 2014) and in the Equatorial Pacific (Turk et al., 2011; Gierach et al., 2012; Radenac et al., 2012; Lee et al., 2014). In coastal regions, phytoplankton production can be modified further by local supply of nutrients through coastal upwelling, riverine input (e.g., Turner et al., 2003) or atmospheric dust deposition (Abram et al., 2003; Jickells et al., 2005). In Polar regions, changes in phytoplankton production are tightly coupled to variations in sea-ice extent and timing of retreat, which can affect light and nutrient availability (Kahru et al., 2011, 2016; Arrigo and van Dijken, 2015).

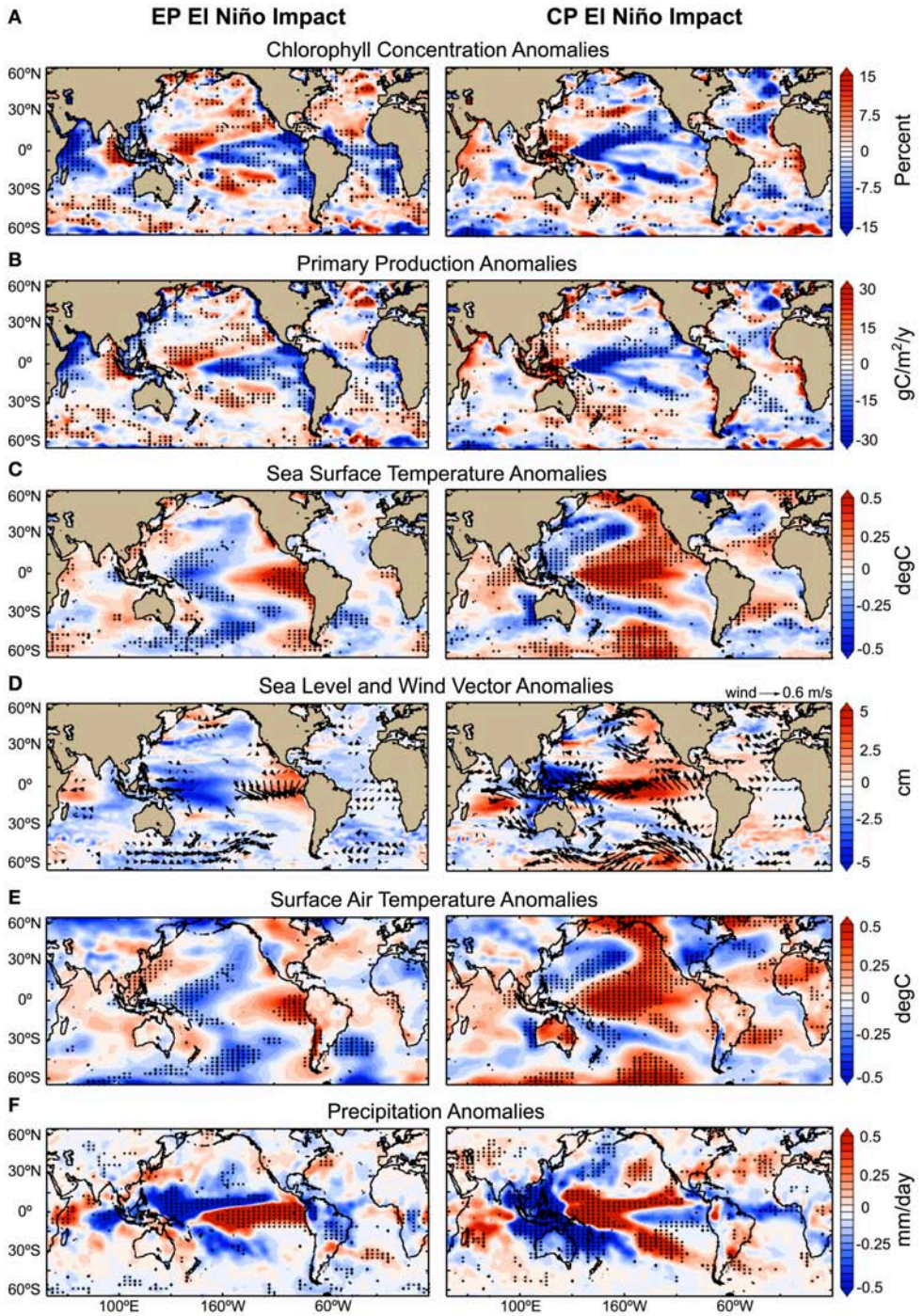


FIGURE 2 | Observed impacts of Eastern and Central Pacific El Niño on biological and physical variables during 1997–2012. Annual mean anomalies are regressed onto the (left) EP and (right) CP El Niño indices. Increase and decrease are indicated by positive (red) and negative (blue) anomalies respectively. **(A)** Chlorophyll Concentration anomalies; **(B)** Primary Production anomalies; **(C)** SST anomalies; **(D)** SL and wind anomalies; **(E)** Surface Air Temperature anomalies; and **(F)** Precipitation anomalies. In all panels, stippling indicates where the linear regression coefficients are significant at the 90% confidence level over the entire period of 1997–2012. The statistical significance of these regression coefficients was estimated according to Student *t*-test and considering the autocorrelation of the time-series. The impact values are estimated based on EP and CP indices equal to one.

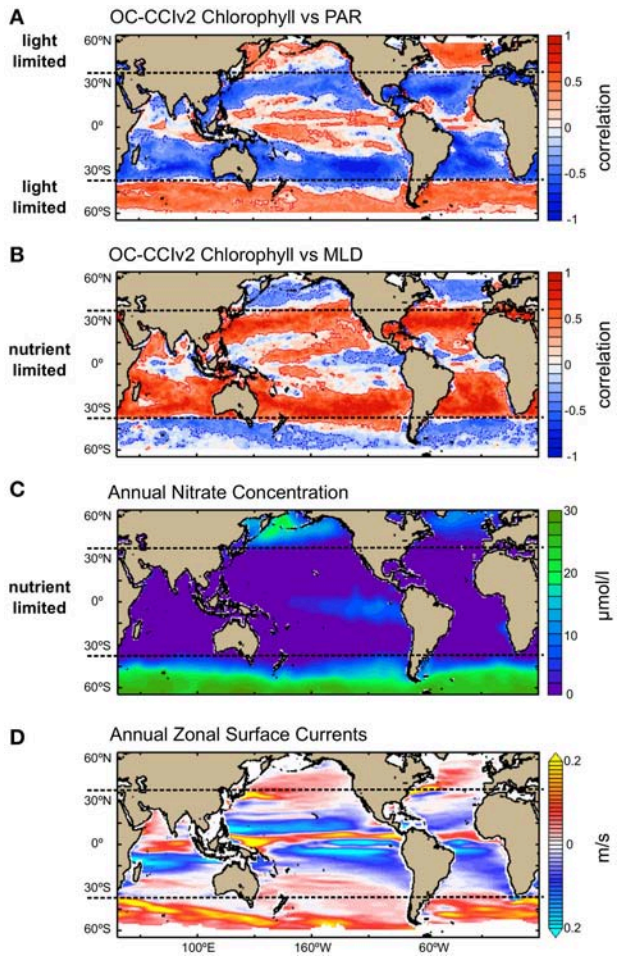


FIGURE 3 | Major factors influencing phytoplankton growth. (A) Correlation map between monthly surface chlorophyll concentration from OC-CCIv2 and photosynthetically active radiation (PAR) over the period Sep. 1997 to Dec. 2012; **(B)** Correlation map between monthly surface chlorophyll concentration from OC-CCIv2 and mixed layer depth (MLD) estimated using SODA vertical profiles of temperature over the period Sep. 1997 to Dec. 2008; **(C)** Annual average surface concentration of nitrate (µmol/l) from World Ocean Atlas Climatology; and **(D)** Annual average zonal surface currents (m/s) from OSCAR NOAA. **(A,B)** Positive correlation is indicated in red color and negative correlation in blue color. Contour lines indicate correlation coefficients that are significant at the 90% confidence level based on Pearson correlation coefficients. **(D)** Positive values indicate eastward currents and negative values westward currents.

Physical Forcing Mechanisms Associated with El Niño Variability

The results of the biological and physical responses to the EP and CP types of El Niño, which are statistically significant, are presented in **Figures 2**, **4**, and with an estimate of uncertainty in the observational product in **Figure 5**. In the Equatorial Pacific Ocean, where ENSO activity is rooted, an EP El Niño event is generated when easterly trade winds weaken in the east and westerlies prevail in the west (**Figures 2D**, **4C**), pushing warmer, nutrient-poor waters to the east (along the coast of Peru and Chile), reducing nutrient availability, leading to a decrease in chlorophyll and PP in the eastern Pacific of −12 ± 5% and −56 ± 21 TgC/y (**Figures 2A,B**, **5**). In contrast, a CP El Niño event is generated when easterly trade winds in the east and westerlies in the west are enhanced (**Figures 2D**,**4C**), pushing warmer, nutrient-poor waters to the central Equatorial Pacific, reducing nutrient availability, which is associated with a decrease in chlorophyll and PP of −14 ± 5% and −68 ± 22 TgC/y (**Figures 2A,B**, **5**). In both cases, regional decreases in phytoplankton are caused by variations in horizontal and vertical advective fluxes responsible for the transport of nutrients to the surface layer, which are driven by perturbation in the wind forcing (Ashok and Yamagata, 2009; Gierach et al., 2012; Messié and Chavez, 2012, 2013; Radenac et al., 2012). Enhanced advection is also observed during CP El Niño in the tropical Atlantic Ocean as the Equatorial easterlies intensify in the east (**Figures 2D**, **4C**), bringing warmer, nutrient-poor waters to around 15°N (Richter et al., 2012), and leading to decreases in chlorophyll and PP of −8 ± 3% and −10 ± 5 TgC/y (**Figures 2A,B**, **5**). In the Indian Ocean, Equatorial easterlies are found to intensify during EP El Niño, promoting horizontal advection of warmer and nutrient-poor waters to the western-side of the basin (**Figure 2C**) (Webster et al., 1999), resulting in decreases in chlorophyll and PP of −11 ± 4% and −82 ± 31 TgC/y (**Figures 2A,B**, **5**), whereas in the eastern-side, the observed increases in chlorophyll and PP of +7 ± 3% and +13 ± 5 TgC/y (**Figures 2A,B**, **5**) are likely to be driven by enhanced upwelling (**Figure 2C**) and atmospheric fallout from Indonesian fires and, further north in the basin, by enhanced nutrient supply from the Ganges and Brahmaputra rivers. The latter processes are consistent with the strong increase in fires (Wooster et al., 2012; Huijnen et al., 2016) and dust deposition reported off the west coast of Sumatra (Murtugudde et al., 1999; Abram et al., 2003), and the increased precipitation patterns observed over the Himalayas (**Figure 4F**). Furthermore, in some regions, these processes may not be sufficient to explain the observed variations in chlorophyll and PP, and other processes may be involved, such as atmospheric dust deposition from desert (e.g., EP El Niño impact in the Cape Verde Sea), extent and duration of sea-ice cover in the Artic (e.g., EP El Niño impact in the Bering and Labrador Seas), and iron limitation in HNLC regions (e.g., EP and CP El Niño impact in the Pacific sector of the Southern Ocean). Further information would be required to validate these forcing mechanisms.

The estimation of EP impact relies heavily on the El Niño event of 1997–1998, which was the single, important EP event that occurred within the relatively short time span of 15 years for which we have the OC-CCI data. To evaluate the impact that this single event had on our results, the correlation analyses have been rerun without the 1997–1998 event. In this case, the influence of EP El Niño on phytoplankton chlorophyll concentration remained significant in the Eastern Pacific Ocean and Western Indian Ocean, but not in the Eastern Indian Ocean region (**Figure 5**), indicating further that other regional climate oscillations are important drivers at basin scale (please see discussion in Section Implications for Climate Impact Research). In addition, the analyses of the EP and CP impact

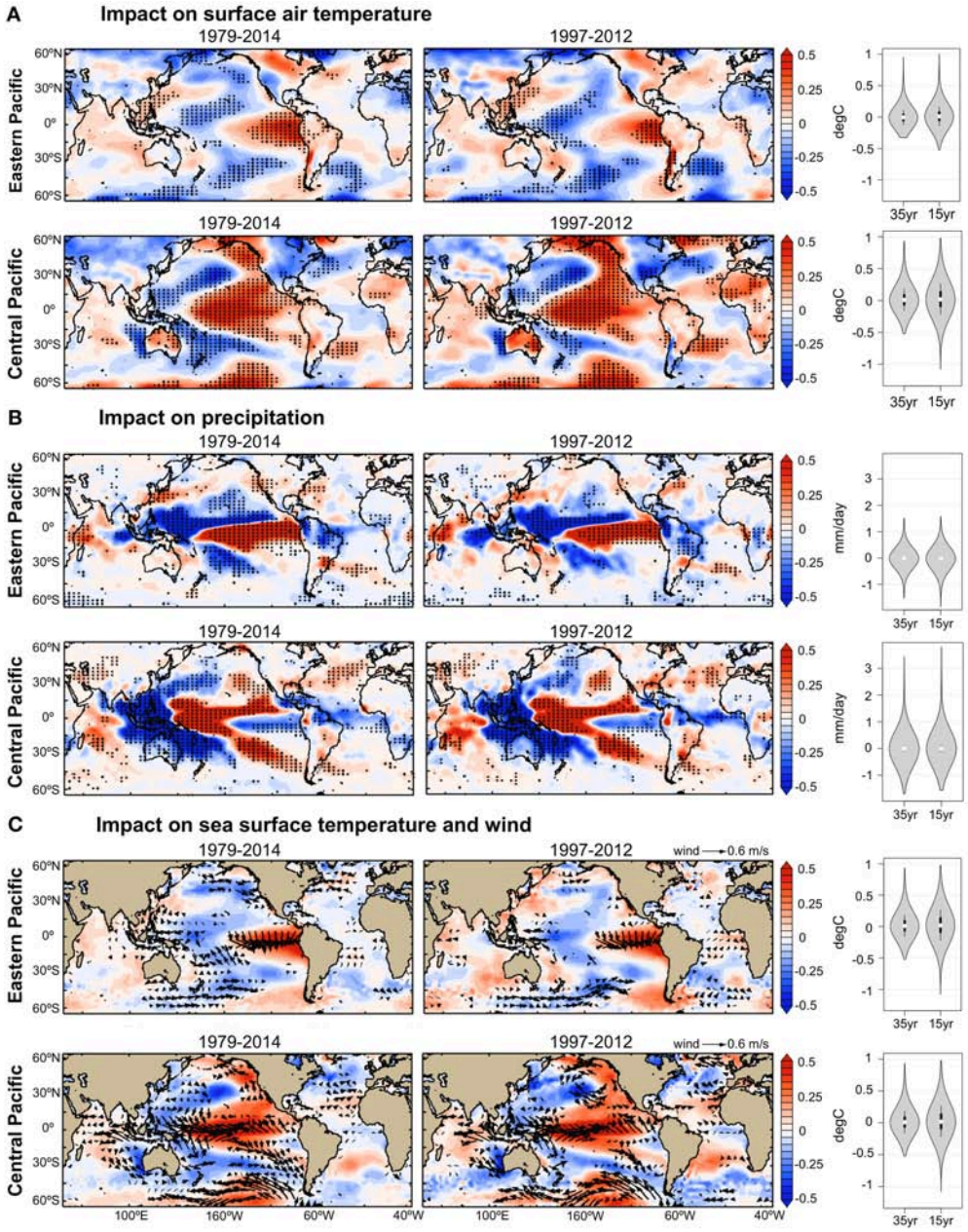


FIGURE 4 | Observed impacts of Eastern and Central Pacific El Niño on physical variables during the periods 1979–2014 and 1997–2012. Annual mean anomalies are regressed onto the annual mean EP and CP El Niño indices. Increase and decrease are indicated by positive (red) and negative (blue) anomalies respectively. In all panels, stippling indicates where the regression coefficients are significant at the 90% confidence level over the entire 35-year period (1979–2014; left panels) and 15-year period (1997–2012; central panels). The statistical significance of these regression coefficients was estimated according to Student *t*-test and considering the autocorrelation of the time-series. The probability density distributions of the regression coefficients are shown in gray shading for the 35 and 15-year periods (right panels). The white dot indicates the position of the median, and the upper and lower ends of the black rectangle indicate the upper and lower quartiles respectively. Surface air temperature and precipitation datasets are from NCEP/NCAR reanalysis, and sea surface temperature and wind datasets are from ECMWF reanalysis (see **Table 1**). The impact values are estimated based on EP and CP indices equal to one.

on physical processes are further validated on interannual to decadal timescales in **Figure 4**. The regression coefficients of the EP and CP impact estimated for the two periods 1979–2014 and 1997–2012 show similar spatial patterns and frequency distributions for the physical variables studied. This indicates that the results presented in **Figure 2** (period 1997–2012) are not skewed to the 1997–1998 EP El Niño event, and that the impact patterns are stable over multi-decadal time scales (35-year), at least for the physical variables. Since phytoplankton dynamics are at the mercy of these physical conditions, we postulate that the

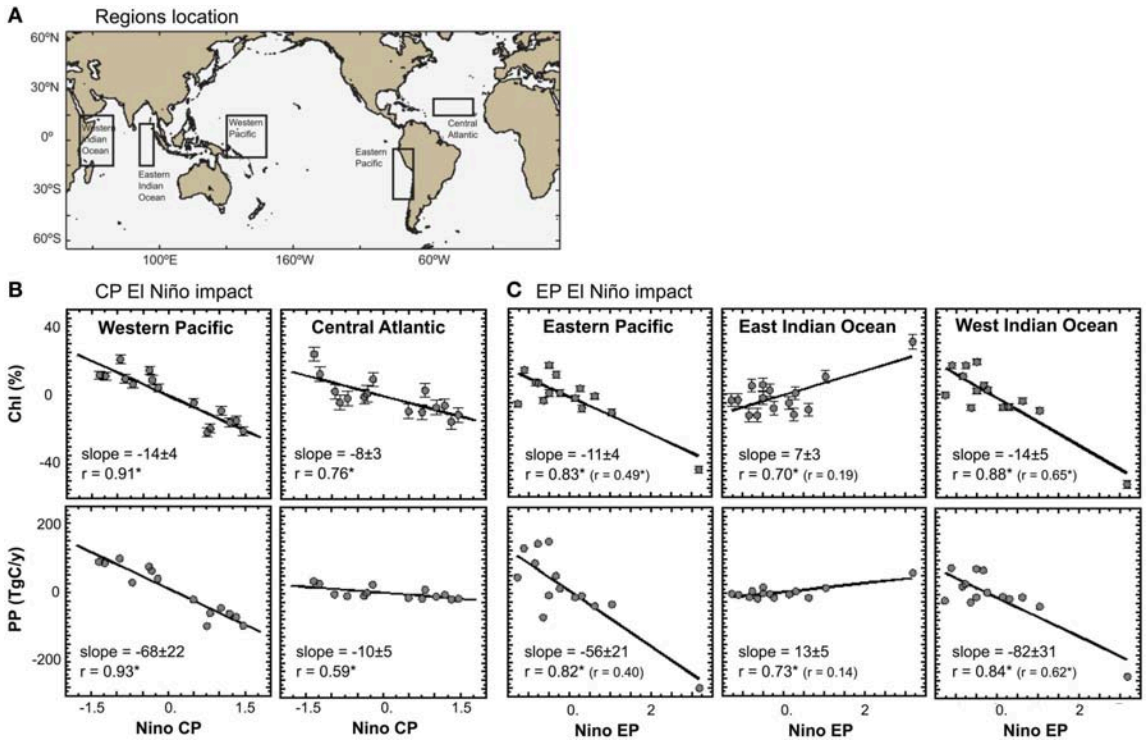


FIGURE 5 | Regional impacts of ENSO climatic perturbations during the period 1997–2012. (A) Location of the areas used in the estimation of regional EP and CP El Niño impact during the period 1997–2012; **(B)** Annual mean relative anomalies of Chlorophyll (in %) vs. EP index, and annual mean anomalies of Primary Production (in TgC/y) versus EP index; and **(C)** Annual mean relative anomalies of Chlorophyll (in %) versus CP index, and annual mean anomalies of Primary Production (in TgC/y) vs. CP index. The slope is provided ±Standard Error. Significant Pearson correlation coefficients at the 90% confidence level are indicated with a star. **(B)** The Pearson correlation coefficients shown in parenthesis are based on the analyses run without including 1997–1998 El Niño event. **(B,C)** Error bars for each point indicate the standard error based on the total number of observations in each corresponding box region. The standard error is calculated based on the root-mean-square-difference and bias observations provided in OC-CCIv2.

inference may also hold for the biological variables studied here (e.g., **Figures 2**, **5**).

Implications for Climate Impact Research

Phytoplankton have a high turn-over rate, responding to changes in their environment at scales ranging from seconds to days, and illustrating well the first-level biological response to environmental changes. At the same time, because of decadal-scale variabilities in the physical forcing fields, it is generally understood that multi-decadal, uninterrupted data are needed to evaluate the impact of climate change on marine ecosystems. Such data are only rarely available from limited *in situ* time series stations (mostly coastal). Furthermore, satellite ocean-color sensors have provided barely two decades of uninterrupted data that can be used for climate research (Sathyendranath and Krasemann, 2014). In this context, El Niño variability, together with other large-scale inter-annual variations, provides an important vehicle to study how phytoplankton in the ocean (and hence the organisms at higher trophic levels) respond to climate variability and identify the driving processes. In turn, monitoring and analysis of long-term changes in these driving processes would help us to improve understanding of projected impact of long-term climate changes on the marine ecosystem.

In the present study, we have addressed potential issues related to the collection of continuous ocean-color time-series and processing of climate-quality products, to the study of biophysical interactions, El Niño remote-forcing mechanisms and their propagation, and the diversity of El Niño events by: (i) using the longest, error-characterized, biased-corrected, climate-quality controlled, global scale merged satellite ocean-color data product from ESA Ocean-Color Climate Change Initiative project; (ii) analyzing in synergy satellite ocean-color data record and reanalysis datasets to identify the dominant mechanisms driving the biophysical interactions; (iii) characterizing local and remote influences of El Niño types on key driving variables of SST, Sea Level, wind, and precipitation; (iv) analyzing annual mean signal centered around the peak timing of El Niño activity in boreal winter; and (v) selecting EP and CP indices, which are computed to enhance differences in SST anomalies from the two most eastern and central Pacific Niño1+2 and Niño 4 regions respectively.

Our work highlights the importance of maintaining a long time series of consistent ocean-color products, to be able to evaluate the impact of climate variability on the biological fields. For example, our results on the impact of EP events could be improved when additional EP events can be incorporated

into the analyses, such that the results would no longer be so heavily dependent on a single EP event, as was the case here. More data from longer time series are also essential to explore non-linearities in biological responses, which could not be investigated here because of limited data availability.

Our analysis shows that the modification of global oceanic phytoplankton under climate change cannot be forecast with respect to changes in a single ocean property. Rather, a range of environmental properties may be involved (e.g., advection in three dimensions, wind, riverine input, atmospheric dust deposition, stratification) whose intensity may vary on a regional basis. The statistical approach to study El Niño impact applied here has permitted us to characterize a complex mosaic of biological responses illustrating that different forcing dominates in different regions. The biophysical processes driving phytoplankton production are summarized in **Figure 6** in the form of an atlas of EP and CP El Niño impact. The influence of CP and EP El Niño events can be felt in the global oceans, although the affected regions are predominantly located in the tropics and subtropics encompassing 66–67% of the total areas affected, and the remaining 33–34% are areas located in high-latitudes. In the tropics and subtropics, 35–39% of the total affected areas showed a decrease in PP associated with reduced nutrient availability during CP and EP El Niño respectively, whereas in higher latitudes 19–20% of the affected areas showed an increase in PP associated with reduced light limitation (**Figure 3**). Even though, the percent of total affected areas are relatively similar between CP and EP El niño events, the regional differentiation is marked, and may be of opposing sign (e.g., along the coast of Peru and Chile, the Benguela upwelling, the Great Barrier Reef), or affected during an EP event but not during a CP event (e.g., in the tropical eastern and western Pacific). Several process-orientated studies have further highlighted the important role played by horizontal processes (together with vertical processes) in the supply of nutrients in the surface layer, and specifically demonstrated significant impacts in Winter new primary production in the North Pacific transition zone (Ayers and Lozier, 2010), inter-annual variations of chlorophyll concentration in the Equatorial Pacific (Gierach et al., 2012; Messié and Chavez, 2013; Dave and Lozier, 2015) and the Red Sea (Raitsos et al., 2015), and decadal variations in phytoplankton abundance in the North Atlantic Subpolar Gyre (Martinez et al., 2015). Thus, both the development of statistical methods to study climate impact, and the assessment of the future evolution of regional physical forcing processes may help us to understand phytoplankton responses to climate change and improve confidence in our projection of future ecosystem state (Bopp et al., 2013; Boyd et al., 2014). The first assessment based on a biogeochemical and ecosystem model output of chlorophyll response to the EP and CP types of El Niño was shown to compare well with remotely-sensed observations in the Equatorial Pacific (Lee et al., 2014). However, the response to El Niño variability projected from biogeochemical and ecosystem models is yet to be investigated at the global scale.

Notwithstanding the dominant influence of El Niño on global climate patterns, other driving factors may enhance or weaken the observed biophysical impact. Examining links between El Niño and inter-annual and decadal climate oscillations (Di Lorenzo et al., 2010; Izumo et al., 2010) may provide further insights toward improving projection of environmental properties and associated phytoplankton responses to climate forcing at global and regional scales. The regional variations associated with El Niño may be superimposed on long-term warming trends (Bopp et al., 2013; IPCC Climate Change, 2013; Boyd et al., 2014; Kumar et al., 2016) and regional-scale oscillations at sub-seasonal and seasonal scales associated with other large-scale climate modes of variability, such as the Atlantic Multidecadal Oscillation (AMO) and Pacific Decadal Oscillation (PDO; Martinez et al., 2009), the North Pacific Gyre Oscillation (NPGO; Di Lorenzo et al., 2008, 2010; Messié and Chavez, 2013), the monsoon and Indian Ocean Dipole (IOD; Saji et al., 1999; Ashok et al., 2007; Izumo et al., 2010; Brewin et al., 2012; Currie et al., 2013). As a result, the regional climate response is not a simple function of the strength and centroid location of an El Niño event. Further, the regional patterns observed using EP and CP El Niño indices may be sufficient to explain only a fraction of all the regional variations on a year-to-year basis (except perhaps where El Niño is likely to dominate the variability of the system such as in the Equatorial Pacific region). For instance, in the Indian Ocean, the ENSO and IOD indices can account for ∼30% and 12% respectively of the regional variations in SST (Saji et al., 1999), and years of co-occurrence of positive IOD and El Niño events may provide positive feedbacks to the SST (Kumar et al., 2016). Therefore, some apparent differences will show between the observed impact of El Niño on biological and physical variables, and the corresponding anomalies.

Implications for the Oceanic Ecosystem and Carbon Cycle

Phytoplankton are at the base of the food chain and transfer energy to higher trophic levels. This transfer of energy has a knock-on effect on fisheries and dependent human societies especially in highly productive and coastal upwelling regions, as well as coral reef ecosystems. The larvae of many marine species graze on phytoplankton during this most vulnerable stage of their lives. Hence, changes in phytoplankton population associated with climate variability may propagate rapidly up the marine food chain and profoundly alter the functioning of marine ecosystems (Platt et al., 2003; Edwards and Richardson, 2004; Lo-Yat et al., 2011). In addition, changes in environmental conditions associated with EP and CP El Niño events have been shown to impact mesozooplankton community with variable time lags in the northern California Current, which in turn can affect top down control on phytoplankton, and disrupt the pelagic food chain (Fisher et al., 2015). Following EP and CP El Niño events, quite different impact on commercially important fisheries have been reported in anchovy catches in the Humboldt Current Large Marine Ecosystem (Jackson et al., 2011) and tuna catches in the Indian Ocean (Kumar et al., 2014). In coral reef ecosystems, changes in phytoplankton population and mass bleaching following an El Niño event can critically affect fisheries, recreation, and tourism services (Hoegh-Guldberg, 1999; Abram et al., 2003; Lo-Yat et al., 2011). Recent analysis in the Andaman Sea, southeast Bay of Bengal, has further

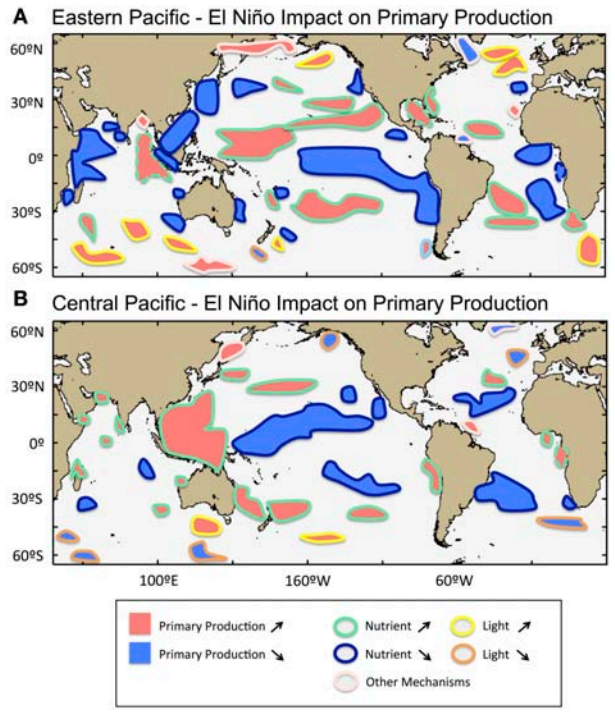


FIGURE 6 | Schematic atlas of the influence of El Niño variability on oceanic phytoplankton in the global oceans. (A) Eastern Pacific El Niño influence, and **(B)** Central Pacific El Niño influence. Increase and decrease in primary production (PP) are indicated in red and blue filled colors respectively. The color contours provide information about the controlling biophysical mechanisms (which are described in Section Physical Forcing Mechanisms Associated with El Niño Variability and Mapped in **Figure 3**): (yellow) PP is nutrient-limited; (dark blue) PP is enhanced by nutrient availability; (orange) PP is light-limited; (turquoise) PP is enhanced by light availability; (light pink or dashed contour) PP may be further controlled by other mechanisms, such as sea-ice melting, atmospheric dust deposition and availability of trace nutrients. The contour delineation of the influence of EP and CP El Niño is generated based on information displayed in **Figure 2** and Supplementary Figure 4.

demonstrated that differences both in intensity and timing of SST warming associated with EP and CP El Niño events, can determine the extent of mass coral bleaching (Lix et al., 2016). In this context, regional differentiation of the impact of each type of El Niño events (**Figure 6**) may provide important information to delineate and establish protected coral reef and fishing areas to facilitate their recovery.

The oceanic carbon sink is part of a very active, natural cycle, in which phytoplankton in the surface layer of the ocean fix, by photosynthesis, dissolved CO_2 in the water into organic matter, some of which subsequently sinks below the mixed layer. Through the associated decrease in the partial pressure of CO_2 in the surface ocean, phytoplankton contribute to the drawdown of dissolved CO_2 from the ocean surface layer (Hauck et al., 2015), which in turn help to modulate the increase in anthropogenic atmospheric CO_2. The estimated El-Niño-driven changes in PP at the regional scale can be considerable, reaching values of -57 ± 21 and -68 ± 22 TgC/y in the Eastern and Central Equatorial Pacific Ocean during EP and CP types of El Niño respectively (**Figure 5**). However, to provide a more complete picture on the influence on the carbon cycle, further investigations are required to quantify the impact of El Niño on carbon export and associated changes in air-sea CO_2 fluxes. The buffering action of the ocean in the carbon cycle is non-linear—it varies with the water temperature (solubility pump), alkalinity (carbonate pump), biological productivity and demineralization (biological pump); the impact on environmental and ecosystem properties must be evaluated at the appropriate scale to allow investigation of the underlying mechanisms driving the variability in the ocean carbon cycle.

As the frequency of extreme El Niño events and the relative frequency of occurrence of CP-El Niño/EP-El Niño are projected to increase under climate warming (Yeh et al., 2009; Lee and McPhaden, 2010; Cai et al., 2015), it is essential to refine our regional assessment of climate impact associated with El Niño variability. The atlas of impact of CP and EP types of El Niño on oceanic phytoplankton (**Figure 6**) can be used for societal benefit. It provides key climate impact information that can allow us to better inform fisheries management on possible risks and opportunities associated with El Niño events, and support more effectively mitigation and adaptation plans for local fisheries-dependent societies. The atlas information can also provide observational basis to test model predictions of the impact of climate change on the marine ecosystem. Finally, from a biogeochemical perspective, such insights on El Niño variability impact are needed to improve our understanding of the buffering capacity of the oceanic carbon cycle under climate change.

AUTHOR CONTRIBUTIONS

MFR designed and implemented the research. MFR, SS, RB, TJ provided materials and analysis tools. MFR, SS, RB, TJ, DR, and TP discussed the results and contributed to the writing of the manuscript.

FUNDING

MFR is funded through a European Space Agency Living Planet Fellowship Grant Ref Number (CCI-LPF-EOPS-MM-16-0078). SS, TJ, and TP are funded through ESA Ocean Color Climate Change Initiative program. SS, MFR, RB, and DR are funded through the NERC's UK National Centre for Earth Observation.

ACKNOWLEDGMENTS

This work is a contribution to the European Space Agency Ocean Color Climate Change Initiative (ESA OC-CCI), the European Space Agency Living Planet Fellowship program (CLIMARECOS), and to the NERC National Centre for Earth Observation (NCEO). The authors thank the ESA CCI teams for providing OC-CCI chlorophyll data, SL-CCI sea level data, and NCEO-ESA-SST-CCI sea surface temperature data. The authors further acknowledge TWAP for providing primary production data; NASA for providing SeaWiFS and MODIS PAR data; NOAA for providing NCEP/NCAR air surface

temperature and precipitation reanalysis data; ECMWF for providing ERA Interim wind and sea surface temperature reanalysis data; and SODA for providing ocean temperature reanalysis data. The authors thank James Dingle for technical support with the OC-CCI data processing, Stéphane Saux-Picart for help in processing the mixed layer depth, Sang-Wook Yeh for discussion about El Niño phenomenon, and Eleni Papathanasopoulou for discussion about socio-economic impacts. The authors wish to acknowledge use of the Ferret program for analysis and graphics in this paper—Ferret is a product of NOAA's Pacific Marine Environmental Laboratory (information is available at http://ferret.pmel.noaa.gov/Ferret/). The authors would also like to acknowledge the Nature Method on-line web-tool BoxPlotR (http://boxplot.tyerslab.com/), which was used to generate the boxplots, and Dimitrios Kleftogiannis for information about boxplot analysis tools. We acknowledge the two reviewers for providing constructive comments on our manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmars.2017.00133/full#supplementary-material

REFERENCES

Ablain, M., Cazenave, A., Larnicol, G., Balmaseda, M., Cipollini, P., Faugère, Y., et al. (2015). Improved sea level record over the satellite altimetry era (1993–2010) from the Climate Change Initiative project. *Ocean Sci.* 11, 67–82. doi: 10.5194/os-11-67-2015

Abram, N. J., Gagan, M. K., McCulloch, M. T., Chappell, J., and Hantoro, W. S. (2003). Coral reef death during the 1997 Indian Ocean Dipole linked to Indonesian wildfires. *Science* 301, 952–955. doi: 10.1126/science.1083841

Antoine, D., André, J.-M., and Morel, A. (1996). Oceanic primary production: 2. Estimation at global scale from satellite (Coastal Zone Color Scanner) chlorophyll. *Glob. Biogeochem. Cycles* 10, 57–69. doi: 10.1029/95GB02832

Arrigo, K. R., and van Dijken, G. L. (2015). Continued increases in Arctic Ocean primary production. *Prog. Oceanogr.* 136, 60–70. doi: 10.1016/j.pocean.2015.05.002

Ashok, K., Behera, S. K., Rao, S. A., Weng, H., and Yamagata, T. (2007). El Niño Modoki and its possible teleconnection. *J. Geophys. Res.* 112:C11007. doi: 10.1029/2006JC003798

Ashok, K., and Yamagata, T. (2009). The El Niño with a difference. *Nature* 461, 481–484. doi: 10.1038/461481a

Ayers, J. M., and Lozier, M. S. (2010). Physical controls on the seasonal migration of the North Pacific transition zone chlorophyll front. *J. Geophys. Res.* 115:C05001. doi: 10.1029/2009JC005596

Behrenfeld, M. J., Boss, E., Siegel, D. A., and Shea, D. M. (2005). Carbon-based ocean productivity and phytoplankton physiology from space. *Glob. Biogeochem. Cycles* 19:GB1006:1-14. doi: 10.1029/2004gb002299

Behrenfeld, M. J., O'Malley, R. T., Siegel, D. A., McClain, C. R., Sarmiento, J. L., Feldman, G. C., et al. (2006). Climate-driven trends in contemporary ocean productivity. *Nature* 444, 752–755. doi: 10.1038/nature05317

Behrenfeld, M. J., Randerson, J. T., McClain, C. R., Feldman, G. C., Los, S. O., Tucker, C. J. et al. (2001). Biospheric primary production during an ENSO transition. *Science* 291, 2594–2595. doi: 10.1126/science.1055071

Bopp, L., Resplandy, L., Orr, J. C., Doney, S. C., Dunne, J. P., Gehlen, M., et al. (2013). Multiple stressors of ocean ecosystems in the 21st century: projections with CMIP5 models. *Biogeosciences* 10, 6225–6245. doi: 10.5194/bg-10-6225-2013

Boyd, P. W., Lennartz, S. T., Glover, D. M., and Doney, S. C. (2014). Biological ramifications of climate-change-mediated oceanic multi-stressors. *Nat. Clim. Change* 5, 71–79. doi: 10.1038/nclimate2441

Boyer, T. P., Antonov, J. I., Baranova, O. K., Coleman, C., Garcia, A., Grodsky, D. R., et al. (2013). *NOAA Atlas NESDIS 72.* World Ocean Database 2013. Silver Spring, MD.

Brewin, R. J. W., Hirata, T., Hardman-Mountford, N. J., Lavender, S. J., Sathyendranatha, S., and Barlow, R. (2012). The influence of the Indian Ocean Dipole on interannual variations in phytoplankton size structure as revealed by Earth Observation. *Deep Sea Res. II 77–80*, 117–127. doi: 10.1016/j.dsr2.2012.04.009

Brewin, R. J. W., Mélin, F., Sathyendranath, S., Steinmetz, F., Chuprin, A., and Grant, M. (2014). On the temporal consistency of chlorophyll products derived from three ocean-colour sensors. *ISPRS J. Photogramm. Remote Sens.* 97, 171–184. doi: 10.1016/j.isprsjprs.2014.08.013

Cai, W., Santoso, A., Wang, G., Yeh, S.-Y., An, S.-I., et al. (2015). ENSO and greenhouse warming. *Nature* 5, 849–859. doi: 10.1038/nclimate2743

Capotondi, A., Wittenberg, A. T., Newman, M., Di Lorenzo, E., Yu, J.-Y., et al. (2015). Understanding ENSO diversity. *Bull. Am. Meteorol. Soc.* 96, 921–938. doi: 10.1175/BAMS-D-13-00117.1

Carton, J. A., and Giese, B. S. (2008). A reanalysis of ocean climate using Simple Ocean Data Assimilation (SODA). *Monthly Weather Rev.* 136, 2999–3017. doi: 10.1175/2007MWR1978.1

Chavez, F. P., Messié, M., and Pennington, J. T. (2011). Marine primary production in relation to climate variability and change. *Ann. Rev. Mar. Sci.* 3, 227–260. doi: 10.1146/annurev.marine.010908.163917

Couto, A. B., Holbrook, N. J., and Maharaj, A. M. (2013). Unravelling eastern pacific and central pacific ENSO contributions in south pacific chlorophyll-a variability through remote sensing. *Remote Sens.* 5, 4067–4087. doi: 10.3390/rs5084067

Currie, J. C., Lengaigne, M., Vialard, J., Kaplan, D. M., Aumont, O., Naqvi, S. W. A., et al. (2013). Indian Ocean dipole and El Niño/southern oscillation impacts on regional chlorophyll anomalies in the Indian Ocean. *Biogeosciences* 10, 6677–6698.

Dave, A. C., and Lozier, M. S. (2015). The impact of advection on stratification and chlorophyll variability in the equatorial Pacific. *Geophys. Res. Lett.* 42, 4523–4531. doi: 10.1002/2015GL063290

de Boyer Montégut, C., Madec, G., Fischer, A. S., Lazar, A., and Ludicone, D. (2004). Mixed layer depth over the global ocean: an examination of profile data and a profile-based climatology. *J. Geophys. Res.* 109:C12003. doi: 10.1029/2004jc002378

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., et al. (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* 137, 553–597. doi: 10.1002/qj.828

Di Lorenzo, E., Cobb, K. M., Furtado, J. C., Schneider, N., Anderson, B. T., et al. (2010). Central Pacific El Niño and decadal climate change in the North Pacific Ocean. *Nat. Geosci.* 3, 762–765. doi: 10.1038/ngeo984

Di Lorenzo, E., Schneider, N., Cobb, K. M., Franks, P. J. S., Chhak, K., et al. (2008). North Pacific Gyre Oscillation links ocean climate and ecosystem change. *Geophys. Res. Lett.* 35:L08607. doi: 10.1029/2007GL032838

Edwards, M., and Richardson, A. J. (2004). Impact of climate change on marine pelagic phenology and trophic mismatch. *Nature* 430, 881–884. doi: 10.1038/nature02808

Fisher, J. L., Peterson, W. T., and Rykaczewski, R. R. (2015). The impact of El Niño events on the pelagic food chain in the northern California Current. *Glob. Change Biol.* 21, 4401–4414. doi: 10.1111/gcb.13054

Frouin, R., McPherson, J., Ueyoshi, K., and Franz, B. A. (2012). A time series of photosynthetically available radiation at the ocean surface from SeaWiFS and MODIS data. *Remote Sens. Mar. Environ. II Proc. SPIE* 8525, 852519. doi: 10.1117/12.981264

Gierach, M. M., Lee, T., Turk, D., and McPhaden, M. J. (2012). Biological response to the 1997-98 and 2009-10 El Niño events in the equatorial Pacific Ocean. *Geophys. Res. Lett.* 39:L10602. doi: 10.1029/2012GL051103

Gordon, R. M., Coale, K. H., and Johnson, K. S. (1997). Iron distributions in the equatorial Pacific: Implications for new production. *Limnol. Oceanogr.* 42, 419–431. doi: 10.4319/lo.1997.42.3.0419

Gregg, W., and Rousseaux, C. S. (2014). Decadal trends in global pelagic ocean chlorophyll: a new assessment integrating multiple satellites, *in situ* data, and models. *J. Geophys. Res. Oceans* 119, 5921–5933. doi: 10.1002/2014JC010158

Hauck, L., Völker, C., Wolf-Gladrow, D. A., Laufkötter, C., Vogt, M., et al. (2015). On the Southern Ocean CO2 uptake and the role of the biological carbon pump in the 21st century. *Glob. Biogeochem. Cycles* 29, 1451–1470. doi: 10.1002/2015GB005140

Hoegh-Guldberg, O. (1999). Climate change, coral bleaching and the future of the world's coral reefs. *Mar. Freshwater Res.* 50, 839–866. doi: 10.1071/MF99078

Hoegh-Guldberg, O., Cai, R., Brewer, P. G., Fabry, V. J., Hilmi, K., et al. (2014). "IPCC Climate Change 2014: Impacts, Adaptation, and Vulnerability Chapter 30 – *The Ocean*," in *Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge, UK; New York, NY: Cambridge University Press).

Huijnen, V., Wooster, M. J., Kaiser, J. W., Gaveau, D. L. A., Flemming, J., Parrington, M., et al. (2016). Fire carbon emissions over maritime southeast Asia in 2015 largest since 1997. *Sci. Rep.* 6:26886. doi: 10.1038/srep26886

IPCC Climate Change (2013). "The Physical Science Basis Summary for Policymakers," in *Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, eds T. F. Stocker, D. Quin, G. -K. Plattner, M. Tignor, S. K. Allen, et al. (Cambridge, UK; New York, NY: Cambridge University Press).

Izumo, T., Vialard, J., Lengaigne, M., De Boyer Montegut, C., Behera, S. K., Luo, J. J., et al. (2010). Influence of the state of the Indian Ocean Dipole on the following year's El Niño. *Nat. Geosci.* 3, 168–172. doi: 10.1038/ngeo760

Jackson, T., Bouman, H. A., Sathyendranath, S., and Devred, E. (2011). Regional-scale changes in diatom distribution in the Humboldt upwelling system as revealed by remote sensing: implications for fisheries. *ICES J. Mar. Sci.* 68, 729–736. doi: 10.1093/icesjms/fsq181

Jickells, T. D., An, Z. S., Andersen, K. K., Baker, A. R., Bergametti, G., et al. (2005). Global iron connections between desert dust, ocean biogeochemistry, and climate. *Science* 308, 67–71. doi: 10.1126/science.1105959

Kahru, M., Brotas, V., Manzano-Sarabia, M., and Mitchell, B. G. (2011). Are phytoplankton blooms occurring earlier in the Arctic? *Glob. Change Biol.* 17, 1733–1739. doi: 10.1111/j.1365-2486.2010.02312.x

Kahru, M., Lee, Z., Mitchell, B. G., and Nevison, C. D. (2016). Effects of sea ice cover on satellite-detected primary production in the Arctic Ocean. *Biol. Lett.* 12:20160223. doi: 10.1098/rsbl.2016.0223

Kalnay, E., Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., et al. (1996). The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteor. Soc.* 77, 437–470.

Kao, H.-Y., and Yu, J.-Y. (2009). Contrasting Eastern-Pacific and Central-Pacific types of ENSO. *J. Clim.* 22, 615–632. doi: 10.1175/2008JCLI2309.1

Kim, H.-M., Webster, P. J., and Curry, J. A. (2009). Impact of shifting patterns of Pacific Ocean warming on North Atlantic tropical cyclones. *Science* 325, 77–80. doi: 10.1126/science.1174062

Kug, J.-S., Jin, F.-F., and An, S.-I. (2009). Two types of El Niño events: cold tongue El Niño and warm pool El Niño. *J. Clim.* 22, 1499–1515. doi: 10.1175/2008JCLI2624.1

Kumar, P. K. D., Steeven Paul, Y., Muraleedharan, K. R., Murty, V. S. N., and Preenu, P. N. (2016). Comparison of long-term variability of Sea Surface Temperature in the Arabian Sea and Bay of Bengal. *Reg. Stud. Mar. Sci.* 3, 67–75. doi: 10.1016/j.rsma.2015.05.004

Kumar, P. S., Pillai, G. N., and Manjusha, U. (2014). El Niño Southern Oscillation (ENSO) impact on tuna fisheries in Indian Ocean. *SpringerPlus* 3:591. doi: 10.1186/2193-1801-3-591

Larkin, N. K., and Harrison, D. E. (2005). On the definition of El Niño and associated seasonal average U.S. weather anomalies. *Geophys. Res. Lett.* 32:L13705. doi: 10.1029/2005GL022738

Lee, K.-W., Yeh, S.-W., Kug, J.-S., and Park, J.-Y. (2014). Ocean chlorophyll response to two types of El Niño events in an ocean-biogeochemical coupled model. *J. Geophys. Res. Oceans* 119, 933–952. doi: 10.1002/2013JC009050

Lee, T., and McPhaden, M. J. (2010). Increasing intensity of El Niño in the central-equatorial Pacific. *Geophys. Res. Lett.* 37:L14603. doi: 10.1029/2010GL044007

Le Quéré, C., Moriarty, R., Andrew, R. M., Peters, G. P., Ciais, P., et al. (2015). Global carbon budget 2014. *Earth Syst. Sci. Data* 7, 47–85. doi: 10.5194/essd-7-47-2015

L'Heureux, M., Takahashi, K., Watkins, A., Barnston, A., Becker, E., Di Liberto, T., et al. (in press). Observing and predicting the 2015-16 El Niño. *Bull. Amer. Meteor. Soc.* doi: 10.1175/BAMS-D-16-0009.1

Li, G., Ren, B., Yang, C., and Zheng, J. (2010). Indices of el niño and el niño modoki: an improved el niño modoki index. *Adv. Atmos. Sci.* 27, 1210–1220. doi: 10.1007/s00376-010-9173-5

Lin, H., and Derome, J. A. (1998). Three-year lagged correlation between the North Atlantic Oscillation winter conditions over the North Pacific and North America. *Geophys. Res. Lett.* 25, 2829–2832. doi: 10.1029/98GL52217

Lix, J. K., Venkatesan, R., Grinson, G., Rao, R. R., Jineesh, V. K., et al. (2016). Differential bleaching of corals based on El Niño type and intensity in the Andaman Sea, southeast Bay of Bengal. *Environ. Monit. Assess.* 188:175. doi: 10.1007/s10661-016-5176-8

Longhurst, A. R. (1998). *Ecological Geography of the Sea, 2nd Edn*. San Diego, CA: Academic Press.

Longhurst, A., Sathyendranath, S., Platt, T., and Caverhill, C. (1995). An estimate of global primary production in the ocean from satellite radiometer data. *J. Plankton Res.* 17, 1245–1271. doi: 10.1093/plankt/17.6.1245

Lo-Yat, A., Simpson, S. D., Meekan, M., Lecchini, D., Martinez, E., and Galzin, R. (2011). Extreme climatic events reduce ocean productivity and larval supply in a tropical reef ecosystem. *Glob. Change Biol.* 17, 1695–1702. doi: 10.1111/j.1365-2486.2010.02355.x

Martinez, E., Antoine, D., D'Ortenzio, F., and Gentili, B. (2009). Climate-driven basin-scale decadal oscillations of oceanic phytoplankton. *Science* 326, 1253–1256. doi: 10.1126/science.1177012

Martinez, E., Raitsos, D. E., and Antoine, D. (2015). Warmer, deeper and greener mixed layers in the north Atlantic subpolar gyre over the last 50 years. *Glob. Change Biol.* 22, 604–612. doi: 10.1111/gcb.13100

McPhaden, M. J., Zebiak, S. E., and Glantz, M. H. (2006). ENSO as an integrating concept in Earth science. *Science* 314, 1740–1745. doi: 10.1126/science.1132588

Merchant, C. J., Embury, O., Roberts-Jones, J., Fiedler, E., Bulgin, C. E., et al. (2014). Sea surface temperature datasets for climate applications from Phase 1 of the European Space Agency Climate Change Initiative (SST CCI). *Geosci. Data J.* 1, 179–191. doi: 10.1002/gdj3.20

Messié, M., and Chavez, F. P. (2012). A global analysis of ENSO synchrony: the oceans' biological response to physical forcing. *J. Geophys. Res.* 117:C09001. doi: 10.1029/2012jc007938

Messié, M., and Chavez, F. P. (2013). Physical-biological synchrony in the global ocean associated with recent variability in the central and western equatorial Pacific. *J. Geophys. Res.* 118, 3782–3794. doi: 10.1002/jgrc.20278

Moore, C. M., Mills, M. M., Arrigo, K. R., Berman-Frank, I., Bopp, L., et al. (2013). Processes and patterns of oceanic nutrient limitation. *Nat. Geosci.* 6, 701–710. doi: 10.1038/ngeo1765

Murtugudde, R. G., Signorini, S. R., Christian, J. R., Busalacchi, A. J., McClain, C. R., and Picaut, J. (1999). Ocean color variability of the tropical Indo-Pacific basin observed by SeaWiFS during 1997-1998. *J. Geophys. Res.* 104, 18351–18366. doi: 10.1029/1999JC900135

Paek, H., Yu, J.-Y., and Qian, C. (2017). Why were the 2015/2016 and 1997/1998 extreme El Niños different? *Geophys. Res. Lett.* 44, 1848–1856. doi: 10.1002/2016GL071515

Platt, T., Fuentes-Yaco, C., and Frank, K. (2003). Spring algal bloom and larval fish survival. *Nature* 423, 398–399. doi: 10.1038/423398b

Platt, T., and Sathyendranath, S. (1988). Oceanic primary production: estimation by remote sensing at local and regional scales. *Science* 241, 1613–1620. doi: 10.1126/science.241.4873.1613

Platt, T., and Sathyendranath, S. (2008). Ecological indicators for the pelagic zone of the ocean from remote sensing. *Remote Sens. Environ.* 112, 3426–3436. doi: 10.1016/j.rse.2007.10.016

Racault, M.-F., Le Quéré, C., Buitenhuis, E., Sathyendranath, S., and Platt, T. (2012). Phytoplankton phenology in the global ocean. *Ecol. Indic.* 14, 152–163. doi: 10.1016/j.ecolind.2011.07.010

Racault, M.-F., Raitsos, D. E., Berumen, M., Sathyendranath, S., Platt, T., and Hoteit, I. (2015). Phytoplankton phenology indices in coral reef ecosystems: application to ocean-color observations in the Red Sea. *Remote Sens. Environ.* 160, 222–234. doi: 10.1016/j.rse.2015.01.019

Racault, M.-F., Sathyendranath, S., Menon, N., and Platt, T. (2017). Phenological responses to ENSO in the global oceans. *Surveys Geophys.* 38, 277–293. doi: 10.1007/s10712-016-9391-1

Racault, M.-F., Sathyendranath, S., and Platt, T. (2014). Impact of missing data on the estimation of ecological indicators from satellite ocean-colour time-series. *Remote Sens. Environ.* 152, 15–28. doi: 10.1016/j.rse.2014.05.016

Radenac, M.-H., Léger, F., Singh, A., and Delcroix, T. (2012). Sea surface chlorophyll signature in the tropical Pacific during eastern and central Pacific ENSO events. *J. Geophys. Res.* 117:C04007. doi: 10.1029/2011JC007841

Raitsos, D. E., Yi, X., Platt, T., Racault, M.-F., Brewin, R. J. W., et al. (2015). Monsoon oscillations regulate fertility of the Red Sea. *Geophys. Res. Lett.* 42, 855–862. doi: 10.1002/2014GL062882

Richter, I., Behera, S. K., Masumoto, Y., Taguchi, B., Sasaki, H., and Yamagata, T. (2012). Multiple causes of interannual sea surface temperature variability in the equatorial Atlantic Ocean. *Nat. Geosci.* 6, 43–47. doi: 10.1038/ngeo1660

Saji, N. H., Goswami, B. N., Vinayachandran, P. N., and Yamagata, T. (1999). A dipole mode in the tropical Indian Ocean. *Nature* 401, 360–363. doi: 10.1038/43854

Sathyendranath, S., and Krasemann, H. (2014). *Climate Assessment Report: Ocean Colour Climate Change Initiative (OC-CCI) – Phase One.* Available online at: http://www.esa-oceancolour-cci.org/?q=documents

Steinmetz, F., Deschamps, P., and Ramon, D. (2011). Atmospheric correction in presence of sun glint: application to MERIS. *Opt. Express* 19, 571–587. doi: 10.1364/OE.19.009783

Takahashi, K., Montecinos, A., Goubanova, K., and Dewitte, B. (2011). ENSO regimes: Reinterpreting the canonical and Modoki El Niño. *Geophys. Res. Lett.* 38:l10704. doi: 10.1029/2011GL047364

Topping, J. (1972). *Errors of Observation and Their Treatment, 4th Edn.*, ed. J. Topping (Whistable: Chapman & Hall), 72–114.

Trenberth, K. E., and Stepaniak, D. P. (2001). Indices of el niño evolution. *J. Climate* 14, 1697–1701. doi: 10.1175/1520-0442(2001)014

Turk, D., Meinen, C. S., Antoine, D., McPhaden, M. J., and Lewis, M. R. (2011). Implications of changing El Niño patterns for biological dynamics in the equatorial Pacific Ocean. *Geophys. Res. Lett.* 38:L23603. doi: 10.1029/2011GL049674

Turner, R. E., Rabalais, N. N., Justic, D., and Dortch, Q. (2003). Global patterns of dissolved N, P and Si in large rivers. *Biogeochemistry* 64, 297–317. doi: 10.1023/A:1024960007569

Vantrepotte, V., and Mélin, F. (2011). Inter-annual variations in the SeaWiFS global chlorophyll a concentration (1997–2007). *Deep Sea Res. I* 58, 429–441. doi: 10.1016/j.dsr.2011.02.003

Webster, P. J., Moore, A. M., Loschnigg, J. P., and Leben, R. R. (1999). Coupled ocean-atmosphere dynamics in the Indian Ocean during 1997-98. *Nature* 401, 356–360. doi: 10.1038/43848

Wilson, C., and Coles, V. J. (2005). Global climatological relationships between satellite biological and physical observations and upper ocean properties. *J. Geophys. Res.* 110:C10001. doi: 10.1029/2004JC002724

Wolter, K., and Timlin, M. S. (1993). "Monitoring ENSO in COADS with a seasonally adjusted principal component index," in *Proceedings of the 17th Climate Diagnostics Workshop* (Norman, OK: NOAA/N MC/CAC, NSSL, Oklahoma Climate Survey, CIMMS and the School of Meteorology, University Oklahoma), 52–57.

Wooster, M. J., Perry, G. L. W., and Zoumas, A. (2012). Fire, drought and El Niño relationships on Borneo (Southeast Asia) in the pre-MODIS era (1980-2000). *Biogeosciences* 9, 317–340. doi: 10.5194/bg-9-317-2012

Xie, P., and Arkin, P. A. (1997). Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Amer. Meteor. Soc.* 78, 2539–2558.

Yeh, S., Kug, J.-S., Dewitte, B., Kwon, M.-H., Kirtman, B. P., and Jin, F. F. (2009). El Niño in a changing climate. *Nature* 461, 511–514. doi: 10.1038/nature08316

Yoder, J., and Kennelly, M. (2003). Seasonal and ENSO variability in global ocean phytoplankton chlorophyll derived from 4 years of SeaWiFS measurements. *Glob. Biogeochem. Cycles* 17:1112. doi: 10.1029/2002GB001942

Yu, J.-Y. (2016). *Monthly CP Index and EP Index.* Available online at: https://www.ess.uci.edu/∼yu/2OSC/Retrieved date March 2016.

Yu, J.-Y., and Kao, H.-Y. (2007). Decadal changes of ENSO persistence barrier in SST and ocean heat content indices: 1958–2001. *J. Geophys. Res.* 112, D13106. doi: 10.1029/2006JD007654

Yu, J.-Y., and Zou, Y. (2013). The enhanced drying effect of Central-Pacific El Niño on US winter. *Environ. Res. Lett.* 8:014019. doi: 10.1088/1748-9326/8/1/014019

Yu, J.-Y., Zou, Y., Kim, S. T., and Lee, T. (2012). The changing impact of El Niño on US winter temperatures. *Geophys. Res. Lett.* 39, L15702. doi: 10.1029/2012GL052483

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Racault, Sathyendranath, Brewin, Raitsos, Jackson and Platt. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Marine Science

ORIGINAL RESEARCH
published: 11 May 2017
doi: 10.3389/fmars.2017.00140

The OLCI Neural Network Swarm (ONNS): A Bio-Geo-Optical Algorithm for Open Ocean and Coastal Waters

Martin Hieronymi[1]*, *Dagmar Müller*[1†] and *Roland Doerffer*[1,2]

[1] Department of Remote Sensing, Institute of Coastal Research, Helmholtz-Zentrum Geesthacht, Geesthacht, Germany,
[2] Brockmann Consult GmbH, Geesthacht, Germany

The processing scheme of a novel in-water algorithm for the retrieval of ocean color products from Sentinel-3 OLCI is introduced. The algorithm consists of several blended neural networks that are specialized for 13 different optical water classes. These comprise clearest natural waters but also waters reaching the frontiers of marine optical remote sensing, namely extreme absorbing, or scattering waters. Considered chlorophyll concentrations reach up to 200 mg m^{-3}, non-algae particle concentrations up to 1,500 g m^{-3}, and the absorption coefficient of colored dissolved organic matter at 440 nm is up to 20 m^{-1}. The algorithm generates different concentrations of water constituents, inherent and apparent optical properties, and a color index. In addition, all products are delivered with an uncertainty estimate. A baseline validation of the products is provided for various water types. We conclude that the algorithm is suitable for the remote sensing estimation of water properties and constituents of most natural waters.

Keywords: ocean color, remote sensing, Sentinel-3, OLCI, extreme Case-2 waters, neural network, fuzzy logic classification

OPEN ACCESS

Edited by:
Tiit Kutser,
University of Tartu, Estonia

Reviewed by:
Tim Moore,
University of New Hampshire, USA
Jenni Attila,
Finnish Environment Institute, Finland
Mark Matthews,
CyanoLakes, South Africa

***Correspondence:**
Martin Hieronymi
martin.hieronymi@hzg.de

†***Present Address:***
Dagmar Müller,
Brockmann Consult GmbH,
Geesthacht, Germany

Specialty section:
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

Received: *15 December 2016*
Accepted: *26 April 2017*
Published: *11 May 2017*

Citation:
Hieronymi M, Müller D and Doerffer R (2017) The OLCI Neural Network Swarm (ONNS): A Bio-Geo-Optical Algorithm for Open Ocean and Coastal Waters. Front. Mar. Sci. 4:140. doi: 10.3389/fmars.2017.00140

INTRODUCTION

The Sentinel-3 Ocean and Land Colour Instrument (OLCI) was developed by the European Space Agency as part of the Copernicus Earth observation program (Donlon et al., 2012). The first of a row of consecutive satellites, Sentinel-3A, was launched early in 2016. Mission objectives include measuring of the ocean reflectance (color) as well as monitoring of sea-water quality and pollution. OLCI is based on the heritage of the Medium Resolution Imaging Spectrometer (MERIS) on board ENVISAT (mission between 2002 and 2012), but with six additional spectral bands. OLCI operates in full resolution mode with a spatial resolution of approximately 300 m and a swath width of 1,270 km. Thus, the instrument images wide sea areas including details of coastal waters, e.g., estuaries, intertidal mudflats, and lagoons, but also inland waters. The challenge is to extract extensively reliable ocean color products such as chlorophyll concentration, *Chl*, from such wide-scale satellite observations, which cover the high natural variability of optical water properties.

The spectral water-leaving reflectance or remote sensing reflectance, R_{rs}, is characterized by absorption and scattering properties of four main components: sea-water, phytoplankton (together with small organisms), colored dissolved organic matter (CDOM), and inorganic particulate material (Mobley, 1994). In addition, wind-dependent air bubbles and boundary conditions may influence the color signal. The composition of water constituents varies considerably, both temporally and regionally. At the open ocean, inherent optical properties (IOPs) of water are determined primarily by phytoplankton and related CDOM and detritus degradation products.

In accordance with the classical (and not unambiguous) bipartite differentiation, these are the so called "Case-1" (C1) waters and all other water types correspond to "Case-2" (Morel and Prieur, 1977; Mobley et al., 2004). Coastal and inland waters can be significantly influenced by other constituents whose concentrations do not covary with the phytoplankton concentration, e.g., due to CDOM and mineral runoff from adjacent land areas or resuspension of bottom material in shallow waters. In extreme cases concentrations of CDOM or inorganic particles can be exceptionally high; those are defined as (Case-2) extremely absorbing (C2AX) and extremely scattering (C2SX) waters respectively (Hieronymi et al., 2016). Absorbing waters are characterized by very low marine reflectance and a shift of the R_{rs} maximum toward the red spectral range. Typically, the CDOM absorption at 440 nm is $>1\ m^{-1}$ in C2AX waters. There are "black lakes," e.g., many boreal lakes, where the reflectance is negligible in almost the entire visible part of spectrum (VIS: 400–700 nm); signal from chlorophyll is—if at all—only detectable in the near infrared (Kutser et al., 2016; NIR in the sensor response division scheme: 700–1,000 nm). Observations from remote sensing face similar challenges for extreme turbid C2SX waters, because non-algae particles mask optical properties of algae particles over large parts of the visible spectrum. But in general, the water appears much brighter; the water-leaving reflectance spectrum has still significant amplitudes in the NIR (Ruddick et al., 2006) and measurably non-zero reflectance at the last OLCI band at 1,020 nm (Knaeps et al., 2012). Typically, the concentration of inorganic suspended matter, *ISM*, is $>100\,g\ m^{-3}$ in C2SX waters (Hieronymi et al., 2016). An overview of water type sub-classification, used for differentiation in this work, is provided in **Table 1**.

Great variability of IOPs causes ambiguousness and therefore a significant degree of uncertainty in the interpretation of the remote sensing signal. We have to deal with a nonlinear and multivariate problem and the ocean color algorithm must be designed accordingly. The capability of bio-(geo)-optical algorithms strongly varies on global, regional, and very small scales and algorithms generally face more difficulties in Case-2 waters (e.g., Blondeau-Patissier et al., 2004; Darecki and Stramski, 2004; Gregg and Casey, 2004; Reinart and Kutser, 2006; Attila et al., 2013; Beltrán-Abaunza et al., 2014; Harvey et al., 2015). Indeed, it is a challenge to bridge the different scales with a high degree of reliability of the ocean color products. And we should not forget that marine atmospheric correction (AC), which is necessary to derive R_{rs} at the sea surface from satellite imagery and thus, provides input for in-water algorithms, is a complex task with additional uncertainties, in particular for extreme waters.

An artificial neural network (NN) is an appropriate regression technique to parametrize the inverse relationship between optical properties and reflectances. It has been proven in the last years that NNs produce reasonable approximations of ocean color products from optically complex (Case-2) waters. NNs have been applied to different satellite sensors in order to derive concentrations of water constituents, inherent and apparent optical properties (IOPs and AOPs), and photosynthetically available radiation (PAR), or to discriminate algae species (Gross et al., 1999; Schiller and Doerffer, 1999; D'Alimonte and Zibordi, 2003; Zhang et al., 2003; Tanaka et al., 2004; Schiller, 2006; Bricaud et al., 2007; Schroeder et al., 2007; Ioannou et al., 2011; Jamet et al., 2012; Chen et al., 2014; Hieronymi et al., 2015; D'Alimonte et al., 2016). Due to their speed, NN-based ocean color algorithms are deployed for operational and near-real time satellite observations, e.g., the MERIS Case-2 water algorithm (Doerffer and Schiller, 2007) and C2RCC (Brockmann et al., 2016).

TABLE 1 | Water case sub-classification that characterize the database in view of concentration ranges of chlorophyll, CDOM, and inorganic suspended matter.

Case	Description	*Chl* [mg m^{-3}]	a_{cdom} (440) [m^{-1}]	*ISM* [g m^{-3}]
C1	Open ocean and algae bloom	0: 200	$X\ Chl^{Y}$	<1.5
C2A	Moderately to strongly absorbing	0: 200	0.1: 1	<10
C2AX	Extremely absorbing	0: 200	> 1	<10
C2S	Moderately to strongly scattering	0: 200	< 0.5	1: 100
C2SX	Extremely scattering	0: 200	< 0.5	>100

In Case-1 (C1), CDOM is related to chlorophyll concentration with arbitrary parameters X and Y.

The objective of this study is to introduce a new in-water processing scheme designed for OLCI ocean color observations called OLCI Neural Network Swarm (ONNS). The distinctive feature of this algorithm is its wide range of applicability in terms of optical water properties ranging from oligotrophic ocean waters to extremely turbid (scattering) or dark (absorbing) waters. The specific goals of the study are: (1) to reference the fundamental processing scheme, (2) to provide the scientific background, (3) to introduce the derived ocean color products, and (4) to evaluate the basic suitability of the algorithm for oceanic and coastal waters, i.e., C1, C2A, C2AX, C2S, and C2SX waters (**Table 1**).

ONNS BASIS AND ALGORITHM DESCRIPTION

ONNS is an in-water processor, which retrieves ocean color (OC) products from Sentinel-3 OLCI satellite scenes. Inputs to the algorithm are normalized remote sensing reflectances (just above the sea surface). Atmospheric correction is not part of the in-water processing scheme, and thus, ONNS fully relies on proper atmospheric correction (see Section Retrieval Accuracy). The processor logic is illustrated in **Figure 1** and documented in the following.

Neural Network Algorithm

As it is the case for all ocean color algorithms, NNs are valid for a certain range of constituents and their concentrations, and some parameters may be deduced more accurately than others, e.g., retrieval of suspended matter is usually the least critical, whereas

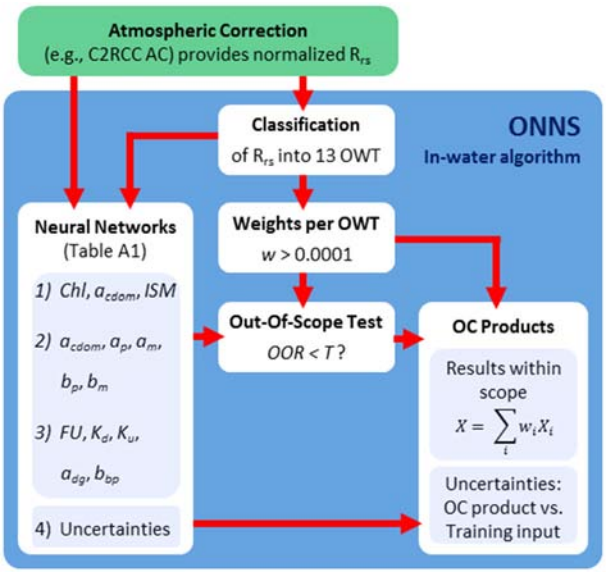


FIGURE 1 | Flowchart of the OLCI Neural Network Swarm in-water ocean color processor.

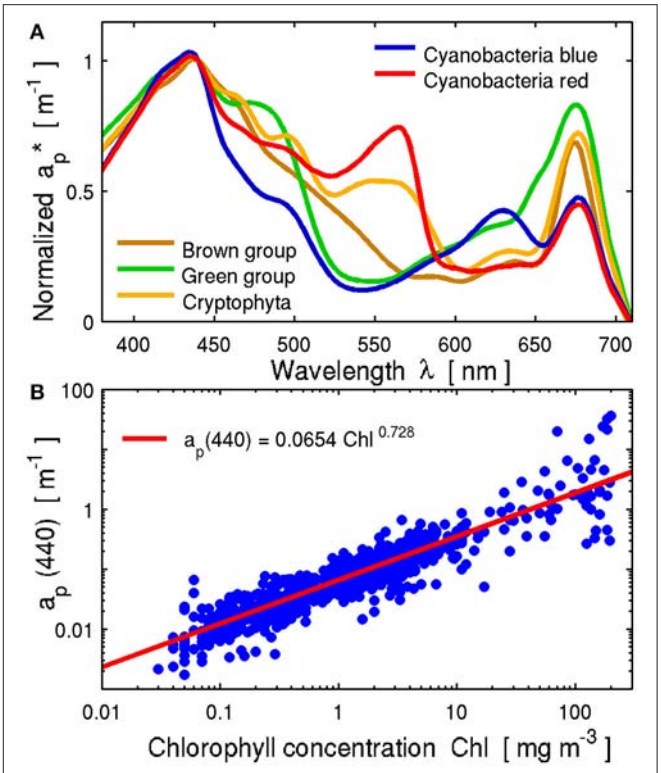


FIGURE 2 | Bio-optical model of phytoplankton absorption. **(A)**, Fundamental absorption spectra of five algae groups. **(B)**, Modelled variability of the absorption coefficient of phytoplankton at 440 nm as function of chlorophyll concentration. The solid line shows the regression line of observations (Bricaud et al., 2004).

CDOM retrievals are the most challenging (Odermatt et al., 2012; Brewin et al., 2015). Our approach proposes to blend various NN algorithms, each optimized for a specific scope. This swarm of neural networks therefore, covers the largest possible variability of water properties including oligotrophic and extreme waters (**Table 1**).

Neural Network Data Basis

Basis for NN training is knowledge of the relationship between water constituents, i.e., their optical activity, and the spectral remote sensing reflectance, R_{rs}. The latter is defined as ratio between water-leaving (upwelling) radiance, L_w, and downwelling irradiance, E_d, both just above the water surface. For training and validation (test) purposes, a large ($>10^5$) dataset has been simulated using the commercial radiative transfer software Hydrolight (version 5.2; Sequoia Scientific, USA; Mobley, 1994). Hydrolight is a forward model to compute R_{rs} and many other light field-related quantities from optical specifications of the water body, such as specific absorption and scattering properties. Considered concentration ranges are defined in **Table 1**. Basis for estimating distributions, ranges, and covariances of optical parameters in the model are different *in situ* datasets: (1) primarily our data from the North and Baltic Sea (HZG), (2) OC-CCI (ESA, worldwide; Valente et al., 2016), (3) HELCOM (Baltic Sea 1997–2013; ICES, 2011), and (4) NOMAD (NASA, worldwide; Werdell and Bailey, 2005). The simulations cover the spectral range from 380 to 1,100 nm in 2.5 nm steps (hyperspectral over full VIS and NIR). Resulting reflectances and AOPs refer to a solar irradiation from zenith direction and nadir viewing angle, i.e., they are fully normalized. Many standard settings of Hydrolight are utilized (Mobley, 1994; Mobley and Sundman, 2013); specific inputs are defined in the following.

The total absorption contains fractions of absorption of pure water, phytoplankton pigments, minerals (also inorganic detritus or non-algae particles), and colored dissolved organic matter (CDOM, also referred as yellow substance or *gelbstoff*). The same distinction is made for scattering; only to CDOM no scattering is attributed. The absorption and scattering coefficients of pure water depend on temperature, salinity, and wavelength (data from WOPP v2 by Röttgers et al., 2016).

Phytoplankton absorption is determined by the composition and concentration of pigments, e.g., chlorophyll-a, *Chl*, which is generally used to quantify the marine biomass concentration. This means that different algae species have unique absorption spectra. Xi et al. (2015) showed the impact of chlorophyll-specific absorption spectra on R_{rs}. They identified five fundamental absorption shapes from which an inversion of algae species from remote sensing reflectance is possible. **Figure 2A** illustrates the basic chlorophyll-specific absorption, a_p^*, spectra normalized at 440 nm that are utilized in this work (from Xi et al., 2015). Mixtures of these spectra represent the variability of spectral shapes that are found in measured data. It has been decided to combine two types of spectra, whereby one component dominates the signal with 80%. The globally most common spectral shape is labeled with "brown group"; it is very similar to the standard absorption spectrum used in Hydrolight and

summarizes Heterokontophyta, Dinophyta, Haptophyta, and others that have a similar spectral shape. The "green group" includes Chlorophyta. Cyanobacteria are separated into blue (e.g., Aphanothece clathrata) and red (e.g., Synechococcus red) species. The two spectra for cyanobacteria are derived from *in situ* absorption measurements in the Baltic Sea. The three other spectra are taken from cultures. The phytoplankton (particle) absorption, a_p, is related to the spectral chlorophyll-specific absorption and chlorophyll concentration, $a_p(\lambda) = a_p^*(\lambda)\ Chl$. The natural variability of phytoplankton absorption is very high (e.g., Bricaud et al., 2004) and included in the simulations (**Figure 2B**). Thus, when assessing *Chl* retrieval performance, this must be kept in mind.

The shape of CDOM absorption is nearly exponential. Exponential functions have been used for C1 water simulations (**Table 1**). Here, CDOM absorption coefficients, a_{cdom}, and exponential slopes are varied strongly in order to display the natural variability (e.g., Valente et al., 2016). In addition, the present work uses modeled absorption spectra that are fitted to spectral measurements (by Rüdiger Röttgers, HZG). Based on this, further CDOM spectra are extrapolated toward ultra-extreme absorption with $a_{cdom}(440) = 20\ m^{-1}$. The exponential slope of these spectra (between 300 and 400 nm) is approximately $0.014\ nm^{-1}$.

Fournier-Forand volume scattering functions have been applied for algae and non-algae particles (see Mobley and Sundman, 2013). The particle backscatter fraction, which actually correlates poorly with *Chl* and mineral (*ISM*) concentrations, is needed for the selection of appropriate phase functions. The corresponding formula of Twardowski et al. (2001) has been used for *Chl*-bearing particles. For inorganic particles, the mineral backscattering-*ISM* relationship of Zhang et al. (2010) has been utilized. The spectral mass-specific scattering coefficients are approximated by exponential functions, following the natural variability shown in measurements of organic-dominated and mineral-dominated waters (Woźniak et al., 2010).

The atmospheric and surface boundary conditions in Hydrolight are set constant, i.e., usage of the semi-empirical sky radiance model, assuming dry air with a marine aerosol type and moderate wind speed of $5\ m\ s^{-1}$. The refractive index of water (as it is the case for absorption and scattering) is a function of water temperature (0–30°C) and salinity (0–35 PSU). The water is (virtually) infinitely deep. Effects of light polarization are not taken into account in Hydrolight.

All simulations have been carried out with and without inelastic scattering, i.e., Raman scattering, CDOM and *Chl* fluorescence, but without internal sources, i.e., no bioluminescence. In the end, data without inelastic scattering have been used for ONNS development. This is unproblematic in the selected setup with the 11 OLCI bands. Seen over the visible spectral range, differences mostly play no role, except for extreme absorbing waters, where high CDOM fluorescence is present. During algae bloom events, very high chlorophyll fluorescence peaks can be observed in nature (e.g., Fawcett et al., 2006), but modeling a certain quantum yield efficiency holds in itself great uncertainties (see Section ONNS Design). However, the simulations have been compared with observations and we generally have found a good agreement (Hieronymi et al., 2016). But we also have found some discrepancies partly related to plausible measuring uncertainties and possibly due to model simplifications.

NN Training

One part of the simulated dataset is put aside for later quasi-independent test purposes (see Section ONNS Application to Validation Data). The rest of the R_{rs} data is optically classified (Section Out-of-Scope Test) and grouped together. The scopes of concentrations together with median values are given in **Table 2**.

The usable wavebands of the in-water algorithm are determined by the atmospheric correction at OLCI bands. We selected 11 (out of 21) OLCI wavebands for NN input (bands 1–8, 12, 16, and 17, i.e., at 400, 412.5, 442.5, 490, 510, 560, 620, 665, 755, 777.5, and 865 nm). The instrument's band widths vary and to be precise, the centers of bands 12 and 16 are actually at 753.75 and 778.75 nm respectively (Donlon et al., 2012). In contrast to other NN algorithms (e.g., Doerffer and Schiller, 2007), sun-viewing geometry is no input to the present NNs, instead input reflectances are normalized.

The selected output parameters are mostly common ocean color products (e.g., Nechad et al., 2015; Valente et al., 2016):

(1) Concentration of chlorophyll, *Chl* [$mg\ m^{-3}$],
(2) Concentration of inorganic suspended matter (minerals), *ISM* [$g\ m^{-3}$],
(3) Absorption coefficient of CDOM at 440 nm, $a_{cdom}(440)$ [m^{-1}],
(4) Absorption coefficient of phytoplankton particles at 440 nm, $a_p(440)$ [m^{-1}],
(5) Absorption coefficient of minerals at 440 nm, $a_m(440)$ [m^{-1}],
(6) Absorption coefficient of detritus plus *gelbstoff* at 412 nm, $a_{dg}(412)$ [m^{-1}],
(7) Scattering coefficient of phytoplankton particles at 440 nm, $b_p(440)$ [m^{-1}],
(8) Scattering coefficient of minerals at 440 nm, $b_m(440)$ [m^{-1}],
(9) Total backscattering coefficient of all particles (organic and inorganic) at 510 nm, $b_{bp}(510)$ [m^{-1}],
(10) Downwelling diffuse attenuation coefficient at 490 nm, $K_d(490)$ [m^{-1}],
(11) Upwelling diffuse attenuation coefficient at 490 nm, $K_u(490)$ [m^{-1}], and.
(12) Forel-Ule number, *FU* [-].

The 12 parameters are results of three independent sets of NNs, one that computes concentrations, one gives IOPs at 440 nm, and the last provides different IOPs and AOPs (see Appendix **Table A1**). Concentrations can be directly derived with NNs or alternatively, they can be estimated using IOPs, e.g., *Chl* from $a_p(440)$ or *ISM* from $b_m(440)$ (Doerffer and Schiller, 2007). The latter approach allows better adaptation of empirical relationships by means of *in situ* match-up data (which in case of OLCI is not yet available at present). All absorption and scattering contributions are retrieved at the reference wavelength 440 nm (pure water IOPs are known). Thereby, it is possible

TABLE 2 | Chlorophyll, CDOM, and inorganic suspended matter concentrations for the 13 optical water types.

OWT	*Chl* [mg m^{-3}]			a_{cdom}(440) [m^{-1}]			*ISM* [g m^{-3}]			Fractions of cases [%]				
	Min	Median	Max	Min	Median	Max	Min	Median	Max	C1	C2A	C2S	C2AX	C2SX
1	0.03	1.4	195	0	2.8	8.3	0	0.8	60	0	0.15	19.51	80.34	0
2	0.03	1.5	195	4.3	8.1	20	0	0.7	10	0	0	0	100	0
3	0.03	1.5	200	0	0.156	20	0	750	1500	0.01	0.01	0.01	0.34	99.63
4	0.03	1.8	195	0	0.164	17	0	110	300	2.41	2.88	25.42	5.24	64.05
5	0.03	2.3	195	0	0.224	2.1	0	2.5	25	6.3	31.42	58.7	3.59	0
6	0.03	1.1	195	0	0.162	0.5	0	0.9	6	8.91	63.31	27.79	0	0
7	70	185	200	0.002	0.262	20	0	2.6	750	20.64	23.49	23.49	23.13	9.25
8	70	185	200	0.008	0.223	18	0	2.75	1000	20.22	16.18	26.84	22.06	14.71
9	0.03	2	195	0	0.096	0.3	0	0.2	5	48.4	42.24	9.36	0	0
10	0.03	1.3	195	0	0.04	0.272	0	0.08	2	90.63	8.83	0.54	0	0
11	0.03	0.33	5	0.002	0.016	0.222	0	0.1	0.5	99.63	0.37	0	0	0
12	0.03	0.2	0.64	0.002	0.01	0.028	0	0.06	0.4	100	0	0	0	0
13	0.03	0.12	0.53	0.002	0.006	0.016	0	0.04	0.2	100	0	0	0	0

The composition of an optical water type with reference to the sub-classification in ***Table 1*** *is additionally shown.*

to estimate the total absorption, total scattering, and total attenuation coefficients. Mineral particles have usually lower absorption characteristics than CDOM, but shape-wise both are very similar. Following this reasoning, semi-analytical models, designed to retrieve IOPs from satellite data, often combine absorption by detritus (in this work only inorganic fraction) and *gelbstoff* (all water constituents which pass a filter pore size of 0.2 μm, which is often synonymous with CDOM). This absorption coefficient often corresponds to 412 nm. A similar idea holds true for the backscattering parameter. It is usually the backscattering coefficient of all marine particles together, which is measured in the field (at 510 nm). The diffuse attenuation coefficients are used to describe the attenuation of irradiance as a function of depth in water. It can be used to compute the depth of the euphotic zone. The Forel-Ule color scale was used for natural water classification long before the satellite era. In open ocean regions, the *FU* number is closely related to *Chl* concentration. Thus, the index can support ocean color trend analysis in the pre-satellite age and afterwards (Wernand et al., 2013; van der Woerd and Wernand, 2015). The color scale visualizes the color of the water body above a white Secchi disk that is hold at half Secchi depth.

A subsequent set of neural nets serves to evaluate the divergence of final OC products from the original training basis, i.e., Hydrolight simulations. The results are part of an uncertainty estimate (see Section Uncertainty Analysis).

The actual NN training procedure is described in Schiller and Doerffer (1999). The utilized multilayer feedforward-backpropagation neural net program is documented in Schiller (2000). The code was embedded in a program to test many NN architectures, i.e., varying numbers of hidden layers and neurons, and to optimize the learning process.

NN Scoring and Selection

Several hundreds of nets per water class and task with much different architecture have been produced. Afterwards, a ranking system has been applied in order to determine the optimal nets without over-training. In principle, statistical parameters such as root-mean-square error and goodness of fit are transformed into relative scores, which evaluate the quality of individual nets (Müller et al., 2015a). The best performing neural network architectures per water class are specified in **Table A1**. Inputs and outputs for the NNs are $\log_{10}(X + 0.001)$, where X stands for R_{rs} or an ocean color product. The only exception is the Forel-Ule number, which is an integer between 1 and 21 and not logarithmized. The logarithmic form of input/output enables a distribution of values, which is closer to a uniform distribution within the range of input data, and therefore better approximation of outputs. The addition of 0.001 allows consideration of zero-values as input.

Fuzzy Logic Classification

Optical water type, OWT, classification based on remote sensing reflectance spectra has been developed to overcome the simplifications of Case-1 and Case-2 waters (Moore et al., 2001). It can bridge the gap between regionalized optical models and the global scale by combining several models according to their respective membership to a certain water type (Moore et al., 2009, 2012, 2014).

The classification is based on the simulated R_{rs} spectra (at 11 OLCI wavebands), but it is used for atmospheric corrected satellite data afterwards. Negative reflectances can occur after AC sometimes, while the spectral shape is still realistic. In order to avoid conflict with negative reflectances, the spectra are therefore transformed by $\log_{10}(R_{rs} + 1)$ (note that R_{rs} is treated differently during classification and the NN application). Before the clustering, these transformed spectra are normalized by their brightness (sum of log-transformed reflectances), so that the classification is based on the shape of the spectrum alone. As the goal is to derive representative spectra, which have different spectral shapes and in particular a different spectral maximum, the sample from the simulation

database does not take into account the frequency of natural occurrence of spectra. Spectra with their maximum at 510, 620, and 777.5 nm come in two distinctive shapes and there is no spectrum with maximum at 865 nm, so that during the agglomerative clustering 13 classes are selected. The 13 OWT classes are described by their class mean and standard deviation per wavelength of the brightness-scaled R_{rs}, which are used for the classification of spectra furthermore. The mean reflectance spectra of the 13 OWT classes are plotted in **Figure 3**.

The five water type categories (**Table 1**) are defined by combinations of concentrations and thresholds. The water classes are designed to represent spectra, which have their maximum in the spectral shape at different wavelengths, independent of their brightness. Combining the water classes and the concentration categories is a test, which spectral shapes can be found in certain concentration ranges (**Table 2**).

Fuzzy set theory allows an element to have membership to one or more OWT classes (Moore et al., 2001). The weight of a class (membership function) is altered to allow for graded memberships, i.e., $0 \leq w_i \leq 1$, and to express partial class membership to the ith class. For constrained fuzzy sets the sum of all 13 weights equals 1. However, the class membership had to be above a minimum threshold, which was set at 0.0001. The membership to a class (weight) is calculated by determining the Mahalanobis distance between the given spectrum and the class means using the classes' covariance matrix respectively. Reconstruction of R_{rs} spectra by means of the fuzzy classification inversion yields mostly satisfying results. However, R_{rs} inversion from different atmospheric corrections reveals expected uncertainties in the violet-blue, which can be the case, if the satellite-acquired spectrum provided by the AC is distinctly different from modeled R_{rs} (which is basis of the classification).

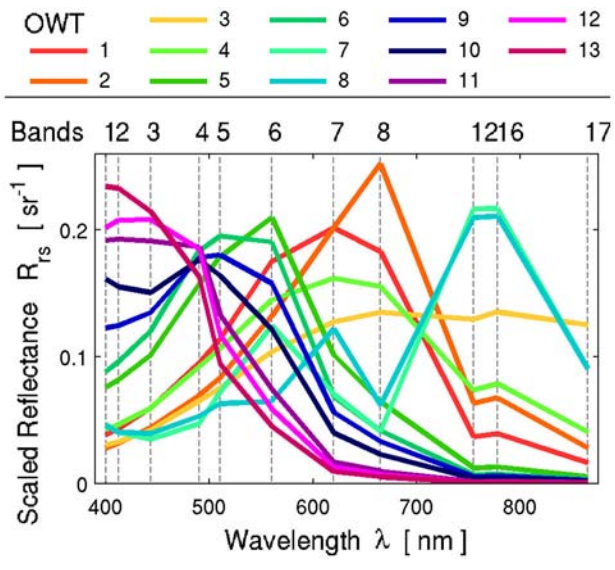


FIGURE 3 | Brightness-scaled remote sensing reflectances for 13 classes of optical water types. Utilized OLCI bands are marked.

Within the ONNS framework (**Figure 1**), the fuzzy logic classification scheme is used to assess the atmospheric corrected R_{rs}, and to determine the corresponding class memberships. The final blended retrieval for each pixel and each ocean color product is a weighted sum of the retrievals of all class-specific NNs (Appendix **Table A1**).

Out-of-Scope Test

Well-constructed NNs have good interpolation properties but produce unpredictable output when forced to extrapolate (Doerffer and Schiller, 2000). Therefore, measures have to be taken to recognize NN input not foreseen in the NN training phase and thus out of scope of the algorithm. Regarding simulated data, the fuzzy classification is well-constructed; maximum (or high) membership of a water class usually correlates well with the scopes of the corresponding NNs. However, it may happen that the classification yields a broad distribution of weights or that all memberships are such low that the spectrum is not classifiable. In the latter case, the satellite image pixel is flagged out. Despite the memberships, a quality measure is applied that evaluates the deviation of the NN input from the NN training range. The out-of-range parameter, *OOR*, is zero if the input is within the range but increases with increasing deviation. The assessment treats the input-reflectances spectrally differently; wavebands in the green spectral range have highest weights. The varying signal-to-noise specifications of OLCI are one argument for this (Donlon et al., 2012). Uncertainties in the fluorescence quantum yield efficiency of phytoplankton are another argument. Furthermore, we observe higher uncertainties in violet and blue wavebands generally shown in atmospheric correction validations (Müller et al., 2015a), but also from *in situ* determinations of R_{rs} due to the variable surface reflectance factor (Hieronymi, 2016; Zibordi, 2016). The allowance for $OOR > 0$ is one of the fine-tuning techniques to gain better spatial homogeneity of an OLCI scene and to adapt the algorithm to *in situ* observations.

Uncertainty Analysis

The determination of uncertainties of OC products is similar to the procedure applied in the C2RCC algorithm (Brockmann et al., 2016). All NNs per water class were reapplied to their training datasets to estimate the OC products. The uncertainty nets compare the estimated value, X_E, with the initial training value, X_T. The uncertainty per product is given as approximation (percent) error:

$$\varepsilon = 100 \frac{X_E - X_T}{X_T}. \quad (1)$$

As it is the case for all OC products, the final approximation error is a weighted sum of the retrievals of all class-specific uncertainty NNs.

Test Data

Remote sensing reflectance data at OLCI wavelengths in conjunction with bio-geo-optical properties of the top water layer are used to evaluate the capacity of ONNS. For this purpose, different statistical parameters have been utilized. The degree

of deviation is presented by the absolute root-mean-square error, *RMSE*. The *Bias* shows the average difference and is a measure for systematic over- or underestimation. Furthermore, the correlation coefficient, *r*, is calculated.

Simulated Data

Hydrolight-simulated data have been used to develop the classification scheme and to train neural nets. From the same data source ($>10^5$), a quasi-independent set of 23,445 reflectance spectra is used for testing and validation; these particular data are not used for ONNS development. The test data contain all water types from **Table 1** and are shared approximately equally.

Simulated CCRR Data

A second synthetic dataset has been used to evaluate the performance of ONNS for the retrieval of water quality parameters. The "CoastColour Round Robin" (CCRR) dataset by Nechad et al. (2015) compiles inputs and results from 5,000 Hydrolight simulation. Atmospheric boundary conditions and simulation setup are comparable with above mentioned simulations. The used data refer to a sun zenith angle of 0°. Remote sensing reflectance at the 11 needed OLCI bands is interpolated from hyperspectral water-leaving reflectances between 350 and 900 nm with 5 nm steps. Corresponding *Chl* and *ISM* concentrations as well as CDOM absorption at 443 nm are tested with the ONNS retrieval (note that ONNS CDOM absorption coefficient refers to 440 nm).

In situ Data

Complete *in situ* datasets for the evaluation of OLCI-specific algorithms like ONNS are not freely available. The accessible data of CCRR (Nechad et al., 2015) and OC-CCI (Valente et al., 2016), which compiles data from several sources (e.g., MOBY, BOUSSOLE, AERONET-OC, NoMAD, MERMAID), lack of several R_{rs} bands. The ONNS algorithm needs only 11 out of 21 OLCI bands, but coinstantaneous data at 400, 755, 777.5, and 865 nm are not available. However, many of these sub-datasets are actually measured hyper-spectrally. Ramses radiometers, for example, that are deployed during our *in situ* campaigns, measure between 320 and 950 nm (TriOS optical sensors, Germany). Extracted multi-spectral reflectances together with *Chl*, *ISM*, and CDOM data are included in the CCRR *in situ* dataset; the corresponding measurement protocols are described in (Nechad et al., 2015). Our 48 data were collected between 2005 and 2006 (but not in winter) onboard a ferry from Cuxhaven to the island Helgoland in the German Bight (see **Figure 5**). Radiometric measurements were conducted under optimal sun-viewing angles (e.g., Zibordi, 2016), but strictly speaking, ONNS requires angle-normalized R_{rs} (with the sun in zenith) as input. Nonetheless, these data are used to test ONNS as well.

Sentinel-3 OLCI Scene

One Sentinel-3 OLCI scene is shown with permission to illustrate the qualitative and spatial application of ONNS (**Figure 5**). The tripartite scene was captured on 20 July 2016 between 9:30 and 9:36 UTC and shows large parts of the North and Baltic Sea. Thus, the scene images many different water types including different algae blooms. The satellite image indicates transparent cirrus clouds over the German Bight and Gulf of Finland, broken clouds over the Skagerrak and Kattegat, and cloud shadows. In comparison with MERIS, OLCI's view is slightly tilted in order to reduce the impact of sun glint, which is somewhat visible at the right edge of the image. Level-1 data of the first OLCI reprocessing are utilized for this work (IPF-OL-1-EO version 06.06). Atmospheric correction of the scene is provided by the C2RCC algorithm ("Case-2 Regional CoastColour," version 0.15, Brockmann et al., 2016). An additional cloud mask was applied using the provided path radiance and viewing angles. At the same time of the satellite data acquisition, we measured R_{rs} in the German Bight for Sentinel-3 validation purposes (**Figure 5**), but these data are not used in this work.

RESULTS

ONNS Application to Validation Data

The classification of simulated test data reveals that a maximum of four classes contribute to the inversion of R_{rs} spectra (the classes have non-zero weights). In principle, all water types (clear to extremely turbid) can be assigned properly. The classification failed on <3% of validation data; of those 56% are absorbing waters (C2A, C2AX) and approximately 70% have high *Chl* concentrations ($Chl > 10\,\mathrm{mg\ m^{-3}}$). The classification of the *in situ* and simulated CCRR data yields no plausible results in approximately 10% of cases. The classifiable 4512 CCRR spectra exhibit maximum memberships in OWTs 1 (9%), 2 (0.5%), 4 (0.16%), 5 (55.2%), 6 (14.9%), 9 (10.6%), 10 (6.9%), 11 (1.7%), 12 (0.3%), and 13 (0.5%). Thus, a high percentage of these data correspond to the Case-1 or moderately to strongly scattering waters (**Tables 1, 2**). The 43 *in situ* data points, which are captured in coastal waters of the German Bight (**Figure 5**), have maximum memberships in OWT 1 (10.4%) and 5 (89.6%).

Examples of the retrieval capabilities of ONNS in comparison with validation data are illustrated in **Figure 4**. Estimates of concentration of *Chl*, *ISM*, and CDOM are shown for different water types, namely Case-1, extreme absorbing, and extreme scattering waters (**Table 1**). In addition, ONNS retrieval tests are shown for simulated data from the CCRR dataset and our *in situ* data. The colors characterize the estimated uncertainty in terms of the percent error. Green marks the generic ±5% uncertainty target for satellite ocean color products (defined for oligotrophic and mesotrophic Case-1 waters), orange and red colors signify an overestimation of the retrieved value in comparison with the expected (trained) value, and blue stands for an underestimation respectively. The uncertainty can be high in ambiguous cases with significant masking effects (in extreme waters) or if the NN data basis already provides high (natural) variability, as for example for *Chl* concentration (compare **Figures 4A,D,G** with **Figure 2B**). Despite high *Chl* variability, the uncertainty target can be achieved for all magnitudes of concentrations (varying over five orders of magnitudes), but with different occurrence in the water types: approximately 30% in C1, 10% in C2AX, and 5% in C2SX waters. An acceptance level of ±50% can be achieved in >97% of cases for C1, >80% in C2AX, and >70% in C2SX respectively. In all the cases, mean and median percentage

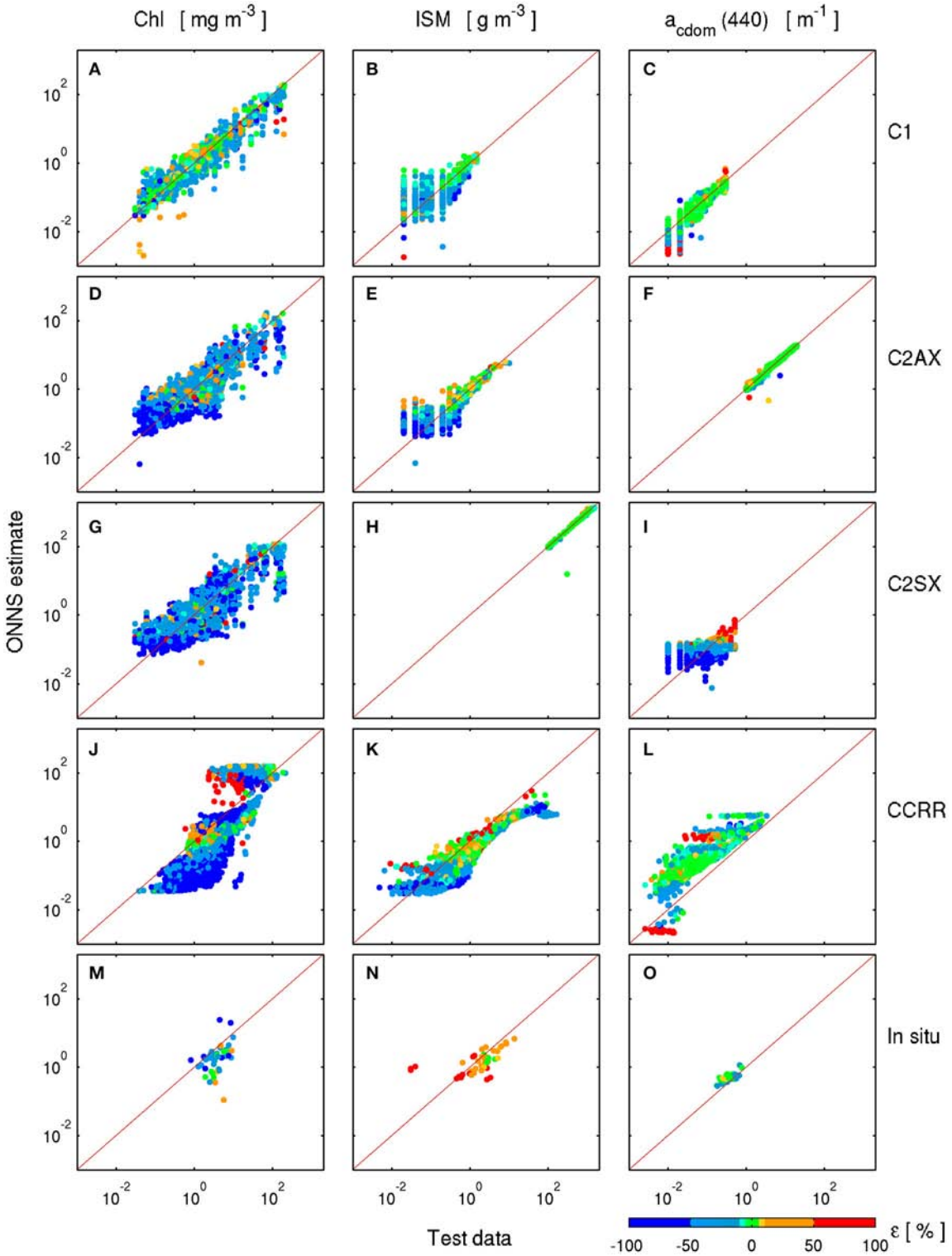


FIGURE 4 | ONNS retrieval capacity for chlorophyll concentration (left), concentration of inorganic suspended matter (center column), and CDOM absorption at 440 nm (right) in comparison with simulated and *in situ* validation data. (A–C): Case-1 data from database, **(D–F)**: Case-2 extreme absorbing waters, (**G–I)**: Case-2 extreme scattering waters, **(J–L)**: simulated data from CoastColour Round Robin (Nechad et al., 2015), **(M–O)**: HZG *in situ* data. Colors indicate the retrieved uncertainty.

errors are slightly negative, i.e., ONNS *Chl* retrieval shows a tendency for underestimation of expected values. Even if the test value is overestimated by ONNS, the uncertainty estimate may point to underestimation. This may be due to blending of NN from different OWT classes with distinctive different ranges of concentrations. In contrast to the simulated validation data, the ONNS *Chl* retrieval of CCRR and *in situ* data yields stronger deviations from the one-to-one line (**Figures 4J,M**).

With regards to *ISM*, the retrieval performance is less skilled if the optical signal of minerals is weak due to low mineral concentrations—as it is the case in oligotrophic waters (**Figure 4B**). The 5% uncertainty target is reached within approximately 20% of all cases in C1 waters, 27% in C2AX, and >87% in C2SX. Thus, the more non-algae particles are present, the better ONNS performs. A similar trend can be observed for CDOM retrieval (**Figures 4C,F,I**). Lowest concentrations vanish in the noise, whereas high concentrations can be retrieved accurately. Approximately 60% target-retrievals can be achieved in C1 and >94% in extreme absorbing waters. **Figure 4I** illustrates the difficulties to separate the absorption signal due to CDOM and minerals; only 6% of estimates fall in the target-uncertainty range. In comparison to the *Chl* retrieval, *ISM* and CDOM retrievals of CCRR and *in situ* data show better agreement (**Figures 4K,L,N,O**).

The NN-estimated uncertainties and corresponding color distributions in **Figure 4** reflect the comparative statistics that are tabulated for all water types (**Table 3**). Additional statistics of all OC products are listed in the (**Table A2**). With reference to the simulated test data and seen over all water types, the smallest differences between estimated and test data occur for the Forel-Ule number and both "mixed" IOPs, $a_{dg}(412)$ and $b_{bp}(510)$. In comparison, larger deviations occur for low-concentration mineral-related values. We found weak water type-independent underestimation for the direct phytoplankton-related quantities [*Chl*, $a_p(440)$, and $b_p(440)$] and for *FU*, $K_d(490)$, $K_u(490)$, $a_{dg}(412)$, and $b_{bp}(510)$. But again, largest *Biases* are observed for the retrieval of non-algae properties in clear oceanic (almost mineral-free) C1 waters. The correlation coefficient reveals strong linear relationship for all cases exclusive of CDOM in extremely scattering waters, here the relation is weak (CDOM retrieval correlation is >0.95 for C2S and the other cases). The statistical values of the comparison with independent CCRR and *in situ* data paint a somewhat different picture with generally lower correlation coefficients (**Table 3**). Both datasets include turbid Case-2 waters that are predominantly characterized by one optical water type, i.e., OWT 5 (see **Table 2**). Most of the other water types are not independently evaluated.

TABLE 3 | Statistics of ONNS retrievals vs. test data.

Statistics	Dataset	*Chl* [mg m^{-3}]	*ISM* [g m^{-3}]	$a_{cdom}(440)$ [m^{-1}]
RMSE	C1	10.2577	0.0856	0.0174
	C2A	10.5276	0.1116	0.0234
	C2S	11.2224	1.4917	0.0290
	C2AX	10.3555	0.1429	0.4356
	C2SX	16.7236	35.6890	0.0971
	CCRR	21.4129	8.5279	0.5067
	In situ	4.3860	1.7515	0.1760
Bias	C1	−1.3144	0.0170	−0.0023
	C2A	−1.1262	0.0110	−0.0049
	C2S	−0.5576	−0.3131	−0.0014
	C2AX	−1.0591	0.0162	−0.0201
	C2SX	−2.1005	−4.2457	−0.0355
	CCRR	1.9023	−2.3966	0.2990
	In situ	−1.2786	−0.8777	0.1410
r	C1	0.8702	0.9212	0.9515
	C2A	0.8495	0.9880	0.9855
	C2S	0.8163	0.9978	0.9540
	C2AX	0.6856	0.9777	0.9932
	C2SX	0.7017	0.9961	0.3120
	CCRR	0.5855	0.7496	0.8030
	In situ	0.3658	0.8616	0.7286

Datasets marked with C1, C2A, C2S, C2AX, and C2SX refer to simulated data that are not used for NN training (numbers of points for comparison are 5392, 4699, 4049, 4526, and 4082 respectively). The independent Hydrolight-simulated CoastColour Round Robin (CCRR) dataset contains 4512 data. 43 match-ups are basis for the in situ data comparison. The corresponding plots in ***Figure 4*** *are shown in log form; the statistical values here are not in log form.*

ONNS Application to OLCI Scene

Application of the new ONNS algorithm to the satellite image is illustrated in **Figure 6**. Again, up to four optical water type classes are needed for the inversion. In this particular scene, water classes 3, 7, and 8 have no contribution to the products (all rather extreme turbid cases, see **Table 2**); all other classes give spatially dependent contributions (**Figure 6A**). Comparison with the mean shapes of the R_{rs} (**Figure 3**) meets regional expectations. In the western part of the Baltic Sea, including the Western Gotland Basin and the Bothnian Sea, the spectra show the strongest resemblance to classes 6, 9, and 10. Spectra of the Eastern Gotland Basin and Gulf of Finland fall into class 5 mostly and the Lagoons behind the Bay of Gdansk have some spectra with the shape of class 1. In contrast, we have found maximum memberships of classes 9, 10, and 11 in the clear open North Sea and Norwegian Sea and classes 1, 2, 4, 5, and 9 along the German and Dutch coasts. In some clear water cases (OWT 9, 10, and 11), the out-of-range warning flag for input spectra raises; these cases are mostly in spatial conjunction with transparent cirrus clouds (**Figure 5**). The Forel-Ule number that is estimated with ONNS provides an intuitively color impression and reconfirms expected geographic characteristics of the sea areas (**Figure 6B**).

Concentrations of *Chl*, *ISM*, and CDOM together with their accompanied uncertainty estimates are shown in **Figure 7**. Only valid sea pixels are shown; land areas and clouds are masked out. However, in spatial vicinity to clouds and coasts, apparently wrong assessments of OC products are possible; here, the predictions are overestimating the true values for the most part. Some areas are very shallow, e.g., the Curonian and Vistula Lagoons, and therefore, bottom reflections cannot be ruled out. This again would lead to possible overestimation of (particle) concentrations. All in all, the ranges of derived concentrations

are reasonable (the colors on the left side of **Figure 7** correspond to the respective units). Previous match-up analyses showed that most of the measured Baltic *Chl* values range between ∼1 and 10 mg m^{-3} with somewhat smaller values in the Skagerrak-Kattegat region in comparison with the Central Baltic Sea (Pitarch et al., 2016). The ONNS-retrieved *Chl* values are in this range and reflect the geographic expectations (**Figure 7A**). But from the filamentous patterns it can be assumed that a cyanobacteria bloom including floating vegetation has developed in the Gotland Basin. In surface blooms, the concentrations are typically much higher, but the estimated biomass concentration seems too low. On the other hand, in some of these cases with visible algae structures, the out-of-range warning flag is raised (mostly OWT 9). Unfortunately, no *in situ* validation data of this OLCI image are available. The spatial distribution of *ISM* yields partly implausible features (**Figure 7B**). Commonly, significant concentrations of non-algae particles are not expected at the open sea. The concentration ranges, however, fit to observations by Berthon and Zibordi (2010). The regional distribution of CDOM, including the east-west gradient and high concentrations in the northern Baltic and Gulf of Finland, is plausible (e.g., Kowalczuk, 1999; Berthon and Zibordi, 2010; Ylöstalo et al., 2016). The corresponding uncertainty maps partly mirror the boundaries of the dominant water classes. Again, the high *Chl* uncertainties reflect the likewise high modeled variability (**Figures 2B, 4, 7D**).

ONNS application to contemporaneous R_{rs} measurements in the German Bight yield plausible results. All measured spectra exhibit maximum membership in OWT 5, the same as derived from the OLCI image for the transect (**Figure 6A**) and from 90% of the *in situ* data from the same area. The results are entirely in the same magnitudes as our previously measured in this area. ONNS estimates *Chl* along the transect between 1.7 and 4.5 mg m^{-3}, *ISM* from 0.7 to 3.4 g m^{-3}, and CDOM absorption (at 440 nm) between 0.38 and 0.68 m^{-1}. Due to tides and hydrologic changes of the Elbe river plume, *ISM* can be higher than 10 g m^{-3} near the coast.

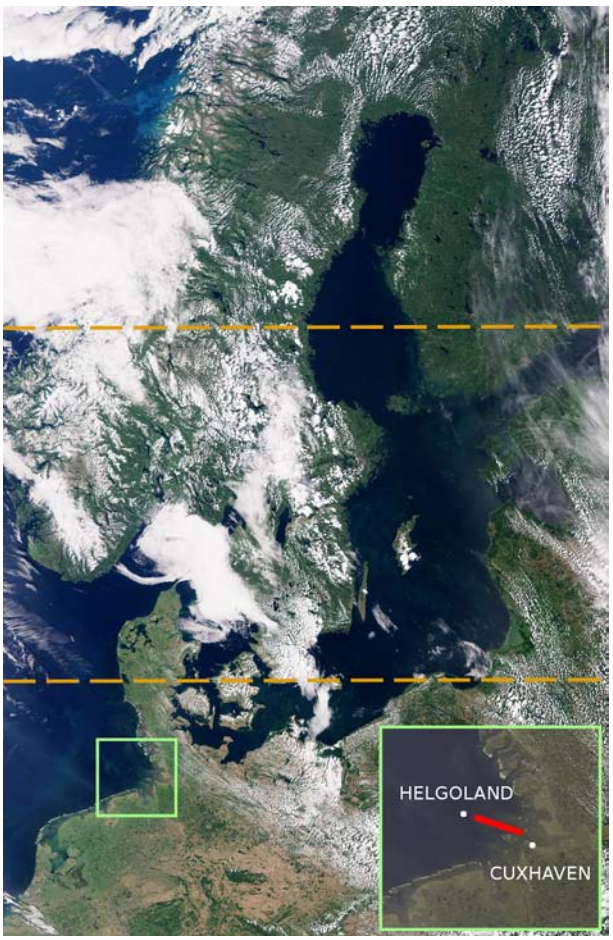


FIGURE 5 | Sentinel-3 OLCI (top-of-atmosphere) scene of 20 July 2016 (contains modified Copernicus Sentinel data [2016] processed by ESA/EUMETSAT/HZG). The boundaries of individual scenes are marked with dashed lines. The picture detail shows the route with reflectance measurements in the German Bight.

DISCUSSION

Retrieval Accuracy

In general, the retrieval statistics of ONNS (**Tables A1**, **A2**, **Figure 4**) display the general problems of OC algorithms in the various water types (e.g., Blondeau-Patissier et al., 2004; Gregg and Casey, 2004; IOCCG, 2010; Odermatt et al., 2012; Brewin et al., 2015; Pitarch et al., 2016). All the more, one has to critically assess the quality of the R_{rs} input spectra, which is influenced by two factors: the sensor calibration (over which we have no control) and the atmospheric correction (AC). A careful atmospheric correction is of systematic importance for the success of the in-water algorithm—in particular for extreme Case-2 waters. Most of the light arriving at the satellite has been scattered by the atmosphere or reflected at the sea surface. The atmospheric path radiance is typically >85% of the total signal in C1 waters, >60% in C2SX, and >94% in C2AX waters (IOCCG, 2010). Existing AC processors address the various modeling aspects quite differently, e.g., the treatment of subvisible cirrus clouds or aerosol properties, and therefore, have strengths and weaknesses for specific water types (Müller et al., 2015a). In view of the new algorithm ONNS, which relies on normalized reflectances, angle-dependent AC processes such as "smile correction" and sun glint handling are important as well to ensure spatial homogeneity of satellite data. OLCI's viewing direction is slightly shifted in comparison with MERIS in order to reduce sun glint contaminated areas. However, some AC processors incorporate sun glint contributions in their reflectance models and derive normalized R_{rs} in this condition. These AC yield much larger coverage of data (Müller et al., 2015b), but areas with high glint should nevertheless be considered cautiously. One of the AC processors is Polymer (Steinmetz et al., 2011), which reveals good performance in comparison with MERIS match-ups (Müller et al., 2015a). Another processor is C2RCC (Case-2 Regional CoastColour; Brockmann et al., 2016), which is an evolution of the precursors "Case-2 Regional," "ForwardNN," and the "MERIS Case-2 water" algorithm (Doerffer and Schiller, 2007). C2RCC is available through ESA's Sentinel toolbox SNAP and it is used in the Sentinel-3 OLCI ground segment processor

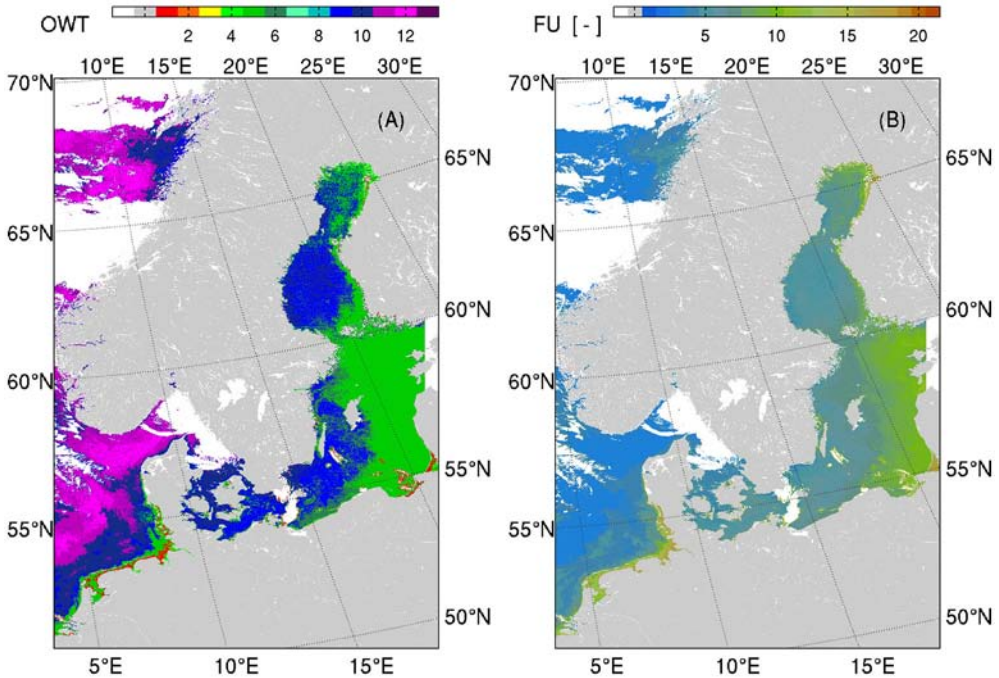


FIGURE 6 | ONNS application to the OLCI scene (20 July 2016, Figure 5). (A): Optical water type classes with maximum membership. The gray color marks land areas, white shows the cloud mask above water and inland waters, and the other colors correspond to the spectra in **Figure 3**. **(B)**: Retrieved Forel-Ule colors with true color impression.

of ESA for generating Case-2 water products. Both Case-2 AC algorithms, Polymer and C2RCC, provide usable normalized R_{rs} at OLCI bands and give comparable memberships of OWT classes. Some differences between C2RCC and Polymer-derived R_{rs} are visible, mainly in the shape of the reflectance spectrum in the violet-blue spectral range. The sensor calibration especially for shorter wavelengths is subject to current investigations. Future versions of AC algorithms may incorporate the specific sensor properties. For this reason, we have to keep in mind that the results of ONNS to a certain degree rely on the applied atmospheric correction and data reprocessing version (subject to ongoing research).

One of the most important ocean color quantity is chlorophyll concentration. It is our general impression that ONNS delivers *Chl* in the expected orders of magnitude. Future tests must show the suitability of ONNS in comparison with other algorithms, globally and for the specific region (e.g., Blondeau-Patissier et al., 2004; Darecki and Stramski, 2004; Gregg and Casey, 2004; Attila et al., 2013). However, the Baltic Sea for example is known for intense cyanobacteria blooms with small-scale patches and extreme high biomass conditions ($Chl > 200\,\mathrm{mg\ m^{-3}}$) partly associated with surface scums and floating algae. Under these conditions, results from different satellite sensors are very variable, values of *Chl* may exceed processing limits and atmospheric correction often fails (Reinart and Kutser, 2006). Optical properties of floating (also air bubble containing) material can be distinctly different from the data basis assumed in this work, e.g., higher backscattering and also higher reflectance in the NIR. Consequently, R_{rs} resembles dry vegetation rather than water (e.g., Kutser, 2004; Matthews et al., 2012), or—in terms of the defined optical water types—looks like (extreme) scattering waters. This means, as a corollary, that high biomass (*Chl*) is rather interpreted as non-algae particles (*ISM*). The out-of-range warning is notified for some of the affected areas. But here, it can be useful to raise an additional flag for surface scum conditions as it is suggested by Matthews et al. (2012). With regards to the possible misinterpretation of algae vs. non-algae particles, we must concede a potential weakness of the bio-geo-optical model assumptions underlying the simulated data basis. For example, the data include high variability of scattering properties but do not take scattering properties of different species into account (only chlorophyll-specific absorption), but it is likely that they are different (e.g., Harmel et al., 2016).

The selected OLCI scene is a good example for phytoplankton diversity. It is not well visible in **Figure 5**, but different algae blooms occur (none of them are confirmed). Very likely, a cyanobacteria bloom occurred in the Gotland Basin of the Baltic Sea. The bright water top left of the image along the Norwegian coast points to the occurrence of blooming coccolithophore. Moreover, west of the island Sylt in the German Bight fingerlike structures related to enhanced biomass are recognizable. The fact that we have to deal with different species within predominant water types increases the uncertainties. Chlorophyll-specific variability is included in the database for ONNS (**Figure 2A**).

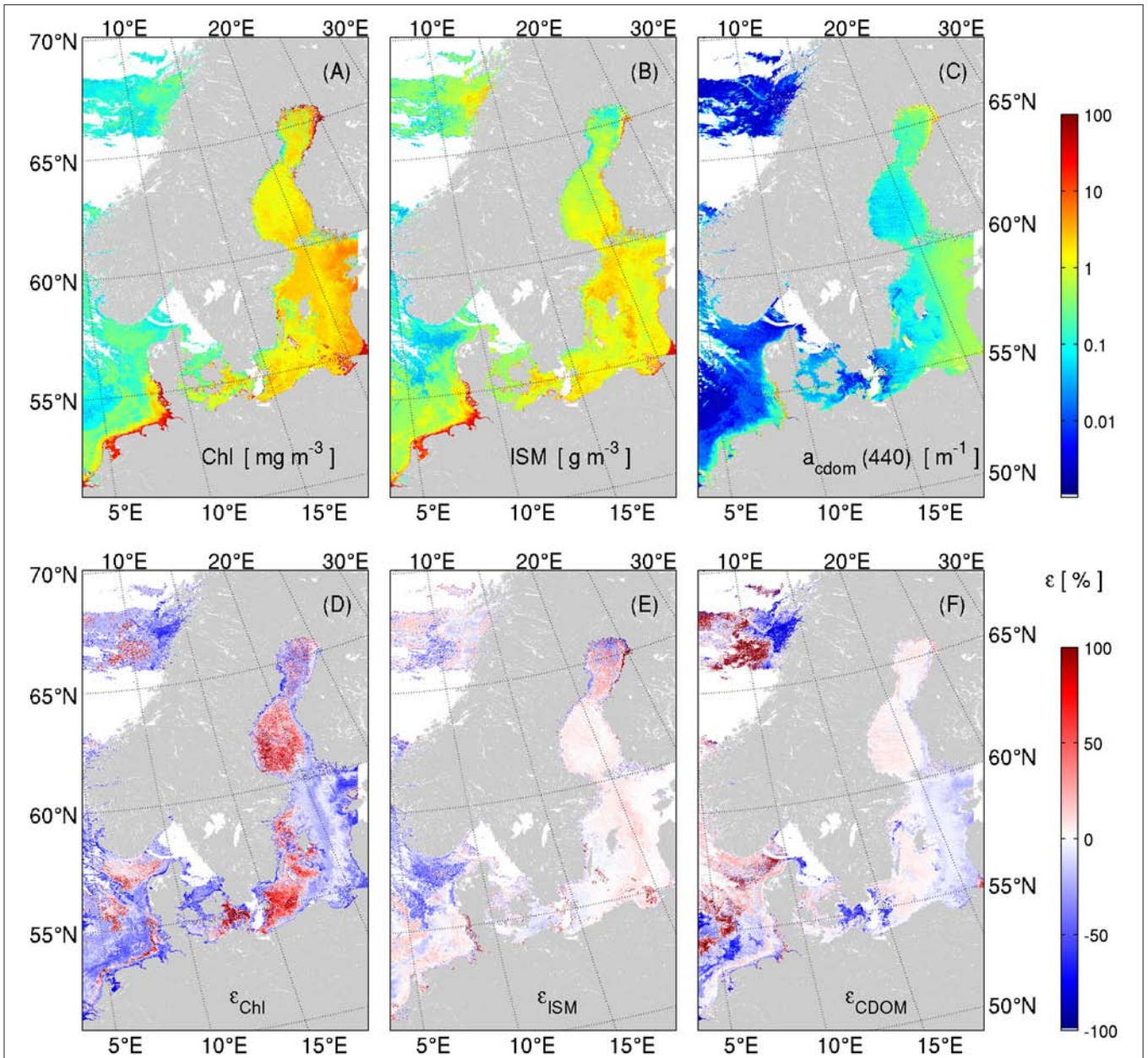


FIGURE 7 | ONNS application to the OLCI scene (20 July 2016, Figure 5). (A–C): Upper panels show concentrations of chlorophyll, inorganic suspended matter, and CDOM (absorption coefficient at 440 nm). **(D–F)**: The panels below show the corresponding retrieval uncertainties. Land and clouds are masked out.

This and the high natural variability of phytoplankton absorption vs. *Chl* concentration are reflected in the uncertainty estimates of ONNS. On this basis, future developments of ONNS may be directed into optical differentiation of diversity with corresponding traceability of uncertainties (Bracher et al., 2017; Mouw et al., 2017).

It is a frequent practice to derive inherent optical properties from ocean color and from this create an empirical relationship to observed concentrations (e.g., Doerffer and Schiller, 2007). Besides the directly retrieved concentrations, ONNS provides a number of IOPs and AOPs. The nets which derive concentrations (**Table A1**, second column) must balance their estimates indirectly by means of the relationship between absorption and scattering properties. NNs that retrieve IOPs or AOPs can rather focus on either spectral reflectance reduction or enhancement, i.e., absorption or scattering. This is demonstrated by the statistical analysis shown in **Table 3**, **A2**. The correlation coefficients of estimated IOPs and AOPs are generally very high, even if we have to deal with cases of significant pigment absorption masking due to the influence of sediments. Once the spectrum is properly classified, the anticipated values are significantly restricted, which leads to high correlation. The

derived sets of IOPs and AOPs form closure in a similar manner as the Hydrolight simulations, e.g., if we compare the sum of absorption and scattering coefficients at 440 nm (which is the attenuation coefficient) with the diffuse attenuation coefficient of downwelling irradiance, $K_d(490)$, the correlation coefficient yields 0.856 for the test data and 0.841 for the ONNS-retrieved values, both seen over all water types. Due to the high input-variability of phytoplankton absorption vs. *Chl* concentration, we expect higher uncertainties related to the biomass concentration than related to other quantities as for example CDOM absorption (**Figures 4**, **7**). The exploitation of various IOPs, e.g., absorption coefficient $a_{dg}(412)$ or backscattering coefficient $b_{bp}(510)$, may lead to more accurate and regionalized OC products. Additionally, certain absorption or scattering properties may help identifying oceanographic features such as water masses or (sub-) mesoscale eddies and frontal systems (structures are best visible in total particulate backscattering). ONNS already is a regionally employable algorithm that delivers plausible outputs, and is thus in line with new multi-water type ocean color algorithms (e.g., D'Alimonte et al., 2014; Moore et al., 2014).

ONNS Design

The present version of the bio-geo-optical processing scheme applies 11 (from 21) OLCI bands (namely at 400, 412.5, 442.5, 490, 510, 560, 620, 665, 755, 777.5, and 865 nm). Wavebands that are affected by phytoplankton fluorescence (at 673.75, 681.25, and to only a minor degree at 708.75 nm) are not utilized. In principle, the inclusion of these bands could help the classification and *Chl* retrieval capacity, in particular for highly eutrophic waters. Admission of the three additional bands, also in the combined form of a fluorescence line height, slightly increases the accuracy of the *Chl* retrieval with respect to the simulated dataset, where inelastic scattering features with the standard settings of Hydrolight are included. But we have to keep in mind that fluorescence (quantum yield efficiency) is subject to strong fluctuations and potential false assessment; it has diurnal variability, depends on nutrient- and light-availability and algae species (e.g., Greene et al., 1994). The retrieval accuracy slightly decreases if the present fluorescence line height mismatches the expected range from the training dataset. Our tests show that this is less of a problem in case of *in situ* measured remote sensing reflectance, but deviations can be higher in case of atmospheric corrected satellite data (tested with C2RCC and Polymer). A proper atmospheric correction for these bands is difficult to achieve. For this reason, the *Chl* fluorescence bands are not used in the present version of the ONNS algorithm.

The main purpose of other OLCI NIR bands is atmospheric correction, e.g., due to oxygen and water vapor absorption and optical features of aerosols. Thus, satellite-derived R_{rs} is not provided for all of the NIR bands (Steinmetz et al., 2011). In contrast, many radiometers that are deployed for *in situ* R_{rs} determination measure hyper-spectrally in the VIS and NIR range (e.g., Ramses sensors). Hence, 20 OLCI bands (or more bands) could be theoretically used for a bio-geo-optical algorithm. The last OLCI band at 1,020 nm is not covered by many radiometers. For this reason and because of the little information gain in most waters, the 1,020 nm band was also not selected for ONNS input. However, we must bear in mind that available spectral bands in compiled *in situ* datasets (e.g., Nechad et al., 2015; Valente et al., 2016) are limited too, making meaningful validation difficult.

The new algorithm deploys R_{rs} that are angle-normalized, i.e., the sun is at zenith and the viewing direction is perpendicular. All sun and viewing angle-related effects must be eliminated by the atmospheric correction prior to ONNS application. The approach simplifies for example comparisons of different satellite sensors. The first step of the processing scheme is to transform the input R_{rs} into brightness-scaled reflectances. The advantage of this approach is that the classification is less sensitive to the amplitude of R_{rs} spectra, which can be shifted by various scattering processes, e.g., due to wind-dependent micro-bubbles in water (white scatterer), marine particle aggregation, particle size, or just under-estimation of the measured total scattering (e.g., McKee et al., 2013).

OWT Classification

The classification of synthetic validation data with same data source shows general good performance for most of the water types. Occasionally, in <3% of the validation data, the fuzzy classification yields no plausible memberships of the classes and thus no ONNS-retrieval values. Classification of very weak remote sensing reflectance signals, for example, is still challenging but mostly possible. The reason is that, e.g., in CDOM-rich lakes, the reflectance is near zero in almost the entire VIS, but nevertheless, significant phytoplankton biomass can be present (e.g., Kutser et al., 2016).

Fuzzy logic classification of the *in situ* and simulated CCRR validation data yields no significant memberships in approximately 10% of cases, i.e., in 5 and 488 cases respectively. Hence, the spectra were not considered plausible and the final blended retrieval delivers no results. One possible explanation is that spectral shapes of R_{rs} appear which not occur in the database with 10^5 spectra. Moore et al. (2001) propose a minimum threshold for class memberships, which was arbitrarily set at 10^{-4}. If this threshold is lowered to 10^{-5}, the non-plausible cases reduce to <1% of both datasets. Further tests are needed in the framework of an all-water-type-embracing validation.

Applied to a satellite scene, all marine spectra are classifiable, but water classification can be problematic and spatially heterogeneous in association with cloud and adjacency effects. Nonetheless, the optical water type classification of the scene basically yields geographically expected results. Three of the water classes (OWT 3, 7, and 8) gained never significant weights. Therefore, they were not used for blending. Those cases include extreme absorbing or scattering cases with very high biomass, e.g., like the mentioned "black lakes" (Kutser et al., 2016) or the Gulf of Finland (Ylöstalo et al., 2016), which is mostly flagged out due to clouds in the example scene. In global terms, they are restricted very locally. However, the three cases each represent a spectral R_{rs} with maxima in one of the three selected NIR bands (**Figure 3**); therefore, they have an essential function in the fuzzy logic classification scheme.

Application to Radiometric *In situ* Data

The OC processor ONNS can be applied to *in situ* measurements as well. In this case, an atmospheric correction is not needed. Remote sensing reflectance can be determined from above- or in-water radiometric measurements. Nechad et al. (2015) assembled reflectance measurements that are gained in five different manners. In case of above-water measurements for example, downwelling irradiance, sky radiance, and upwelling radiance are measured under specific sun-viewing angle conditions. From this, R_{rs} is determined using a surface reflectance factor, which depends on sun-viewing geometry and wind speed and which is usually between 2 and 5% for optimal viewing angles (Hieronymi, 2016; Zibordi, 2016). The surface reflectance factor determines the shape of R_{rs} mainly in the violet-blue range. Thus, there are some uncertainties for the application of the classification scheme and the overall processor. Optimally, the measured R_{rs} is adapted to the input criteria for ONNS, i.e., fully normalized, a moderate wind speed, no micro-bubbles in water, etc., In our tests, the fuzzy logic classification scheme is applicable and yields useable inputs for ONNS. Furthermore, ONNS retrieves OC products in the expected orders of magnitude. However, this cannot hide the fact that the retrieval statistic for the comparison with *in situ* data could be better (**Table 3**). Thus, more and all-water-type-embracing validation is needed.

Outlook

Proper validation of ONNS products using *in situ* data and OLCI match-ups will be a future task. Every OWT class must be validated (and possibly readjusted) independently, knowing that it is the balance between the water constituents (phytoplankton, minerals, and CDOM), represented in the training data, that decides on the quality of the OC products (D'Alimonte et al., 2016). Furthermore, higher validation uncertainties must be expected in extreme Case-2 waters and heterogeneous waters in coastal areas or during algae blooms (e.g., Kutser, 2004; Pahlevan et al., 2016). These *in situ* validation uncertainties must be incorporated into the delivered uncertainty products. The aim of this paper is to provide the scientific background description of the processor together with a baseline validation of the present ONNS version (v0.4). Once the processor has been validated, it is planned to make it freely accessible via ESA's Sentinel toolbox SNAP. After that, the algorithm will be compared with other bio-geo-optical algorithms.

In principle, ONNS can provide results in near-real time. The computational time depends on the (in our case high) number of neurons of the NNs and a swarm of 4×13 NNs obviously takes more time. However, single NNs are fast and the processing can occur in parallel. Thus, OC products can be disseminated in near real time mode, which usually comprises the time up to one day after satellite acquisition.

CONCLUSIONS

This study presents a novel in-water algorithm for the retrieval of ocean color remote sensing products from atmospheric corrected OLCI-like satellite imagery or *in situ* radiometric measurements. The algorithm consists of several specialized neural networks with task-optimized architectures (OLCI Neural Network Swarm). The products contain concentrations of water constituents (*Chl* and *ISM*), inherent and apparent optical properties [$a_{cdom}(440)$, $a_p(440)$, $a_m(440)$, $a_{dg}(412)$, $b_p(440)$, $b_m(440)$, $b_{bp}(510)$, $K_d(490)$, and $K_u(490)$], and a sea color index (*FU*). In addition, all products are delivered with an uncertainty estimate that describes the deviation of the product from the original data basis. The algorithm makes use of a comprehensive fuzzy logic classification scheme. Thirteen optical water type classes have been identified based on Hydrolight simulated and brightness-scaled remote sensing reflectances at 11 OLCI bands (400, 412.5, 442.5, 490, 510, 560, 620, 665, 755, 777.5, and 865 nm). The corresponding water types range from clearest sea waters to extreme Case-2 waters (**Table 1**). This includes chlorophyll concentrations up to 200 mg m^{-3}, non-algae particle concentrations up to 1,500 g m^{-3}, and an absorption coefficient of colored dissolved organic matter up to 20 m^{-1} at 440 nm. A baseline validation of ONNS products for the various water types is provided, showing principle strengths and weaknesses of the algorithm. With simulated test data the algorithm performs generally well within the wide range of optical properties of the water. Additional tests have been conducted using simulated data from the independent CCRR database and a few *in situ* data; both datasets contain mostly turbid Case-2 waters, which are classified in few optical water type classes. As might be expected, these comparisons revealed somewhat worse correlation but are overall encouraging, for example regarding *ISM* and CDOM retrieval. An appropriate full validation for all OWT classes and all provided ocean color products is still to be done. Conclusions on the performance using OLCI Earth observation data can be drawn after throughout validation against field measurements or other bio-geo-optical algorithms. The shown example demonstrates that ONNS-estimated ocean color products are mostly within the range of observed concentrations (e.g., Kowalczuk, 1999; Berthon and Zibordi, 2010; Pitarch et al., 2016; Ylöstalo et al., 2016). From our present point of view, we conclude that the new ONNS in-water algorithm is suited for the remote sensing estimation of water properties and constituents of most natural waters.

AUTHOR CONTRIBUTIONS

MH and DM developed the concept of the processor ONNS with consultancy of RD. MH prepared the synthetic data basis, trained the neural networks, and wrote the paper. DM developed the water type classification, processed data, and contributed text modules. RD helped with the processor development and fed the discussion.

FUNDING

This work is a contribution to the European Space Agency (ESA) funded Ocean Colour Climate Change Initiative (OC-CCI: AO-1/6207/09/I-LG), Case-2 Extreme Water project (C2X: 4000113691/15/I-LG), and Living Planet Fellowship Programme (LowSun-OC: 4000112803/15/I-SBo).

ACKNOWLEDGMENTS

This paper is an outcome of the CLEO Workshop "Colour and Light in the Ocean from Earth Observation," held in Frascati, Italy in September 2016. The authors would like to acknowledge data processing and discussions with colleagues Rüdiger Röttgers, Hajo Krasemann, Kerstin Heymann, and Wolfgang Schönfeld. In addition, the authors thank the C2X project and science support team for valuable comments throughout the algorithm development, in particular Carsten Brockmann, Kerstin Stelzer, Ana Ruescas, François Steinmetz, Kevin Ruddick, Bouchra Nechad, Gavin Tilstone, Stefan Simis, and Peter Regner. We thank ESA/ EUMETSAT/ EU Copernicus for providing Sentinel-3 data and for permission to use them. Finally, the detailed comments of three reviewers and of the guest associate editor Tiit Kutser are highly appreciated.

REFERENCES

Attila, J., Koponen, S., Kallio, K., Lindfors, A., Kaitala, S., and Ylöstalo, P. (2013). MERIS Case II water processor comparison on coastal sites of the northern Baltic Sea. *Rem. Sens. Environ.* 128, 138–149. doi: 10.1016/j.rse.2012.07.009

Beltrán-Abaunza, J. M., Kratzer, S., and Brockmann, C. (2014). Evaluation of MERIS products from Baltic Sea coastal waters rich in CDOM. *Ocean Sci.* 10, 377–396. doi: 10.5194/os-10-377-2014

Berthon, J. F., and Zibordi, G. (2010). Optically black waters in the northern Baltic Sea. *Geophys. Res. Let.* 37:L09605. doi: 10.1029/2010gl043227

Blondeau-Patissier, D., Tilstone, G. H., Martinez-Vicente, V., and Moore, G. F. (2004). Comparison of bio-physical marine products from SeaWiFS, MODIS and a bio-optical model with *in situ* measurements from Northern European waters. *J. Optics A* 6:875. doi: 10.1088/1464-4258/6/9/010

Bracher, A., Bouman, H., Brewin, R. J., Bricaud, A., Brotas, V., Ciotti, A. M., et al. (2017). Obtaining phytoplankton diversity from ocean color: a scientific roadmap for future development. *Front. Mar. Sci.* 4:55. doi: 10.3389/fmars.2017.00055

Brewin, R. J., Sathyendranath, S., Müller, D., Brockmann, C., Deschamps, P.-Y., Devred, E., et al. (2015). The ocean colour climate change initiative: III. a round-robin comparison on in-water bio-optical algorithms. *Rem. Sens. Environ.* 162, 271–294. doi: 10.1016/j.rse.2013.09.016

Bricaud, A., Claustre, H., Ras, J., and Oubelkheir, K. (2004). Natural variability of phytoplanktonic absorption in oceanic waters: influence of the size structure of algal populations. *J. Geophys. Res.* 109:C11010. doi: 10.1029/2004jc002419

Bricaud, A., Mejia, C., Blondeau-Patissier, D., Claustre, H., Crepon, M., and Thiria, S. (2007). Retrieval of pigment concentrations and size structure of algal populations from their absorption spectra using multilayered perceptrons. *Appl. Opt.* 46, 1251–1260. doi: 10.1364/AO.46.001251

Brockmann, C., Doerffer, R., Peters, M., Stelzer, K., Embacher, S., and Ruescas, A. (2016). "Evolution of the C2RCC neural network for Sentinel 2 and 3 for the retrieval of ocean colour products in normal and extreme optically complex waters," in *Proceeding of Living Planet Symposium* (Prague: ESA SP-740).

Chen, J., Quan, W., Cui, T., Song, Q., and Lin, C. (2014). Remote sensing of absorption and scattering coefficient using neural network model: development, validation, and application. *Rem. Sens. Environ.* 149, 213–226. doi: 10.1016/j.rse.2014.04.013

D'Alimonte, D., Kajiyama, T., and Saptawijaya, A. (2016). Ocean color remote sensing of atypical marine optical cases. *IEEE Trans. Geosci. Rem. Sens.* 54, 6574–6586. doi: 10.1109/TGRS.2016.2587106

D'Alimonte, D., and Zibordi, G. (2003). Phytoplankton determination in an optically complex coastal region using a multilayer perceptron neural network. *IEEE Trans. Geosci. Rem. Sens.* 41, 2861–2868. doi: 10.1109/TGRS.2003.817682

D'Alimonte, D., Zibordi, G., Kajiyama, T., and Berthon, J. F. (2014). Comparison between MERIS and regional high-level products in European seas. *Rem. Sens. Environ.* 140, 378–395. doi: 10.1016/j.rse.2013.07.029

Darecki, M., and Stramski, D. (2004). An evaluation of MODIS and SeaWiFS bio-optical algorithms in the Baltic Sea. *Rem. Sens. Environ.* 89, 326–350. doi: 10.1016/j.rse.2003.10.012

Doerffer, R., and Schiller, H. (2000). Neural network for retrieval of concentrations of water constituents with the possibility of detecting exceptional out of scope spectra. in *Proceedings IGARSS 2000. IEEE 2000 International,* Vol. 2 (Honolulu), 714–717.

Doerffer, R., and Schiller, H. (2007). The MERIS Case 2 water algorithm. *Int. J. Rem. Sens.* 28, 517–535. doi: 10.1080/01431160600821127

Donlon, C., Berruti, B., Buongiorno, A., Ferreira, M. H., Féménias, P., Frerick, J., et al. (2012). The global monitoring for environment and security (GMES) sentinel-3 mission. *Rem. Sens. Environ.* 120, 37–57. doi: 10.1016/j.rse.2011.07.024

Fawcett, A., Bernard, S., Pitcher, G. C., Probyn, T. A., and du Randt, A. (2006). Real-time monitoring of harmful algal blooms in the southern Benguela. *Afri. J. Mar. Sci.* 28, 257–260. doi: 10.2989/18142320609504158

Greene, R. M., Kolber, Z. S., Swift, D. G., Tindale, N. W., and Falkowski, P. G. (1994). Physiological limitation of phytoplankton photosynthesis in the eastern equatorial Pacific determined from variability in the quantum yield of fluorescence. *Limnol. Oceanogr.* 39, 1061–1074. doi: 10.4319/lo.1994.39.5.1061

Gregg, W. W., and Casey, N. W. (2004). Global and regional evaluation of the SeaWiFS chlorophyll data set. *Rem. Sens. Environ.* 93, 463–479. doi: 10.1016/j.rse.2003.12.012

Gross, L., Thiria, S., and Frouin, R. (1999). Applying artificial neural network methodology to ocean color remote sensing. *Ecol. Mod.* 120, 237–246. doi: 10.1016/S0304-3800(99)00105-2

Harmel, T., Hieronymi, M., Slade, W., Röttgers, R., Roullier, F., and Chami, M. (2016). Laboratory experiments for inter-comparison of three volume scattering meters to measure angular scattering properties of hydrosols. *Opt. Exp.* 24, A234–A256. doi: 10.1364/oe.24.00a234

Harvey, E. T., Kratzer, S., and Philipson, P. (2015). Satellite-based water quality monitoring for improved spatial and temporal retrieval of chlorophyll-a in coastal waters. *Rem. Sens. Environ.* 158, 417–430. doi: 10.1016/j.rse.2014.11.017

Hieronymi, M. (2016). Polarized reflectance and transmittance distribution functions of the ocean surface. *Opt. Exp.* 24, A1045–A1068. doi: 10.1364/oe.24.0a1045

Hieronymi, M., Krasemann, H., Müller, D., Brockmann, C., Ruescas, A., Stelzer, K., et al. (2016). "Ocean Colour Remote Sensing of Extreme Case-2 Waters," in *Proceedings of Living Planet Symposium* (Prague: ESA SP-740).

Hieronymi, M., Müller, D., Krasemann, H., Schönfeld, W., Röttgers, R., and Doerffer, R. (2015). "Regional ocean colour remote sensing algorithm for the Baltic Sea," in *Proceedings of Sentinel-3 for Science Workshop* (Venice: ESA SP-734).

ICES (2011). *Historical Plankton Dataset*. Copenhagen: ICES.

Ioannou, I., Gilerson, A., Gross, B., Moshary, F., and Ahmed, S. (2011). Neural network approach to retrieve the inherent optical properties of the ocean from observations of MODIS. *Appl. Opt.* 50, 3168–3186. doi: 10.1364/A.O.50.003168

IOCCG (2010). "Atmospheric correction for remotely-sensed ocean-colour products," in *Reports and Monographs of the International Ocean-Colour Coordinating Group*, ed M. Wang (Dartmouth: IOCCG).

Jamet, C., Loisel, H., and Dessailly, D. (2012). Retrieval of the spectral diffuse attenuation coefficient K_d (λ) in open and coastal ocean waters using a neural network inversion. *J. Geophys. Res.* 117:C10023. doi: 10.1029/2012JC008076

Knaeps, E., Dogliotti, A. I., Raymaekers, D., Ruddick, K., and Sterckx, S. (2012). *In situ* evidence of non-zero reflectance in the OLCI 1020 nm band for a turbid estuary. *Rem. Sens. Environ.* 120, 133–144. doi: 10.1016/j.rse.2011.07.025

Kowalczuk, P. (1999). Seasonal variability of yellow substance absorption in the surface layer of the Baltic Sea. *J. Geophys. Res.* 104, 30047–30058. doi: 10.1029/1999JC900198

Kutser, T. (2004). Quantitative detection of chlorophyll in cyanobacterial blooms by satellite remote sensing. *Limnol. Oceanogr.* 49, 2179–2189. doi: 10.4319/lo.2004.49.6.2179

Kutser, T., Paavel, B., Verpoorter, C., Ligi, M., Soomets, T., Toming, K., et al. (2016). Remote sensing of black lakes and using 810 nm reflectance peak for retrieving water quality parameters of optically complex waters. *Rem. Sens.* 8:497. doi: 10.3390/rs8060497

Matthews, M. W., Bernard, S., and Robertson, L. (2012). An algorithm for detecting trophic status (chlorophyll-a), cyanobacterial-dominance, surface scums and floating vegetation in inland and coastal waters. *Rem. Sens. Environ.* 124, 637–652. doi: 10.1016/j.rse.2012.05.032

McKee, D., Piskozub, J., Röttgers, R., and Reynolds, R. A. (2013). Evaluation and improvement of an iterative scattering correction scheme for *in situ* absorption and attenuation measurements. *J. Atmos. Oceanic Tech.* 30, 1527–1541. doi: 10.1175/JTECH-D-12-00150.1

Mobley, C. D. (1994). *Light and Water: Radiative Transfer in Natural Waters.* New York, NY: Academic.

Mobley, C. D., Stramski, D., Bissett, W. P., and Boss, E. (2004). Optical modeling of ocean waters: is the Case 1-Case 2 classification still useful? *Oceanography* 17, 60–67. doi: 10.5670/oceanog.2004.48

Mobley, C. D., and Sundman, L. K. (2013). *Hydrolight 5.2, Ecolight 5.2, Technical Documentation*. Bellevue: Sequoia Scientific Inc.

Moore, T. S., Campbell, J. W., and Dowell, M. D. (2009). A class-based approach to characterizing and mapping the uncertainty of the MODIS ocean chlorophyll product. *Rem. Sens. Environ.* 113, 2424–2430. doi: 10.1016/j.rse.2009.07.016

Moore, T. S., Campbell, J. W., and Feng, H. (2001). A fuzzy logic classification scheme for selecting and blending satellite ocean color algorithms. *IEEE Trans. Geosci. Rem. Sens.* 39, 1764–1776. doi: 10.1109/36.942555

Moore, T. S., Dowell, M. D., Bradt, S., and Verdu, A. R. (2014). An optical water type framework for selecting and blending retrievals from bio-optical algorithms in lakes and coastal waters. *Rem. Sen. Environ.* 143, 97–111. doi: 10.1016/j.rse.2013.11.021

Moore, T. S., Dowell, M. D., and Franz, B. A. (2012). Detection of coccolithophore blooms in ocean color satellite imagery: a generalized approach for use with multiple sensors. *Rem. Sens. Environ.* 117, 249–263. doi: 10.1016/j.rse.2011.10.001

Morel, A., and Prieur, L. (1977). Analysis of variations in ocean color. *Limnol. Oceanogr.* 22, 709–722. doi: 10.4319/lo.1977.22.4.0709

Mouw, C. B., Hardman-Mountford, N. J., Alvain, S., Bracher, A., Brewin, R. W., Bricaud, A., et al. (2017). A consumer's guide to satellite remote sensing of multiple phytoplankton groups in the global ocean. *Front. Mar. Sci.* 4:41. doi: 10.3389/fmars.2017.00041

Müller, D., Krasemann, H., Brewin, R. J., Brockmann, C., Deschamps, P. Y., Doerffer, R., et al. (2015a). The ocean colour climate change initiative: I. A methodology for assessing atmospheric correction processors based on *in-situ* measurements. *Rem. Sens. Environ.* 162, 242–256. doi: 10.1016/j.rse.2013.11.026

Müller, D., Krasemann, H., Brewin, R. J., Brockmann, C., Deschamps, P. Y., Doerffer, R., et al. (2015b). The ocean colour climate change initiative: II. spatial and temporal homogeneity of satellite data retrieval due to systematic effects in atmospheric correction processors. *Rem. Sens. Environ.* 162, 257–270. doi: 10.1016/j.rse.2015.01.033

Nechad, B., Ruddick, K., Schroeder, T., Oubelkheir, K., Blondeau-Patissier, D., Cherukuru, N., et al. (2015). CoastColour Round Robin data sets: a database to evaluate the performance of algorithms for the retrieval of water quality parameters in coastal waters. *Earth Syst. Sci. Data* 7, 319–348. doi: 10.5194/essd-7-319-2015

Odermatt, D., Gitelson, A., Brando, V. E., and Schaepman, M. (2012). Review of constituent retrieval in optically deep and complex waters from satellite imagery. *Rem. Sens. Environ.* 118, 116–126. doi: 10.1016/j.rse.2011.11.013

Pahlevan, N., Sarkar, S., and Franz, B. A. (2016). Uncertainties in coastal ocean color products: impacts of spatial sampling. *Rem. Sens. Environ.* 181, 14–26. doi: 10.1016/j.rse.2016.03.022

Pitarch, J., Volpe, G., Colella, S., Krasemann, H., and Santoleri, R. (2016). Remote sensing of chlorophyll in the Baltic Sea at basin scale from 1997 to 2012 using merged multi-sensor data. *Ocean Sci.* 12, 379–389. doi: 10.5194/os-12-379-2016

Reinart, A., and Kutser, T. (2006). Comparison of different satellite sensors in detecting cyanobacterial bloom events in the Baltic Sea. *Rem. Sens. Environ.* 102, 74–85. doi: 10.1016/j.rse.2006.02.013

Röttgers, R., Doerffer, R., McKee, D., and Schönfeld, W. (2016). *The Water Optical Properties Processor (WOPP): Pure Water Spectral Absorption, Scattering, and Real Part of Refractive Index Model, ATBD,* Issue 1.8, ESA Water Radiance project.

Ruddick, K., De Cauwer, V., Park, Y. J., and Moore, G. (2006). Seaborne measurements of near infrared water-leaving reflectance: the similarity spectrum for turbid waters. *Limnol. Oceanogr.* 51, 1167–1179. doi: 10.4319/lo.2006.51.2.1167

Schiller, H. (2000). *Feedforward-Backpropagation Neural Net Program ffbp1.* 0, Report GKSS 2000/37.

Schiller, H., and Doerffer, R. (1999). Neural network for emulation of an inverse model operational derivation of Case II water properties from MERIS data. *Int. J. Rem. Sens.* 20, 1735–1746. doi: 10.1080/014311699212443

Schiller, K. (2006). Derivation of photosynthetically available radiation from METEOSAT data in the German Bight with neural nets. *Ocean Dyn.* 56, 79–85. doi: 10.1007/s10236-006-0058-1

Schroeder, T., Schaale, M., and Fischer, J. (2007). Retrieval of atmospheric and oceanic properties from MERIS measurements: a new Case-2 water processor for BEAM. *Int. J. Rem. Sens.* 28, 5627–5632. doi: 10.1080/01431160701601774

Steinmetz, F., Deschamps, P. Y., and Ramon, D. (2011). Atmospheric correction in presence of sun glint: application to MERIS. *Opt. Exp.* 19, 9783–9800. doi: 10.1364/OE.19.009783

Tanaka, A., Kishino, M., Doerffer, R., Schiller, H., Oishi, T., and Kubota, T. (2004). Development of a neural network algorithm for retrieving concentrations of chlorophyll, suspended matter and yellow substance from radiance data of the ocean color and temperature scanner. *J. Oceanogr.* 60, 519–530. doi: 10.1023/B:JOCE.0000038345.99050.c0

Twardowski, M. S., Boss, E., Macdonald, J. B., Pegau, W. S., Barnard, A. H., and Zaneveld, J. R. V. (2001). A model for estimating bulk refractive index from the optical backscattering ratio and the implications for understanding particle composition in case I and case II waters. *J. Geophys. Res.* 106, 14129–14142. doi: 10.1029/2000JC000404

Valente, A., Sathyendranath, S., Brotas, V., Groom, S., Grant, M., Taberner, M., et al. (2016). A compilation of global bio-optical *in situ* data for ocean-colour satellite applications. *Earth Syst. Sci. Data* 8, 235–252. doi: 10.5194/essd-8-235-2016

van der Woerd, H. J., and Wernand, M. R. (2015). True colour classification of natural waters with medium-spectral resolution satellites: seaWiFS, MODIS, MERIS and OLCI. *Sensors* 15, 25663–25680. doi: 10.3390/s151025663

Werdell, P. J., and Bailey, S. W. (2005). An improved *in-situ* bio-optical data set for ocean color algorithm development and satellite data product validation. *Rem. Sens. Environ.* 98, 122–140. doi: 10.1016/j.rse.2005.07.001

Wernand, M. R., van der Woerd, H. J., and Gieskes, W. W. (2013). Trends in ocean colour and chlorophyll concentration from 1889 to 2000, worldwide. *PLoS ONE* 8:e63766. doi: 10.1371/journal.pone.0063766

Woźniak, S. B., Stramski, D., Stramska, M., Reynolds, R., Wright, V. M., Miksic, E. Y., et al. (2010). Optical variability of seawater in relation to particle concentration, composition, and size distribution in the nearshore marine environment at Imperial Beach, California. *J. Geophys. Res.* 115:C08027. doi: 10.1029/2009jc005554

Xi, H., Hieronymi, M., Röttgers, R., Krasemann, H., and Qiu, Z. (2015). Hyperspectral differentiation of phytoplankton taxonomic groups: a comparison between using remote sensing reflectance and absorption spectra. *Rem. Sens.* 7, 14781–14805. doi: 10.3390/rs71114781

Ylöstalo, P., Seppälä, J., Kaitala, S., Maunula, P., and Simis, S. (2016). Loadings of dissolved organic matter and nutrients from the Neva River into the Gulf of Finland–Biogeochemical composition and spatial distribution within the salinity gradient. *Mar. Chem.* 186, 58–71. doi: 10.1016/j.marchem.2016.07.004

Zhang, M., Tang, J., Song, Q., and Dong, Q. (2010). Backscattesring ratio variation and its implications for studying particle composition: a case study in Yellow and East China seas. *J. Geophys. Res.* 115:C12014. doi: 10.1029/2010jc006098

Zhang, T., Fell, F., Liu, Z. S., Preusker, R., Fischer, J., and He, M. X. (2003). Evaluating the performance of artificial neural network techniques for pigment retrieval from ocean color in Case I waters. *J. Geophys. Res.* 108:3286. doi: 10.1029/2002jc001638

Zibordi, G. (2016). Experimental evaluation of theoretical sea surface reflectance factors relevant to above-water radiometry. *Opt. Exp.* 24, A446–A459. doi: 10.1364/oe.24.00a446

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Hieronymi, Müller and Doerffer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

APPENDIX

TABLE A1 | Numbers of neurons from selected neural network architectures (input, hidden, and output layers) for all 13 optical water type classes.

NN input	11 (R_{rs} at 400, 412.5, 442.5, 490, 510, 560, 620, 665, 755, 777.5, and 865 nm)	11 (R_{rs} at 400, 412.5, 442.5, 490, 510, 560, 620, 665, 755, 777.5, and 865 nm)	11 (R_{rs} at 400, 412.5, 442.5, 490, 510, 560, 620, 665, 755, 777.5, and 865 nm)	12 [NN outputs: Chl, $a_{cdom}(440)$, ISM, $a_p(440)$, $a_m(440)$, $b_p(440)$, $b_m(440)$, FU, $K_d(490)$, $K_u(490)$, $a_{dg}(412)$, $b_{bp}(510)$]
OWT 1	23 × 76 × 55 × 36	23 × 76 × 55 × 36	97 × 77 × 37	37 × 77 × 97
OWT 2	23 × 76 × 55 × 36	23 × 41 × 59 × 43	23 × 76 × 55 × 36	37 × 77 × 97
OWT 3	17 × 97 × 47	17 × 97 × 47	23 × 76 × 55 × 36	37 × 77 × 97
OWT 4	97 × 77 × 37	23 × 41 × 59 × 43	23 × 76 × 55 × 36	37 × 77 × 97
OWT 5	23 × 41 × 59 × 43	37 × 77 × 97	97 × 77 × 37	97 × 77 × 37
OWT 6	23 × 41 × 59 × 43	23 × 76 × 55 × 36	23 × 76 × 55 × 36	37 × 77 × 97
OWT 7	17 × 97 × 47	17 × 97 × 47	23 × 41 × 59 × 43	97 × 77 × 37
OWT 8	17 × 97 × 47	23 × 41 × 59 × 43	23 × 76 × 55 × 36	23 × 41 × 59 × 43
OWT 9	23 × 41 × 59 × 43	23 × 76 × 55 × 36	97 × 77 × 37	97 × 77 × 37
OWT 10	23 × 41 × 59 × 43	23 × 76 × 55 × 36	23 × 76 × 55 × 36	37 × 77 × 97
OWT 11	23 × 47 × 22 × 7	97 × 77 × 37	97 × 77 × 37	37 × 77 × 97
OWT 12	97 × 77 × 37	97 × 77 × 37	23 × 76 × 55 × 36	37 × 77 × 97
OWT 13	23 × 76 × 55 × 36	23 × 41 × 59 × 43	23 × 41 × 59 × 43	37 × 77 × 97
NN output	**3 (Chl, $a_{cdom}(440)$, ISM)**	**5 [$a_{cdom}(440)$, $a_p(440)$, $a_m(440)$, $b_p(440)$, $b_m(440)$]**	**5 [FU, $K_d(490)$, $K_u(490)$, $a_{dg}(412)$, $b_{bp}(510)$]**	**12 [Training inputs: Chl, $a_{cdom}(440)$, ISM, $a_p(440)$, $a_m(440)$, $b_p(440)$, $b_m(440)$, FU, $K_d(490)$, $K_u(490)$, $a_{dg}(412)$, $b_{bp}(510)$]**

Three sets of NNs deliver selected concentrations, IOPs, AOPs, and the Forel-Ule color code. A fourth set of NNs (right column) estimates the uncertainties of the NN outputs. Inputs and outputs for the NNs are $\log_{10}(X + 0.001)$, where X stands for R_{rs} or an ocean color product (this applies not for FU).

TABLE A2 | Correlation coefficients of additional ONNS retrievals vs. simulated validation data subdivided by water type.

OC product	C1	C2A	C2S	C2AX	C2SX
$a_p(440)$	0.8548	0.9977	0.9990	0.9937	0.9370
$a_m(440)$	0.9228	0.9872	0.9975	0.9817	0.9818
$b_p(440)$	0.9105	0.8974	0.8528	0.8030	0.7878
$b_m(440)$	0.8412	0.9062	0.9838	0.8901	0.8274
FU	0.9854	0.9805	0.9802	0.9581	0.7230
$K_d(490)$	0.8488	0.9993	0.9989	0.9945	0.9991
$K_u(490)$	0.8464	0.9992	0.9989	0.9881	0.9992
$a_{dg}(412)$	0.9867	0.9943	0.9990	0.9936	0.9984
$b_{bp}(510)$	0.9877	0.9981	0.9990	0.9960	0.9984

Statistics are based on $N_{C1} = 5392$, $N_{C2A} = 4699$, $N_{C2S} = 4049$, $N_{C2AX} = 4526$, and $N_{C2SX} = 4082$.

frontiers in Marine Science

ORIGINAL RESEARCH
published: 15 May 2017
doi: 10.3389/fmars.2017.00114

Bio-Optical Properties of Two Neigboring Coastal Regions of Tropical Northern Australia: The Van Diemen Gulf and Darwin Harbour

David Blondeau-Patissier[1, 2, 3], Thomas Schroeder[3], Lesley A. Clementson[4], Vittorio E. Brando[5], Diane Purcell[1, 6], Phillip Ford[7], David K. Williams[6], David Doxaran[8], Janet Anstee[7], Nandika Thapar[9] and Miguel Tovar-Valencia[2]*

[1] North Australia Marine Research Alliance (NAMRA), Darwin, NT, Australia, [2] Research Institute for the Environment and Livelihoods (RIEL), Charles Darwin University, Darwin, NT, Australia, [3] Oceans & Atmosphere, Commonwealth Scientific and Industrial Research Organization (CSIRO), Brisbane, QLD, Australia, [4] Oceans & Atmosphere, Commonwealth Scientific and Industrial Research Organization (CSIRO), Hobart, TAS, Australia, [5] Institute of Atmospheric Sciences and Climate, National Research Council (CNR), Rome, Italy, [6] Australian Institute of Marine Science (AIMS), Darwin, NT, Australia, [7] Oceans & Atmosphere, Commonwealth Scientific and Industrial Research Organization (CSIRO), Canberra, ACT, Australia, [8] Laboratoire d'Océanographie de Villefranche-sur-Mer (LOV), Centre national de la recherche scientifique (CNRS), Villefranche-sur-mer, France, [9] Agriculture, Commonwealth Scientific and Industrial Research Organization (CSIRO), Canberra, ACT, Australia

OPEN ACCESS

Edited by:
Tiit Kutser, Estonian Marine Institute, University of Tartu, Estonia

Reviewed by:
Oliver Zielinski, University of Oldenburg, Germany
James Acker, NASA, USA

***Correspondence:**
David Blondeau-Patissier
david.blondeau-patissier@csiro.au

Specialty section:
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

Received: *01 December 2016*
Accepted: *10 April 2017*
Published: *15 May 2017*

Citation:
Blondeau-Patissier D, Schroeder T, Clementson LA, Brando VE, Purcell D, Ford P, Williams DK, Doxaran D, Anstee J, Thapar N and Tovar-Valencia M (2017) Bio-Optical Properties of Two Neigboring Coastal Regions of Tropical Northern Australia: The Van Diemen Gulf and Darwin Harbour. Front. Mar. Sci. 4:114. doi: 10.3389/fmars.2017.00114

This study focuses on the seasonal and spatial characterization of inherent optical properties and biogeochemical concentrations in the Van Diemen Gulf and Darwin Harbour, two neighboring tropical coastal environments of Northern Australia that exhibit shallow depths (~20 m), large (>3 m) semi-diurnal tides, and a monsoonal climate. To gain insight in the functioning of these optically complex coastal ecosystems, a total of 23 physical, biogeochemical, and optical parameters were sampled at 63 stations during three field campaigns covering the 2012 wet and dry seasons, and the 2013 dry season. The total light absorption budget in the Van Diemen Gulf was dominated by non-algal particles (a_{NAP}; >45%) during the dry season (May–October) and colored dissolved organic matter (a_{CDOM}; 60%) during the wet season (November–April). The combined absorption by a_{NAP} and a_{CDOM} generally exceeded ~80% of the total absorption budget from 400 to 620 nm, with phytoplankton, a_{Phy}, accounting for <20%. In Darwin Harbour, where only the dry season conditions were sampled, the total absorption budget was dominated by an equivalent contribution of a_{CDOM}, $a_{NAP,}$ and phytoplankton. The major processes explaining the seasonal variability observed in the Van Diemen Gulf are resuspension from seasonal south-easterly trade winds in combination with the tidal energy and shallow bathymetry during the dry season months, and mostly terrestrial river runoff during the monsoon which discharge terrestrial CDOM from the surrounding wetlands. Due to light-limited conditions all year round, the particulate scattering coefficient [$b_p(555)$] contributed significantly (90%) to the beam attenuation coefficient c(555), thus strongly limiting phytoplankton growth (Chlorophyll a ~1 mg.m^{-3}). Spatially, the Van Diemen Gulf had higher total suspended solids and nutrient concentrations than Darwin Harbour, with dissolved organic carbon and a_{CDOM} subjected to photobleaching during the dry season. Key bio-optical relationships derived from this comprehensive

set of parameters, the first ever to be collected in this tropical coastal environment, were successfully used for a region-specific seasonal parameterization a region-specific seasonally parameterized ocean color algorithm. Challenges related to the parameterization, and the use, of ocean color remote sensing algorithms for these optically complex waters are discussed.

Keywords: coastal waters, tropical waters, Northern Australia, optical properties, water quality, seasonal variability

INTRODUCTION

Tropical coastal systems represent <1% of the world's ocean volume, yet they host some of the most productive and diverse ecosystems on earth (Jennerjahn, 2012; Bowen et al., 2013). Although, recent studies suggest that tropical marine regions will experience less drastic environmental changes than temperate regions in the future (Acevedo-Trejos et al., 2014), they remain largely understudied. This is particularly true in the southern hemisphere where very few datasets that capture the seasonal variations in bio-optical properties exist. The literature available on the tropical coastal waters of Far North Australia, referred to as the North Marine Region (NMR, **Figure 1**), is limited to marine phytoplankton productivity of this region (Hallegraeff and Jeffrey, 1984; Ilahude and Mardanis, 1990), its geology (Woodroffe et al., 1993) and the physical processes that influence its oceanography (Condie, 2011; Li et al., 2014). Across the NMR there is a need to develop better coastal ecosystem-based management strategies, which require quantitative assessment of the processes controlling the coastal marine environment. Combined bio-optical and biogeochemical measurements are key to quantify the ecosystem response to physical and chemical drivers. However, the remote location of the NMR makes the acquisition of *in situ* data by conventional sampling methods difficult (**Figure 1**), thus the use of ocean color and sea surface temperature (SST) satellite datasets for the study of the spatio-temporal patterns of biogeochemical processes in such remote areas is ideal. Recent studies using satellite-derived sub-surface light attenuation [$K_d(\lambda)$], Chlorophyll-a (Chl-a) and total suspended solids (TSS) concentrations have shown the distinct seasonal cycles of those parameters in the NMR (Schroeder et al., 2009; Blondeau-Patissier et al., 2011, 2014). Standard ocean color algorithms were used however, thus likely limiting the accuracy of the quantitative retrievals of the above-mentioned products due to the optical complexity of the North Australian shelf waters (IOCCG, 2000). Recommendations for future applications include the implementation of ocean color algorithms regionally tuned to the NMR using *in situ* optical measurements collected in this region for their parameterization. The success of such approach has been previously shown in other coastal regions (e.g., Brando et al., 2012; Tilstone et al., 2012; Roy et al., 2013). The optical properties and concentrations presented in this study are the first to have been collected in the region, and may be used as a baseline in any future bio-optical assessments guiding environmental management.

Located within the NMR and bounded to the east by the Gulf of Carpentaria, the Van Diemen Gulf (VDG) is a semi-enclosed bay (~16,000 km^2) with two narrow openings (~25–30 km wide), one to the North into the Arafura Sea and a second to the West into the Beagle Gulf (**Figure 1A**). The VDG, our first study region, is a dynamic marine environment that is characterized by shallow depths (<20 m) and strong tidal forcing. Six major river catchments surround the VDG: to the east are the Wildman River and the West, South (10,000 km^2), and East Alligator Rivers; to the west are the Mary (8,000 km^2) and Adelaide Rivers (638 km^2; **Figure 1A**). While the western catchments have been actively used for agricultural purposes, mainly cattle grazing, the eastern catchments are primarily for conservation (indigenous lands and national parks; CSIRO, 2009). Our second study region, Darwin Harbour (DH) (**Figure 1B**), is a semi-enclosed, shallow macro-tidal estuarine system (~3,000 km^2) connected to the VDG via the Clarence Strait (e.g., Williams et al., 2006; Andutta et al., 2014). Home to more than half of the Northern Territory's population (~200,000), Darwin Harbour is an isolated coastal region with the closest other cities being located >1,000 km to the east (Cairns) or west (Broome) (**Figure 1**). While there are no major ocean currents in the NMR, tidal currents play a significant role in water movement in this region (Condie, 2011). The monsoonal climate of the NMR region is characterized by a wet season that extends from November to April, during which more than three-quarters of the yearly rainfall occurs* (e.g., Story et al., 1969). Monsoonal rainfall (~1,700 $mm \cdot yr^{-1}$)[1] generates large quantities of freshwater that enter the coastal waters of the VDG via the surrounding catchments (**Figure 1**). Monsoonal winds are mostly northerly or north-westerly, with episodic cyclones (e.g., Acker et al., 2009; Lyon, 2010), while south-easterly trade winds predominate during the dry season (May–October). The rainfall, wind speed and direction vary dramatically between the wet and dry seasons, thus seasonal differences in water column mixing, turbidity, salinity, wave patterns, and wind-driven surface currents are to be expected. Overall a limited amount of light penetrates the water column of our study regions all year around, thus phytoplankton biomass and productivity are affected. In the neighboring Gulf of Carpentaria, Burford et al. (2012c) found that sediment resuspension from the tidal energy was the most important physical process limiting light penetration, and consequently phytoplankton primary production.

Surface inherent optical properties (IOPs) and biogeochemical concentrations, as well as nutrients for some stations, were sampled in the VDG and DH shelf waters during one wet

[1]BOM Bureau of Met., 73 year statistics (1941–1974) http://www.bom.gov.au, Climate statistics for Australian locations: Darwin Airport.

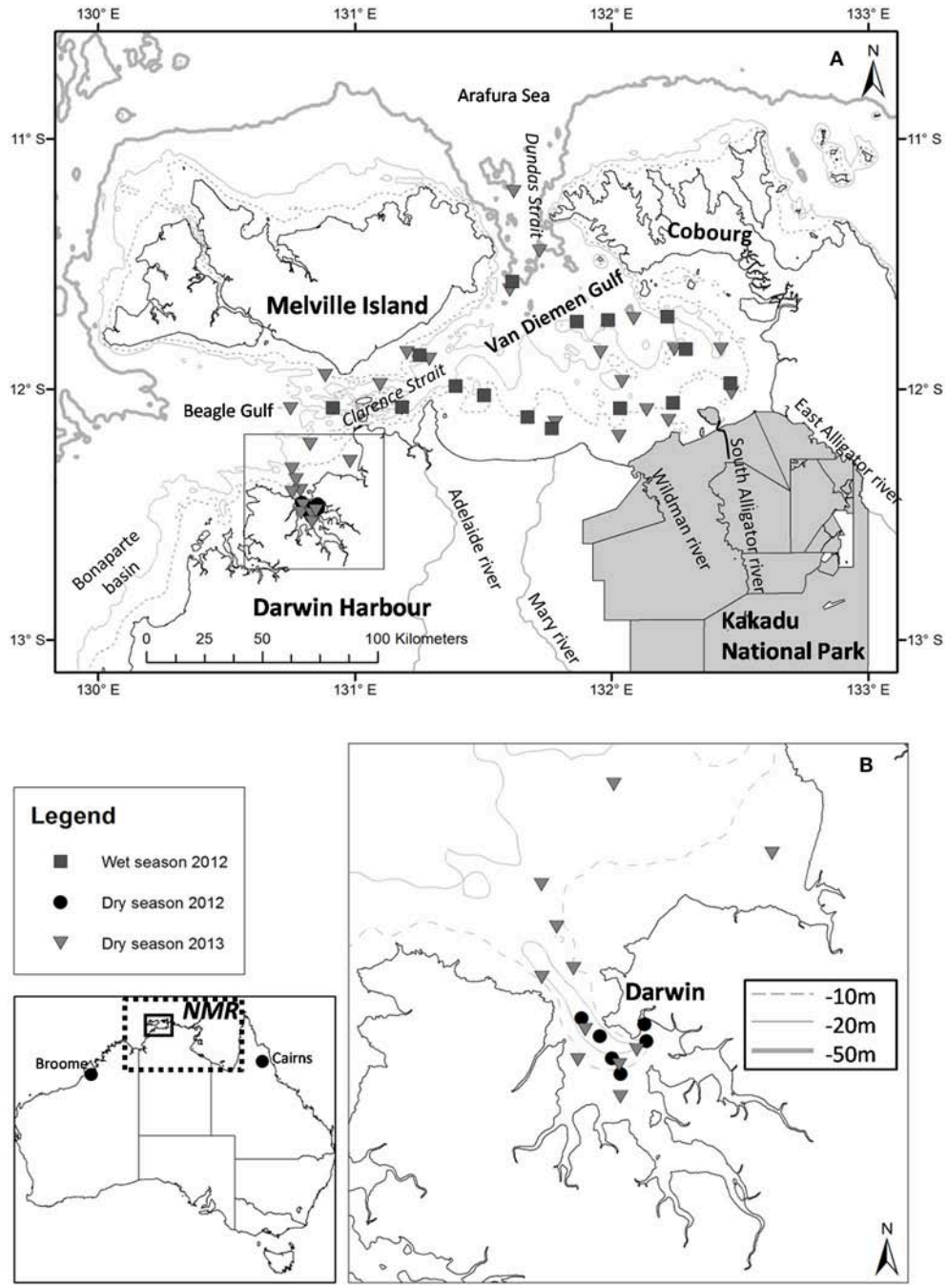


FIGURE 1 | (A) Locations of the dry and wet season sampling stations for the 2012 and 2013 field campaigns; **(B)** Darwin Harbour (inset). The extent of the Northern Marine Region (NMR; dotted line) is shown. Bathymetry contours are indicated in light gray [isobaths −10 and −50 m for **(A)**, and isobaths −10, −20, −50 m for **(B)**].

season (March 2012) and two dry seasons (September 2012, 2013; **Tables 1A,B**). The specific objectives of this study are (1) to characterize the spatial and seasonal bio-optical variability of DH and VDG to understand the functioning of these complex coastal ecosystems by identifying the major controlling processes and (2) to examine the relationships that can be derived between parameters to provide recommendations for their use in biogeochemical modeling and the parameterization of satellite ocean color algorithms specific to this dynamic coastal region of the NMR.

METHODS

Characteristics of the Sampling Effort

The dataset presented in this study comprises 23 metadata, biogeochemical and optical parameters (**Table 1**) sampled at

TABLE 1 | (A) Details of the seasonal field sampling and (B) the variables measured, (C) Notations: abbreviations, symbols and units of all the parameters referred to in this study.

A

Field sampling	Region	Dates	N	Season
1	Van Diemen Gulf	29–31 March 2012	16	Wet
2	Darwin Harbour	24, 26 September; 8 October 2012	15	Dry
3	Van Diemen Gulf	10–13 September 2013	22	Dry
3	Darwin Harbour	13–15 September 2013	10	Dry

B

Field sampling	IOP	Biogeochemical concentrations	POC	DOC	Phytoplankton counts, EEM, nutrients
1	X	X		X	
2	X	X			
3	X	X	X	X	X
3	X	X	X	X	X

C

Variable or symbol	Description	Units
Salinity	n/a	PSU
Secchi depth	n/a	m
Temperature	n/a	°C
Chl-a	Chlorophyll-a	$mg.m^{-3}$
DOC	Dissolved organic carbon	$mg.L^{-1}$
CDOM	Chromophoric dissolved organic matter	m^{-1}
NAP	Non-algal particulate	m^{-1}
POC	Particulate organic carbon	$mg.L^{-1}$
PSC	Photosynthetic carotenoids	$mg.m^{-3}$
PPC	Photoprotective carotenoids	$mg.m^{-3}$
TSS	Total suspended solids	$mg.L^{-1}$
Sr	EEM spectral slope ratio (276–295 nm/350–400 nm)	nm^{-1}
$a_{CDOM}(\lambda)$	Absorption coefficient of CDOM	m^{-1}
$a_{NAP}(\lambda)$	Absorption coefficient of non-algal particulate matter	m^{-1}
$a_p(\lambda)$	Absorption coefficient of particulate matter ($a_{phy}+a_{NAP}$)	m^{-1}
$a_{phy}(\lambda)$	Absorption coefficient of phytoplankton	m^{-1}
$a_{tot}(\lambda)$	Total absorption coefficient ($a_{phy}+a_{NAP}+a_{CDOM}$)	m^{-1}
S_{CDOM}, S_{NAP}	Spectral slope coefficient of CDOM or NAP	nm^{-1}
$a_{NAP^*}(\lambda)$	Specific absorption coefficient of NAP	$m^2.g^{-1}$
$a_{Phy^*}(\lambda)$	Specific absorption coefficient of Phytoplankton	$m^2.mg^{-1}$
$b(\lambda)$	Scattering coefficient	m^{-1}
$c(\lambda)$	Beam attenuation coefficient [$c(\lambda) = a(\lambda)+ b(\lambda)$]	m^{-1}
$b_b(\lambda)$	Backscattering coefficient [$b_b(\lambda) = b_{bp}(\lambda)+ b_{bw}(\lambda)$]	m^{-1}
$b_{\widetilde{bp(\lambda)}}$	Particulate backscattering to scattering ratio (b_{bp}/b_p)	unitless
γ	Slope of the backscattering coefficient	nm^{-1}
λ	Wavelength lambda	nm

63 stations over three field campaigns in the VDG and DH (**Figure 1**). Overall, the sampling effort corresponds to 13 days of field sampling—the distance from Darwin to the furthest sampling site in the VDG is 125 nautical miles, equivalent to >12 h of travel time for a vessel equipped for water sampling. Two research campaigns were undertaken in the VDG during (1) the wet season of 2012 (29–31 March 2012; $N = 16$) and (2) the dry season of 2013 (10–15 September 2013; $N = 22$; **Figure 1A**; **Table 1A**). Additionally, DH was also sampled on two occasions but only during the dry season: first, an intensive field campaign that solely focused on DH was undertaken during 3 days at the end of the 2012 dry season (24, 26 September and 8 October 2012; $N = 15$) and second, several stations ($N = 10$) were sampled in DH at the end of the September 2013 dry season field campaign when returning from the VDG (**Figure 1**; **Table 1A**). At the time of their sampling, the bottom depths of all stations ranged from 5 to 60 m (median: 21 m; $N = 63$). Water sampling was carried out 3–6 times a day to cover various tidal conditions.

The tidal regime in DH and VDG is semi-diurnal. While the tides in DH are macro-tidal with a mean spring tidal range of 6 m, the tidal range in DH can reach up to 8 m during spring tides, which is large compared to the mean depth (<20 m; **Figure 1B**). The tides are meso-tidal in the VDG, with a mean spring tidal range of 3 m and some tidal amplification up to 6 m in the south east of the Gulf, near the mouths of the Alligator Rivers (**Figure 1**). As a result the tidal phasing between DH and VDG can be between 1.5 and 2 h, with tides in the VDG occurring before those in DH. The tide tables and tide charts that are publicly available are from the Darwin City tide gauge only, and thus are not applicable to the entire study region. As our field sampling occurred across an area that covered both VDG and DH, the tidal properties at each station were computed for the sampling dates, times and locations using a two dimensional, depth-averaged, finite element, hydrodynamic Resources Modeling Associates (RMA) numerical model (Williams, 2009). The tidal range, the vertical difference between succeeding (or preceding) high and low tides depending on the time of sampling, was computed for each station.

The wet season of 2012 was characterized by above-average rainfall (1,661 mm > long $term_{1941-2014}$ mean: 1,282 mm), with March 2012 (570 mm) receiving 80% more rain than the long-term average monthly rainfall. Similarly, the September 2012 rainfall (21 mm) exceeded the long-term average by 35%, while September 2013 was exceptionally dry (0.2 mm) (source: BOM*).

Optical Measurements

At each station, vertical profiles of temperature, salinity and density (WET Labs Water Quality Monitor, WQM) as well as light absorption, beam attenuation (WET Labs ac-s with a 10-cm path-length), and light backscattering (WET Labs ECO BB-9) coefficients were measured (**Tables 1A,B**). The BB-9 backscattering meter measures the light backscattered at an angle of 124° with a fixed gain at nine wavelengths, which were set at 412, 440, 488, 510, 532, 595, 650, 676, and 715 nm. The ac-s measures the light absorption [$a(\lambda)$] and beam attenuation [$c(\lambda)$] at multiple wavelengths ($N = 82$) from 401.6 to 739.7 nm. All instruments were deployed together in an optics cage to allow for simultaneous collection of the measurements. Corrections for temperature and salinity effects on water optical properties were

applied to the ac-s data using the WQM measurements (Pegau et al., 1997). The incomplete recovery of the scattered light in the ac-s absorption tube was rectified using the proportional method described in Zaneveld et al. (1994). The BB-9 backscattering dataset was corrected for light loss due to absorption over the path length at each angle and wavelength using light absorption and scattering values from the ac-s (Boss et al., 2004). All vertical profiles were binned at 0.2 m depth intervals. In addition, the water transparency was visually estimated by the same observer at all stations by lowering a black and white disk (Secchi disk). The Secchi depth reported in this study is the depth at which the disk is no longer visible.

Assuming that dissolved organic matter in water has a negligible effect on scattering, the particulate scattering coefficient $b_p(\lambda)$ was derived by subtracting the scattering by pure water, $b_w(\lambda)$ from the difference between the total light attenuation $c(\lambda)$ and total light absorption (particulate+dissolved), $a_{Tot}(\lambda)$ [i.e., $b_p(\lambda) = (c(\lambda) - a_{Tot}(\lambda)) - b_w(\lambda)$]. For the scattering coefficient, the particulate backscattering coefficient, $b_{bp}(\lambda)$, was obtained by subtracting the backscattering component of pure water, $b_{bw}(\lambda)$, from $b_b(\lambda)$. The spectral slope of the backscattering coefficient, γ, was computed as a power-law (Whitmire et al., 2007).

For the determination of the particulate absorption coefficient, the sum of the phytoplankton and the non-algal particle (NAP) components [$a_P(\lambda) = a_{phy}(\lambda)+a_{NAP}(\lambda)$], a volume of 0.4–1 L of surface seawater was filtered through a 25 mm Whatman GF/F glass-fiber filters. Filters were stored flat in liquid nitrogen until analysis. Optical densities were measured over the 250–800 nm spectral range with 0.9 nm increments, using a Cintra 404 UV/VIS dual beam spectrophotometer equipped with an integrating sphere. Pigmented material was extracted from the sample filter using the method of Kishino et al. (1985) to determine the optical density of the non-algal matter, $OD_{NAP}(\lambda)$. The optical density due to phytoplankton was obtained by difference [$OD_{phy}(\lambda) = OD_P(\lambda)$-$OD_{NAP}(\lambda)$]. The pathlength amplification effect due to the filter was corrected by using the algorithm of Mitchell (1990). Finally, absorption coefficients of phytoplankton [$a_{phy}(\lambda)$], non-algal particles [$a_{NAP}(\lambda)$], and its slope S_{NAP} were computed as described in Clementson et al. (2004).

All phytoplankton absorption spectra were plotted and quality controlled to avoid the inclusion of those contaminated by non-algal material.

Biogeochemical Measurements

For all measurements, water samples were collected at the surface ($\leq$2 m), using either Niskin bottles (via a Rosette system) or by lowering a clean, polyethylene bucket from the side of the ship.

CDOM, POC, and DOC Concentrations

Water samples were filtered through a Whatman Anodisc membrane (0.22 μm; 47 mm) for CDOM and dissolved organic carbon (DOC) analysis. Filters were pre-rinsed with Milli-Q water prior to filtration. To track possible contamination of the glass filtration unit by the filters, the initial filtrate (100 ml) of Milli-Q water was discarded, and the subsequent filtrate of Milli-Q water was stored as a blank for the first and last samples collected during the campaign. The final filtrate was transferred to a SCHOTT glass bottle pre-rinsed with the same filtered seawater; the same was done for the DOC samples which were then preserved with 0.5 ml of 50% H_3PO_4. The samples were stored at 4°C. The analysis of DOC and Particulate Organic Carbon (POC), derived from subtracting DOC (water filtered through 0.22 μm) from the unfiltered water samples, is further described in the method section of MacIejewska and Pempkowiak (2014b). DOC and POC were not measured for 15 stations sampled in Darwin Harbour in the 2012 dry season, and POC ($N = 29$) was only collected during the 2013 dry season field campaign. The CDOM absorbance of each filtrate, after equilibrating to room temperature, was measured from 250 to 800 nm in a 10 cm pathlength quartz cell using a Cintra 404 UV/VIS spectrophotometer, with fresh Milli-Q water (Millipore) as a reference. The CDOM absorption coefficient (m^{-1}) was calculated using the equation a CDOM = 2.3(OD(λ)/l) where l is the cell path length in meters. Finally, an exponential function was fitted to the CDOM spectra over the wavelength range 350–680 nm from which the slope, S_{CDOM}, was derived as described in Clementson et al. (2004).

During the dry season sampling of September 2013, a second set of CDOM samples ($N = 26$) was analyzed by fluorescence Excitation-Emission Matrix spectroscopy (EEM). EEM is often used to trace photochemical and microbial reactions associated with fluorescent dissolved organic matter (FDOM). In this study, it was used to assess CDOM sources (terrestrial or marine). Water samples were filtered through 0.45-μm syringe filters into SCHOTT bottles pre-washed with HCl. The samples were stored at 4°C while at sea. They were then stored at −20°C once in the laboratory and were allowed to reach room temperature before analysis. Each sample was analyzed using a Horiba Jobin Yvon Aqualog Excitation-Emission spectro-fluorometer (240–600 nm) and a Starna 1-cm quartz cell. The EEM spectra were recorded under standard instrumental conditions and were subsequently corrected for internal absorbance effects. Both the first and second order Raman and Rayleigh lines were removed, and the intensity was expressed in Quinine Sulfate Units (QSUs) using the appropriate instrument normalization for the integration time (Watson and Zielinski, 2013). The spectral slope ratios Sr, were computed separately from the same data by least squares fitting of the log transformed raw absorbance in the ranges 275–296 and 350–400 nm (Helms et al., 2008).

Total Suspended Solids

For the determination of TSS concentrations, water samples (0.50–3.50 L) were filtered through pre-weighed Whatman GF/F glass microfiber filters (0.7 μm; 47 mm) pre-combusted at 450°C. A blank filter with filtered seawater was used as a reference. The filter was then rinsed with ∼50 ml of distilled water to remove any salt from the filter and dried to constant weight at 65°C to determine the TSS (see Figure 1 of Neukermans et al., 2012). The filters were then placed in a furnace at 450°C for 3 h, allowed to cool and weighed to determine the amount of inorganic material remaining on the filter. To quantify sample variability between TSS samples, triplicates were taken at each station. Overall, the

relative standard deviation between all triplicates was 8% (N = 57), varying the least for the DH dry season samples (2%) and the most for the VDG dry and wet season samples (13%). Laboratory-estimated particulate inorganic (PIM) and organic (POM) matter fractions of TSS were only available for the 2012 wet season samples and were extracted as described in Oubelkheir et al. (2014).

Determination of Phytoplankton Pigment Concentrations and Cell Counts

For phytoplankton pigment concentration and composition, water volumes of 0.10–1.20 L were filtered through Whatman GF/F glass microfiber filters (0.7 μm; 25 mm) and stored in liquid nitrogen until analysis. Phytoplankton pigments, extracted by High-performance Liquid Chromatography (HPLC) following Clementson (2013), were grouped into five categories: (1) Chl-a (Chlorophyll-a, divinyl-Chl-a, chlorophyllide-a and pheopigments); (2) Chl-b (Chl-b and divinyl Chl-b); (3) Chl-c (Chl-c1, -c2); (4) photosynthetic carotenoids (PSC; fucoxanthin, peridinin, 19–HF and 19–BF); and (5) photoprotective carotenoids (carotenoids (PPC, zeaxanthin, diadinoxanthin, alloxanthin, lutein, α- and β-carotene). The relative contribution of picophytoplankton (<2 μm), nanophytoplankton (2–20 μm), and microphytoplankton (>20 μm) was estimated for each sample using diagnostic pigments as described in Uitz et al. (2006), improved from Vidussi et al. (2001).

Surface water samples for microscopic phytoplankton cell counts were collected in 100-ml darkened bottles and preserved with paraformaldehyde for a final concentration of ∼1% paraformaldehyde (2013 dry season campaign only). Using the Utermöhl technique, sub-samples were settled in 100-ml settling chambers for 24 h and examined using an inverted microscope (Nikon Eclipse Ti-S) following the methodology outlined by Hasle (1978). Large and numerically rare taxa were counted during full examination of the settling chamber (×100), while small and numerically dominant taxa were counted on 1–2 transects of the chamber (×400), or from cumulative counts of 5–10 fields of view. Diatoms and dinoflagellates were identified to the genera or species level, based on Hallegraeff et al. (2010). For picophytoplankton cell counts, samples were collected in 2.5-ml cryovials, preserved in 1% paraformaldehyde, kept in liquid nitrogen while at sea and stored at −80°C until analysis. An Accuri C6 Flow Cytometer was used, following the procedure described in Zubkov et al. (2007). Cells' viability was assessed through microscope counts, where cells that were broken or lysed were considered to be non-viable cells. When cell numbers were compared (i.e., between the phytoplankton community >2 μm and picoplankton <2 μm), phytoplankton cell counts amounted to <0.001% of the total (picoplankton + phytoplankton).

Specific Inherent Optical Properties

Mass-specific inherent properties (SIOPs), the absorption and backscattering coefficients normalized to their respective concentrations, are necessary for the parameterization of (semi-) analytical bio-optical models for complex coastal waters (e.g., Brando et al., 2012; Tilstone et al., 2012; Le et al., 2015). However, SIOP datasets are seldom available for most of the world's coastal ocean. To further characterize our optical dataset and to help the development of ocean color algorithms for this sub-region of the NMR, we computed the specific absorption coefficients for phytoplankton, $a_{phy*}(\lambda)$, obtained by normalizing $a_{ph}(\lambda)$ to Chl-a, as well as the mass-specific non-algal particulate absorption, $a_{NAP*}(\lambda)$, and backscattering, $b_{bp*}(\lambda)$, coefficients by normalizing a_{NAP} and b_{bp} by their TSS concentrations. Further, we acknowledge that the algal contribution to TSS may not have been negligible, during the wet season in particular.

Nutrient Concentrations

Water samples for nutrients were collected at surface (0.5 m; 2013 dry season campaign only) and filtered through pre-combusted (500°C) 25 mm Whatman GF/F glass fiber filters. The samples were frozen and stored at −20°C prior to analysis. Nutrient species nitrite+nitrate ($NO_2^- + NO_3^-$), ammonium (NH_4^+) and phosphate (PO_4^{3-}) were analyzed using a segmented flow analysis system following Ryle et al. (1981).

Statistics

The effect of seasons (i.e., dry vs. wet season) or locations (DH vs. VDG) on the measured variables were tested using a one-way analysis of variance (ANOVA). The normality of the distributions was verified using the Shapiro–Wilk test statistics at $p > 0.05$ and the homogeneity of variance was tested with the Bartlett's test prior to analysis. If none of the transformations used led to normally distributed data, the non-parametric Mann–Whitney–Wilcoxon test was applied.

RESULTS

Metadata: Salinity, Secchi Depths, and Tides

The range of a selected set of metadata, concentrations and IOPs for the dry and wet seasons is shown in **Table 2**. Most (13) of the variables showed significant seasonal variations due to the monsoon (**Tables 2, 3**). Surface waters were found to be significantly cooler (median dry/wet: 28.35/29.83°C; $p < 0.001$; $N = 63$) and more saline (median dry/wet: 34.64/29.43 PSU; $p < 0.005$; $N = 63$) during the dry season (**Table 2**). Surface salinity varied by 33% during the wet season ($N = 16$), with stations featuring differences of up to 4 PSU between the surface and the bottom of the cast (**Table 2**). In comparison, dry season stations showed little variation (median: 0.1 PSU; $N = 47$).

Significant differences in seasonal and regional Secchi depths were found (**Table 3**). Most of the Secchi depths sampled ranged from 1.0 to 3.0 m (67%), with 45% of the stations having a Secchi depth <2 m. The clearest waters were sampled in DH (Secchi depth >3 m), in the straits (>4 m), and in the Beagle Gulf (4 m), while most ($N = 17$) of the VDG dry stations had Secchi depths <2 m (**Figure 1**). Amongst the metadata, the Secchi depth was the parameter most correlated, albeit mildly, with the tidal range ($R^2 = 0.23$, $N = 56$) and temperature ($R^2 = 0.22$, $N = 50$), but not with salinity ($R^2 < 0.1$, $N = 58$). Secchi depths were strongly correlated with a_{NAP} ($R^2 = 0.73$, $N = 56$) and a_{Tot} ($R^2 = 0.73$, $N = 46$), while there was a mild co-variation with TSS (R^2 =

TABLE 2 | Range of surface (<2 m) values in metadata, concentrations and IOPs in the Van Diemen Gulf and Darwin Harbour during the 2012–2013 dry and wet seasons field samplings.

Variable	*N*	Min–Max	Mean ± SD	*N*	Min–Max	Mean ± SD
		Wet season			Dry season	
Temperature	16	27.31–30.62	29.71 ± 0.95	47	26.37–30.01	28.21 ± 1.10
Salinity	16	24.72–32.75	29.32 ± 2.55	47	34.12–35.05	34.62 ± 0.23
Secchi depth	16	0.31–5.53	2.78 ± 1.30	47	0.25–4.05	2.26 ± 1.00
TSS	13	1.73–27.1	5.99 ± 6.64	44	2.59–143.1	30.41 ± 29.5
Chl-a	11	0.33–2.33	0.98 ± 0.57	45	0.33–2.78	0.96 ± 0.44
Size index	11	11.23–36.21	24.9 ± 8.50	45	13.81–46.12	27.53 ± 9.24
DOC	13	0.92–1.85	1.39 ± 0.29	29	0.87–2.35	1.15 ± 0.28
POC	n/a	–	–	29	1.18–2.23	1.38 ± 0.21
Sr	n/a	–	–	29	1.59–3.37	2.08 ± 0.34
$a_{CDOM}(440)$	10	0.07–0.27	0.17 ± 0.06	28	0.05–0.37	0.13 ± 0.07
S_{CDOM}	10	0.010–0.017	0.014 ± 0.002	28	0.008–0.017	0.013 ± 0.002
$a_{phy}(440)$	10	0.02–0.08	0.05 ± 0.01	42	0.02–0.10	0.05 ± 0.01
$a_{phy440/676}$	10	1.86–2.77	2.28 ± 0.28	42	1.45–2.96	2.06 ± 0.32
$a_{NAP}(440)$	11	0.02–0.79	0.14 ± 0.22	45	0.04–5.42	0.36 ± 0.88
S_{NAP}	11	0.01–0.01	0.01 ± 0.00	45	0.01–0.01	0.01 ± 0.00
$a_{NAP}/a_p(440)$	9	0.38–0.73	0.56 ± 0.13	41	0.42–0.89	0.69 ± 0.12
$a_{tot}(440)$	11	0.16–2.32	0.51 ± 0.61	35	0.13–6.86	0.67 ± 1.22
$b_{p\,555}$	15	0.46–4.14	1.55 ± 0.87	33	0.81–4.97	2.14 ± 1.30
c(555)	15	0.50–4.24	1.61 ± 0.88	33	1.05–5.07	2.38 ± 1.36
$b_{bp}(555)$	15	0.01–0.10	0.03 ± 0.02	25	0.02–0.73	0.14 ± 0.21
γ	15	0.44–0.93	0.68 ± 0.17	25	0.23–1.90	0.82 ± 0.58
$b_{bp}{}^{*}(555)$	11	0.001–0.02	0.012 ± 0.005	27	0.00–0.03	0.026 ± 0.06
$a_{NAP}{}^{*}(440)$	11	0.02–0.03	0.025 ± 0.006	45	0.002–0.03	0.010 ± 0.01
$\widetilde{b}_{bp\,555}$	14	0.01–0.04	0.02 ± 0.00	25	0.02–0.06	0.03 ± 0.00

The number of samples, N, varies due to quality control. Dry season includes samples from both the Van Diemen Gulf and Darwin Harbour, Wet season includes samples from the Van Diemen Gulf only. SD; Standard Deviation.

0.34, $N = 57$). Temperature was moderately correlated with TSS ($R^2 = 0.47$, $N = 47$). Tidal range was moderately correlated with a_{NAP} slopes ($R^2 = 0.45$, $N = 50$), and weakly with any of the concentrations.

Biogeochemical Concentrations and Inherent Optical Properties

The overall average Chl-a concentrations sampled during these three campaigns was 1.0 mg.m^{-3} ($N = 56$). The highest Chl-a concentrations measured were 2.8 mg.m^{-3} during the dry season and 2.3 mg.m^{-3} during the wet season, both sampled at the same VDG station located <15 km from the mouth of the Mary River (**Figure 1A**). Although, no significant difference was found for Chl-a between seasons or locations (**Table 3**), higher Chl-a concentrations (>1.3 mg.m^{-3}) were found along the VDG coast and within DH (with a gradual decrease from the inner ∼1.63 mg.m^{-3} to the outer ∼<1 mg.m^{-3} of the Harbour), while lower Chl-a concentrations were found in the middle of the VDG (0.6–1.25 mg.m^{-3}) and at stations sampled in the Beagle Gulf and Dundas Strait (<0.6 mg.m^{-3}; **Figures 1**, **2A**; **Table 2**). TSS was found to be significantly higher for stations sampled during the dry season and located in the VDG (54.2 ± 40.0 mg.L^{-1}; $N = 20$) in comparison to stations sampled in DH (15.6 ± 9.7 mg.L^{-1}; $N = 24$) or during the wet season (6.0 ± 6.6 mg.L^{-1}; $N = 13$; **Tables 2, 3**). Higher TSS concentrations were typically found along the coast, while lower TSS concentrations were found in the inner Harbour (**Figure 2B**). There was no interdependence between TSS and Chl-a ($R^2 < 0.1$; $N = 55$; **Figure 3A**; **Table 5**), while TSS concentrations seem to increase with increasing tidal range, albeit with no significant relationship ($R^2 < 0.1$; $p > 0.1$, $N = 51$; **Figure 3B**). The wet season TSS samples (VDG wet) were found to be mostly (75%) composed of inorganic material ($N = 13$).

Organic carbon is divided into a particulate (POC) and a dissolved (DOC) fraction, both playing major roles in the ocean carbon cycle (Bauer et al., 2013) with the majority (∼95%) of the ocean organic carbon being composed of DOC (e.g., Hansell and Carlson, 2001). The concentration in organic carbon varies according to the distance from land, with open ocean waters having less organic carbon than those in coastal regions. For instance, POC was found to be in the order ∼0.1 mg.L^{-1} in Pacific Ocean waters (Claustre et al., 1999; Fabiano et al., 1999), 0.3 mg.L^{-1} in shallow shelf waters of the Northwest Atlantic (Bauer et al., 2002), 1.4 mg.L^{-1} in the Baltic Sea coastal waters (MacIejewska and Pempkowiak, 2014b), and up to 1.8 mg.L^{-1} in Chesapeake Bay (Fisher et al., 1998). For this study, POC

TABLE 3 | One-way ANOVA results for selected concentrations, IOPs and metadata. The factor "Season" refers to dry vs. wet seasons (Van Diemen Gulf stations only), and "Location" refers to Van Diemen Gulf vs. Darwin Harbour (during the dry season only).

Variable	Season		Location	
	df;F	*p*	*df;F*	*p*
Chl-a	1; 0.05	0.81 (NS)	1;0.94	0.33 (NS)
TSS	1; 65.18	**0.001*****	1;25.96	**0.001*****
DOC	1;18.58	**0.001*****	1;12.63	**0.002*****
POC		–		0.92 (NS)
$a_{CDOM}(440)$	1; 2.39	0.134 (NS)	1;0.33	0.58 (NS)
$a_{phy}(440)$	1; 0.22	0.644 (NS)	1;0.33	0.57 (NS)
$a_{phy}(440/676)$	1; 2.24	0.148 (NS)	1;0.03	0.88 (NS)
$a_{phy^*}(440)$	1; 1.29	0.266 (NS)	1;8.32	**0.006*****
$a_{Tot}(440)$	1; 0.98	0.328 (NS)	1;6.81	**0.021****
$a_p(440)$	1; 12.10	**0.002*****	1;9.10	**0.005*****
$a_{NAP}/a_p(440)$	1; 0.111	**0.012****	1;16.0	**0.001*****
$b_p(555)$	1;12.83	**0.002*****	1;14.43	**0.001*****
$b_{bp}(555)$	1; 25.28	**0.001*****	1;37.58	**0.001*****
Υ		0.475 (NS)		∼1 (NS)
$b_{bp}(555)$	1; 8.30	**0.008*****	1; 4.55	**0.045****
$b_{bp}{}^*(555)$		**0.006*****		0.08 (NS)
$a_{NAP}{}^*(440)$		**0.001*****		0.15 (NS)
S_{NAP}		1.00 (NS)		**0.001*****
S_{cdom}		0.18 (NS)		0.22 (NS)
c(555)		**0.003*****		**0.008*****
Fucoxanthin	1;1.89	0.18 (NS)	1;7.71	**0.011****
Salinity		**0.001*****	1;33.81	**0.001*****
Temperature		**0.001*****		**0.001*****
Secchi		**0.004****		**0.001*****

*If only the p-value is indicated, a non-parametric test was applied on the dataset. p*** < 0.01; ** < 0.05; NS (i.e., not significant; p > 0.05).*

(only available for the dry season samples) was found to be as high as 2.23 $mg.L^{-1}$, with a mean concentration of 1.42 ± 0.26 $mg.L^{-1}$ ($N = 19$) in the VDG and 1.33 ± 0.06 $mg.L^{-1}$ ($N = 10$) in DH (**Table 2**). No significant difference in POC between the two locations was found (**Table 3**). In open ocean waters, Morel (1988) proposed the empirical relationship POC = 90 $Chl\text{-}a^{0.57}$ ($R^2 = 0.68$; $N = 409$), but in coastal waters, such robust relationships between POC and Chl-a may not be observed because of the higher non-phytoplankton contribution (e.g., Sathyendranath et al., 2009). For our dataset, after converting POC from $mg.L^{-1}$ to $mg.m^{-3}$, we obtained POC = 1,444 $Chl\text{-}a^{0.19}$ ($R^2 = 0.40$; $N = 29$; $p < 0.005$; **Figure 3C**; **Tables 2**, **5**). TSS samples were also examined in relation to their POC content to assess their relative fraction of organic material but this was not possible due to a poor covariation of POC with TSS ($R^2 = 0.14$; $p < 0.05$; $N = 29$) (**Figure 3D**).

Understanding the dynamics of DOC in coastal systems is key to accurately assessing the important role coastal regions play in the global carbon cycle (Fichot and Benner, 2014; Fichot et al., 2014). While POC is the sum of the masses of all organic particles, largely composed of phytoplankton and organic detritus (e.g., fecal pellets; Romero-Ibarra and Silverberg, 2011), DOC is directly related to micro-organism activities ranging from photosynthesis to virus lysis (Agustí and Duarte, 2013). The use of CDOM as a direct and reliable proxy for DOC in coastal waters has been demonstrated in previous studies. For instance, Fichot and Benner (2011) used multiple linear regressions to successfully ($R^2 > 80\%$) retrieve DOC from $a_{CDOM}(275)$ and S_{CDOM} in the Beaufort Sea and the Northern Gulf of Mexico. More recently, Vantrepotte et al. (2015) retrieved DOC from $a_{CDOM}(412)$ using linear regressions and a dataset of contrasting coastal waters, comprising samples from the English Channel ($R^2 = 0.72$), Vietnam ($R^2 = 0.81$), and French Guiana ($R^2 = 0.78$). For our dataset, there was no correlation between DOC and $a_{CDOM}(400)$ ($R^2 \sim 0.1$; $p > 0.1$; $N = 55$; **Figure 4C**). DOC was found to be significantly different between both seasons and locations (**Tables 2, 3**). It was higher during the wet season (1.39 ± 0.29 $mg.L^{-1}$; $N = 13$; **Table 2**) and varied spatially from lower concentrations in the embayment of the VDG (∼0.8–1.0 $mg.L^{-1}$) increasing toward Clarence Strait (>1.0 $mg.L^{-1}$) and reaching up to 2.35 $mg.L^{-1}$ in DH (**Figure 1**). Overall DOC (1.34 ± 0.37 $mg.L^{-1}$; $N = 10$) in DH was found to be much lower than that observed by Burford et al. (2012a) in Buffalo Creek (2.9 ± 1.8 $mg.L^{-1}$), 20 Km North of Darwin. This difference can be partly explained by the seasonal effect: samples from DH were from the dry season, while Burford et al.'s Buffalo Creek DOC values were sampled during the wet season (December 2008; Burford et al., 2012a). No correlation was found between DOC and salinity ($R^2 = 0.10$; $N = 38$) (**Figure 4D**).

The absorption of CDOM at 440 nm was not found to be statistically significantly different between seasons (**Table 3**). Yet $a_{CDOM}(440)$ was characterized by higher values during the wet season (median dry/wet: 0.11/0.18 m^{-1}; **Table 2**), and stations located in the VDG had slightly lower a_{CDOM} (440) values (0.11 ± 0.09 m^{-1}; $N = 21$) than those in DH (0.14 ± 0.03 m^{-1}; $N = 10$; **Figure 2C**). Overall, $a_{CDOM}(440)$ was found to weakly increase with Chl-a ($R^2 = 0.20$, $N = 37$; **Figure 4A**). Also a proxy for salinity, a_{CDOM} can be used as a marker of freshwater influence during flooding conditions (e.g., Schroeder et al., 2012), but for our dataset, there was a very weak relationship between salinity and a_{CDOM} (**Figure 4B**): albeit with a poor relationship, CDOM was found to increase with salinity during the dry season ($R^2 = 0.14$; $N = 26$; **Figure 4B**), while it decreased with salinity during the wet season ($R^2 = 0.34$; $N = 11$) possibly due to its terrestrial source (thus in higher concentration near the rivers). Its spectral slope, S_{CDOM}, was not seasonally or spatially uniform (0.008–0.017; $N = 38$) but statistically, no seasonal or spatial difference was found (**Tables 2**, **3**; **Figure 5A**). S_{CDOM} was within the set range described by Blough and Del Vecchio (2002). The spectral slope of NAP did not change between seasons but unlike S_{CDOM}, S_{NAP} was found to differ spatially ($p < 0.001$; $N = 45$) with the dry season samples located in the VDG (median: 0.014 m^{-1}; $N = 20$) having steeper slopes than those from DH (0.012 m^{-1}; $N = 25$) (**Figure 5B**). From the EEM, we found Sr to be between 1.5 in DH and 3.5 in the VDG (**Figure 13A**). The slope ratios are in accordance with coastal environments, as sampled in our study, where Sr ranged from wetlands (0.69) to oceanic (9.02) (Helms et al., 2008). In addition, $a_{CDOM}(350)$ was found to exponentially decrease with $S_{275-295}$ (**Figure 13B**), as per

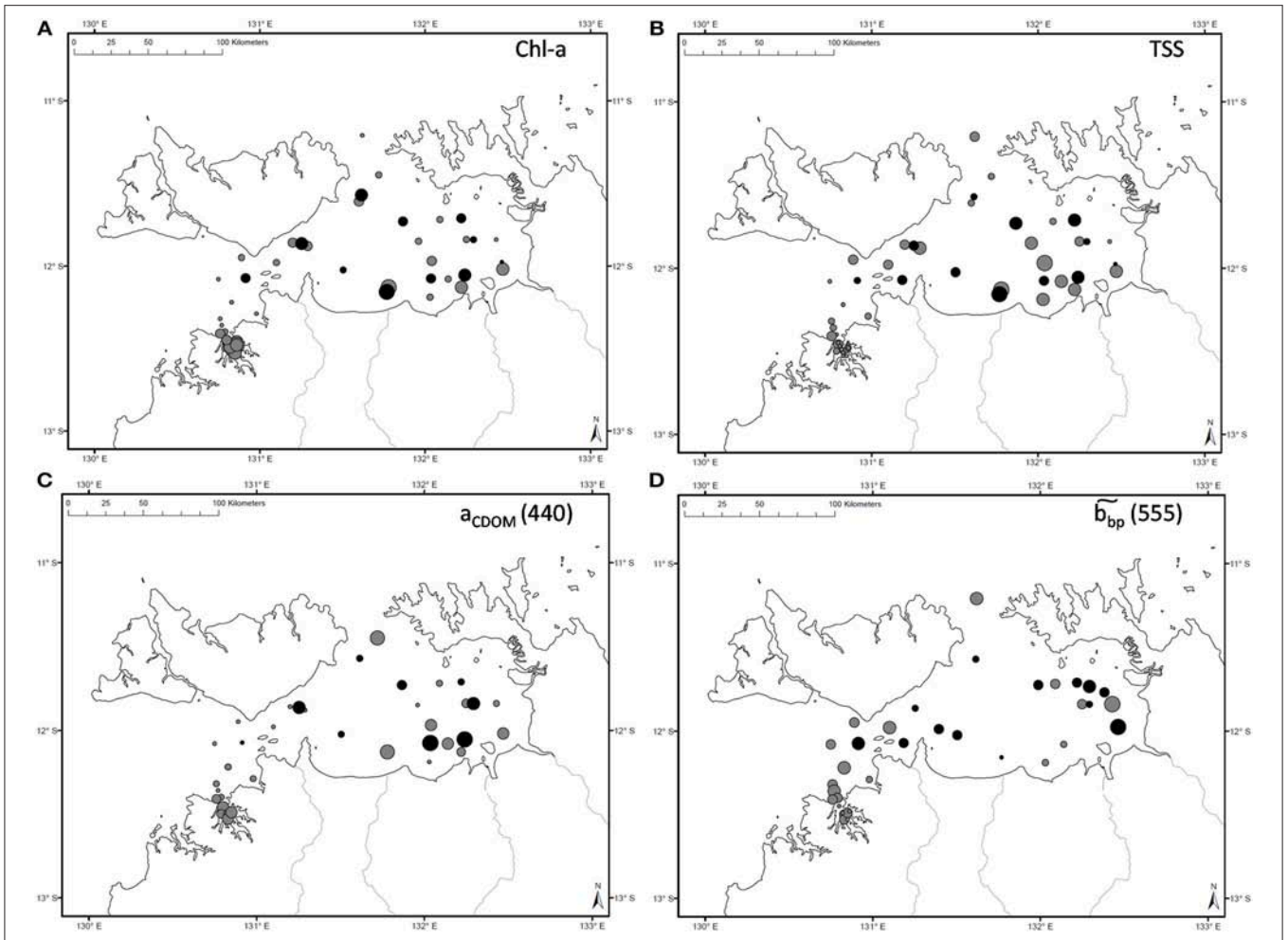


FIGURE 2 | Spatial distribution of measured (A) Chl-a, **(B)** TSS, **(C)** $a_{CDOM}(440)$, and **(D)** $\widetilde{b_{bp\,555}}$ during the dry (gray) and wet (black) seasons. Symbol sizes are representative of the concentration range sampled during a specific season.

Figure 6 of Fichot and Benner (2012), thus indicating possible effects of photobleaching for the VDG-dry season samples.

The magnitude of the particulate backscattering coefficient varied broadly [$0.01 < b_{bp}(555) < 0.73\ m^{-1}$; $N = 40$; **Table 2**]. Its associated spectral slopes ($0.2 < \gamma < 1.9\ m^{-1}$) were well within the range reported by previous studies (Snyder et al., 2008; **Figure 5C**) but $b_{bp}(555)$ slopes did not feature any seasonal or spatial variations ($\gamma \sim -0.6\ m^{-1}$; **Table 3**). While $b_{bp}(555)$ was evidently strongly correlated with $b_p(555)$ ($R^2 = 0.84$; $N = 37$; $p < 0.001$; **Table 5**), there was also a significant correlation between $b_{bp}(555)$ and TSS ($R^2 = 0.53$; $p < 0.001$; $N = 36$; **Figures 7A,B**; **Table 5**). Overall, the mass-specific non-algal particulate coefficient at 440 nm, $a_{NAP*}(440)$, varied between 0.002 and 0.035 $m^2.g^{-1}$ with a mean value of 0.025 $m^2.g^{-1}$ during the wet season decreasing to 0.010 $m^2.g^{-1}$ during the dry season. The wet season average is comparable to other coastal waters around the world such as that found in French Guiana (i.e., 0.023 $m^2.g^{-1}$; Loisel et al., 2009) but it is much lower than that of the North Sea (0.033 $m^2.g^{-1}$; Babin et al., 2003a) for instance; the extremely low dry season average confirms that the dry season samples are largely dominated by inorganic particles, which is in accordance with our hypotheses of a NAP-dominated system at that time of the year. A significantly strong correlation between aNAP*(555) and bbp*(555) and b_{bp}*(555) ($R^2 = 0.77$, $p < 0.001$, $N = 36$) is shown in **Figure 7C**. Such strong relationship between the two parameters was previously reported in the Great Barrier Reef, albeit at a different wavelength [aNAP*(440) & $R^2 = 0.82$; see Figure 13 of Blondeau-Patissier et al. (2009)], and in the Gulf of Mexico (D'Sa et al., 2007).

Temporal and spatial differences were found in the beam attenuation coefficient, c_{555}, which was significantly higher during the dry season ($p < 0.05$; $N = 33$), particularly in the Gulf (VDG_{dry}: $3.33 \pm 1.42\ m^{-1}$; $p = 0.01$; $N = 11$), compared with the wet season (VDG_{wet}:$1.61 \pm 0.89\ m^{-1}$; $N = 15$) and DH ($1.91 \pm 1.08\ m^{-1}$; $N = 22$; **Figure 6**; **Table 2**). In the VDG, independent of the season, the total absorption at 555 nm, $a_{Tot}(555)$, was a negligible portion of c_{555} ($4 \pm 1\%$; $N = 26$), while in DH it played a more important role ($15 \pm 8\%$; $N = 22$). The beam

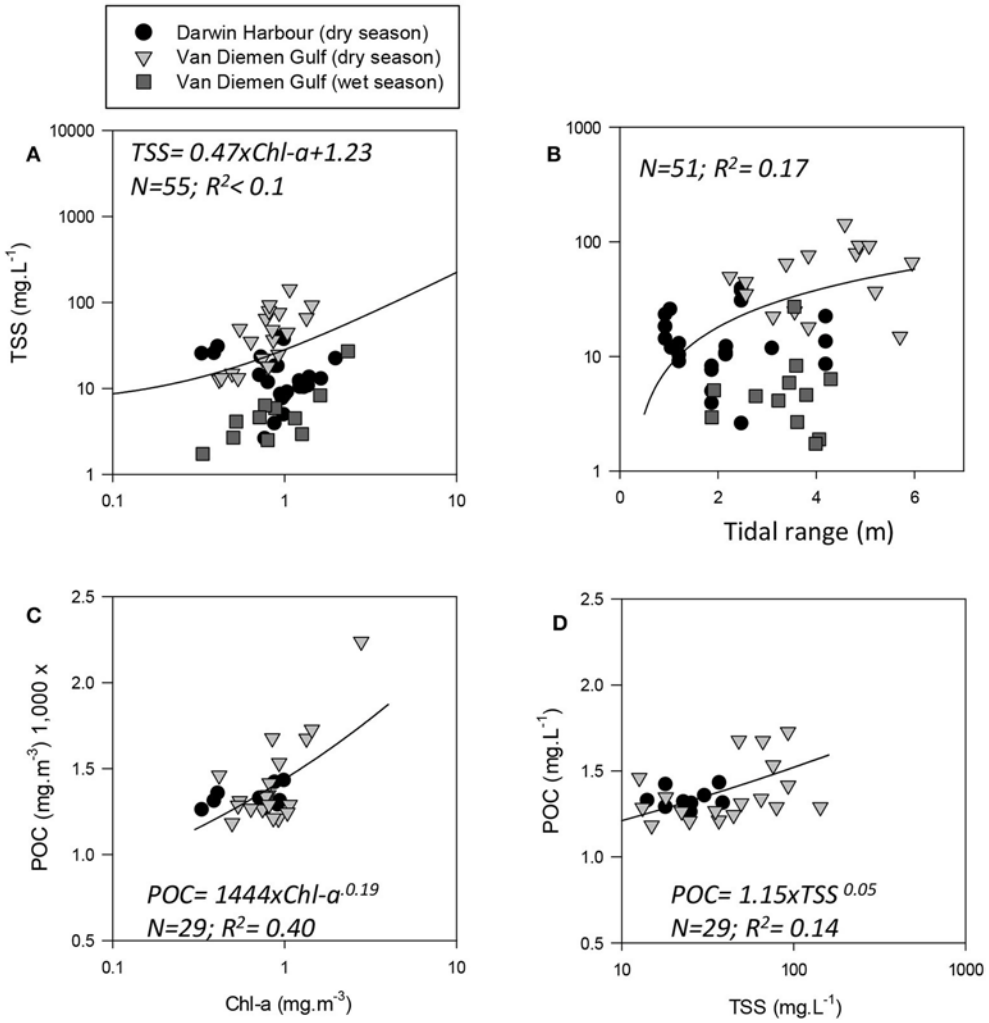


FIGURE 3 | Relationships between: (A) TSS and Chl-a, **(B)** TSS and tidal range (note: × is linear), and POC as a function of **(C)** Chl-a and **(D)** TSS.

attenuation also displayed a strong relationship with $b_{bp}(555)$ ($R^2 = 0.80$; $p < 0.001$; $N = 36$; **Figure 6A**). While Secchi depths were found to be relatively correlated with c(555) ($R^2 = 0.43$; $N = 47$; $p < 0.005$; **Figure 6C**), c(555) was not correlated with the particulate backscattering to scattering ratio $b_{bp}^{\sim}(555)$ ($N = 26$; $R^2 = 0.20$; $p > 0.1$; **Figure 6B**). This latter parameter, $b_{bp}^{\sim}$, has been linked to the composition of the particle assemblage (e.g., Loisel et al., 2007), from the particle size distribution to its refraction index (e.g., Twardowski et al., 2001; Boss et al., 2004). For our dataset, $b_{bp}^{\sim}(555)$ surface values varied three-fold (0.02–0.07; $N = 36$; **Figure 7D**; **Table 2**). We found a significant difference in $b_{bp}^{\sim}(555)$ between locations and seasons (**Tables 2, 3**). Although, its mean value of 0.03 was above that found by Whitmire et al. (2007) (i.e., $0.01 < b_{bp}^{\sim}(555) < 0.02$) over various coastal, oceanic, and freshwater environments of the US, it was in accordance with measurements found by McKee et al. (2009) in the shallow, macro-tidal estuary of the Bristol Channel where $b_{bp}^{\sim}(532)$ was found to reach up to 0.07. Albeit at a different, shorter wavelength, the bulk of their $b_{bp}^{\sim}$ measurements had a bi-modal distribution at 0.01 and 0.03. Spatially, $b_{bp}^{\sim}(555)$ was found to differ ($p < 0.05$; $N = 23$), increasing from DH (0.027 ± 0.005; $N = 12$) to the VDG (0.033 ± 0.009; $N = 11$; **Figures 2D**, **7D**). The scattering-to-attenuation ratio, b/c(555), averaged 0.91 ± 0.08 ($N = 48$) (**Figure 9A**), with the scattering coefficient, b(555), contributing >90% ($N = 26$) of the beam attenuation coefficient in the VDG independently of the seasons (**Figures 6**, **9D**). In DH however, this contribution decreased to 79% during the dry season sampling in 2012 ($N = 14$) but was equivalent to that of the VDG (>90%) during the dry season sampling in 2013 ($N = 8$).

The relative contributions of phytoplankton, NAP and CDOM absorptions to the total absorption budget for our samples ($N = 49$) are displayed in **Figure 8**. Overall, the waters of this region were largely (60%) dominated by CDOM during the wet season and mostly (>45%) by NAP during the dry season, with phytoplankton contributing very little to the total absorption at

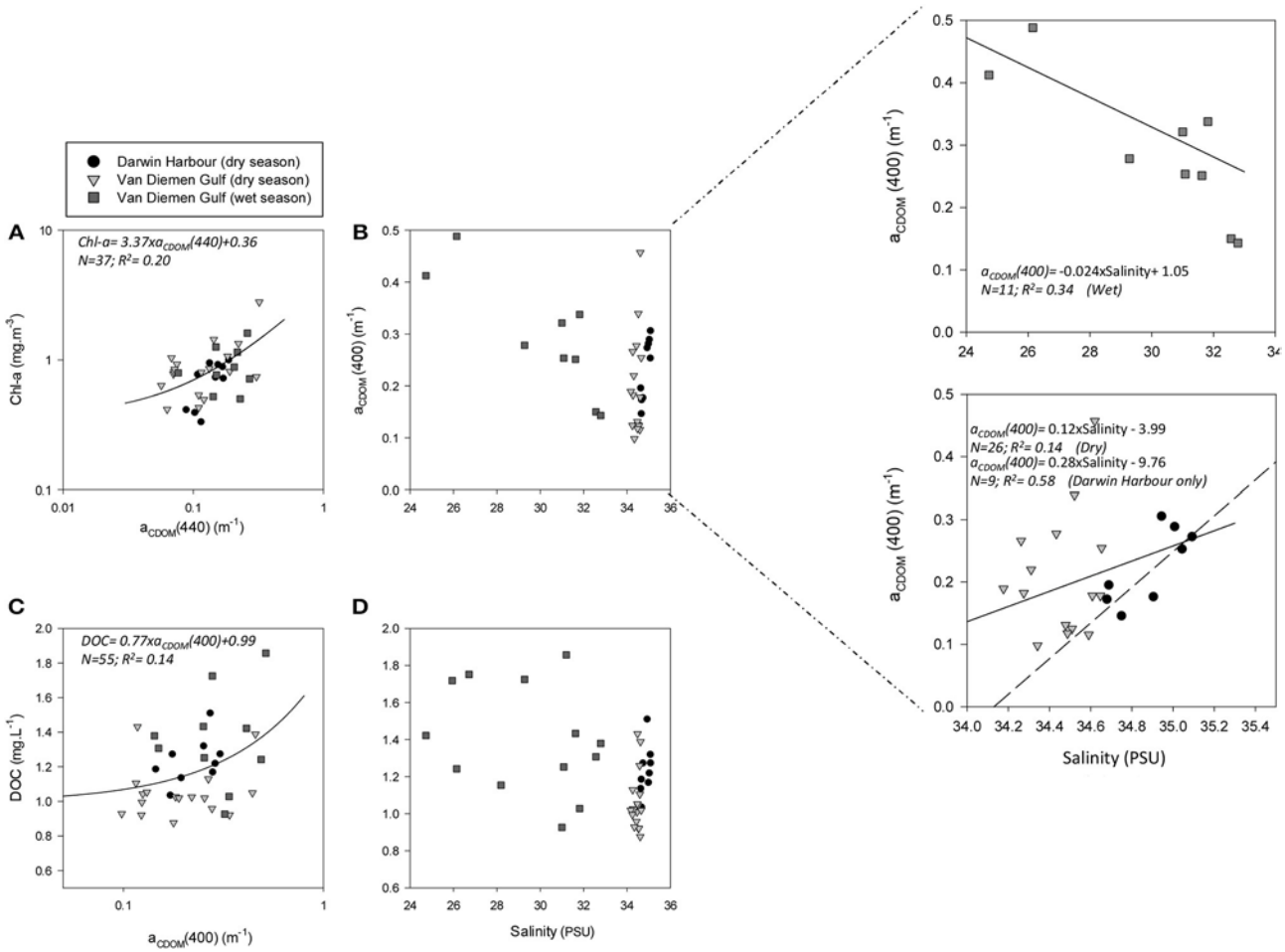


FIGURE 4 | Relationships between: (A) Chl-a and $a_{CDOM}(440)$, **(B)** $a_{CDOM}(440)$ and salinity, **(C)** DOC and $a_{CDOM}(400)$, and **(D)** DOC and salinity. The dashed line indicates the regression line for the Darwin Harbour samples only ($N = 9$).

440 nm (**Figure 8D**). The combined absorption by NAP and CDOM generally exceeded 70% of the total absorption from 400 to 620 nm. The second highest combination was CDOM and phytoplankton (>60%), while NAP and phytoplankton contributed 50–60% for the same wavelengths. The median contribution of phytoplankton at 440 nm was 18% overall (**Figure 8D**), increasing to >55% at 665 nm. Darwin Harbour showed a mixed CDOM and NAP assemblage.

The particulate absorption coefficient, $a_P(440)$, was found to be equally correlated with both TSS ($R^2 = 0.30$; $p < 0.001$; $N = 51$; **Figure 9B**) and Chl-a ($R^2 = 0.28$; $p < 0.001$; $N = 50$ not shown). A significant seasonal and local (VDG) variation was also found for $a_{NAP}/a_P(440)$ ($p < 0.05$; $N = 50$) (**Figure 9C**; **Table 3**).

The quality of the optical closure between ac-s and filter pad total absorptions [$a_{Tot}(\lambda)$] was assessed at the selected wavelengths of 412, 440, 510, 532, and 676 nm for the VDG dry and wet season field campaign measurements. There was a very good agreement ($R^2 > 0.9$) between the two methods in the blue and green spectral regions [i.e., $a_{Tot\,(ac-s)} = 0.96xa_{Tot\,(Filters)}$; $R^2 = 0.99$; $N = 26$ at 412 nm and $a_{Tot\,(ac-s)} = 0.79xa_{Tot\,(Filters)}$; $R^2 = 0.97$; $N = 26$ at 532 nm] but the optical closure largely degraded in the near-infrared [i.e., $a_{Tot\,(ac-s)} = 0.31xa_{Tot\,(Filters)}$; $R^2 < 0.1$; $N = 26$ at 676 nm]. The absolute relative errors between the two measurements increased from 3% in the blue to 16% in the green (532 nm) to 85% in the NIR. The correction for residual scattering in the reflecting tube of an ac-s (or ac-9) varies spectrally (Röttgers et al., 2013), while it is considered wavelength independent in the proportional correction method (as selected for this study). This may lead to high discrepancies (Leymarie et al., 2010; Pitarch et al., 2016) and although we acknowledge that the choice of the proportional correction method significantly affected our ac-s estimates in the NIR, no ac-s or bb-9 data beyond 555 nm was used in this study.

Phytoplankton: Absorption Coefficients, Pigments, Cell Counts, and Nutrients

Phytoplankton cell sizes in DH were mostly (>50%) dominated by picoplankton (<2 μm). In the VDG, the dry season samples were dominated (>60%) by microplankton (>20 μm), and the wet seasons samples were a mixed population assemblage

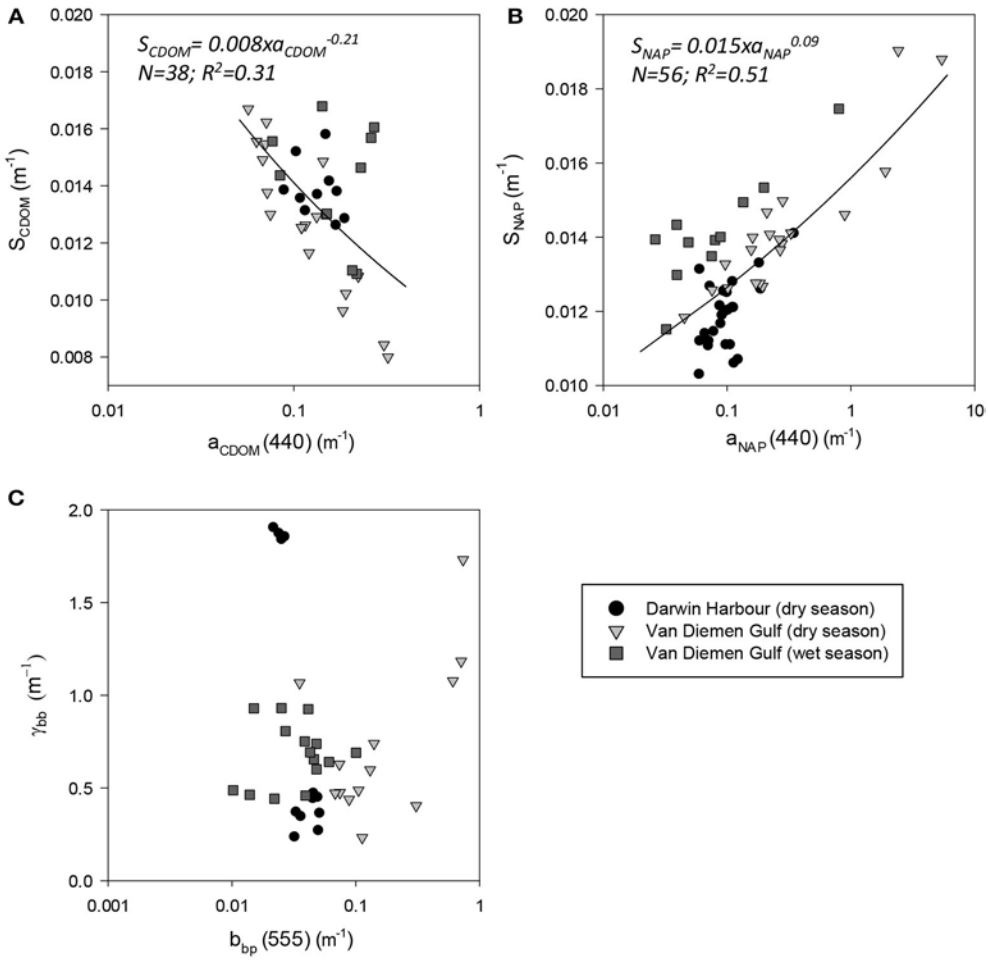


FIGURE 5 | Spectral slopes. Slopes of **(A)** CDOM, **(B)** NAP, and **(C)** b_{bp} as a function of their respective coefficients.

(**Figure 10**). The two major carotenoid pigments, fucoxanthin and zeaxanthin, are diagnostic pigments for diatoms and cyanobacteria, respectively (**Figures 11E,F**). Overall, fucoxanthin was the most abundant pigment (60% of the 57 stations) and was mostly present in the VDG dry season samples (30%) (**Table 3**), while zeaxanthin characterized the DH samples (26%). As expected, fucoxanthin strongly co-varied with Chl-a ($N = 57$; $R^2 = 0.70$; $p < 0.001$; **Figure 11C**). From the phytoplankton cell counts (2013 dry season campaign only), we found that the dominance of diatoms increased spatially, from DH to the Gulf. For the VDG dry season samples, fucoxanthin/Chl-a ratios were typically one order of magnitude higher than those of zeaxanthin, confirming that diatoms were the most abundant phytoplankton group (>70%; $N = 36$).

In the VDG, Chl-c, an accessory pigment in diatoms, was found to be more abundant than Chl-b (53%), a pigment generally associated with e.g., picophytoplankton Prochlorococcus which was mostly present in DH (72%). The pigment ratio Chl-b/Chl-a was low in the waters along Melville Island and higher in Darwin Harbour, while Chl-c/Chl-a increased from the VDG coast to Melville Island and was higher in Darwin Harbour but lower in the VDG wet samples. The VDG dry season dataset was characterized by higher PSC (**Figures 11A–D**), while the VDG wet and DH samples were higher in PPC. The PPC/PSC ratio provides a photo-physiologic index for phytoplankton cells. Environmental stresses, such as high light or low nutrient availability, are usually associated with higher PPC, thus resulting in higher PPC/PSC ratios. A low PPC/PSC ratio would in turn be associated with phytoplankton cells receiving low light levels in high-nutrient surface waters. However, the PPC/PSC ratio also varies between phytoplankton taxonomic groups, as phytoplankton cells exhibit a variety of tolerances and adaptations to light and nutrient exposure. Overall, the VDG samples had lower PPC/PSC ratios (∼0.68) during the dry season than during the wet season (∼1.50), which is due to limited light penetration from high turbidity levels. In contrast, the PPC/PSC ratios were generally high (∼1.55) for the Darwin Harbour samples, likely due to a better light penetration through the water column in this system at the time of our sampling (**Figure 6**).

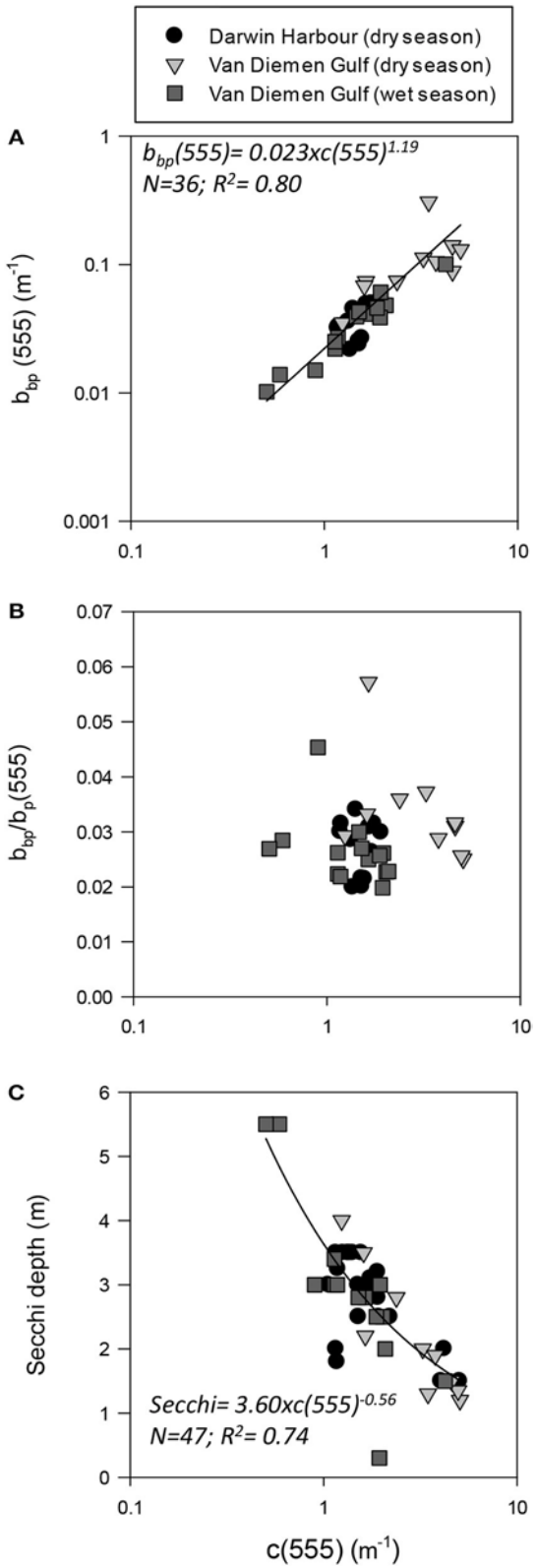


FIGURE 6 | Beam attenuation coefficient (c_{555}) as a function of (A) backscattering ($b_{bp\,555}$); **(B)** the backscattering-to-scattering ratio ($b_{b\,555}$); and **(C)** Secchi depth.

Cyanobacterium *Trichodesmium* sp. inhabits tropical waters, and in low wind stress conditions it produces large surface blooms that can be monitored from space (McKinna, 2014). Blooms of *Trichodesmium* sp. are known to occur in the region, although the dynamic coastal environment of the VDG may not favor its optimum growth, due to water column mixing. *Trichodesmium* sp. patches were seen at five stations over 3 days during the September 2013 field campaign, and high counts (>200 units/L) of cyanobacterial cells were found at one station in the eastern embayment of the Gulf, thus confirming previous findings from remote sensing observations of Blondeau-Patissier et al. (2014) that these cyanobacteria blooms occur regularly between August and October in the VDG-DH region.

The absorption coefficient of phytoplankton, $a_{phy}(440)$, was found to decrease from the coast to the center of the Gulf and was higher in the inner Harbour, but there was no significant difference in $a_{phy}(440)$ between the dry and wet season samples. There was a significant correlation between $a_{phy}(440)$ and Chl-a ($R^2 = 0.55$; $p < 0.001$; $N = 50$; **Figure 12B**; **Table 5**), and no significant difference was found between the relationship derived from our dataset and that of Bricaud et al. (2010) ($p > 0.1$) or Bricaud et al. (2004) ($p > 0.1$). There was no significant seasonal variation in $a_{phy}(440/676)$ (**Figure 12A**; **Table 3**), and no significant difference was found between the model derived for this dataset and that of Bricaud et al. (1995) ($N = 50$; $p > 0.05$), therefore inferring that any of these models can be used in the VDG and/or DH (**Figure 12**). A significant difference ($p < 0.005$) however, was found for $a_{phy*}(440)$ between locations (**Figure 12C**; **Table 3**), thus reflecting the difference in phytoplankton composition between the VDG and DH.

Nutrients, in particular nitrate and ammonium, have the greatest potential to limit phytoplankton growth in coastal marine systems (Malerba et al., 2015). For our study regions, nutrient concentrations sampled during the dry season 2013 were much higher in the VDG in comparison to DH: mean ammonium and phosphate concentrations were almost twice those recorded in DH, while nitrate was three times greater in the VDG (**Figure 14**; **Table 4**). DH is a nitrogen depleted environment (Wolanski et al., 2006) and nitrate showed an increasing concentration gradient from the inner Harbour ($\sim$0.06 μmol.L^{-1}; $N = 5$) to the outer Harbour ($\sim$0.1 μmol.L^{-1}; $N = 6$; **Figures 1B, 14**), consistent with the principle that oceanic waters coming into the harbor are richer in nutrient content.

Summary of Results

Poor correlation ($R^2 < 0.2$) was found between DOC and $a_{CDOM}(400)$ or salinity, as well as between salinity and a_{CDOM} or Secchi depths, and between POC and TSS with Chl-a. A mild correlation ($0.2 < R^2 < 0.4$) was found between Secchi depths and tidal range or temperature or TSS, between $a_P(440)$ and TSS and Chl-a and also between $a_{CDOM}(440)$ and Chl-a. A moderate correlation ($0.4 < R^2 < 0.6$) was found between TSS and temperature, and between the tidal range and a_{NAP} slopes. A moderate to strong correlation ($0.6 < R^2 < 1.0$) was found between Secchi depths and $a_{NAP}(440)$ and $a_{Tot}(440)$. A very good optical closure was found between ac-s measurements and the filter pads for the total absorption coefficient [$a_{Tot}(412–555)$;

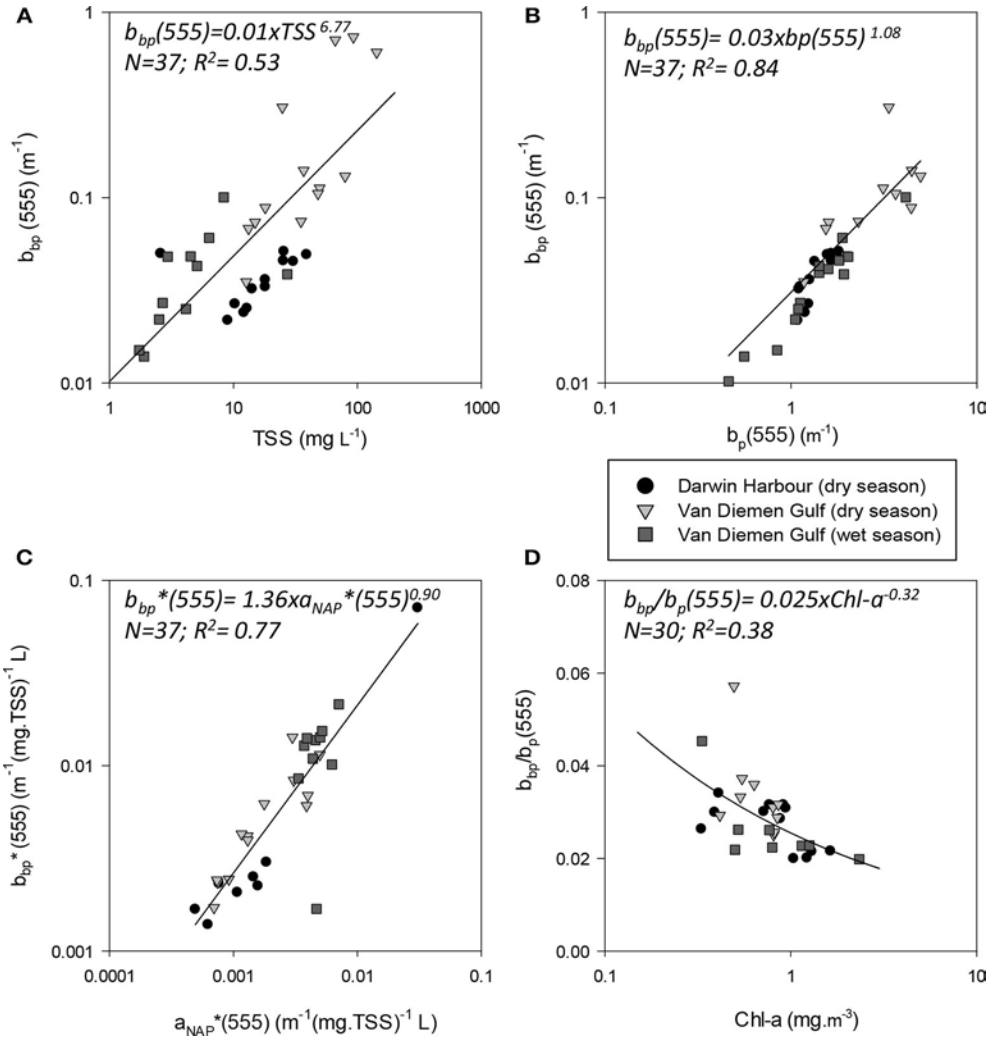


FIGURE 7 | Absorption and scattering properties of particles. Scatterplots of backscattering (b_{bp555}) as a function of **(A)** TSS and **(B)** the scattering coefficient (b_{p555}), **(C)** Mass-specific NAP coefficient and mass-specific backscattering coefficient and **(D)** b_b(555) as a function of Chl-a.

$R^2 = 0.9$]. The backscattering b_{bp}(555) and scattering b_p(555) coefficients were highly correlated ($R^2 > 0.8$). DOC was found to be significantly different between seasons and locations. While CDOM was found to be higher during the wet season, TSS concentrations were found to be higher during the dry season in the VDG, especially for stations along the coast. DH (dry season) was composed of a mixed assemblage of CDOM and NAP. No data was available for the wet season months in DH. Chl-a and the backscattering coefficient b_{bp}(555) did not show any statistical difference between seasons or locations, which is in contrast to c(555) and the backscattering ratio which varied seasonally and spatially.

DISCUSSION

A Complex Environment

Effect of Local and Seasonal Forcing

The major processes controlling the optical properties in DH and the VDG are mainly the wind speed and direction during the dry season, and the tides year round. In addition, the river discharges in this region are controlled by seasonal rainfall and thus add to the complexity of this coastal system by increasing land-sourced CDOM delivery to the coastal waters. The coastal waters of Northern Australia have a tidal range that is amongst the largest in the world for a coastline facing an open ocean. For our dataset, the tidal range varied from 1 to 6 m. Stations located in DH (dry season only; $N = 24$) were mostly sampled during neap tides, with a mean tidal range of 2.05 ± 1.04 m. In the VDG however, the tidal ranges were 20% higher at the time of our dry season sampling (4.05 ± 1.15 m; $N = 16$) than during the wet season sampling (3.30 ± 0.78 m; $N = 12$), partly explaining the higher concentrations in TSS, from resuspension, found during the dry season. The strong tidal currents in combination with the complex, shallow bathymetry caused the high spatial variability observed in most of the surface measurements from site to site (e.g., **Figure 2**) as a result of localized, small scale up- and downwelling processes. Previous modeling studies have shown that Darwin Harbour's

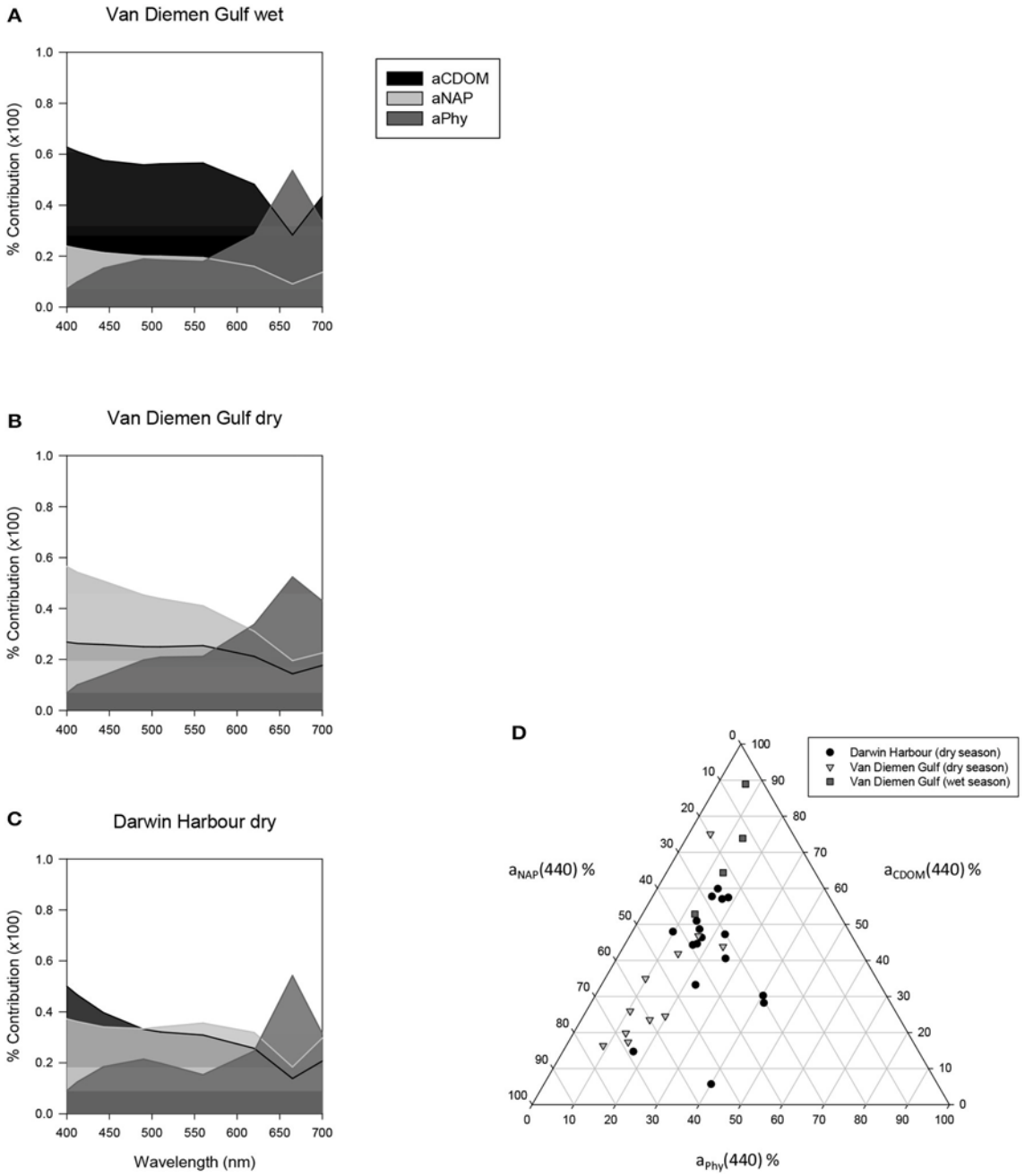


FIGURE 8 | Average relative contribution of CDOM, NAP, and phytoplankton to the total absorption between 400 and 700 nm (A–C). Wavelength-specific ternary plot of their relative absorption at 440 nm **(D)**.

hydrodynamics are driven mainly by tides, with the wind and seasonal river inputs playing somewhat smaller roles (Li et al., 2012). In particular, Li et al. (2014) reported that the dynamics of TSS in DH vary with the spring-neap tidal cycle, with the whole water column being well-mixed during spring tides. For our study, the DH stations were sampled during the clearest conditions (neap tides) and thus our DH dataset may not be the most representative of this environment. However, it provides an interesting comparison with the VDG that is further discussed in the Section Challenges for Remote Sensing of Water Quality in the NMR.

In addition to tidal movement, another physical forcing explaining the higher TSS concentrations in the VDG during the dry season are the prevailing south-easterly winds occurring at this time of the year. These trade winds, when acting in phase with the tidal currents, significantly enhance resuspension.

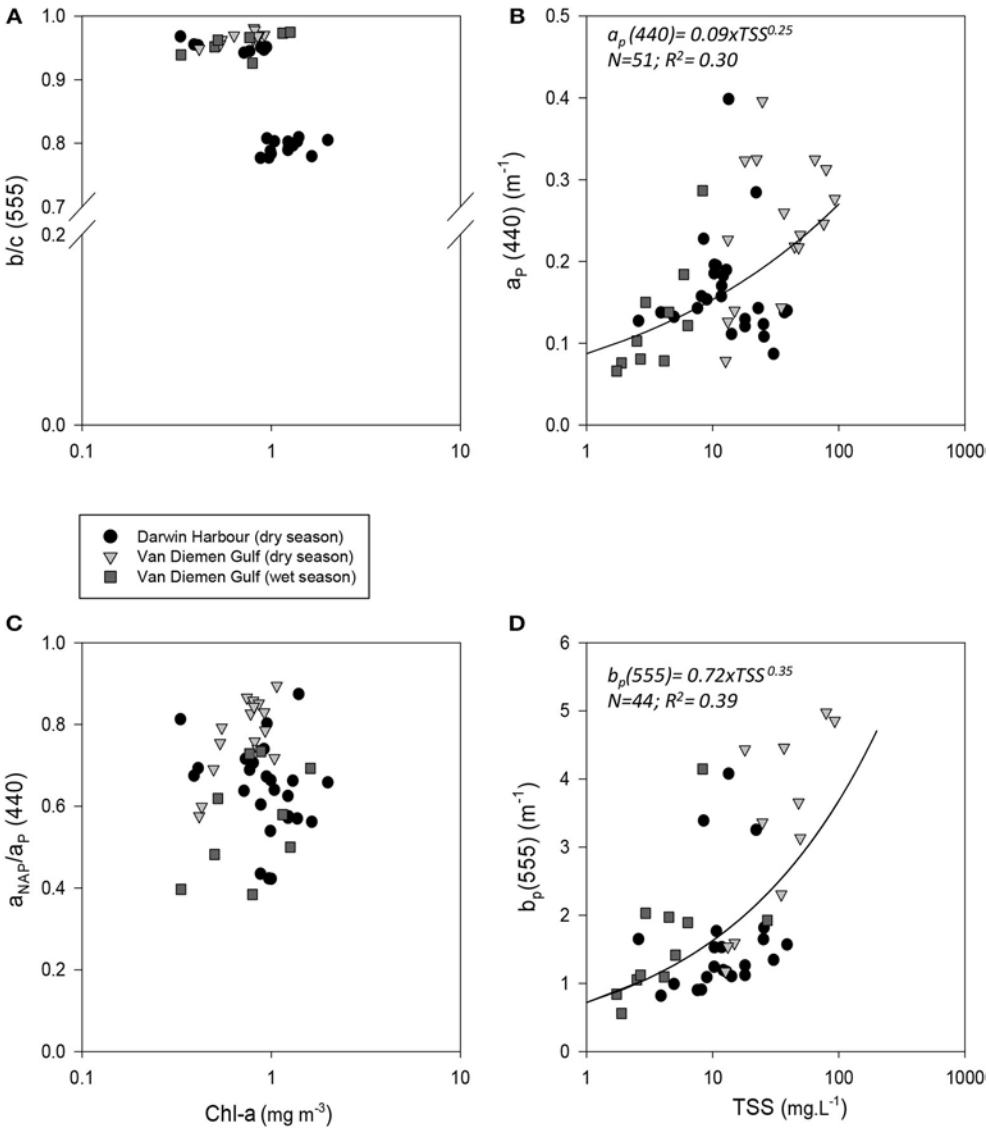


FIGURE 9 | Relationships between (A) b/c(555) and Chl-a; **(B)** a_p(440) and TSS; **(C)** proportion of non-algal particulate absorption to particulate absorption as a function of Chl-a and **(D)** b_p(555) and TSS.

While the recorded wind speeds were very low during the wet season field campaign (<2 m.s^{-1}; W-SW), winds of up to 10 m.s^{-1} occurred during the dry season sampling. Webster and Ford (2010) found that wind-induced waves in combination with tidal currents contributed to higher concentrations of sediment in Keppel Bay, a shallow embayment adjacent to the Fitzroy estuary in sub-tropical Queensland; In the North Sea, Hommersom et al. (2009) also found that TSS showed large short-term spatial variability due to the tidal energy, in combination with winds. These studies align with our findings that the combination of tidal energy, shallow depths, and strong winds is the main physical process controlling the increased TSS concentrations found during the dry season in the VDG.

The Alligator Rivers are directly connected to the VDG (**Figure 1A**), and their freshwater flows are substantial during the first 4 months of the year due to the highly seasonal rainfall. Their discharges, loaded with terrestrial material, are trapped within the coastal boundary layer, a body of turbid inshore water. Very little mixing occurs between the turbid coastal boundary layer and the rest of the Gulf. Hence we sampled much higher concentrations of sediment material along the coast in comparison to the middle of the Gulf (**Figure 2**). In addition the entire coastal area between DH and the Cobourg Peninsula (**Figure 1A**) is dominated by mangroves and tidal flats that are also known to play a key role in the sediment redistribution (Li et al., 2014). An explanation for the surprisingly low contribution of non-algal particles and TSS found for the

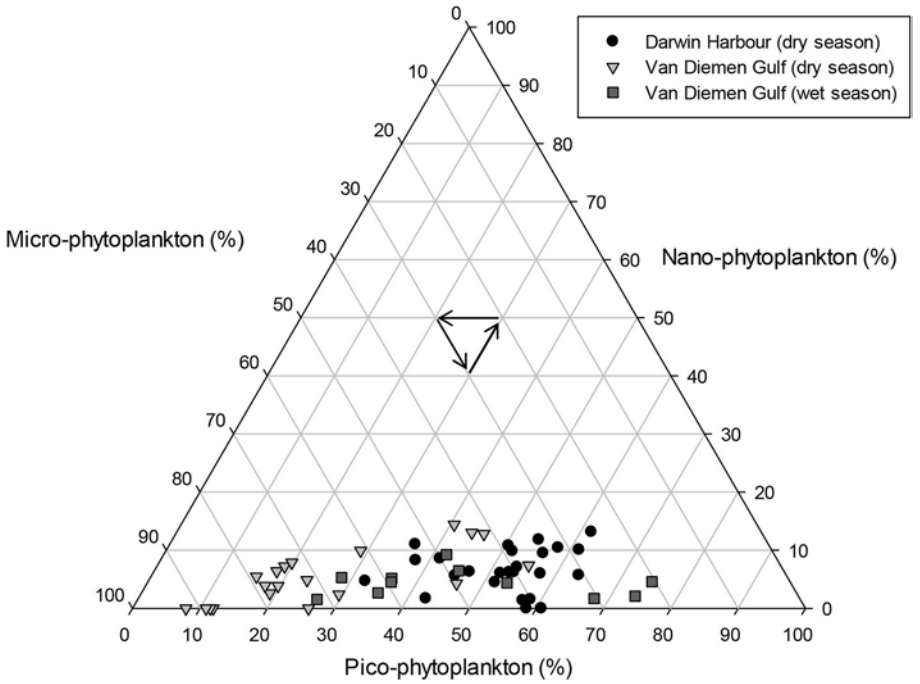


FIGURE 10 | Proportions of phytoplankton cell sizes for the Darwin Harbour dry ($N = 25$), Van Diemen Gulf dry ($N = 20$), and Van Diemen Gulf wet ($N = 11$) samples. Size classes are defined as micro >20 μm, nano, 2–20 μm, and pico, <2 μm.

wet season samples in the VDG is the river runoff being largely filtered through most of the Kakadu National Park wetlands (20,000 km^2) prior to reaching the coast (**Figure 1A**; e.g., Pusey et al., 2015). The seasonal variability in river runoff also directly results in significant difference in salinity between seasons. The large amount of runoff reaching the coastal waters at the end of the wet season (∼April), which coincided with the timing of our VDG 2012 wet season field campaign (**Table 1A**), explains the lower salinity observed during this seasonal sampling. Terrestrial runoff plays a role in CDOM characterization, by changing its slopes and concentrations. Higher $a_{CDOM}(440)$ values were observed during the wet season but surprisingly, these concentrations were not found to be significantly different from the dry season. Neither was a relationship found between a_{CDOM} and salinity, but it has to be stressed that the CDOM dataset for the wet season was limited ($N = 10$), with only four stations having salinities lower than 28 PSU. Further seasonal sampling, in particular during the wet season and in DH, is required.

Our results showed that there was, statistically, no spatial (VDG vs. DH) or seasonal (dry vs. wet) differences in Chl-a concentrations. But within the VDG however, there was an inshore-offshore Chl-a gradient: stations close to the Alligator Rivers and Mary River featured Chl-a from 1.3 up to 2.8 $mg.m^{-3}$, steadily decreasing to 0.5 $mg.m^{-3}$ in the Dundas Strait (**Figures 1**, **2A**). This distribution, which was not necessarily reflected in the other properties (beside possibly CDOM; i.e., **Figure 2**), is associated with the nutrient stock being concentrated mostly along the coastline and in the eastern embayment of the Gulf (VDG) due to the presence of the boundary layer that allows phytoplankton to grow (**Figure 14**). In DH, a similar mechanism occurs whereby the upper harbor is characterized by a longer water residence time (exceeding 20 days; Williams et al., 2006) in comparison to the inner section (**Figure 14**). It takes up to 22 days for the nutrients located in the upper section of the harbor to reach the sea (DHAC, 2010).

Chl-a is an indicator of phytoplankton biomass and light availability is undeniably playing a central role in the phytoplankton production in this study region: while it was found that phytoplankton production is highest in the wet season (e.g., Blondeau-Patissier et al., 2014), it is limited during the dry season possibly because of the high turbidity from the water column mixing, generated by both the shallow bathymetry and strong, dry seasonal winds, which will limit light penetration. This as reflected by lower PPC/PSC ratios and lower Chl-a concentrations found in the center of the VDG during the dry season (**Figure 2A**). Wet season river discharges release large amounts of nutrients and, together with the release of CDOM, are likely to fuel the primary productivity (e.g., Burford et al., 2011).

Previous studies on subtropical estuarine systems of the South East Australian coast have shown the pronounced effect of short-timescale variability in dissolved and particulate matter (particle size, composition) due to the spring-neap tidal cycle (Oubelkheir et al., 2006), stressing the need to take into account the tidal phase in the sampling strategy. Shi et al. (2011) demonstrated that the magnitude of the spring-neap tidal cycle on the variations of satellite-derived $K_d(490)$ and TSS was comparable to the seasonal

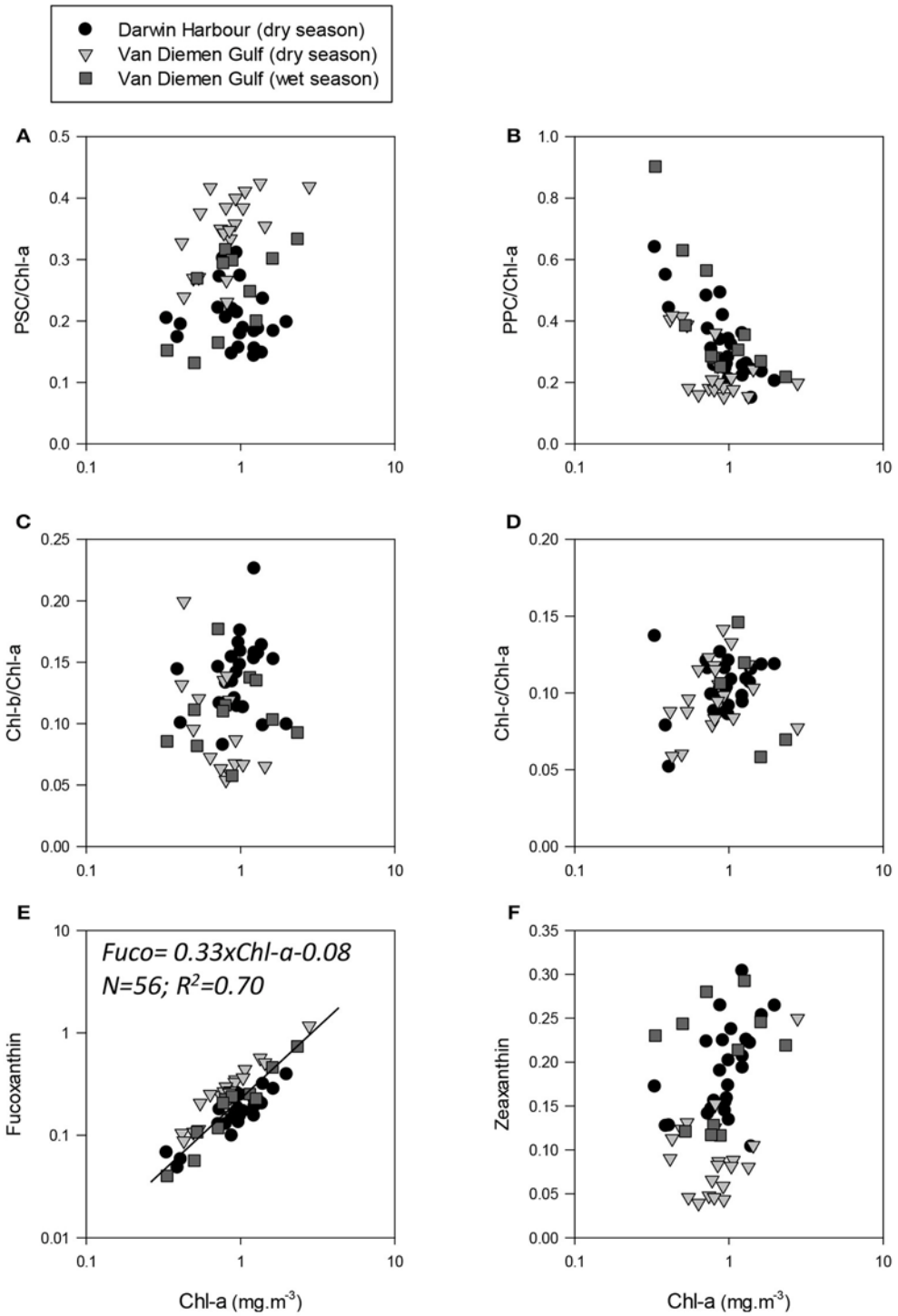


FIGURE 11 | Photo-acclimation and photo-adaptation processes in the NMR. Pigment ratios as a function of Chl-a. **(A)** photosynthetic carotenoids to Chl-a ratios; **(B)** non-photosynthetic carotenoids/Chl-a ratios; **(C)** Chl-b to Chl-a ratios; and **(D)** Chl-c to Chl-a ratios. Fuxoxanthin and zeaxanthin as a function of Chl-a are shown in **(E,F)**, respectively.

effect observed in their coastal region of SE Asia. The tidal range in their study region was comparable (∼3 m) to that observed during our study of the VDG. Oubelkheir et al. (2014) also emphasized the role of short-term processes, such as wind stress and tides, as key drivers of the dissolved and particulate material in the shallow and dynamic subtropical environment of Moreton

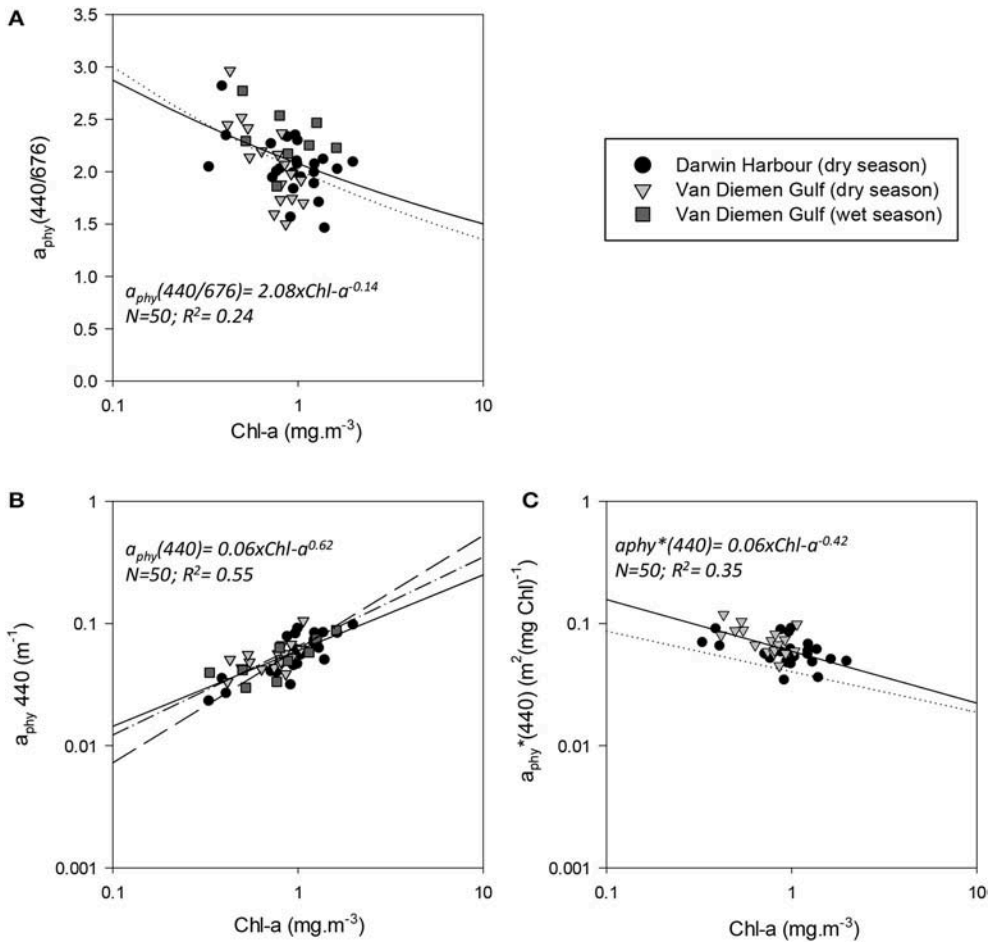


FIGURE 12 | Relationships between phytoplankton and Chl-a. Reference models from Bricaud et al. (1995) **(A–C)** dotted line; Bricaud et al. (2004) **(B)** dash-dot line, and Bricaud et al. (2010) long dash line are shown.

Bay (SE Queensland). In our study, the sampling strategy was primarily to ensure our stations had a large spatial distribution rather than to sample the small-scale temporal variability due to the tides, thus the tidal phase was not taken into account. This should be addressed in future field campaigns in this region, as the spring-neap tidal cycles affect both the properties sampled in the field and satellite-derived ocean color products.

A System Driven by CDOM and NAP

Characteristic of coastal waters, most optical properties and some of the biogeochemical concentrations sampled during these three field campaigns cover a large range of variability (**Table 2**). While $a_{NAP}(440)$ (coefficient of variation = 252%), $a_{tot}(555)$ (CV = 159%), and $b_{bp}(555)$ (CV = 167%) are the IOPs that varied most across seasons and locations, $a_{phy}(440/676)$ (15%) and Chl-a (45%) were amongst those parameters that varied the least. Our results highlight that the optical properties of our coastal system (VDG in particular) was mostly driven by NAP during the dry season and CDOM during the wet season, with little influence of phytoplankton on the total absorption budget at 440 nm. Sources of NAP include phytoplankton bio-products and non-algal detritus and we concluded that NAP in the VDG did not originate from phytoplankton but rather from sediment. This is supported by the weak correlation of NAP with both Chl-a ($R^2 = 0.33$; $N = 55$) and a_{phy} ($R^2 < 0.10$; $N = 52$), while being significantly correlated with TSS ($R^2 \sim 0.50$; $N = 55$). The important (∼84%) contributions of $a_{CDOM}(440)$ and $a_{NAP}(440)$ masked the contribution due to phytoplankton, which at ∼16% was below the contribution of $a_{phy}(440)$ to the total absorption budget (20–60%) obtained in European coastal waters by Babin et al. (2003b).

For our study, it was found that only ∼10% of the phytoplankton cells counted during the 2013 dry field campaign were viable, thus indicating that the productivity must be due to a contribution from the microbial population in this system. The UV spectroscopic characteristics of the DOC sampled in the VDG during the dry season of 2013 (low absorbance at 350 nm) are consistent with organic matter that has been subjected to bacterial degradation (**Figure 13**). Our results therefore suggest that the DOC of the dry season samples may have a different chemical character from the DOC delivered during the wet season. The size of the DOC fraction

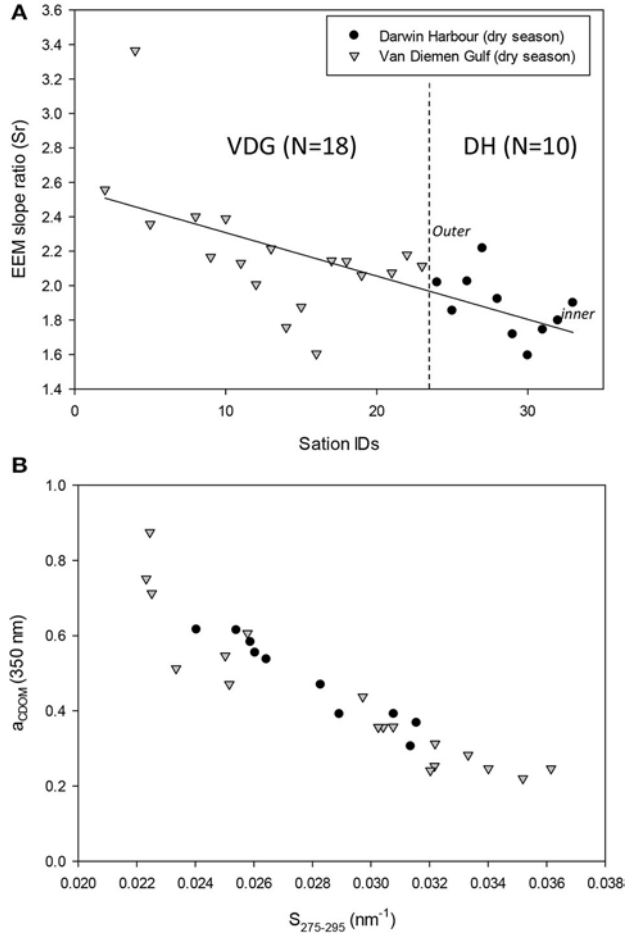


FIGURE 13 | (A) EEM spectral slope ratios (Sr) showing a decline through the course of the 2013 dry season field campaign from the VDG to DH and **(B)** relationship between aCDOM and the EEM spectral slope in the UV.

in the samples remained stable, indicating possible shifts in microbial community structure, which would be expected in a tropical environment. Analysis of the FDOM from the EEM also suggested that during the dry season, the material was composed mostly (∼70%) of aromatic organic carbon (Weishaar et al., 2003) with a relatively constant spectral slope (0.029 ± −0.003), therefore inferring that only a small fraction of the DOC was derived from terrestrial sources (Fichot and Benner, 2012). The underlying assumption that CDOM properties can be inferred from DOC concentrations (Shanmugam, 2011; Hestir et al., 2015) was, however, not verified in this dataset. Other recent studies also reported the complete lack of relationship between DOC and CDOM (e.g., Nelson and Siegel, 2013) because the controlling factors for DOC and CDOM were different (Yamashita et al., 2013).

Regional Differences during the Dry Season

The waters of the VDG were found to have a higher scattering component than the waters of DH, largely contributing to higher beam attenuation during the dry season in particular. This result is consistent with the higher TSS concentrations found in the VDG (dry) in comparison to DH (e.g., **Figure 2B**). The two environments were sampled over different tidal conditions, as the sampling in DH only occurred during neap tides. The fetch length—much larger in the VDG—and the seasonal forcing are most likely explaining the significant differences between the two locations. The VDG is ∼130 km wide from the shore to the tip of the Cobourg Peninsula (**Figure 1**), thus offering much more ocean surface for the wind to generate sea surface roughness and resuspension. The harbor in comparison, is 15 km wide, mostly protected from the wind and its entrance is the section the most influenced by incoming waves from the ocean.

There were distinct differences in phytoplankton pigments, composition and cell counts between the samples collected in the VDG and those collected in DH (**Figure 1**). From cell counts and identification during the 2013 dry season field campaign, diatoms were present at 23 of the 33 stations sampled in the VDG during the dry season. This microphytoplankton dominance was also observed by Burford et al. (1995) in the neighboring Gulf of Carpentaria, suggesting possible similarities between the two gulfs; picophytoplankton however, was predominantly present in DH. It can be inferred that at the time of the 2013 field sampling, the higher number of cyanobacteria cells indicates a marine influence in DH. There was a low number of phytoplankton cell counts for most samples (∼250 cells/L) overall, possibly due to the low productivity at this time of year (e.g., Blondeau-Patissier et al., 2014). In contrast, a greater diversity in phytoplankton cells was found in the Gulf. This is explained by the higher levels of nutrients being available in the VDG while DH's primary production is limited, mostly by nitrogen (Burford et al., 2008; **Figure 14**; **Table 4**). Differences in nutrient concentrations between the two locations was expected because the oceanic and VDG waters are known to be richer in nutrients in comparison to DH (Wolanski et al., 2006). Nitrate and phosphate concentrations are relatively low in the VDG (**Table 4**) but silicate is mostly present, explaining the large dominance of diatoms in the VDG (Condie and Dunn, 2006).

CDOM was not found to be different between the two locations, yet it has an inshore-offshore pattern (**Figure 2C**) with an evident decrease from the inner (0.17 m^{-1}) to the outer (0.08 m^{-1}) harbor for DH. We hypothesize that during the wet season, the larger amount of CDOM delivered from river runoff would compare differently to that of DH, as the amount discharged into the Gulf (e.g., East Alligator River discharge: 7 GL/yr) is much larger than the amount of CDOM that would discharged in DH (e.g., Elizabeth River discharge: 3 GL/yr). The relative contributions of the three absorption components, namely $a_{phy}(\lambda)$, $a_{NAP}(\lambda)$, and $a_{CDOM}(\lambda)$, were investigated in this study and showed that DH was mostly characterized by approximately equal contributions of $a_{NAP}(\lambda)$ and $a_{CDOM}(\lambda)$, while $a_{NAP}(\lambda)$ was predominant in the VDG. The role of S_{NAP} is poorly understood but this parameter is often found to be stable across marine environments (Matsuoka et al., 2011). This was not the case in this study as $S_{NAP}(\lambda)$ was significantly different between the two regions with significantly higher slopes in the VDG. This difference is likely related to the

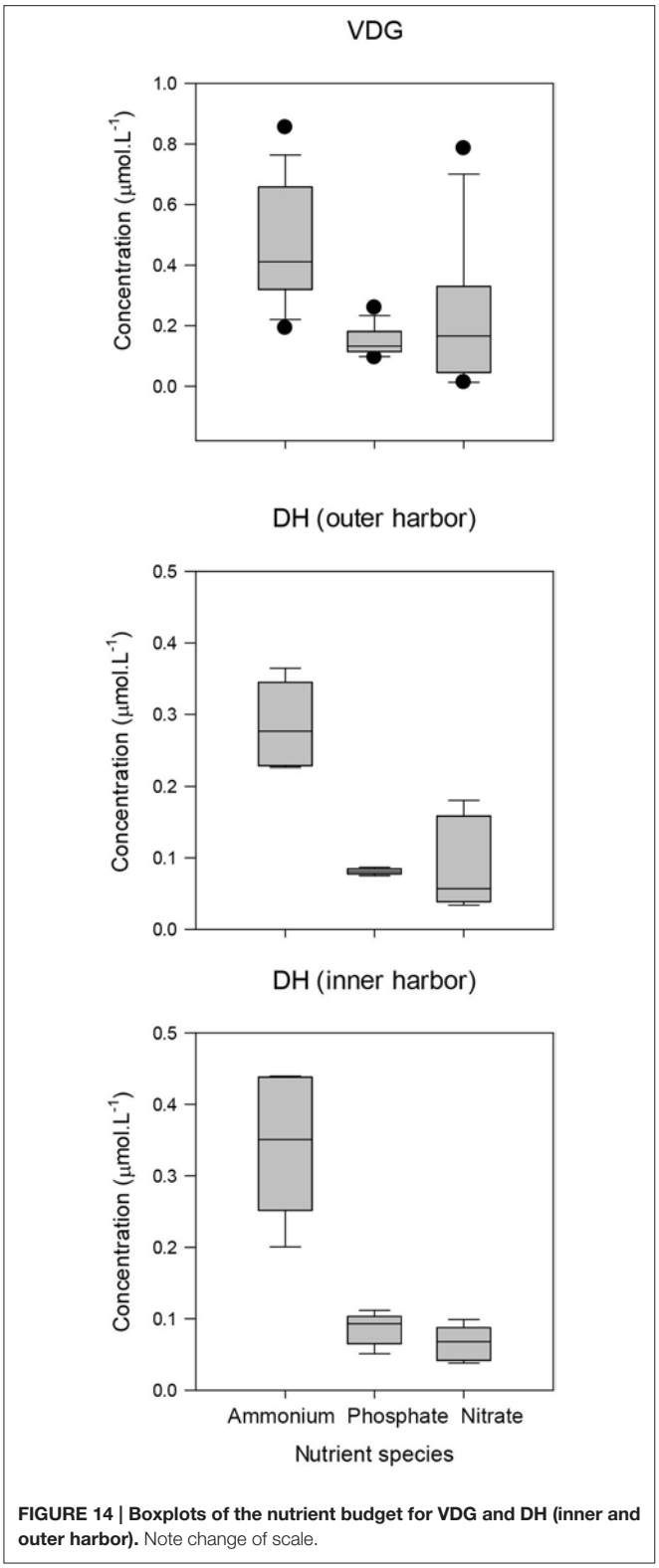


FIGURE 14 | Boxplots of the nutrient budget for VDG and DH (inner and outer harbor). Note change of scale.

particle assemblage characterizing the two regions, also reflected in $b_{bp}^{\sim}(555)$. In contrast, the dominance of organic material during the wet season months is consistent with an enhanced phytoplankton productivity, as reported in Blondeau-Patissier et al. (2014). In the Gulf of Carpentaria, Burford and Rothlisberg (1999) also reported higher integrated primary production (~955 mg.C.m^{-2}.day^{-1}) during the wet season months, therefore supporting this finding. Both b_{p555} ($R^2 = 0.97$; $N = 48$) and $b_{bp\,555}$ ($R^2 = 0.80$; $N = 36$) were found to strongly co-vary with c_{555} in DH and VDG, with the high b/c ratio (>90%) emphasizing that standard ocean color algorithms would likely fail in these waters, hence supporting the necessity to use a regional algorithm with an (S)IOP-based parameterization for the derivation of accurate water quality products in VDG and DH. **Table 5** provides a list of selected relationships between key (S)IOP and concentrations.

TABLE 4 | Nutrient budget for the two regions (in µmol.L^{-1}) as sampled during the 2013 dry season field campaign.

Nutrient species	Van diemen Gulf ($N = 18$)	Darwin harbour ($N = 10$)
Ammonium	0.47 ± 0.20	0.32 ± 0.08
Phosphate	0.15 ± 0.05	0.08 ± 0.02
Nitrate	0.24 ± 0.24	0.08 ± 0.05

Challenges for Remote Sensing of Water Quality in the NMR

The effect of tidal currents on the spatial variability of suspended sediments can easily be observed from space using satellite observations. **Figure 15** shows the two MODIS-Aqua (NASA) images covering the study region at a 250-m resolution during the 2013 dry season field campaign. Tidal currents are stronger during spring tides resulting in a more contrasted spatial distribution of suspended sediments, in comparison to neap cycles. This highlights yet another challenge in these highly dynamic environments: the scaling difference—while the satellite integrates over a larger area, the ground measurements represent point observations and their direct comparability, with respect to the satellite observations, may be questioned.

A previous study on the dynamics of phytoplankton blooms in the VDG from the MERIS mission (Blondeau-Patissier et al., 2014) found that an increase in TSS was occurring predominantly during the dry season. This observation is supported by the findings of the present work (**Tables 2, 3**). MERIS sensor estimates of dry season TSS concentrations from the 2014 satellite study were much lower (4 ~ 10 mg.L^{-1}) when compared to the results from these field samples. A total of 10 stations (with associated biogeochemical measurements) from the dry ($N = 4$) and wet ($N = 6$) season field trips in the VDG were located within the water mass cluster selected for the satellite study (see Figure 5B of Blondeau-Patissier et al., 2014). For these station locations, *in situ* TSS was found to increase from an overall average of 4 mg.L^{-1} during the wet season to 30 mg.L^{-1} during the dry season. This ~10-fold increase in TSS concentrations between seasons was observed in the MERIS satellite study, but it is important to recall that only few ($N = 4$) dry season stations in the VDG were used for this comparison. Conversely, *in situ* Chl-a concentrations weakly increased from

TABLE 5 | Summary of selected relationships.

x	*y*	*N*	Model	R^2
Chl-a	TSS	55	$y = 0.47x + 1.23$	<0.1
Chl-a	POC	29	$y = 1444x^{0.19}$	0.40
POC	TSS	29	$y = 1.15x^{0.05}$	0.14
$b_{bp}(555)$	TSS	36	$y = 0.01x^{6.77}$	0.53
$b_{bp}(555)$	$b_p(555)$	37	$y = 0.03x^{1.08}$	0.84
$b_p(555)$	TSS	44	$y = 0.72x^{0.35}$	0.39
$a_p(440)$	TSS	51	$y = 0.09x^{0.25}$	0.30
$b_{bp}(555)$	c(555)	36	$y = 0.023x^{1.19}$	0.80
S_{CDOM}	$a_{CDOM}(440)$	38	$y = 0.008x^{-0.21}$	0.31
S_{NAP}	$a_{NAP}(440)$	56	$y = 0.015x^{0.09}$	0.51
Chl-a	$a_{CDOM}(440)$	37	$y = 3.37x+0.36$	0.20
DOC	$a_{CDOM}(400)$	55	$y = 0.77x+0.99$	0.14
$a_{phy}(440/676)$	Chl-a	50	$y = 2.08x^{-0.14}$	0.24
$a_{phy}(440)$	Chl-a	50	$y = 0.06x^{0.62}$	0.55
$aphy^*(440)$	Chl-a	50	$y = 0.06x^{-0.42}$	0.35

The number of samples, N, varies due to quality control.

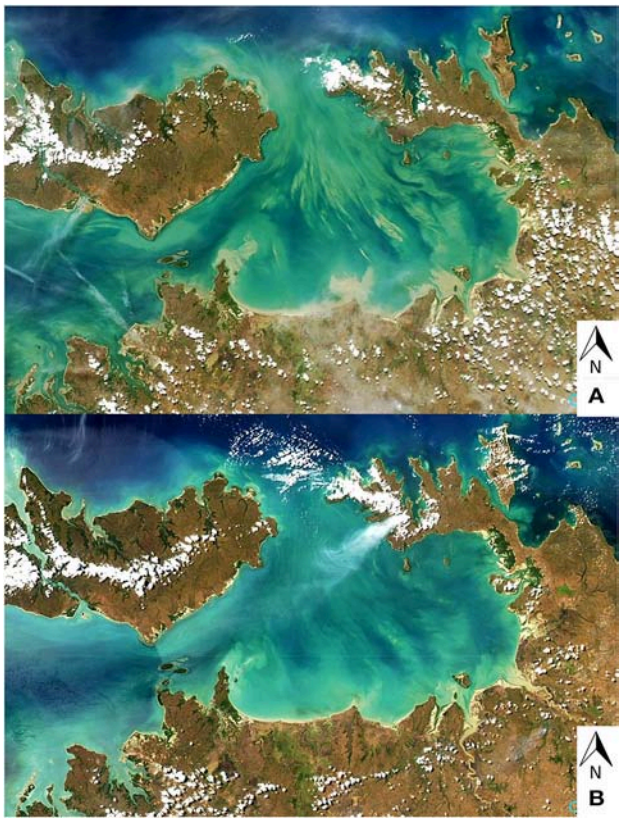


FIGURE 15 | Large spatial variability of suspended sediment depending on the tidal energy: MODIS-Aqua 250-m resolution images during the dry season 2013 field campaign. (A) 12 September 2013 (spring tide) and **(B)** 14 September 2013 (neap tide).

0.6 to ∼0.75 $mg.L^{-1}$. We can only interpret this comparison with caution for at least two reasons: first, because of the limited *in situ* dataset used for this exercise, and second because of the seasonal dominance of CDOM and NAP in this system which will inevitably hamper accurate satellite retrievals, of Chl-a in particular. This discrepancy also highlights the need to parameterize a region-specific remote sensing algorithm for the NMR. A seasonal parameterization, based on the wet and dry season *in situ* optical observations presented in this study, has been applied to MODIS-Aqua imagery of the VDG (Schroeder et al., 2015).

In recent years, ocean color remote sensing has provided a powerful means for studying ocean biogeochemistry and ecosystems over large spatial scales (Gardner et al., 2006; Tang, 2011; Swirgon and Stramska, 2015). Both the dissolved and particulate fractions of organic carbon can affect light penetration, and thus, optical properties may be used as proxies for DOC and POC (Pan et al., 2014), but satellite retrievals of these parameters may be a major challenge in the NMR because of their poor correlation with any of the variables measured. It is known that the use of optical proxies for satellite retrieval of POC is not straightforward because of the highly variable relationships between parameters (e.g., Cetinić et al., 2012). In the VDG, the maximum POC concentrations were found to be much higher than in other coastal systems, such as Chesapeake Bay (Fisher et al., 1998). For this study, POC was found to be strongly correlated with Chl-a ($R^2 = 0.40$; $N = 29$) and less so with TSS ($R^2 = 0.14$; $N = 29$), probably due to the composition of both the POC and the phytoplankton community (Zhu et al., 2006; Wang et al., 2009; MacIejewska and Pempkowiak, 2014a,b).

Of particular interest from a coastal ecosystem-based management point of view is the development of ecological indicators from ocean color remote sensing, such as seasonal cycle of phytoplankton biomass, spatial distribution of phytoplankton types or the delineation of ecological provinces (Platt and Sathyendranath, 2008). These require accurate *in situ* bio-optical measurements of phytoplankton absorption and chlorophyll to be used for remote sensing algorithm parameterization and validation. Long-term (>10 years) satellite-time series are now commonly used to assess trends in the productivity of coastal and ocean regions, and it is always stressed that well-calibrated ocean color sensors and algorithms are paramount to these estimates (e.g., Signorini et al., 2015). First and foremost, sea-truth measurements must be performed as closely as possible—both spatially and temporally—to the satellite observations if they are to be used for the validation of remote sensing algorithms. In highly dynamic coastal environments such as those of this study, it is recommended that *in situ* measurements be used only from samples collected within less than ±30 min of the satellite overpass to minimize bias (Martinez-Vicente et al., 2003). Independent of the match-up, the choice of the satellite algorithm, as well as its parameterization, is also an important factor. Aurin and Dierssen (2012) tested the performance of four semi-analytical algorithms at retrieving optical and biogeochemical properties in the complex waters of Long Island Sound (North Atlantic), a coastal system very similar to our region because of its comparable high proportion of CDOM

and NAP and its low (<20%) phytoplankton contribution to the total absorption. The quasi-analytical algorithm (QAA; e.g., Lee et al., 2010) was found to perform better [in comparison to the three other models tested, namely, a linear matrix inversion (LMI)-type algorithm, GSM and C99] because both S_{NAP} and S_{CDOM} were relatively stable in space and time. This is not the case for S_{NAP} in our study.

CONCLUDING REMARKS

The continental shelf of the NMR is the widest on the Australian continent, it is also one of the most pristine coastal environments worldwide (Morrison and Delaney, 1996; Halpern et al., 2008). Currently, the Northern Territory has no specific coastal management legislation or coastal climate change policy in place. Catchments of this region and the coastal marine environment are highly connected, yet the VDG lacks a coastal environmental monitoring program extending beyond the boundary of the well-managed Kakadu National Park. Therefore, long-term monitoring of key bio-physical coastal water quality parameters of the VDG, from *in situ* and satellite observations, would be of benefit to the regional environmental management. But, to date, the lack of *in situ* optical datasets in the NMR has limited the derivation of regional ocean color satellite algorithms for this region.

The dataset presented in this study fills this gap and is the first collected in the tropical, optically complex coastal waters of VDG and DH. Results from this study generally show that these two regions are separate coastal environments with different optical characteristics. The VDG was found to be mostly dominated by CDOM during the wet season and NAP during the dry season, while DH has a mixed absorption budget. The phytoplankton populations are also different, with the VDG being characterized by bigger phytoplankton cells (diatoms) in comparison to DH, due primarily to differences in nutrient stocks (**Table 4**) and light availability. The strong, south-easterly winds and the tidal energy are a combination that increases water turbidity in the Gulf during the dry season, thus limiting light penetration and hence phytoplankton growth. This study did not allow for a detailed assessment of the seasonal effect on optical properties and concentrations in DH. Additional field observations during the wet season would be necessary to seasonally characterize this environment.

From a remote sensing point of view, algorithm developments should focus on the synergistic use of new geostationary satellites such as the recently launched (October 2014) Advanced Himawari Imager (AHI) onboard Himawari-8 (Japan Meteorological Agency, JMA) in combination with polar-orbiting sensors. Geostationary instruments offer far greater temporal imaging resolution—up to 10 min intervals using Himawari-8—and have therefore the potential to resolve the bio-optical variability due to semi-diurnal tidal cycles. The performance of a seasonally parameterized MODIS-Aqua algorithm at retrieving CDOM, NAP, and Chl concentrations in the VDG is evaluated in Schroeder et al. (2015). This regional algorithm is based on the aLMI approach of Brando et al. (2012) and the dataset presented in this study was used for its parametrization.

AUTHOR CONTRIBUTIONS

DB and TS conceived the idea of the study, collected the samples, designed, organized, and participated to the 2012 wet and 2013 dry season campaigns; DB designed and collected the samples of the 2012 dry season campaign, designed and wrote the manuscript, created all the plots, and performed all the numerical data analysis; TS provided **Figure 15**; LC helped organize the logistic for all the field sampling, expertly analyzed the datasets of spectrophotometric absorptions and bio-geo concentrations in the laboratory, and provided comments on the manuscript; TS and VB provided valuable comments on all versions of the manuscript; DP contributed to the design of the 2013 dry season campaign, collected, and analyzed (a) the nutrient samples and (b) the phytoplankton cell counts, and contributed to the interpretation of the results related to phytoplankton and nutrient budgets; PF analyzed the EEM dataset and provided very valuable comments on the manuscript; DW provided the tidal heights modeled for each station and contributed to some of the interpretation of the results from his expert knowledge of the region; DD provided guidance on the IOP analysis and valuable comments to earlier versions of this manuscript; JA, NT, and MT participated, and helped organize the logistic of at least one of the three field campaigns, and expertly deployed the complex optical instrumentation necessary to collect the dataset used in this paper. All co-authors have approved the manuscript to be published.

FUNDING

This research was funded by the North Australia Marine Research Alliance (NAMRA) and the Australian Government's National Environmental Research Program (NERP). VB was supported by the European Union (FP7-People Co-funding of Regional, National, and International Programmes, GA no. 600407) and the CNR RITMARE Flagship Project. In-kind support was provided by the CSIRO and the Australian Institute of Marine Science (AIMS).

ACKNOWLEDGMENTS

The authors would like to thank Dr. Arnold G. Dekker (CSIRO), for valuable comments on the manuscript prior to its submission, Dr. A. Bricaud (LOV, Villefranche-sur-mer) for assistance with the selection of phytoplankton spectra suitable for this analysis, Dr. Ian Leiper (Charles Darwin University) for help with ArcGIS and some of the field sampling in Darwin Harbour, and Ms. Julie Haines (CSIRO Adelaide, Waite Campus) for the analysis of the DOC and POC samples. This research was funded by the North Australia Marine Research Alliance (NAMRA) and the Australian Government's National Environmental Research Program (NERP). VB was supported by the European Union (FP7-People Co-funding of Regional,

National, and International Programmes, GA no. 600407) and the CNR RITMARE Flagship Project. In-kind support was provided by the CSIRO and the Australian Institute of Marine Science (AIMS). We also would like to thank Dr. Edward Butler (AIMS, Darwin Office) for facilitating our participation to the dry season field campaign aboard RV Solander. We thank the captain (Mr. Christopher Davis), cruise leader (Mr. Marcus Stowar) and the entire crew of the RV Solander (AIMS), as well as Mr. Steve Compain (Arafura Bluewaters Charters), for providing great working conditions and route flexibility during our field sampling. We also wish to thank Dr. Edward Butler (AIMS), Dr. Simon Townsend (AHU) and the technical services at CSIRO, Charles Darwin University, AIMS (Darwin office) and the Aquatic Health Unit (Northern Territory Government) for making their laboratory facilities available to us during this research: in particular Mrs. Rebecca Edwards and Mrs. Heidi Franklin (CSIRO), Mrs. Kirsty McAllister (AIMS), Mrs. Ellie Hayward, Mr. Matthew Gray (Fieldwork Support, CDU), and Mrs. Julia Fortune, Mr. Matthew Majid (Fieldwork Support, AHU). A 1 month visit to the Laboratoire d'Océanographie de Villefranche (LOV, CNRS/UPMC) for collaborative work between the primary author and DD has helped us progress in the data processing required for this study. Travels from Australia to France were supported by a travel grant for early career researchers (Scientific Mobility Program 2013; French Embassy, Canberra, Australia), a proposal written by both DB and DD. We acknowledge the use of Rapid Response imagery from the Land, Atmosphere Near real-time Capability for EOS (LANCE) system operated by the NASA/GSFC/Earth Science Data and Information System (ESDIS) with funding provided by NASA/HQ. Finally, the authors are grateful to the two external reviewers, and the Specialty Chief Editor of Frontiers in Marine Science for providing constructive comments to the manuscript.

REFERENCES

Acevedo-Trejos, E., Brandt, G., Steinacher, M., and Merico, A. (2014). A glimpse into the future composition of marine phytoplankton communities. *Front. Mar. Sci.* 1:15. doi: 10.3389/fmars.2014.00015

Acker, J., Lyon, P., Hoge, F., Shen, S., Roffer, M., and Gawlikowski, G. (2009). Interaction of hurricane katrina with optically complex water in the gulf of mexico: interpretation using satellite-derived inherent optical properties and chlorophyll concentration. *IEEE Geosci. Remote Sens. Lett.* 6, 209–213. doi: 10.1109/LGRS.2008.2007658

Agustí, S., and Duarte, C. M. (2013). Phytoplankton lysis predicts dissolved organic carbon release in marine plankton communities. *Biogeosciences* 10, 1259–1264. doi: 10.5194/bg-10-1259-2013

Andutta, F. P., Wang, X. H., Li, L., and Williams, D. (2014). "Hydrodynamics and sediment transport in a macro-tidal estuary: Darwin harbour, Australia," in *Estuaries of Australia in 2050 and Beyond*, ed E. Wolanski (Dordrecht: Springer), 111–129.

Aurin, D. A., and Dierssen, H. M. (2012). Advantages and limitations of ocean color remote sensing in CDOM-dominated, mineral-rich coastal and estuarine waters. *Remote Sens. Environ.* 125, 181–197. doi: 10.1016/j.rse.2012.07.001

Babin, M., Morel, A., Fournier-Sicre, V., Fell, F., and Stramski, D. (2003a). Light scattering properties of marine particles in coastal and oceanic waters as related to the particle mass concentration. *Limnol. Oceanogr.* 48, 843–859. doi: 10.4319/lo.2003.48.2.0843

Babin, M., Stramski, D., Ferrari, G. M., Claustre, H., Bricaud, A., Obolensky, G., et al. (2003b). Variations in the light absorption coefficients of phytoplankton, nonalgal particles, and dissolved organic matter in coastal waters around Europe. *J. Geophys. Res.* 108, 3211. doi: 10.1029/2001JC000882

Bauer, J. E., Cai, W.-J., Raymond, P. A., Bianchi, T. S., Hopkinson, C. S., and Regnier, P. A. G. (2013). The changing carbon cycle of the coastal ocean. *Nature* 504, 61–70. doi: 10.1038/nature12857

Bauer, J. E., Druffel, E. R. M., Wolgast, D. M., and Griffin, S. (2002). Temporal and regional variability in sources and cycling of DOC and POC in the northwest Atlantic continental shelf and slope. *Deep Sea Res. Part II Top. Stud. Oceanogr.* 49, 4387–4419. doi: 10.1016/S0967-0645(02)00123-6

Blondeau-Patissier, D., Brando, V. E., Oubelkheir, K., Dekker, A. G., Clementson, L. A., and Daniel, P. (2009). Bio-optical variability of the absorption and scattering properties of the Queensland inshore and reef waters, Australia. *J. Geophys. Res.* 114, C05003. doi: 10.1029/2008JC005039

Blondeau-Patissier, D., Dekker, A. G., Schroeder, T., and Brando, V. E. (2011). *Thematic Findings: Surface Phytoplankton Blooms and Phytoplankton Biomass in Coastal Waters*. State of the Environment 2011 Committee, Australia State of the Environment 2011. Independent Report to the Australian Government Minister for Sustainability, Environment, Water, Population and Communities, DSEWPaC, Canberra, ACT, 410–411.

Blondeau-Patissier, D., Schroeder, T., Brando, V. E., Maier, S. W., Dekker, A. G., and Phinn, S. (2014). ESA-MERIS 10-year mission reveals contrasting phytoplankton bloom dynamics in two tropical regions of Northern Australia. *Remote Sens.* 6, 2963–2988. doi: 10.3390/rs6042963

Blough, N. V., and Del Vecchio, R. (2002). "Chromophoric DOM in the coastal environment," in *Biogeochemistry of Marine Dissolved Organic Matter*, *1st Edn.*, eds D. A. Hansell and C. A. Carlson (San Diego, CA: Academic Press), 509–546.

Boss, E., Pegau, W. S., Lee, M., Twardowski, M., Shybanov, E., Korotaev, G., et al. (2004). Particulate backscattering ratio at LEO 15 and its use to study particle composition and distribution. *J. Geophys. Res.* 109:C01014. doi: 10.1029/2002JC001514

Bowen, B. W., Rocha, L. A., Toonen, R. J., and Karl, S. A. (2013). The origins of tropical marine biodiversity. *Trends Ecol. Evol.* 28, 359–366. doi: 10.1016/j.tree.2013.01.018

Brando, V. E., Dekker, A. G., Park, Y.-J., and Schroeder, T. (2012). An adaptive semi-analytical inversion of ocean colour radiometry in optically complex waters. *Appl. Opt.* 51, 2808–2833. doi: 10.1364/AO.51.002808

Bricaud, A., Babin, M., Claustre, H., Ras, J., and Tieche, F. (2010). Light absorption properties and absorption budget of Southeast Pacific waters. *J. Geophys. Res.* 115:19. doi: 10.1029/2009JC005517

Bricaud, A., Babin, M., Morel, A., and Claustre, H. (1995). Variability in the chlorophyll-specific absorption coefficients of natural phytoplankton: analysis and parameterization. *J. Geophys. Res.* 100, 13321–13332.

Bricaud, A., Claustre, H., Ras, J., and Oubelkheir, K. (2004). Natural variability of phytoplanktonic absorption in oceanic waters: influence of the size structure of algal populations. *J. Geophys. Res.* 109, C11010. doi: 10.1029/2004JC002419

Burford, M. A., Alongi, D. M., McKinnon, A. D., and Trott, L. A. (2008). Primary production and nutrients in a tropical macrotidal estuary, Darwin Harbour, Australia. *Estuar. Coast. Shelf Sci.* 79, 440–448. doi: 10.1016/j.ecss.2008.04.018

Burford, M. A., Revill, A. T., Palmer, D. W., Clementson, L., Robson, B. J., and Webster, I. T. (2011). River regulation alters drivers of primary productivity along a tropical river-estuary system. *Mar. Freshw. Res.* 62, 141–151. doi: 10.1071/MF10224

Burford, M. A., Revill, A. T., Smith, J., and Clementson, L. (2012a). Effect of sewage nutrients on algal production, biomass and pigments in tropical tidal creeks. *Mar. Pollut. Bull.* 64, 2671–2680. doi: 10.1016/j.marpolbul.2012.10.008

Burford, M. A., and Rothlisberg, P. C. (1999). Factors Limiting phytoplankton production in a tropical continental shelf ecosystem. *Estuar. Coast. Shelf Sci.* 48, 541–549. doi: 10.1006/ecss.1999.0471

Burford, M. A., Rothlisberg, P. C., and Wang, Y. (1995). Spatial and temporal distribution of tropical phytoplankton species and biomass in

the Gulf of Carpentaria, Australia. *Mar. Ecol. Prog. Ser.* 118, 255–266. doi: 10.3354/meps118255

Burford, M. A., Webster, I. T., Revill, A. T., Kenyon, R. A., Whittle, M., and Curwen, G. (2012c). Controls on phytoplankton productivity in a wet–dry tropical estuary. *Estuar. Coast. Shelf Sci.* 113, 141–151. doi: 10.1016/j.ecss.2012.07.017

Cetinić, I., Perry, M. J., Briggs, N. T., Kallin, E., D'Asaro, E. A., and Lee, C. M. (2012). Particulate organic carbon and inherent optical properties during 2008 North Atlantic Bloom Experiment. *J. Geophys. Res.* 117:C06028. doi: 10.1029/2011JC007771

Claustre, H., Morel, A., Babin, M., Cailliau, C., Marie, D., Marty, J. C., et al. (1999). Variability in particle attenuation and chlorophyll fluorescence in the Tropical Pacific: scales, patterns, and biogeochemical implications. *J. Geophys. Res.* 104, 3401–3422. doi: 10.1029/98JC01334

Clementson, L. A. (2013). "The CSIRO method," in *The Fifth SeaWiFS HPLC Analysis Round-Robin Experiment (SeaHARRE-5),* eds S. B. Hooker, L. Clementson, C. S. Thomas, L. Schlüter, M. Allerup, J. Ras, H. Claustre, C. Normandeau, J. Cullen, M. Kienast, W. Kozlowski, M. Vernet, S. Chakraborty, S. Lohrenz, M. Tuel, D. Redalje, P. Cartaxana, C. R. Mendes, V. Brotas, S. G. Prabhu Matondkar, S. G. Parab, A. Neeley, and E. Skarstad Egeland (Greenbelt, MD: Goddard Space Flight Center; NASA).

Clementson, L. A., Parslow, J. S., Turnbull, A. R., and Bonham, P. I. (2004). Properties of light absorption in a highly coloured estuarine system in south-east Australia which is prone to blooms of the toxic dinoflagellate Gymnodinium catenatum. *Estuar. Coast. Shelf Sci.* 60, 101–112. doi: 10.1016/j.ecss.2003.11.022

Condie, S. A. (2011). Modeling seasonal circulation, upwelling and tidal mixing in the Arafura and Timor Seas. *Cont. Shelf Res.* 31, 1427–1436. doi: 10.1016/j.csr.2011.06.005

Condie, S. A., and Dunn, J. R. (2006). Seasonal characteristics of the surface mixed layer in the Australasian region: implications for primary production regimes and biogeography. *Mar. Freshw. Res.* 57, 569–590. doi: 10.1071/MF06009

CSIRO (2009). *Water in the Van Diemen region of the Timor Sea Drainage.* A report to the Australian Government from the CSIRO Northern Australia Sustainable Yields Project, CSIRO Water for a Healthy Country Flagship, 363–452.

DHAC (2010). *Darwin Harbour Advisory Committee: Ecosystem Research Group*. Palmerston, NT: Darwin Harbour Advisory Committee.

D'Sa, E. J., Miller, R. L., and McKee, B. A. (2007). Suspended particulate matter dynamics in coastal waters from ocean color: application to the northern Gulf of Mexico. *Geophys. Res. Lett.* 34:L23611. doi: 10.1029/2007GL031192

Fabiano, M., Povero, P., Donavero, R., and Misic, C. (1999). Particulate organic matter composition in a semienclosed Periantarctic system: the Straits of Magellan. *Sci. Mar.* 63, 89–98. doi: 10.3989/scimar.1999.63s189

Fichot, C. G., and Benner, R. (2011). A novel method to estimate DOC concentrations from CDOM absorption coefficients in coastal waters. *Geophys. Res. Lett.* 38:L03610. doi: 10.1029/2010GL046152

Fichot, C. G., and Benner, R. (2012). The spectral slope coefficient of chromophoric dissolved organic matter (S275–295) as a tracer of terrigenous dissolved organic carbon in river-influenced ocean margins. *Limnol. Oceanogr.* 57, 1453–1466. doi: 10.4319/lo.2012.57.5.1453

Fichot, C. G., and Benner, R. (2014). The fate of terrigenous dissolved organic carbon in a river-influenced ocean margin. *Glob. Biogeochem. Cycles* 28:2013GB004670. doi: 10.1002/2013GB004670

Fichot, C. G., Lohrenz, S. E., and Benner, R. (2014). Pulsed, cross-shelf export of terrigenous dissolved organic carbon to the Gulf of Mexico. *J. Geophys. Res.* 119, 1176–1194. doi: 10.1002/2013jc009424

Fisher, T., Hagy, J., and Rochelle-Newall, E. (1998). Dissolved and particulate organic carbon in Chesapeake Bay. *Estuaries* 21, 215–229. doi: 10.2307/1352470

Gardner, W. D., Mishonov, A. V., and Richardson, M. J. (2006). Global POC concentrations from *in-situ* and satellite data. *Deep Sea Res. Part II Top. Stud. Oceanogr.* 53, 718–740. doi: 10.1016/j.dsr2.2006.01.029

Hallegraeff, G. M., Bolch, C. J. S., Hill, D. R. A., Jameson, I., LeRoi, J. M., McMinn, A., et al. (2010). *Algae of Australia: Phytoplankton of Temperate Coastal Waters.* Canberra, ACT: CSIRO; Australian Biological Resources Study.

Hallegraeff, G. M., and Jeffrey, S. W. (1984). Tropical phytoplankton species and pigments in continental shelf waters of north and north-west Australia. *Mar. Ecol. Prog. Ser.* 20, 59–74. doi: 10.3354/meps020059

Halpern, B. S., Walbridge, S., Selkoe, K. A., Kappel, C. V., Micheli, F., D'Agrosa, C., et al. (2008). A global map of human impact on marine ecosystems. *Science* 319, 948–952. doi: 10.1126/science.1149345

Hansell, D. A., and Carlson, C. A. (2001). Marine dissolved organic matter and the carbon cycle. *Oceanography* 14, 41–49. doi: 10.5670/oceanog.2001.05

Hasle, G. (1978). "The inverted microscope method," in *Phytoplankton Manual,* ed A. Sournia (Paris: UNESCO), 88–96.

Helms, J. R., Stubbins, A., Ritchie, J. D., Minor, E. C., Kieber, D. J., and Mopper, K. (2008). Absorption spectral slopes and slope ratios as indicators of molecular weight, source, and photobleaching of chromophoric dissolved organic matter. *Limnol. Oceanogr.* 53, 955–969. doi: 10.4319/lo.2008.53.3.0955

Hestir, E. L., Brando, V., Campbell, G., Dekker, A., and Malthus, T. (2015). The relationship between dissolved organic matter absorption and dissolved organic carbon in reservoirs along a temperate to tropical gradient. *Remote Sens. Environ.* 156, 395–402. doi: 10.1016/j.rse.2014.09.022

Hommersom, A., Peters, S., Wernand, M. R., and de Boer, J. (2009). Spatial and temporal variability in bio-optical properties of the Wadden Sea. *Estuar. Coast. Shelf Sci.* 83, 360–370. doi: 10.1016/j.ecss.2009.03.042

Ilahude, A. G., and Mardanis, K. (1990). On the hydrology and productivity of the northern Arafura Sea. *Netherland J. Sea Res.* 25, 573–583. doi: 10.1016/0077-7579(90)90079-V

IOCCG (2000). "Remote Sensing of Ocean Colour in Coastal, and Other Optically-Complex, Waters," in *Reports of the International Ocean-Colour Coordinating Group*, ed S. Sathyendranath (Dartmouth, NS: Bedford Institute).

Jennerjahn, T. C. (2012). Biogeochemical response of tropical coastal systems to present and past environmental change. *Earth Sci. Rev.* 114, 19–41. doi: 10.1016/j.earscirev.2012.04.005

Kishino, M., Takahashi, M., Okami, N., and Ichimura, S. (1985). Estimation of the spectral absorption coefficients of phytoplankton in the sea. *Bull. Mar. Sci.* 37, 634–642.

Le, C., Lehrter, J. C., Hu, C., Schaeffer, B., MacIntyre, H., Hagy, J. D., et al. (2015). Relation between inherent optical properties and land use and land cover across Gulf Coast estuaries. *Limnol. Oceanogr.* 60, 920–933. doi: 10.1002/lno.10065

Lee, Z. P., Arnone, R., Hu, C. M., Werdell, P. J., and Lubac, B. (2010). Uncertainties of optical parameters and their propagations in an analytical ocean color inversion algorithm. *Appl. Opt.* 49, 369–381. doi: 10.1364/AO.49.000369

Leymarie, E., Doxaran, D., and Babin, M. (2010). Uncertainties associated to measurements of inherent optical properties in natural waters. *Appl. Opt.* 49, 5415–5436. doi: 10.1364/AO.49.005415

Li, L., Wang, X. H., Andutta, F., and Williams, D. (2014). Effects of mangroves and tidal flats on suspended-sediment dynamics: observational and numerical study of Darwin Harbour, Australia. *J. Geophys. Res.* 119, 5854–5873. doi: 10.1002/2014jc009987

Li, L., Wang, X. H., Williams, D., Sidhu, H., and Song, D. (2012). Numerical study of the effects of mangrove areas and tidal flats on tides: a case study of Darwin Harbour, Australia. *J. Geophys. Res.* 117, C06011–C06012. doi: 10.1029/2011jc007494

Loisel, H., Meriaux, X., Berthon, J.-F., and Poteau, A. (2007). Investigation of the optical backscattering to scattering ratio of marine particles in relation to their biogeochemical composition in the eastern English Channel and southern North Sea. *Limnol. Oceanogr.* 52, 739–752. doi: 10.4319/lo.2007.52.2.0739

Loisel, H., Mériaux, X., Poteau, A., Artigas, L. F., Lubac, B., Gardel, A., et al. (2009). Analyze of the inherent optical properties of French Guiana coastal waters for remote sensing applications. *J. Coast. Res.* 56, 1532–1536.

Lyon, P. E. (2010). "Inherent Optical properties in the wakes of Tropical Cyclones/Storms/Depressions: 1997 through 2006," in *The 2010 AGU Ocean Sciences Meeting AGU* (Washington, DC).

MacIejewska, A., and Pempkowiak, J. (2014a). DOC and POC in the southern Baltic Sea. Part II – evaluation of factors affecting organic matter concentrations using multivariate statistical methods. *Oceanologia.* 57, 168–176. doi: 10.5697/oc.56-3.523

MacIejewska, A., and Pempkowiak, J. (2014b). DOC and POC in the water column of the southern Baltic. Part I. Evaluation of factors influencing sources, distribution and concentration dynamics of organic matter. *Oceanologia* 56, 523–548. doi: 10.5697/oc.56-3.523

Malerba, M. E., Connolly, S. R., and Heimann, K. (2015). An experimentally validated nitrate–ammonium–phytoplankton model including effects of

starvation length and ammonium inhibition on nitrate uptake. *Ecol. Modell.* 317, 30–40. doi: 10.1016/j.ecolmodel.2015.08.024

Martinez-Vicente, V., Tilstone, G. H., Blondeau-Patissier, D., and Menezes, A. (2003). "Validation of MERIS geophysical products in open ocean and coastal waters of the southern North Sea and western English Channel," in *MAVT: ESA* (Frascati), 11.

Matsuoka, A., Hill, V., Huot, Y., Babin, M., and Bricaud, A. (2011). Seasonal variability in the light absorption properties of western Arctic waters: parameterization of the individual components of absorption for ocean color applications. *J. Geophys. Res.* 116, C02007. doi: 10.1029/2009JC005594

McKee, D., Chami, M., Brown, I., Calzado, V. S., Doxaran, D., and Cunningham, A. (2009). Role of measurement uncertainties in observed variability in the spectral backscattering ratio: a case study in mineral-rich coastal waters. *Appl. Opt.* 48, 4663–4675. doi: 10.1364/AO.48.004663

McKinna, L. I. W. (2014). Three decades of ocean-color remote-sensing *Trichodesmium* spp. in the World's oceans: a *review*. *Progr. Oceanogr.* 131, 177–199. doi: 10.1016/j.pocean.2014.12.013

Mitchell, B. G. (1990). "Algorithms for determining the absorption coefficient of aquatic particulates using the quantitative filter technique (QFT)," in *Proceedings of Ocean Optics X, SPIE* (Orlando, FL), 137–148.

Morel, A. (1988). Optical modeling of the upper ocean in relation to its biogenous matter content (Case 1 waters). *J. Geophys. Res.* 93, 10749–10768. doi: 10.1029/JC093iC09p10749

Morrison, R. J., and Delaney, J. R. (1996). Marine pollution in the Arafura and Timor Seas. *Mar. Pollut. Bull.* 32, 327–334. doi: 10.1016/0025-326X(96)00004-5

Nelson, N. B., and Siegel, D. A. (2013). The Global distribution and dynamics of chromophoric dissolved organic matter. *Annu. Rev. Mar. Sci.* 5, 447–476. doi: 10.1146/annurev-marine-120710-100751

Neukermans, G., Ruddick, K., Loisel, H., and Roose, P. (2012). Optimization and quality control of suspended particulate matter concentration measurement using turbidity measurements. *Limnol. Oceanogr.* 10, 1011–1023. doi: 10.4319/lom.2012.10.1011

Oubelkheir, K., Clementson, L., Webster, I., Ford, P., Dekker, A. G., Radke, L., et al. (2006). Using inherent optical properties to investigate biogeochemical dynamics in a tropical macrotidal coastal system. *J. Geophys. Res.* 111, C07021. doi: 10.1029/2005JC003113

Oubelkheir, K., Ford, P. W., Clementson, L. A., Cherukuru, N., Fry, G., and Steven, A. D. L. (2014). Impact of an extreme flood event on optical and biogeochemical properties in a sub-tropical coastal peri-urban embayment (Eastern Australia). *J. Geophys. Res.* 119, 6024–6045. doi: 10.1002/2014JC010205

Pan, D., Liu, Q., and Bai, Y. (2014). Review and suggestions for estimating particulate organic carbon and dissolved organic carbon inventories in the ocean using remote sensing data. *Acta Oceanol. Sin.* 33, 1–10. doi: 10.1007/s13131-014-0419-4

Pegau, W. S., Gray, D., and Zaneveld, J. R. V. (1997). Absorption and attenuation of visible and near-infrared light in water: dependence on temperature and salinity. *Appl. Opt.* 36, 6035–6046. doi: 10.1364/AO.36.006035

Pitarch, J., Volpe, G., Colella, S., Santoleri, R., and Brando, V. (2016). Absorption correction and phase function shape effects on the closure of apparent optical properties. *Appl. Opt.* 55, 8618–8636. doi: 10.1364/AO.55.008618

Platt, T., and Sathyendranath, S. (2008). Ecological indicators for the pelagic zone of the ocean from remote sensing. *Remote Sens. Environ.* 112, 3426–3436. doi: 10.1016/j.rse.2007.10.016

Pusey, B. J., Kennard, M. J., Larson, H. K., Alsop, Q., Hammer, M., and Buckle, D. J. (2015). Estuarine fishes of the South Alligator River, Kakadu National Park, northern Australia. *Mar. Freshw. Res.* 67, 1797–1812. doi: 10.1071/MF15221

Romero-Ibarra, N., and Silverberg, N. (2011). The contribution of various types of settling particles to the flux of organic carbon in the Gulf of St. Lawrence. *Cont. Shelf Res.* 31, 1761–1776. doi: 10.1016/j.csr.2011.08.006

Röttgers, R., McKee, D., and Woźniak, S. B. (2013). Evaluation of scatter corrections for ac-9 absorption measurements in coastal waters. *Methods Oceanogr.* 7, 21–39. doi: 10.1016/j.mio.2013.11.001

Roy, S., Sathyendranath, S., Bouman, H., and Platt, T. (2013). The global distribution of phytoplankton size spectrum and size classes from their light-absorption spectra derived from satellite data. *Remote Sens. Environ.* 139, 185–197. doi: 10.1016/j.rse.2013.08.004

Ryle, V. D., Mueller, H. R., and Gentien, P. (1981). *Automated Analysis of Nutrients in Tropical Sea Waters*. Technical bulletin Oceanography series No. 3. AIMS Oceanography Series Tech. Bulletin, Towsnville, QLD.

Sathyendranath, S., Stuart, V., Nair, A., Oka, K., Nakane, T., Bouman, H., et al. (2009). Carbon-to-chlorophyll ratio and growth rate of phytoplankton in the sea. *Mar. Ecol. Prog. Ser.* 383, 73–84. doi: 10.3354/meps07998

Schroeder, T., Brando, V., Blondeau-Patissier, D., Clementson, L. A., Anstee, J. M., and Thapar, N. (2015). *Remote Sensing Methods to Map and Monitor the Condition of Coastal Habitats and Other Surrogates for Biodiversity, Part B: Water Quality Mapping of the Van Diemen Gulf, Image Processing and Data Analysis*. Brisbane, QLD: CSIRO Oceans & Atmosphere Flagship.

Schroeder, T., Dekker, A., and Rathbone, C. E. (2009). *Remote Sensing for Light Attenuation Mapping in the North Marine Region*. CSIRO Wealth from Oceans Flagship Report to the Department of the Environment, Water, Heritage and the Arts, CSIRO Land and Water, Canberra, ACT.

Schroeder, T., Devlin, M. J., Brando, V. E., Dekker, A. G., Brodie, J. E., Clementson, L. A., et al. (2012). Inter-annual variability of wet season freshwater plume extent into the Great Barrier Reef lagoon based on satellite coastal ocean colour observations. *Mar. Pollut. Bull.* 65, 210–223. doi: 10.1016/j.marpolbul.2012.02.022

Shanmugam, P. (2011). New models for retrieving and partitioning the colored dissolved organic matter in the global ocean: implications for remote sensing. *Remote Sens. Environ.* 115, 1501–1521. doi: 10.1016/j.rse.2011.02.009

Shi, W., Wang, M., and Jiang, L. (2011). Spring-neap tidal effects on satellite ocean color observations in the Bohai Sea, Yellow Sea, and East China Sea. *J. Geophys. Res.* 116:C12032. doi: 10.1029/2011JC007234

Signorini, S. R., Franz, B. A., and McClain, C. R. (2015). Chlorophyll variability in the oligotrophic gyres: mechanisms, seasonality and trends. *Front. Mar. Sci.* 2:1. doi: 10.3389/fmars.2015.00001

Snyder, W. A. R., Arnone, A., Davis, C. O., Goode, W., Gould, R. W., Ladner, S., et al. (2008). Optical scattering and backscattering by organic and inorganic particulates in U.S. coastal waters. *Appl. Opt.* 47, 666–677. doi: 10.1364/AO.47.000666

Story, R., Galloway, R., McAlpine, J., Aldrick, J., and Williams, M. (1969). *Lands of the Alligator Rivers Area, Northern Territory*. Land Research Series. Melbourne, VIC: CSIRO.

Swirgon, M., and Stramska, M. (2015). Comparison of *in situ* and satellite ocean color determinations of particulate organic carbon concentration in the global ocean. *Oceanologia* 57, 25–31. doi: 10.1016/j.oceano.2014.09.002

Tang, D. (2011). *Remote Sensing of the Changing Oceans*. Berlin: Springer.

Tilstone, G., Peters, S., van der Woerd, H., Eleveld, M., Ruddick, K., Schoenfeld, W., et al. (2012). Variability in specific-absorption properties and their use in a semi-analytical Ocean Colour Algorithm for MERIS in North Sea and Western English Channel coastal waters. *Remote Sens. Environ.* 118, 320–338. doi: 10.1016/j.rse.2011.11.019

Twardowski, M. S., Boss, E., MacDonald, J. B., Pegau, W. S., Barnard, A. H., and Zaneveld, J. R. V. (2001). A model for estimating bulk refractive index from the optical backscattering ratio and the implications for understanding particle composition in case I and case II waters. *J. Geophys. Res.* 106, 14129–14142. doi: 10.1029/2000JC000404

Uitz, J., Claustre, H., Morel, A., and Hooker, S. B. (2006). Vertical distribution of phytoplankton communities in open ocean: an assessment based on surface chlorophyll. *J. Geophys. Res.* 111:C08005. doi: 10.1029/2005JC003207

Vantrepotte, V., Danhiez, F.-P., Loisel, H., Ouillon, S., Mériaux, X., Cauvin, A., et al. (2015). CDOM-DOC relationship in contrasted coastal waters: implication for DOC retrieval from ocean color remote sensing observation. *Opt. Express* 23, 33–54. doi: 10.1364/OE.23.000033

Vidussi, F., Claustre, H., Manca, B. B., Luchetta, A., and Marty, J. C. (2001). Phytoplankton pigment distribution in relation to upper thermocline circulation in the eastern Mediterranean Sea during winter. *J. Geophys. Res.* 106, 19939–19956. doi: 10.1029/1999JC000308

Wang, X. J., Behrenfeld, M., Le Borgne, R., Murtugudde, R., and Boss, E. (2009). Regulation of phytoplankton carbon to chlorophyll ratio by light, nutrients and temperature in the Equatorial Pacific Ocean: a basin-scale model. *Biogeosciences* 6, 391–404. doi: 10.5194/bg-6-391-2009

Watson, J., and Zielinski, O. (2013). *Subsea Optics and Imaging*. Sawston: Woodhead Publishing.

Webster, I. T., and Ford, P. W. (2010). Delivery, deposition and redistribution of fine sediments within macrotidal Fitzroy Estuary/Keppel Bay: Southern Great Barrier Reef, Australia. *Cont. Shelf Res.* 30, 793–805. doi: 10.1016/j.csr.2010.01.017

Weishaar, J. L., Aiken, G. R., Bergamaschi, B. A., Fram, M. S., Fujii, R., and Mopper, K. (2003). Evaluation of specific ultraviolet absorbance as an indicator of the chemical composition and reactivity of dissolved organic carbon. *Environ. Sci. Technol.* 37, 4702–4708. doi: 10.1021/es030360x

Whitmire, A. L., Boss, E., Cowles, T. J., and Pegau, W. S. (2007). Spectral variability of the particulate backscattering ratio. *Opt. Express* 15, 7019–7031. doi: 10.1364/OE.15.007019

Williams, D. (2009). *Dredging of Sand from Darwin Harbour, Hydrographic and Marine Life. Part 1 Hydrodynamics and Sediment Transport*. Darwin, NT: Australian Institute of Marine Science.

Williams, D., Wolanski, E., and Spagnol, S. (2006). "Hydrodynamics of Darwin harbour," in *The Environment in Asia Pacific Harbours*, ed E. Wolanski (Dordrecht: Springer), 461–476.

Wolanski, E., McKinnon, D. A., Williams, D., and Alongi, D. M. (2006). "An estuarine ecohydrology model of Darwin Harbour, Australia,"in *The Environment in Asia Pacific Harbours*, ed E. Wolanski (Dordrecht: Springer), 477–488.

Woodroffe, C. D., Mulrennan, M. E., and Chappell, J. (1993). Estuarine infill and coastal progradation, southern van diemen gulf, northern Australia. *Sediment. Geol.* 83, 257–275. doi: 10.1016/0037-0738(93)90016-X

Yamashita, Y., Nosaka, Y., Suzuki, K., Ogawa, H., Takahashi, K., and Saito, H. (2013). Photobleaching as a factor controlling spectral characteristics of chromophoric dissolved organic matter in open ocean. *Biogeosciences* 10, 7207–7217. doi: 10.5194/bg-10-7207-2013

Zaneveld, J. R. V., Kitchen, J. C., and Moore, C. (1994). "The scattering error correction of reflecting-tube absorption meters," in *Proceedings of SPIE Ocean Optics XII,* ed S. Ackleson (Bergen), 44–55.

Zhu, Z. Y., Zhang, J., Wu, Y., and Lin, J. (2006). Bulk particulate organic carbon in the East China Sea: tidal influence and bottom transport. *Prog. Oceanogr.* 69, 37–60. doi: 10.1016/j.pocean.2006.02.014

Zubkov, M. V., Burkill, P. H., and Topping, J. N. (2007). Flow cytometric enumeration of DNA-stained oceanic planktonic protists. *J. Plankton Res.* 29, 79–86. doi: 10.1093/plankt/fbl059

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Blondeau-Patissier, Schroeder, Clementson, Brando, Purcell, Ford, Williams, Doxaran, Anstee, Thapar and Tovar-Valencia. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Marine Science

ORIGINAL RESEARCH
published: 16 May 2017
doi: 10.3389/fmars.2017.00137

Shellfish Aquaculture from Space: Potential of Sentinel2 to Monitor Tide-Driven Changes in Turbidity, Chlorophyll Concentration and Oyster Physiological Response at the Scale of an Oyster Farm

Pierre Gernez[1]*, *David Doxaran*[2] and *Laurent Barillé*[1]

[1] Mer Molécules Santé (MMS EA 2160), Université de Nantes, Nantes, France, [2] Laboratoire d'Océanographie de Villefranche (UMR 7093), Centre Nationnal de la Recherche Scientifique, UPMC, Villefranche sur mer, France

OPEN ACCESS

Edited by:
Tiit Kutser,
University of Tartu, Estonia

Reviewed by:
Emmanuel Devred,
Fisheries and Oceans Canada, Canada
Zhubin Zheng,
Gannan Normal University, China

***Correspondence:**
Pierre Gernez
pierre.gernez@univ-nantes.fr

Specialty section:
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

Received: 15 January 2017
Accepted: 25 April 2017
Published: 16 May 2017

Citation:
Gernez P, Doxaran D and Barillé L (2017) Shellfish Aquaculture from Space: Potential of Sentinel2 to Monitor Tide-Driven Changes in Turbidity, Chlorophyll Concentration and Oyster Physiological Response at the Scale of an Oyster Farm. Front. Mar. Sci. 4:137. doi: 10.3389/fmars.2017.00137

The algorithms of Novoa et al. (2017) and Gons et al. (2005) were recalibrated and applied to Sentinel2 data to retrieve suspended particulate matter (SPM) and chlorophyll *a* (chl *a*) concentration in the environmentally and economically important intertidal zones. Sentinel2-derived chl *a* and SPM concentration distributions were analyzed at the scale of an oyster farm over a variety of tidal conditions. Sentinel2 imagery was then coupled with ecophysiological modeling to analyze the influence of tide-driven chl *a* and SPM dynamics on oyster clearance and chl consumption rates. Within the studied oyster farming site (Bourgneuf Bay along the French Atlantic coast), chl consumption rate mirrored the changes in chl *a* concentration during neap tides, whereas oyster clearance and chl consumption rates were both negatively impacted by high SPM concentration during spring tides.

Keywords: Sentinel2, ocean color, chlorophyll, turbidity, oyster, aquaculture, microphytobenthos, mudflat

INTRODUCTION

One of the most striking features of the intertidal zone is the formation of microphytobenthos (MPB) biofilms at sediment surface during those low tides that occur in daylight (MacIntyre et al., 1996; Paterson et al., 1998; Jesus et al., 2009). In many mudflat MPB biofilms are visible from space, and they have been studied using airborne and satellite remote sensing (Méléder et al., 2003; van der Wal et al., 2010; Kazemipour et al., 2012; Brito et al., 2013). Although MPB main ecological functions are carried out when it is organized in the form of biofilms, benthic microalgae can also be resuspended into the water column together with other sedimentary particles throughout the tidal cycle (Koh et al., 2006; Ubertini et al., 2012). This can result in significant enrichment of nearshore waters with a high concentration of chlorophyll *a* (chl *a*) that becomes available food for suspension feeders such as the Pacific oyster *Crassostrea gigas* and other commercially and ecologically important bivalves (Kang et al., 2006; Choy et al., 2009). In coastal zones, despite the high contribution of tidal flats to primary production (Underwood and Kromkamp, 1999), the spatial distribution and temporal dynamic of chl *a* concentration in intertidal waters has been little studied using ocean color remote sensing so far.

In the optically complex and very diverse coastal zone, separating the contribution of chl *a* from other colored constituents [namely particulate inorganic matter (PIM), and colored dissolved organic matter (CDOM)] in the water column is notoriously difficult due to the rapidly changing concentrations of CDOM and PIM coming from sediment resuspension, river plume, and land runoff (Blondeau-Patissier et al., 2014). In turbid tidal flat and adjacent coastal areas, the main challenge arises from the difficulty to detect chl *a* from the high load of suspended particulate matter (SPM). In estuarine and nearshore waters, algorithms based on the analysis of the chl *a* absorption band in the near-infrared (NIR) spectral region around 675 nm were demonstrated to generally outperform other methods (Le et al., 2013).

Due to its spectral characteristics (namely the red and NIR spectral bands at 665 and 705 nm), we hypothesize that the Multi Spectral Imager (MSI) onboard Sentinel2 has the potential to quantify chl *a* concentration in turbid waters, provided that these waters are exposed to resuspension of benthic microalgae. Besides its relevant spectral characteristics, Sentinel2 also offers the advantage of high spatial resolution (20 m), making it possible to observe narrow bays and estuaries where shellfish farms are usually located. The first objective of the present study is therefore to analyze the potential of Sentinel2 for shellfish aquaculture monitoring, and more specifically to test the retrieval of SPM and chl *a* concentration in a turbid oyster farming ecosystem. The second objective is to analyze the tide-driven influence of SPM and chl *a* variability on oyster ecological response at the scale of an oyster farm. For that purpose, and building on previous studies (Gernez et al., 2014; Thomas et al., 2016), Earth Observation (EO) and shellfish physiological modeling were interconnected in order to remotely quantify the influence of rapidly changing environmental conditions on oyster clearance and chl consumption rates.

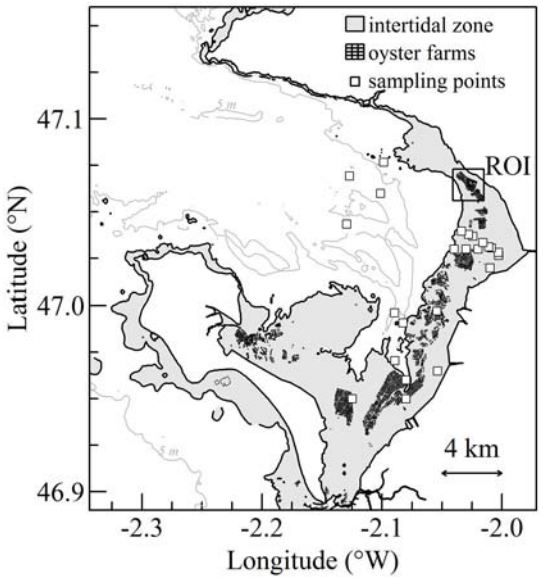


FIGURE 1 | Bourgneuf Bay on the French Atlantic coast. Intertidal zone, oyster farms, and sampling stations are shown in gray, black, and white symbols, respectively. The location of the rectangle corresponds to the oyster farm's region of interest (ROI).

MATERIALS AND METHODS

Study Site

Bourgneuf Bay is a macrotidal bay along the French Atlantic coast, mostly constituted of mudflats, and widely used for shellfish aquaculture (oyster annual yield was 5,330 tons in 2010, Dessinges et al., 2012). In the present study, a focus was made on a shellfish farming site located at the northern limit of the oyster aquaculture zone (**Figures 1**, **2**). Due to tidal resuspension, SPM concentration seldom decreases below 50 g m^{-3} and regularly exceeds 500 g m^{-3}. As a too high SPM concentration impacts oyster clearance rate and other physiological functions (Barillé et al., 1997), oysters grown in this farming site are negatively impacted by high SPM concentration (Gernez et al., 2014). Daily mean chl *a* concentration was reported to vary between 4 and 14 mg m^{-3} (Dutertre et al., 2009), and monthly means between 5 and 30 mg m^{-3} were previously reported at the study site (Barillé-Boyer et al., 1997).

In situ Data

Field data were acquired during two bio-optical cruises in Bourgneuf Bay from 08 to 12 April 2013 and from 12 to 13 April 2016 in the frame of the ANR GIGASSAT and FP7 HIGHROC projects, respectively. During both cruises, water sampling and radiometric measurements were performed following the same protocol. Sampling stations were located nearshore, mostly within the intertidal zone and in the vicinity of farming sites (**Figure 1**). Sampling took place at different times of the tidal cycle in order to acquire reflectance spectra over a wide range of SPM and chl *a* concentration. The same flat-bottomed barge was used during both cruises. This kind of vessel makes it possible to navigate throughout the shallow intertidal waters, even during low tide. Some stations were visited when the water depth was as low as 0.5 m. The bottom was never visible from above surface, even at the shallowest station due to the extremely high turbidity.

Radiometric Data

Above-water radiometric measurements were conducted following standard protocols (Mueller et al., 2000) to determine the spectral water-leaving radiance reflectance (also commonly referred as the marine reflectance), $\rho_w(\lambda)$, defined as:

$$\rho_w(\lambda) = \pi[L_u(\lambda) - \rho_{sky} L_{sky}(\lambda)]/E_d(\lambda) \quad (1)$$

where $L_u(\lambda)$ is the upwelling radiance from the water and air-sea interface measured at a zenith angle of about 37°, $L_{sky}(\lambda)$ is the sky radiance, ρ_{sky} is the air-water radiance reflection coefficient, $E_d(\lambda)$ is the above-water downwelling irradiance, and λ is the wavelength. The barge was oriented away from the sun to avoid shadowing effects. Radiance sensors were pointed at a solar azimuth angle between 90 and 135°. The radiometric data were acquired simultaneously during about 5 min of stable sky

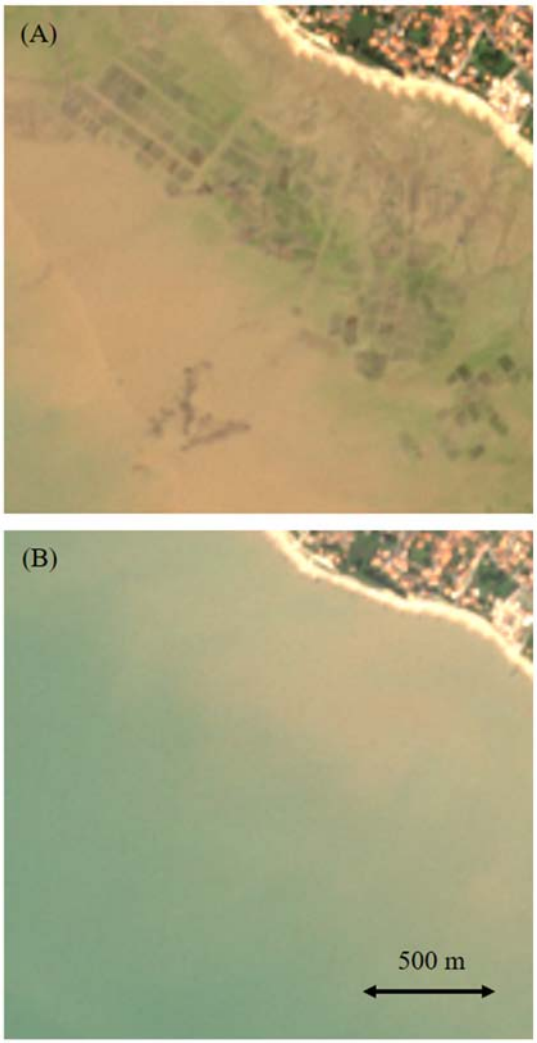


FIGURE 2 | Rayleigh-corrected Sentinel2 Red Blue Green (RGB) image of the oyster farming site during low tide the 30 September 2015 (A), and during high tide the 15 March 2016 **(B)**. The mudflat is emerged during low tide, and the oyster tables are visible from above. During high tide, oyster tables are not visible due to the extremely high turbidity.

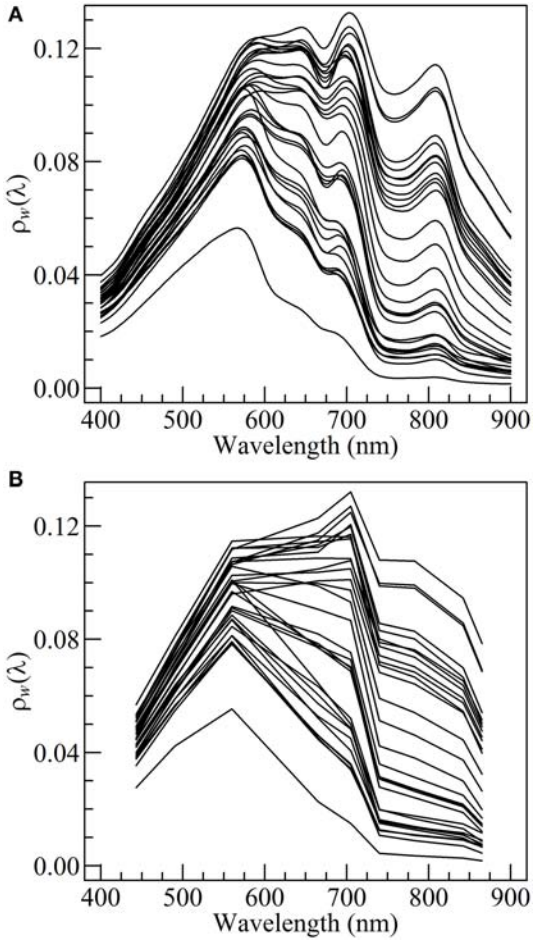


FIGURE 3 | *In situ* marine reflectance $\rho_w(\lambda)$ at TriOS (A), and Sentinel2 **(B)** spectral resolution.

conditions using three TriOS radiometers, two measuring the radiance signal and one measuring the downwelling irradiance.

A thorough quality control was made and only clear sky data were selected. Wave height was <0.5 m and wind speed was <5.0 m s^{-1} during both cruises. The ρ_{sky} coefficient was taken as 0.02 following Austin (1974). The TriOS data were averaged over the time span of the measurement, smoothed over a 10 nm moving-window, cut within 400 and 900 nm (**Figure 3A**), and then spectrally downgraded at the resolution of the MSI onboard Sentinel2 using the spectral response function provided by the European Space Agency (**Figure 3B**).

Seawater Samples

Seawater samples were collected just below the surface concomitantly with radiometric measurements. Seawater samples were stored in 1 l bottles until they were filtered in the laboratory in the evening. The turbidity, T (in Formazin Nephelometric Unit, FNU), of each water sample was determined in triplicate using a 2100Q portable turbidimeter (Hach Company, Loveland, CO, USA) in order to optimize the volume of filtered seawater as in Neukermans et al. (2012). SPM, defined as the dry mass of particles per unit volume of seawater, was then determined using a standard gravimetric technique. Measured volumes of seawater (between 10 and 200 ml depending of the turbidity of the sample) were filtered through 25 mm diameter preweighed Whatman GF/F glass-fiber filters. At the end of filtration, sample filters were rinsed with deionized water to remove sea salt. The filters were frozen and shipped at the Laboratoire d'Océanographie de Villefranche (LOV). The dry mass of particles collected on the filter was then measured with a MT5 microbalance (Mettler-Toledo Intl. Inc.) with a resolution of 0.001 mg. A significant relationship between SPM and T was obtained ($p < 0.01$).

Depending on the turbidity, between 10 and 300 ml of seawater was also filtered through 25 mm GF/F filters for high

performance liquid chromatography (HPLC) pigment analysis. The filters were frozen in liquid nitrogen and shipped for analysis at LOV, where HPLC analysis was performed according to Ras et al. (2008). The total chlorophyll *a* (chl *a*) concentration was computed as the sum of the "true" chlorophyll *a*, divinyl-chlorophyll *a*, and chlorophyllide *a*.

Bio-Optical Algorithms

Chlorophyll *a* Algorithm

Several algorithms are available for Sentinel2/MSI (Beck et al., 2016; Toming et al., 2016) to retrieve chl *a* concentration from $\rho_w(\lambda)$ in coastal waters. For our study site, an intercomparison exercise based on *in situ* measurements demonstrates that the chlorophyll-retrieval algorithm of Gons et al. (2005) provided the most satisfactory results (see Supplementary Information for more details). This algorithm was originally developed for the Medium Resolution Imaging Spectrometer (MERIS) using the bands at 665, 705, and 775 nm (Gons, 1999; Gons et al., 2002, 2005). It was applied here to Sentinel2/MSI using bands B4 (665 nm), B5 (705 nm), and B7 (783 nm). Due to the shift from 775 to 783 nm, a recalibration has been performed to update the algorithm to Sentinel2/MSI. The chlorophyll-retrieval is done in three steps. First, the backscattering coefficient (b_b) is estimated from ρ_w at 783 nm:

$$b_b(783) = 1.56\,\rho_w(783)/[0.082 - 0.6\,\rho_w(783)] \quad (2a)$$

Note that Equation (2a) is specific to Sentinel2/MSI, and replaces the original equation for MERIS:

$$b_b(775) = 1.61\,\rho_w(775)/[0.082 - 0.6\,\rho_w(775)] \quad (2b)$$

Second, the phytoplankton absorption at 665 nm is retrieved from a NIR/red band ratio:

$$a_{phy}(665) = (0.70 + b_b)\rho_w(705)/\rho_w(665) - 0.40 - b_b^p \quad (3)$$

where p is a unitless tuning parameter. Third, chl *a* concentration is computed by division with the chlorophyll-specific absorption coefficient at 665 nm, $a^*_{phy}(665)$:

$$[\text{chl } a] = a_{phy}(665)/a^*_{phy}(665) \quad (4)$$

It is assumed in Equation (3) that $b_b(\lambda)$ is spectrally neutral between 665 and 783 nm, and that at 665 nm the absorption by chlorophyll *a* and by pure seawater is much higher than the absorption by mineral particles and CDOM.

The parameters $a^*_{phy}(665)$ and p were initially estimated using a large dataset of field measurements from diverse inland, estuarine and coastal waters (Gons, 1999; Gons et al., 2002, 2005). For our study site in Bourgneuf Bay, $a^*_{phy}(665)$ was recalibrated to 0.133 m^2 (mg chl $a)^{-1}$ (standard error is 0.002 m^2 (mg chl $a)^{-1}$) and p to 1.02. The recalibration was done using a fitting procedure based on a root mean square error minimization (**Figure 4**).

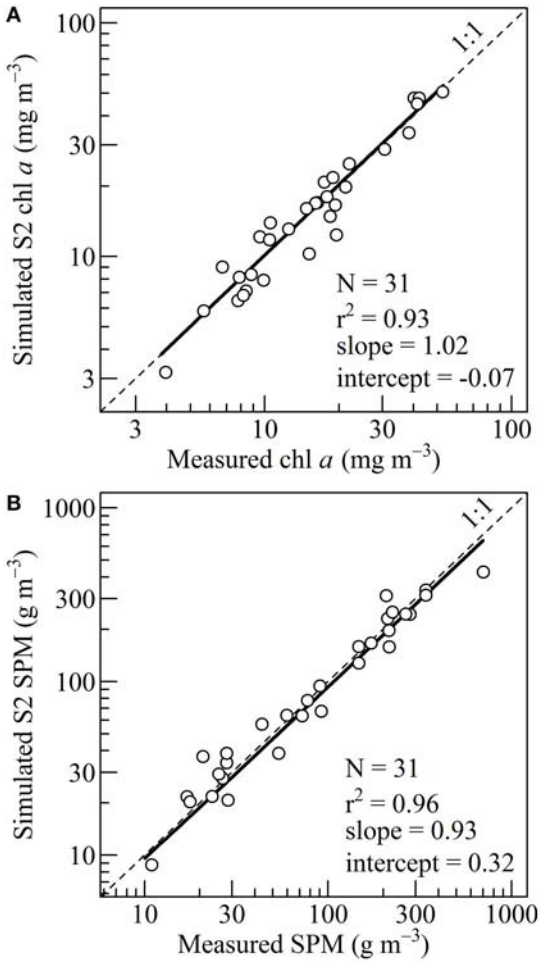


FIGURE 4 | (A) Linear regression between the measured and simulated chlorophyll *a* concentration, obtained from *in situ* measurements. The thick line shows the fit between 4 and 52.5 mg m^{-3}. The fit was used to calibrate the Gons et al. (2005) algorithm. **(B)** Linear regression between the measured and simulated suspended particulate matter concentration, obtained from *in situ* measurements. The thick line shows the fit between 10 and 700 g m^{-3}. The fit was used to calibrate the Novoa et al. (2017) algorithm.

SPM Algorithm

The SPM concentration was computed using a multi-conditional algorithm previously developed for Bourgneuf Bay and the Loire estuary (Novoa et al., 2017). This algorithm has been validated for the Operational Land Imager (OLI) onboard Landsat8 (Novoa et al., 2017). A spectral recalibration has been performed here so that the algorithm could be applied to Sentinel2/MSI using bands B4 (665 nm) and B8A (865 nm). The algorithm is based on a switching method that automatically selects the most relevant SPM vs. ρ_w relationship to avoid saturation effects at high turbidity. The final SPM concentration is computed as a dynamic combination of SPM retrievals in the red and NIR bands:

$$[\text{SPM}] = \alpha[\text{SPM}]_{\text{red}} + \beta[\text{SPM}]_{\text{NIR}} \quad (5)$$

where $[SPM]_{red}$, $[SPM]_{NIR}$, α, and β are defined as:

$$[SPM]_{red} = 297\rho_w(665) \,/\, [1 - \rho_w(665)/0.1238] \quad (6)$$

$$[SPM]_{NIR} = 4302\rho_w(865) \,/\, [1 - \rho_w(865)/0.2115] \quad (7)$$

$$\alpha = \log[0.090/\rho_w(665)] \,/\, \log(0.090/0.046) \quad (8)$$

$$\beta = \log[\rho_w(665)/0.046] \,/\, \log(0.090/0.046) \quad (9)$$

Due to the shift from 655 nm (Landsat8/OLI) to 665 nm (Sentinel2/MSI) the coefficients used in Equation 6 were recalibrated for Sentinel2/MSI using a fitting procedure based on a root mean square error minimization (**Figure 4**). The coefficients used in Equations (7–9) are the initial values computed by Novoa et al. (2017).

Satellite Data and Processing

Atmospheric Correction

Ortho-rectified, geo-located, and radiometrically calibrated top-of-atmosphere (TOA) reflectance Sentinel2 images were downloaded in the SAFE format from the US Geological Survey web portal (https://earthexplorer.usgs.gov). A single scene can contain multiple granules (sub-tiles), but the USGS web portal makes it possible to directly download Sentinel2 data at granule level, thus reducing downloading and processing time. Sentinel2 TOA data was processed using the ACOLITE software (http://odnature.naturalsciences.be/remsem/software-and-data/acolite) to derive the water-leaving radiance. This software proposes two options for the atmospheric correction (AC): (i) the NIR algorithm based on the assumption of spatial homogeneity of the red/NIR ratio for aerosol and marine reflectance (Ruddick et al., 2000; Vanhellemont and Ruddick, 2014) using Sentinel2 spectral bands at 665 and 865 nm, (ii) and the SWIR algorithm based on the assumption of zero water-leaving reflectance in the SWIR, using Sentinel2 spectral bands at 1,610 and 2,190 nm (Vanhellemont and Ruddick, 2015, 2016). ACOLITE establishes a per-tile aerosol type (or epsilon) as the ratio between the Rayleigh corrected reflectance in the two aerosol correction bands, for pixels where the marine reflectance can be assumed to be zero (i.e., where ρ_w(665 nm) <0.005, as defined by Vanhellemont and Ruddick, 2014). The epsilon is then used to extrapolate the observed aerosol reflectance to the NIR and visible bands. For the SWIR algorithm, ACOLITE also provides a choice for aerosol correction using a fixed epsilon over the region of interest (ROI), or a per pixel variable epsilon.

As the NIR AC option is not adapted to turbid waters (Vanhellemont and Ruddick, 2015), we used here the SWIR AC option with a fixed epsilon over the ROI, as recommended by several authors (Van der Zande et al., 2016; Novoa et al., 2017; Tristan Harmel, personal communication). The ROI was taken as Bourgneuf Bay and the Loire estuary (i.e., longitude from −2.35 to −1.95°E, and latitude from 46.85 to 47.35°N). The atmospheric correction is then performed in two steps: (i) a Rayleigh correction for scattering by air molecules using a look-up table generated using 6SV (Vermote et al., 2006), and (ii) an aerosol correction based on the assumption of black water reflectance in the SWIR bands due to the extremely high pure-water absorption, and an exponential spectrum for multiple scattering aerosol reflectance. Due to the low signal in the SWIR wavelengths, a spatial smoothing filtering for these bands was performed (Vanhellemont and Ruddick, 2016)

The final output of ACOLITE software is the $\rho_w(\lambda)$ data in Network Common Data Form (NetCDF). SPM and chl *a* concentrations were then computed from $\rho_w(\lambda)$ using the R project for statistical computing (R Development Core Team, 2008).

Selection, Clustering, and Sorting of Satellite Data

In intertidal waters, the spatio-temporal distribution of in-water suspended constituents is mainly driven by tidal dynamics. A total of 12 clear sky images was selected in order to observe the oyster farming site over a variety of seasonal, hydrological, and tidal conditions (**Table 1**). The time difference between satellite observation and low tide varied from <1 h to more than 5 h, thus providing a set of images acquired from low to high tide. The water height at the nearest reference harbor varied between 0.93 and 4.43 m, and the oyster farming site was observed over a variety of tidal configurations, from almost full emersion to complete submersion.

During the selected days of satellite acquisition the tidal range varied from 2.75 to 6 m, encompassing neap and spring tides (**Table 1**). The dataset was then divided in two subsets according to the tidal amplitude so that images were either clustered into neap tide (tidal amplitude <4 m) or into spring tide (tidal amplitude >4 m).

Irrespective of their acquisition date, Sentinel2 data were tidally sorted from ebb tide to flow tide according to the time difference between satellite observation and low tide, and to the water height at the time of acquisition. Sentinel2-derived SPM and chl *a* concentration maps were thus clustered in 2 composite tidal cycles, either representative of neap or spring tide. In order to investigate the influence of changes in SPM and

TABLE 1 | Sentinel2 data used in the present study.

Date	Time	Water height (m)	Time of low tide	Tidal range (m)	Tide type
20150729	11:06	3.65	08:03	3.81	Neap
20150801	11:16	0.93	10:22	5.49	Spring
20150821	11:16	3.30	13:50	2.97	Neap
20150910	11:16	3.88	14:13	3.48	Neap
20160315	11:01	4.43	14:58	3.31	Neap
20160318	11:15	4.31	06:06	2.75	Neap
20160407	11:12	1.28	09:43	6.00	Spring
20160723	11:07	1.46	12:10	4.59	Spring
20160815	11:08	4.06	07:35	3.05	Neap
20160822	11:05	1.71	12:37	4.98	Spring
20161021	11:03	3.41	13:45	4.08	Spring
20161130	11:04	1.53	10:24	4.11	Spring

*Tide information was taken from the service hydrographique et océanographique de la Marine (SHOM) web portal using Pornic (France) as reference harbor (http://maree.shom.fr/). All times are UT. Data acquired during neap and spring tides were used in **Figures 5**, **6**, **9**, and in **Figures 7**, **8**, **10**, respectively.*

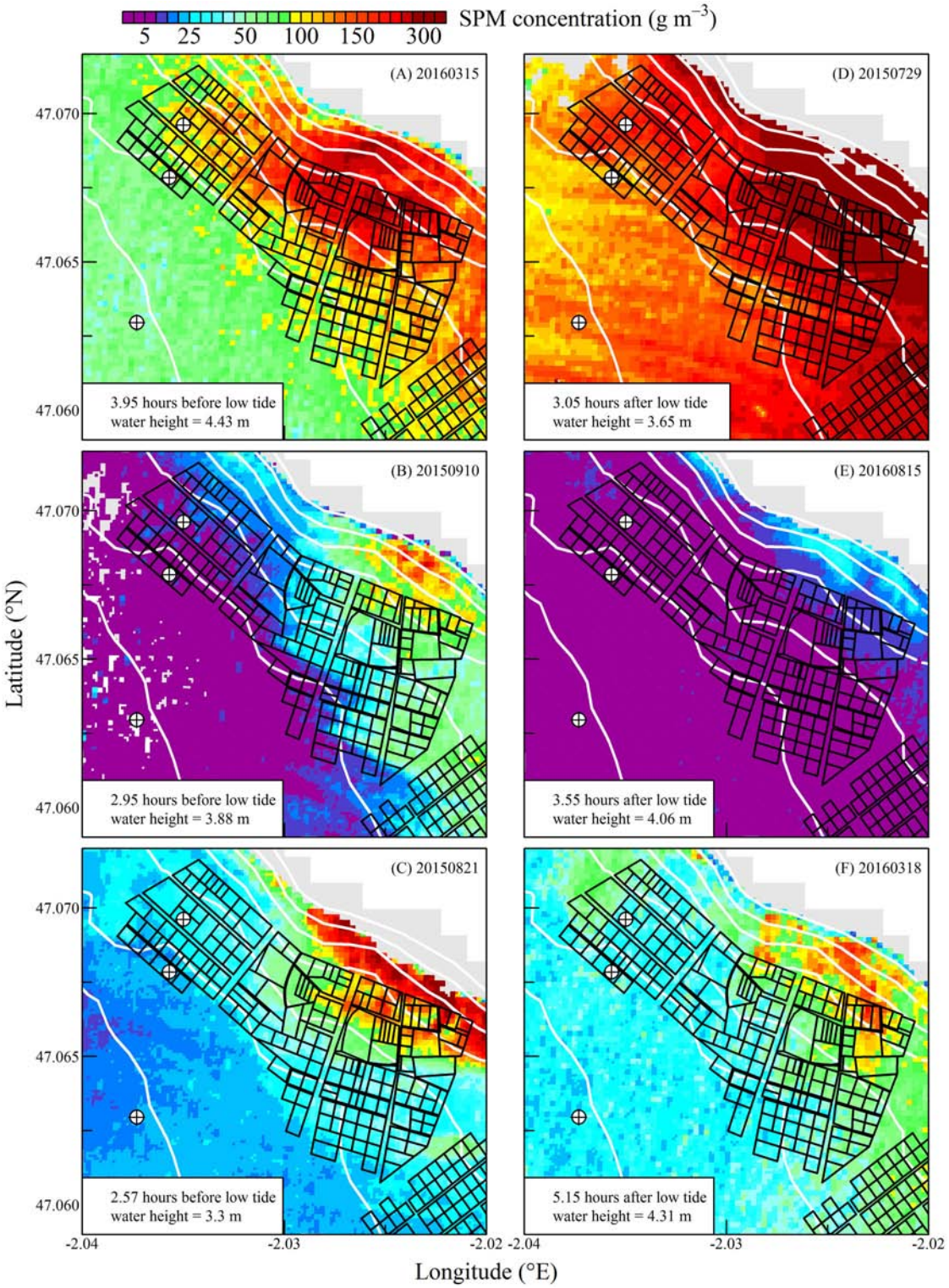


FIGURE 5 | Suspended particulate matter concentration during neap tides. Sentinel2 data were sorted following a composite tidal cycle from ebb tide **(A–C)** to flow tide **(D–F)**. Time difference between Sentinel2 acquisition and low tide is indicated, as well as the water height at the nearest reference harbor. The black polygons show the location of oyster tables. The white lines show the isobaths from 0 to 6 m above chart datum. The emerged part of the intertidal zone is in gray. Circled crosses around isobaths 0, 1, and 2 m show the points for which oyster clearance and consumption rates were computed (see **Figures 9**, **10**).

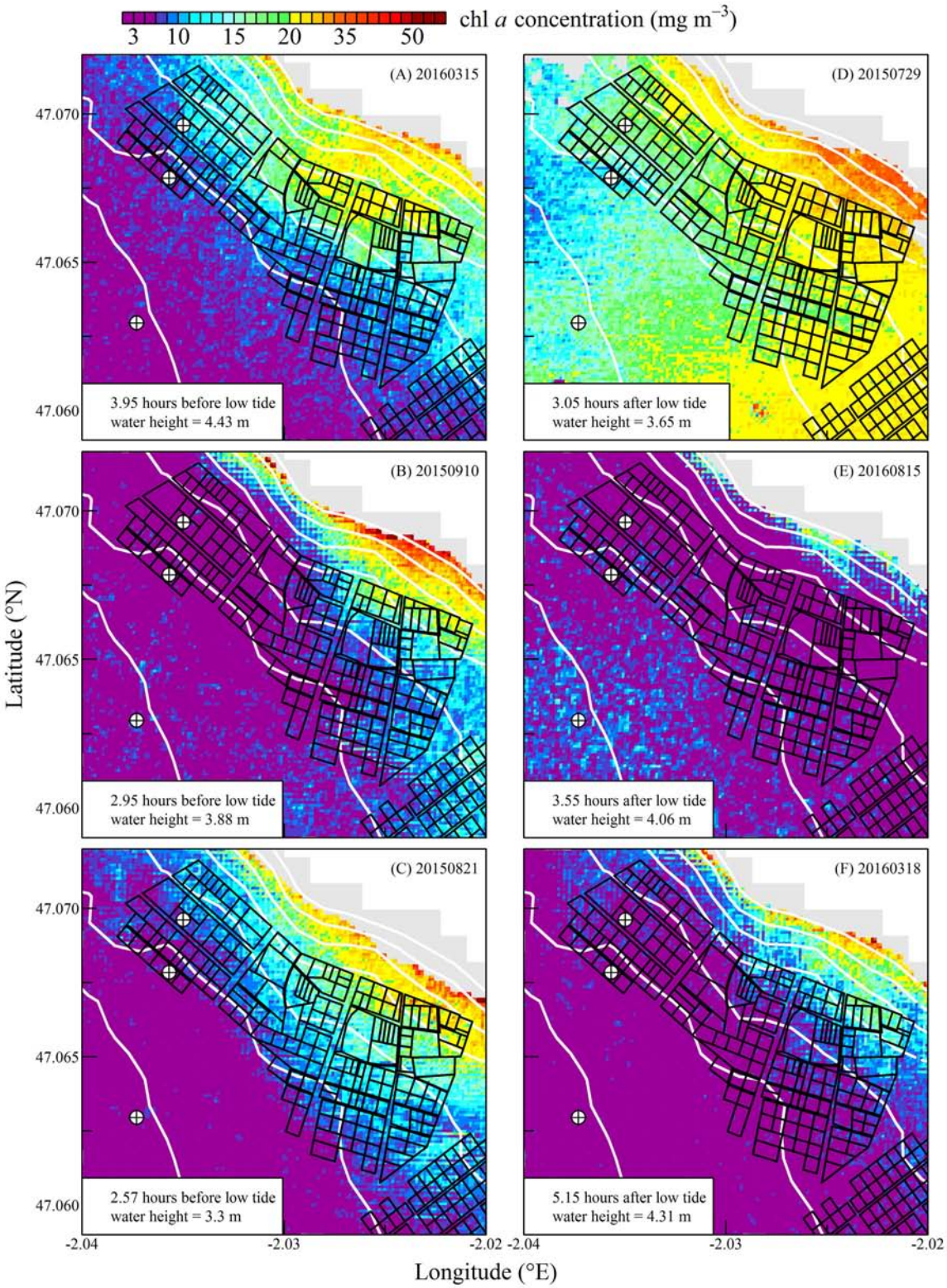


FIGURE 6 | Same as for Figure 5 but for chlorophyll *a* concentration.

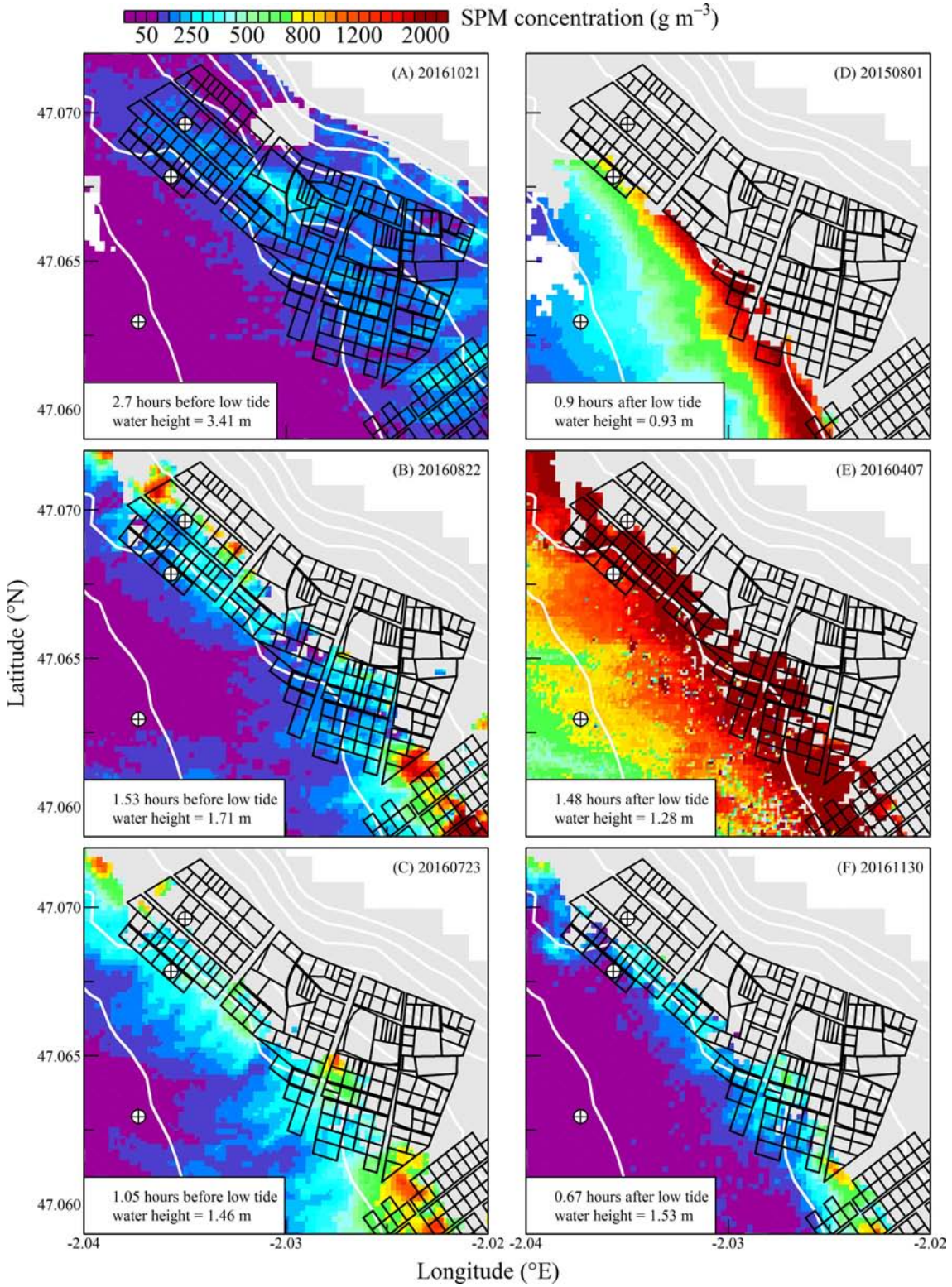


FIGURE 7 | Same as for Figure 5 but during spring tides. Note the change in the color scale.

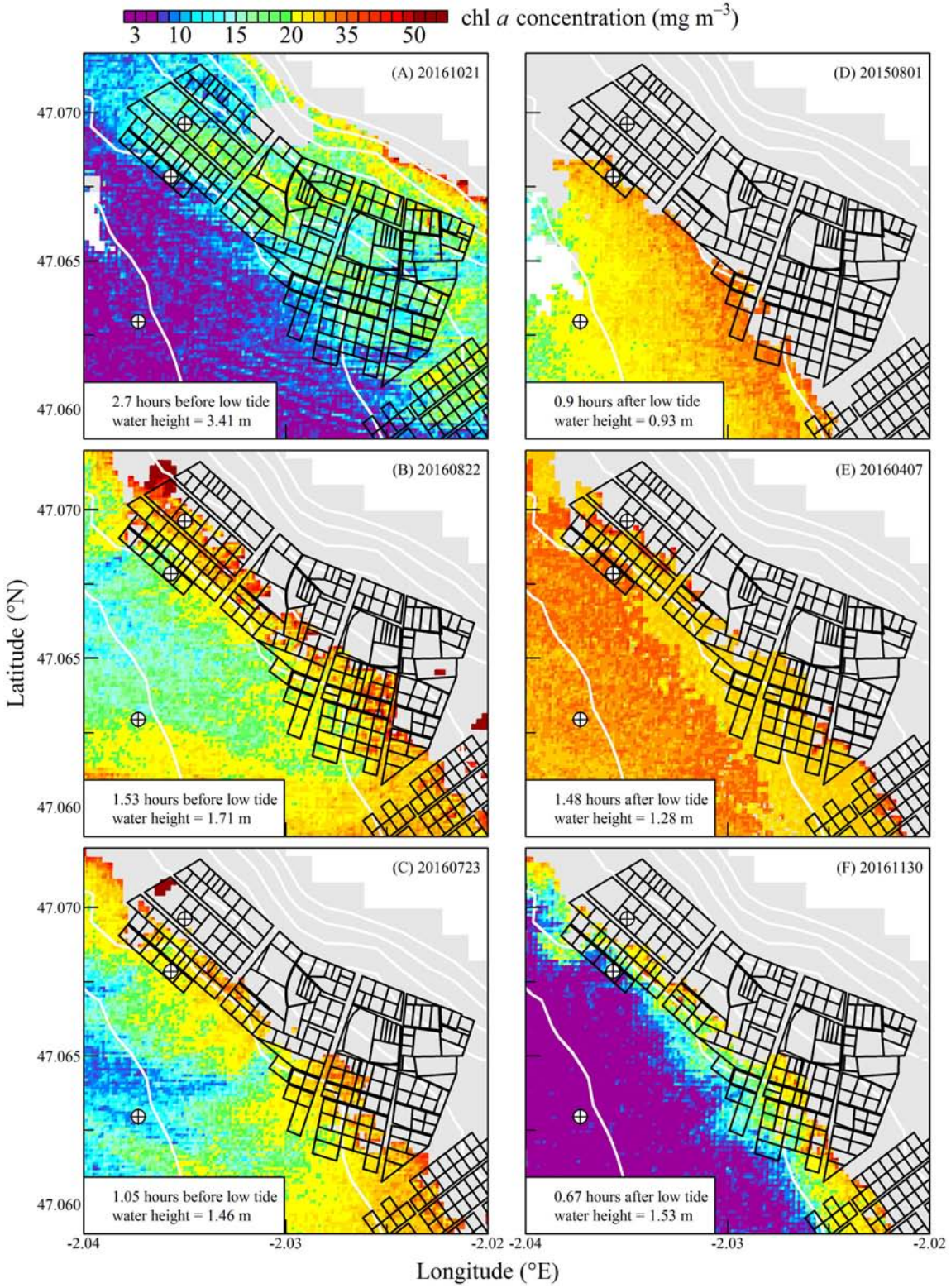


FIGURE 8 | Same as for Figure 6 but during spring tides.

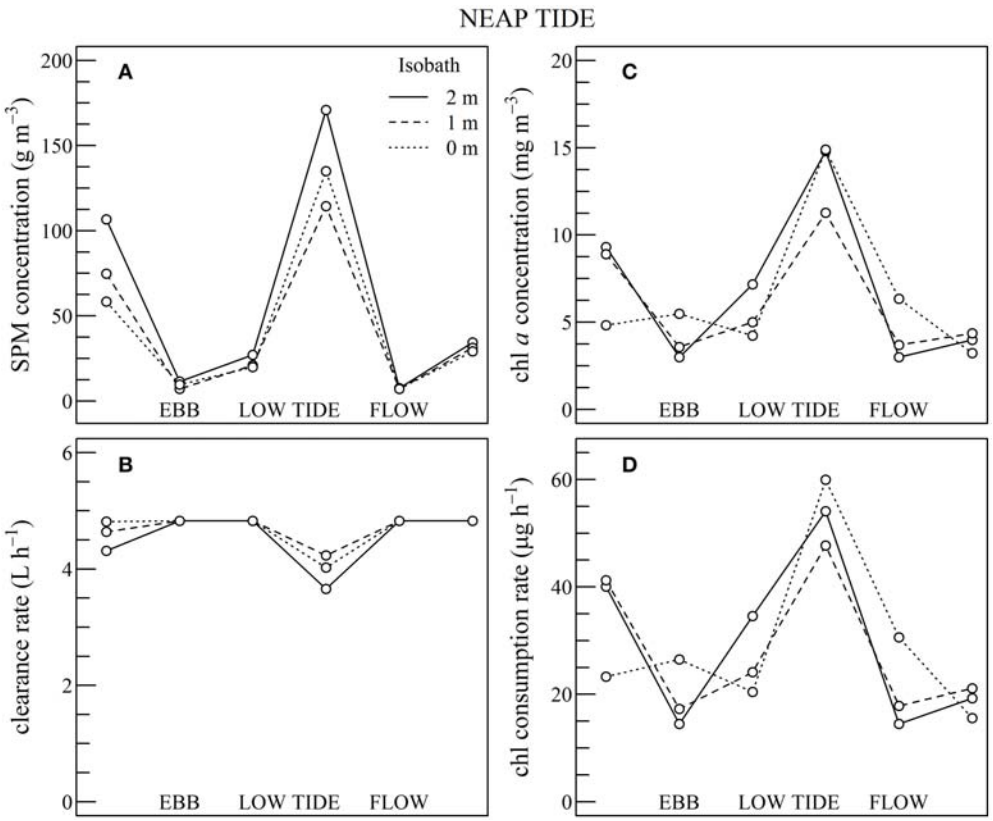


FIGURE 9 | Using Sentinel2 data shown in Figures 5, 6, composite tidal cycle of suspended particulate matter (SPM) concentration (A), oyster clearance rate **(B)**, chlorophyll *a* (chl *a*) concentration **(C)**, and chl consumption rate **(D)** during neap tides at three bathymetric locations, as indicated.

chl *a* concentration on oysters, several physiological functions were directly retrieved from satellite data, as described below.

Simulating Oyster Physiology from Space

Oyster clearance rate was computed from SPM concentration as in Barillé et al. (1997) using a non-linear function response (see also Figure 3 in Gernez et al., 2014). Briefly, oyster clearance rate is constant and equal to 4.8 L h^{-1} when SPM concentration is lower than 60 g m^{-3}. Over this threshold the clearance rate is negatively impacted by the high turbidity. It follows a linear and decreasing trend between 60 and ∼200 g m^{-3}, and exponentially collapses over ∼200 g m^{-3}. This latter point corresponds to the saturation of the oyster gills (Barillé et al., 1997). The chlorophyll consumption rate, defined as the biomass of chl *a* consumed per hour, is then computed as the product of the chl *a* concentration by the clearance rate (Barillé et al., 1997):

$$\mathrm{CONS} = [\mathrm{chl}\ a]\cdot \mathrm{CR} \tag{10}$$

where CR and CONS are, respectively the oyster clearance and chlorophyll consumption rates.

The clearance and chl consumption rates were simulated for each pixel of the Sentinel2 images using satellite-derived SPM and chl *a* data. In order to analyze the influence of the tide-driven chl *a* and SPM dynamic on oyster physiological response, composite tidal cycles of the clearance and chl consumption rates were also computed for the neap tide and spring tide clusters.

RESULTS

In situ Reflectance Spectra

In situ hyperspectral marine reflectance spectra show some typical characteristics of coastal turbid waters (**Figure 3A**). From 400 to 580 nm $\rho_w(\lambda)$ is relatively insensitive to changes in SPM and chl *a* concentration, preventing the use of a blue to green ratio algorithm to derive chl *a* concentration. From 600 to 900 nm the changes in the magnitude and spectral composition of the marine reflectance are mainly driven by variation in SPM concentration, notably as a result of particulate scattering.

In the red and NIR spectral region, the ρ_w spectra also display several features associated with the presence of pigment-bearing particles. An inflection in the reflectance slope is visible around 632 nm, attributable to the absorption by both chl *a* and chl *c*, a pigments association specific to diatoms (Méléder et al., 2005). The most striking feature is however the trough at 675 nm associated with chl *a* absorption, and the resulting reflectance shift between 675 and 700 nm, generally referred to as the NIR/red edge (Gons et al., 2002). Significant chl *a* concentration was confirmed by the analysis of HPLC data.

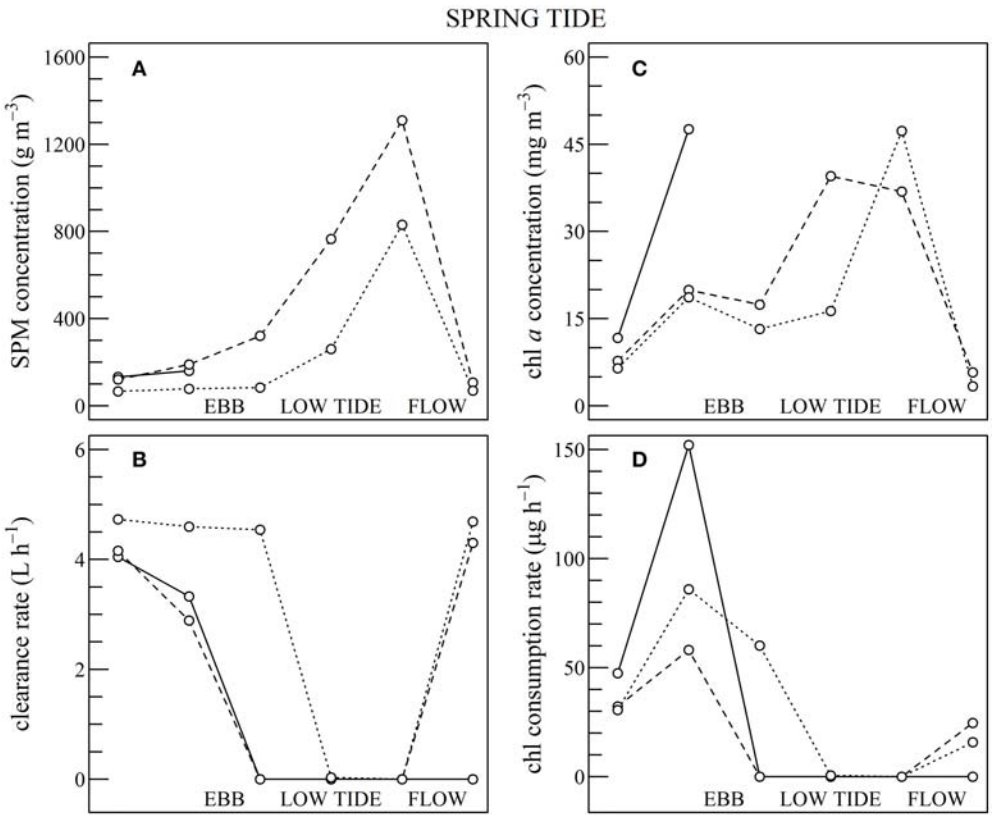


FIGURE 10 | Same as for Figure 9 but using Sentinel2 data shown in Figures 7, 8, during spring tides.

It most likely originates from the tide-driven resuspension of benthic microalgae resuspended together with surface sediments.

Though a significant loss of information results from the downscaling of the TriOS hyperspectral reflectance to S2/MSI spectral resolution, a NIR/red edge between 665 and 705 nm is still noticeable on several S2-simulated ρ_w spectra (**Figure 3B**), making it possible to apply the Gons et al. (2005) algorithm in Bourgneuf Bay's intertidal waters.

In situ Calibration of the SPM and Chl *a* Algorithms

The SPM and chl *a in situ* data acquired concomitantly with the $\rho_w(\lambda)$ were used to calibrate the bio-optical algorithms using a fitting procedure (**Figure 4**). *In situ* SPM concentration ranged from 10.92 to 700.83 g m^{-3}, with a mean of 146.53 g m^{-3}. Over this range, a significant relationship was obtained between simulated and measured SPM concentration (p-value $< 10^{-5}$, correlation coefficient of 0.96, a slope of 0.93 and an intercept of 0.32). For the SPM retrieval, the root mean square error was 56.26 g m^{-3}.

The range of chl *a* concentration was from 3.97 to 52.51 mg m^{-3}, with a mean of 18.48 mg m^{-3}. In the initial algorithm of Gons et al. (2005), the chlorophyll-retrieval parameters were originally set to $p = 1.05$ and $a^*_{phy}(665) = 0.014$ m^2 (mg Chl *a*)$^{-1}$ using observations performed in a variety of inland, estuarine and coastal waters over a range of chl *a* concentration from 1 to 181 mg m^{-3} (Gons, 1999; Gons et al., 2002). For the present study p was recalibrated to 1.02. A mean $a^*_{phy}(665)$ of 0.013 m^2 (mg Chl a)$^{-1}$ was obtained, a value consistent with previously reported specific absorption coefficients for the Saint Laurent Estuary (Bricaud et al., 1995), and for the Baltic and North seas (Babin et al., 2003). The comparison between the marine reflectance- and HPLC-derived chl *a* concentration was satisfactory (**Figure 4A**). The obtained linear regression shows a significant correlation coefficient of 0.93 (p-value $< 10^{-5}$), a slope of 1.02 and an intercept of −0.07. For the chl retrieval, the root mean square error was 3.05 mg m^{-3}.

Chl *a* Concentration within the Oyster Farm

The high spatial resolution (20 m) and spectral characteristics of Sentinel2 made it possible to quantify the distribution of SPM and chl *a* concentration at the scale of an oyster farm (**Figures 5–8**). Three main features characterized SPM and chl *a* spatial distribution in the shellfish farming site. First, both SPM and chl *a* concentrations are generally high, often exceeding 200 g m^{-3} and 10 mg m^{-3}, respectively. Second, their spatial structure depends on the bathymetry. SPM and chl *a* concentrations generally increase coastward, and the changes observed in their spatial distribution are more or less parallel to the isobaths (see

for examples **Figure 6C** where the change in [chl *a*] from <5 to >10 mg m^{-3} occurred around the 1 m isobath, and **Figure 7D** where [SPM] displayed a clear gradient coastward). Third, the tidal cycle is a significant driver of SPM and chl *a* dynamics. Due to the tidal resuspension of benthic microalgae together with the other particles of the sedimentary surface, the chl *a* distribution was generally associated with SPM, a common feature of intertidal mudflats (Koh et al., 2006). Spatial fronts of highest chl *a* concentration generally move seaward during ebb tide (**Figures 6A–C**, **8A–C**), and coastward during flow tide (**Figures 6D–F**, **8D–F**).

SPM and chl *a* concentrations exhibited similar spatial pattern during neap tide and spring tides, but the amplitude of the changes in their concentrations varied markedly between neaps and springs. For example, within the oyster farming zone SPM concentration varied from <10 to 300 g m^{-3} during neap tides and from 50 to >1,000 g m^{-3} during spring tides. Chl *a* concentration varied from <5 to 25 mg m^{-3} during neap tides and from 10 to 40 mg m^{-3} during spring tides.

The tide-driven variability was confirmed by the analysis of SPM and chl *a* concentration at the three selected fixed locations (black circled crosses in **Figures 5–8**). Both SPM and chl *a* concentration increased during ebb tide, reaching their maximum value during low tide and the start of the flow tide, and eventually decreasing during the end of the flow tide (**Figures 9A,C**, **10A,C**). The temporal correlation between SPM and chl *a* concentration is attributable to the tide-driven resuspension of surface sediments, which contain both mineral particles and benthic microalgae. As expected, the amplitude of the tide-driven changes in SPM and chl *a* was higher during spring than during neap tides.

The temporal changes in clearance and chl consumption rates were then analyzed in order to quantify the influence of the tide-driven SPM and chl *a* dynamic on the oyster physiological responses.

Influence of SPM and Chl *a* Variation on Oyster Physiology

During neap tides the simulated chl consumption rate mirrored the changes observed in chl *a* concentration (**Figures 9C,D**). This is attributable to the limited negative impact of SPM concentration on the clearance rate (**Figures 9A,B**). Generally SPM concentration remained below 100 g m^{-3} throughout the composite tidal cycle, except during flow tide where SPM concentration increased up to 175 g m^{-3} due to the erosion of surface sediments by tidal currents. The clearance rate mostly fluctuated between 4 and 5 L h^{-1}, and the decrease which occurred just after low tide was too small to counterbalance the increase in chl consumption.

During spring tides the simulated oyster physiological functions were more complexly affected by SPM and chl *a* tidal variability (**Figure 10**). First, a significant part the intertidal zone rapidly became emerged, and after mid-ebb the oysters could no longer filter seawater nor consume particles (see plain lines in **Figure 10**). In the waters just outside the intertidal zone, SPM concentration rapidly exceeded 200 g m^{-3} due to the strong tidal currents occurring from mid-ebb to the end of flow tide (dashed and dotted lines in **Figure 10A**). Such high SPM concentrations are known to saturate the oyster gills (Barillé et al., 1997), and it resulted in the dramatic collapse of the clearance rate from mid-ebb to the end of flow tide (**Figure 10B**). Meanwhile, the positive effects of the tide-driven increase in chl *a* concentration were rapidly counterbalanced by the collapse of the clearance rate, and the chl consumption rate dropped to zero from mid-ebb to the end of the flow (**Figure 10D**).

In summary, during neap tide, the tidal cycle in SPM concentration does not negatively impact oyster physiological response very much, and the low-tide increase in chl *a* concentration was directly translated into an increase in chl consumption for the farmed oysters. As most of the intertidal zone remains immerged during neap tides, the altitudinal location of the oyster farms has little influence on the tidal cycle of the clearance and chl consumption rate (**Figure 9**). During spring tides on the contrary, a significant fraction of the intertidal zone is emerged. In the waters adjacent to the intertidal zone and in the intertidal areas still under water, the high SPM concentration negatively impacts both clearance and chl consumption rate, whatever the chl *a* concentration, thus further limiting oyster physiological activity during a significant fraction of the tidal cycle (**Figure 10**).

DISCUSSION

Chl *a* Algorithms in Turbid Coastal Waters

The Gons et al. (2005) algorithm, recalibrated to Sentinel2/MSI, was fitted to *in situ* measurement (**Figure 4**) before being applied to Bourgneuf Bay. The fit was satisfactory, as demonstrated by the consistency of the value of $a^*_{phy}(665)$ with the literature (Bricaud et al., 1995; Gons et al., 2002; Babin et al., 2003). Other algorithms could have been used. Besides the Gons et al. (2005) method, various approaches were previously developed for turbid and/or eutrophic coastal and inland waters, including the 2-, and 3-band models (Dall'Olmo et al., 2005; Gitelson et al., 2008; Le et al., 2013), the fluorescence line height (FLH, Gower et al., 1999), the maximum chlorophyll index (MCI, Gower et al., 2005), and the 705 nm peak height (Toming et al., 2016). These algorithms are all based on a NIR/red edge either associated with chl *a* absorption at 675 nm or with sun-induced chl *a* fluorescence. An obvious limitation of these NIR/red algorithms is the lack of sensitivity in waters where the reflectance trough associated with chl *a* absorption around 675 nm is hardly pronounced (see for example the lowest $\rho_w(\lambda)$ spectrum in **Figure 3**). In less eutrophic waters (i.e., chl *a* concentration smaller than ~4 mg m^{-3}), the use of blue-green wavelengths would be more relevant to retrieve chl *a* concentration.

A recent study demonstrated that the 2- and 3-band models (Dall'Olmo et al., 2005) worked well with simulated Sentinel2/MSI-like imagery (Beck et al., 2016). In our algorithm inter-comparison (see Supplementary Information), the most performant method for our study site was the Gons et al. (2005) algorithm. Besides its good performance, another advantage of the Gons et al. (2005) algorithm is that the p and $a^*_{phy}(665)$

parameters seem to be relatively stable over a variety of coastal waters, including the numerous inland, estuarine, and coastal sites initially sampled by Gons (1999), and the intertidal waters of Bourgneuf Bay. Additional field data are however needed to assess the geographical robustness of a common set of parameters, as well as its seasonal stability. While the range of chl *a* concentration retrieved from Sentinel2 data in our study site was consistent with previous field measurements (Barillé-Boyer et al., 1997), the accuracy of the chl *a* concentration maps is more difficult to quantitatively appraise due to the lack of validation data. More *in situ* and match-up data are needed to improve the method.

Advantages and Limitations of Sentinel2 for Aquaculture Applications

The retrieval of chl *a* concentration using a NIR/red algorithm is relevant in eutrophic waters, but at lower chl *a* concentration the accuracy would probably decrease. It is then generally advised to switch toward shorter wavelengths in the blue-green parts of the visible spectrum. As far as Sentinel2 is concerned, the lack of a spectral band at around 412 nm will certainly limit the performance of chl *a* inversion methods, as it has been long demonstrated that such a spectral band improves the deconvolution of chl *a*, PIM, and CDOM absorption (Carder et al., 1999). For sensors equipped with a spectral band at 412 nm such as MERIS, the ocean color 5 (OC5) algorithm (Gohin et al., 2002) has been recently recommended in recent intercomparison studies of the North West European (Tilstone et al., 2017) and Vietnamese (Loisel et al., 2017) coastal waters. As already indicated by Vanhellemont and Ruddick (2016), several other limitations specific to Sentinel2 can also arise, due to its relatively wide bands, low signal-to-noise ratio, and lack of vicarious calibration.

Despite these issues, Sentinel2 offers four main advantages for the remote-sensing of shellfish farming ecosystems. First, its high spatial resolution (20 m) made it possible to observe aquaculture sites located nearshore, in narrow bays and estuaries (Gernez et al., 2014), and to analyze within-farm spatial variability. Second, its relatively small revisit time (which is now 5 days since the launch of Sentinel-2B) increases the probability of acquiring cloud-free data over a given site. Sentinel2 acquisition frequency also limits subsampling and observation biases for the study of rapidly varying environments. There is no doubt that the Sentinel2 time-series will strengthen observation robustness and statistical descriptors of the very dynamic and changing coastal waters. Third, its SWIR spectral band facilitates atmospheric correction over turbid waters (Vanhellemont and Ruddick, 2015). Fourth, its spectral resolution in the red and NIR spectral regions made it possible to apply a variety of chlorophyll inversion algorithms (see previous section). Altogether, these characteristics represent a significant improvement for the remote sensing of turbid oyster farming ecosystems, and more generally for coastal zone observation.

Shellfish Ecology from Space?

The combination of EO and shellfish physiological models opens new perspectives for aquaculture management, shellfish farming ecosystems studies (Gernez et al., 2014), and more broadly for a better understanding of the coastal ocean response to global changes. For example, the poleward extent of the Pacific oyster (a well-known invasive species, Herbert et al., 2016) along the European coasts has been recently quantitatively analyzed using an original coupling of EO with mechanistic physiological oyster modeling (Thomas et al., 2016). In another recent study the EO time-series archive has been used with climatic, biological and energetics models to better understand predicted changes in growth, reproduction and mortality risk for commercially and ecologically important bivalves in the Mediterranean Sea (Montalto et al., 2016).

Concurrently with the increase of EO aquaculture applications, the development of improved satellite products should not be neglected. The detection of phytoplankton species causing harmful algal blooms (HABs) is a major concern for fisheries and shellfish farming management (Sourisseau et al., 2016), and recent algorithm developments have proved useful to provide early warnings (Davidson et al., 2009) or statistical estimation of HAB-related risks (Kurekin et al., 2014). Enhanced characterization of the composition of the particulate assemblage could also be used to improve satellite-derived aquaculture products. For example, as oysters have the ability to preferentially select organic rather than mineral particles before ingestion (Barillé et al., 1997; Dutertre et al., 2009), estimation of the organic fraction of the particulate assemblage (Woźniak et al., 2010) could be used to better constrain shellfish physiological models.

CONCLUSION

In summary, it has been demonstrated that Sentinel2/MSI has the potential to map chlorophyll *a* and SPM concentration in turbid, chlorophyll-rich, intertidal waters. Sentinel2 high spatial resolution (20 m) made it possible to analyze SPM and chl *a* distribution at the scale of an oyster farm, thus opening new opportunities for aquaculture applications. The influence of the tidal dynamic on SPM and chl *a* concentration was highlighted, and its influence on oyster physiological response was analyzed in the shellfish farm and adjacent nearshore waters. During neap tides oysters were little influenced by the high turbidity, whereas during spring tides their clearance and chl consumption rates were significantly impacted by the extremely high SPM concentration during a significant fraction of the tidal cycle. This study confirms the potential of EO for marine spatial planning (Ouellette and Getinet, 2016), and offers a generic framework where the combination of high resolution satellite remote sensing with bivalves ecophysiological model makes it possible to explore the response of cultivated suspension feeders to environmental conditions in many coastal areas, and to optimize site selection for shellfish farming.

AUTHOR CONTRIBUTIONS

All authors contribute to work design, data acquisition, and data interpretation. PG processed the data and wrote the manuscript.

All authors gave their approval to the manuscript final version.

FUNDING

This work has been supported by the "Programme National de Télédétection Spatiale" (PNTS, http://www.insu.cnrs.fr/pnts) in the frame of the TURBO project (grant n° PNTS-2015-07), by the French Research National Agency (ANR) in the frame of the GIGASSAT project (grant n° ANR-12-AGRO-0001, http://www.gigassat.org/), by the European Union's Research and Innovation FP7 program in the frame of the HIGHROC project (grant n° 606797, http://www.highroc.eu/), and by the Tools for Assessment and Planning of Aquaculture Sustainability (TAPAS) project, a Horizon 2020 Research and Innovation Action funded by the European Commission (Grant agreement No: 678396, http://tapas-h2020.eu/).

ACKNOWLEDGMENTS

The European Space Agency is acknowledged for the production and distribution of Sentinel2 data. The US Geological Survey is thanked for maintaining the earth explorer web portal (https://earthexplorer.usgs.gov). Morgane Larnicol and Stéfani Novoa are thanked for their participation to field work. The authors thank Quinten Vanhellemont and Kevin Ruddick for making the ACOLITE software freely available. The organizers of the CLEO workshop and special issue are thanked for their efforts to federate the ocean color scientific community. Emmanuel Devred and Zhubin Zheng are thanked for their comments.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmars.2017.00137/full#supplementary-material

REFERENCES

Austin, R. W. (1974). "Inherent spectral radiance signatures of the ocean surface," in *Ocean Color Analysis*, eds S. Q. Duntley, R. W. Austin, W. H. Wilson, C. F. Edgerton, and S. W. Moran (San Diego, CA: Scripps Inst. of Oceanogr), 195.

Babin, M., Stramski, D., Ferrari, G. M., Claustre, H., Bricaud, A., Obolensky, G., et al. (2003). Variations in the light absorption coefficients of phytoplankton, nonalgal particles, and dissolved organic matter in coastal waters around Europe. *J. Geophys. Res. Oceans* 108, 3211. doi: 10.1029/2001jc000882

Barillé, L., Héral, M., and Barillé-Boyer, A.-L. (1997). Modélisation de l'écophysiologie de l'huître *Crassostrea gigas* dans un environnement estuarien. *Aquat. Living Res.* 10, 31–48. doi: 10.1051/alr:1997004

Barillé-Boyer, A.-L., Haure, J., and Baud, J.-P. (1997). *L'ostréiculture en baie de Bourgneuf, Relation Entre la Croissance des Huîtres Crassostrea Gigas et le Milieu Naturel: Synthèse de 1986 à 1995*. IFREMER Rep. DRV/RA/RST/97-16. Available online at http://archimer.ifremer.fr/doc/00000/1633/.

Beck, R., Zhan, S., Liu, H., Tong, S., Yang, B., Xu, M., et al. (2016). Comparison of satellite reflectance algorithms for estimating chlorophyll-a in a temperate reservoir using coincident hyperspectral aircraft imagery and dense coincident surface observations. *Remote Sens. Environ.* 178, 15–30. doi: 10.1016/j.rse.2016.03.002

Blondeau-Patissier, D., Gower, J. F., Dekker, A. G., Phinn, S. R., and Brando, V. E. (2014). A review of ocean color remote sensing methods and statistical techniques for the detection, mapping and analysis of phytoplankton blooms in coastal and open oceans. *Prog. Oceanogr.* 123, 123–144. doi: 10.1016/j.pocean.2013.12.008

Bricaud, A., Babin, M., Morel, A., and Claustre, H. (1995). Variability in the chlorophyll-specific absorption coefficients of natural phytoplankton: analysis and parameterization. *J. Geophys. Res. Oceans* 100, 13321–13332.

Brito, A. C., Benyoucef, I., Jesus, B., Brotas, V., Gernez, P., Mendes, C. R., et al. (2013). Seasonality of microphytobenthos revealed by remote-sensing in a South European estuary. *Cont. Shelf Res.* 66, 83–91. doi: 10.1016/j.csr.2013.07.004

Carder, K. L., Chen, F. R., Lee, Z. P., Hawes, S. K., and Kamykowski, D. (1999). Semianalytic moderate-resolution imaging spectrometer algorithms for chlorophyll *a* and absorption with bio-optical domains based on nitrate-depletion temperatures. *J. Geophys. Res. Oceans* 104, 5403–5421.

Choy, E. J., Richard, P., Kim, K. R., and Kang, C. K. (2009). Quantifying the trophic base for benthic secondary production in the Nakdong River estuary of Korea using stable C and N isotopes. *J. Exp. Mar. Biol. Ecol.* 382, 18–26. doi: 10.1016/J.JEMBE.2009.10.002

Dall'Olmo, G., Gitelson, A. A., Rundquist, D. C., Leavitt, B., Barrow, T., and Holz, J. C. (2005). Assessing the potential of SeaWiFS and MODIS for estimating chlorophyll concentration in turbid productive waters using red and near-infrared bands. *Remote Sens. Environ.* 96, 176–187. doi: 10.1016/j.rse.2005.02.007

Davidson, K., Miller, P., Wilding, T. A., Shutler, J., Bresnan, E., Kennington, K., et al. (2009). A large and prolonged bloom of *Karenia mikimotoi* in Scottish waters in 2006. *Harmful Algae* 8, 349–361. doi: 10.1016/j.hal.2008.07.007

Dessinges, M., Le Bihan, V., Le Goff, F., and Petit, M. (2012). *Ostréiculture en Pays de la Loire*. Carnet de bord n°5, CRC Pays de la Loire.

Dutertre, M., Beninger, P. G., Barillé, L., Papin, M., Rosa, P., Barillé, A. L., et al. (2009). Temperature and seston quantity and quality effects on field reproduction of farmed oysters, *Crassostrea gigas*, in Bourgneuf Bay, France. *Aquat. Living Res.* 22, 319–329. doi: 10.1051/alr/2009042

Gernez, P., Barillé, L., Lerouxel, A., Mazeran, C., Lucas, A., and Doxaran, D. (2014). Remote sensing of suspended particulate matter in turbid oyster-farming ecosystems. *J. Geophys. Res. Oceans* 119, 7277–7294. doi: 10.1002/2014JC010055

Gitelson, A. A., Dall'Olmo, G., Moses, W., Rundquist, D. C., Barrow, T., Fisher, T. R., et al. (2008). A simple semi-analytical model for remote estimation of chlorophyll-*a* in turbid waters: Validation. *Remote Sens. Environ.* 112, 3582–3593. doi: 10.1016/j.rse.2008.04.015

Gohin, F., Druon, J. N., and Lampert, L. (2002). A five channel chlorophyll concentration algorithm applied to SeaWiFS data processed by SeaDAS in coastal waters. *Int. J. Remote Sens.* 23, 1639–1661. doi: 10.1080/01431160110071879

Gons, H. J. (1999). Optical teledetection of chlorophyll *a* in turbid inland waters. *Environ. Sci. Tech.* 33, 1127–1132.

Gons, H. J., Rijkeboer, M., and Ruddick, K. G. (2002). A chlorophyll-retrieval algorithm for satellite imagery (Medium Resolution Imaging Spectrometer) of inland and coastal waters. *J. Plankton Res.* 24, 947–951. doi: 10.1093/plankt/24.9.947

Gons, H. J., Rijkeboer, M., and Ruddick, K. G. (2005). Effect of a waveband shift on chlorophyll retrieval from MERIS imagery of inland and coastal waters. *J. Plankton Res.* 27, 125–127. doi: 10.1093/plankt/fbh151

Gower, J. F. R., Doerffer, R., and Borstad, G. A. (1999). Interpretation of the 685 nm peak in water-leaving radiance spectra in terms of fluorescence, absorption and scattering, and its observation by MERIS. *Int. J. Remote Sens.* 20, 1771–1786.

Gower, J. F. R., King, S., Borstad, G., and Brown, L. (2005). Detection of intense plankton blooms using the 709 nm band of the MERIS imaging spectrometer. *Int. J. Remote Sens.* 26, 2005–2012. doi: 10.1080/01431160500075857

Herbert, R. J. H., Humphreys, J., Davies, C. J., Roberts, C., Fletcher, S., and Crowe, T. P. (2016). Ecological impacts of non-native Pacific oysters (*Crassostrea gigas*) and management measures for protected areas in Europe. *Biodivers. Conserv.* 25, 2835–2865. doi: 10.1007/s10531-016-1209-4

Jesus, B., Brotas, V., Ribeiro, L., Mendes, C. R., Cartaxana, P., and Paterson, D. M. (2009). Adaptations of microphytobenthos assemblages to sediment type and tidal position. *Cont. Shelf Res.* 29, 1624–1634. doi: 10.1016/j.csr.2009.05.006

Kang, C. K., Lee, Y. W., Choy, E. J., Shin, J. K., Seo, I. S., and Hong, J. S. (2006). Microphytobenthos seasonality determines growth and reproduction in intertidal bivalves. *Mar. Ecol. Prog. Ser.* 315, 113–127. doi: 10.3354/meps315113

Kazemipour, F., Launeau, P., and Méléder, V. (2012). Microphytobenthos biomass mapping using the optical model of diatom biofilms: application to hyperspectral images of Bourgneuf Bay. *Remote Sens. Environ.* 127, 1–13. doi: 10.1016/j.rse.2012.08.016

Koh, C. H., Khim, J. S., Araki, H., Yamanishi, H., Mogi, H., and Koga, K. (2006). Tidal resuspension of microphytobenthic chlorophyll *a* in a Nanaura mudflat, Saga, Ariake Sea, Japan: flood–ebb and spring–neap variations. *Mar. Ecol. Prog. Ser.* 312, 85–100. doi: 10.3354/meps312085

Kurekin, A. A., Miller, P. I., and Van der Woerd, H. J. (2014). Satellite discrimination of *Karenia mikimotoi* and *Phaeocystis* harmful algal blooms in European coastal waters: Merged classification of ocean colour data. *Harmful Algae* 31, 163–176. doi: 10.1016/j.hal.2013.11.003

Le, C., Hu, C., Cannizzaro, J., English, D., Muller-Karger, F., and Lee, Z. (2013). Evaluation of chlorophyll-*a* remote sensing algorithms for an optically complex estuary. *Remote Sens. Environ.* 129, 75–89. doi: 10.1016/j.rse.2012.11.001

Loisel, H., Vantrepotte, V., Ouillon, S., Ngoc, D. D., Herrmann, M., Tran, V., et al. (2017). Assessment and analysis of the chlorophyll-a concentration variability over the Vietnamese coastal waters from the MERIS ocean color sensor (2002–2012). *Remote Sens. Environ.* 190, 217–232. doi: 10.1016/j.rse.2016.12.016

MacIntyre, H. L., Geider, R. J., and Miller, D. C. (1996). Microphytobenthos: the ecological role of the "secret garden" of unvegetated, shallow-water marine habitats. I. Distribution, abundance and primary production. *Estuaries* 19, 186–201.

Méléder, V., Barillé, L., Rincé, Y., Morançais, M., Rosa, P., and Gaudin, P. (2005). Spatio-temporal changes in microphytobenthos structure analysed by pigment composition in a macrotidal flat (Bourgneuf Bay, France). *Mar. Ecol. Prog. Ser.* 297, 83–99. doi: 10.3354/meps297083

Méléder, V., Launeau, P., Barillé, L., and Rincé, Y. (2003). Microphytobenthos assemblage mapping by spatial visible-infrared remote sensing in a shellfish ecosystem. *C. R. Biol.* 326, 377–389. doi: 10.1016/S1631-0691(03)00125-2

Montalto, V., Helmuth, B., Ruti, P. M., Dell'Aquila, A., Rinaldi, A., and Sarà, G. (2016). A mechanistic approach reveals non linear effects of climate warming on mussels throughout the Mediterranean sea. *Climatic Change* 139, 293–306. doi: 10.1007/s10584-016-1780-4

Mueller, J. L., Davis, C., Arnone, R., Frouin, R., Carder, K., Lee, Z. P., et al. (2000). "Above-water radiance and remote-sensing reflectance measurements and analysis protocols," in *Ocean Optics Protocols for Satellite Ocean Color Sensor Validation, Revision 3*, eds G. S. Fargion and J. L. Mueller (Greenbelt, MD: NASA Goddard Space Flight Center), 171–182.

Neukermans, G., Ruddick, K., Loisel, H., and Roose, P. (2012). Optimization and quality control of suspended particulate matter concentration measurement using turbidity measurements. *Limnol. Oceanogr. Meth.* 10, 1011–1023. doi: 10.4319/lom.2012.10.1011

Novoa, S., Doxaran, D., Ody, A., Vanhellemont, Q., Lafon, V., Lubac, B., et al. (2017). Atmospheric corrections and multi-conditional algorithm for multi-sensor remote sensing of suspended particulate matter in low-to-high turbidity levels Coastal Waters. *Remote Sens.* 9:61. doi: 10.3390/rs9010061

Ouellette, W., and Getinet, W. (2016). Remote sensing for marine spatial planning and integrated coastal areas management: achievements, challenges, opportunities and future prospects. *Remote Sens. Appl. Soc. Environ.* 4, 138–157. doi: 10.1016/j.rsase.2016.07.003

Paterson, D. M., Wiltshire, K. H., Miles, A., Blackburn, J., Davidson, I., Yates, M. G., et al. (1998). Microbiological mediation of spectral reflectance from intertidal cohesive sediments. *Limnol. Oceanogr.* 43, 1207–1221.

Ras, J., Uitz, J., and Claustre, H. (2008). Spatial variability of phytoplankton pigment distributions in the Subtropical South Pacific Ocean: comparison between in situ and modelled data. *Biogeosciences* 5, 353–369. doi: 10.5194/bg-5-353-2008

R Development Core Team (2008). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online at: http://www.R-project.org

Ruddick, K. G., Ovidio, F., and Rijkeboer, M. (2000). Atmospheric correction of SeaWiFS imagery for turbid coastal and inland waters. *Appl. Opt.* 39, 897–912. doi: 10.1364/AO.39.000897

Sourisseau, M., Jegou, K., Lunven, M., Quere, J., Gohin, F., and Bryere, P. (2016). Distribution and dynamics of two species of *Dinophyceae* producing high biomass blooms over the French Atlantic Shelf. *Harmful Algae* 53, 53–63. doi: 10.1016/j.hal.2015.11.016

Thomas, Y., Pouvreau, S., Alunno-Bruscia, M., Barillé, L., Gohin, F., Bryère, P., et al. (2016). Global change and climate-driven invasion of the Pacific oyster (*Crassostrea gigas*) along European coasts: a bioenergetics modelling approach. *J. Biogeogr.* 43, 568–579. doi: 10.1111/jbi.12665

Tilstone, G., Mallor-Hoya, S., Gohin, F., Couto, A. B., Sá, C., Goela, P., et al. (2017). Which ocean colour algorithm for MERIS in North West European waters? *Remote Sens. Environ.* 189, 132–151. doi: 10.1016/j.rse.2016.11.012

Toming, K., Kutser, T., Laas, A., Sepp, M., Paavel, B., and Nõges, T. (2016). First experiences in mapping lake water quality parameters with Sentinel-2 MSI imagery. *Remote Sens.* 8:640. doi: 10.3390/rs8080640

Ubertini, M., Lefebvre, S., Gangnery, A., Grangeré, K., Le Gendre, R., and Orvain, F. (2012). Spatial variability of benthic-pelagic coupling in an estuary ecosystem: consequences for microphytobenthos resuspension phenomenon. *PLoS ONE* 7:e44155. doi: 10.1371/journal.pone.0044155

Underwood, G. J. C., and Kromkamp, J. C. (1999). Primary production by phytoplankton and microphytobenthos in Estuaries. *Adv. Ecol. Res.* 29, 93–153. doi: 10.1016/S0065-2504(08)60192-0

van der Wal, D., Wielemaker-van den Dool, A., and Herman, P. M. (2010). Spatial synchrony in intertidal benthic algal biomass in temperate coastal and estuarine ecosystems. *Ecosystems* 13, 338–351. doi: 10.1007/s10021-010-9322-9

Van der Zande, D., Vanhellemont, Q., and Ruddick, K. (2016). "Validation of Landsat-8/OLI for ocean colour applications with AERONET-OC sites in Belgian coastal waters," in *Ocean Optics Conference* (Victoria, BC), 23–28.

Vanhellemont, Q., and Ruddick, K. (2014). Turbid wakes associated with offshore wind turbines observed with Landsat 8. *Remote Sens. Environ.* 145: 105–115. doi: 10.1016/j.rse.2014.01.009

Vanhellemont, Q., and Ruddick, K. (2015). Advantages of high quality SWIR bands for ocean colour processing: Examples from Landsat-8. *Remote Sens. Environ.* 161, 89–106. doi: 10.1016/j.rse.2015.02.007

Vanhellemont, Q., and Ruddick, K. (2016). "ACOLITE For Sentinel-2: Aquatic Applications of MSI imagery," in *ESA Special Publication Presented at the ESA Living Planet Symposium held in Prague* (Czechia), 9–13.

Vermote, E., Tanré, D., Deuzé, J., Herman, M., Morcrette, J., and Kotchenova, S. (2006). *Second Simulation of a Satellite Signal in the Solar Spectrum-Vector* (6SV). *6S User Guide Version 3*. Greenbelt, MD: NASA Goddard Space Flight Center.

Woźniak, S. B., Stramski, D., Stramska, M., Reynolds, R. A., Wright, V. M., Miksic, E. Y., et al. (2010). Optical variability of seawater in relation to particle concentration, composition, and size distribution in the nearshore marine environment at Imperial Beach, California. *J. Geophys. Res. Oceans* 115:C08027. doi: 10.1029/2009jc005554

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Gernez, Doxaran and Barillé. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Marine Science

METHODS
published: 19 May 2017
doi: 10.3389/fmars.2017.00126

Regional Empirical Algorithms for an Improved Identification of Phytoplankton Functional Types and Size Classes in the Mediterranean Sea Using Satellite Data

Annalisa Di Cicco[1]*, *Michela Sammartino*[1], *Salvatore Marullo*[1,2] and *Rosalia Santoleri*[1]

[1] Institute of Atmospheric Sciences and Climate, National Research Council (CNR), Rome, Italy, [2] National Agency for New Technologies, Energy and Sustainable Economic Development, Frascati, Italy

Regional relationships to estimate the main Phytoplankton Functional Types (PFTs) and Size Classes (PSCs) from satellite data are presented. Following the abundance-based approach and selecting the Total Chlorophyll *a* (TChl*a*) as descriptor of the trophic status of the environment, empirical relations between the TChl*a* concentration and seven accessory pigments, marker for the main algal groups, have been developed for the Mediterranean Sea. Using only *in-situ* data acquired in this basin, firstly a previous regional diagnostic pigment analysis has been conducted to evaluate the specific pigment ratios featuring the phytoplankton assemblage that occurs in the Mediterranean Sea. Secondly, the new regional PFT and PSC algorithms have been calibrated and validated on the *in-situ* dataset. The statistical analysis showed a very good predictive power for all the new regional models. A quantitative comparison with global abundance-based models applied to our validation dataset showed that the regionalization improves the uncertainty and the spread of about one order of magnitude for all the classes (e.g., in the nano class, where the mean bias error improves from −0.056 to 0.001 mg m^{-3}). These results highlighted that a regionalization for the PSC and PFT estimates are required, to take into account the peculiar bio-optical properties of the Mediterranean Sea. Finally, the new regional equations have been applied to the Mediterranean TChl*a* satellite (1998–2015) time series to estimate annual and monthly PFT and PSC climatology. The analysis of the climatological maps, relative to the phytoplankton assemblage distribution patterns, reveals that all the three size classes reach their maxima in the higher nutrient areas, with absolute values >3 mg m^{-3} of TChl*a* for micro-, and about 1.6 and 0.4 mg m^{-3} for nano- and pico-phytoplankton, respectively. Moreover, the nano component shows intermediate percentage values in the whole basin, ranging from 30 to 40% of the TChl*a* in the western basin, up to 45% in the more productive areas. In terms of chlorophyll concentration, in the coastal areas we find the predominance of the Diatoms and Haptophytes, while in the ultra-oligotrophic waters Prokaryotes predominates on the other groups, constituting the principal component of the pico-phytoplankton.

Keywords: PFTs, PSCs, Mediterranean Sea, regional algorithms, ocean color

OPEN ACCESS

Edited by:
Shubha Sathyendranth, Plymouth Marine Laboratory, UK

Reviewed by:
Emmanuel Devred, Fisheries and Oceans Canada, Canada
Jochen Wollschläger, University of Oldenburg, Germany

***Correspondence:**
Annalisa Di Cicco
annalisa.dicicco@artov.isac.cnr.it

Specialty section:
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

Received: *16 January 2017*
Accepted: *18 April 2017*
Published: *19 May 2017*

Citation:
Di Cicco A, Sammartino M, Marullo S and Santoleri R (2017) Regional Empirical Algorithms for an Improved Identification of Phytoplankton Functional Types and Size Classes in the Mediterranean Sea Using Satellite Data. Front. Mar. Sci. 4:126. doi: 10.3389/fmars.2017.00126

INTRODUCTION

Phytoplankton have a key role in the biogeochemistry of the Earth, with a predominant position in several ecological processes as climate regulation, food webs, fossil fuel formation, and correlated economic human activities (Falkowski et al., 2003). The biogeochemical functions performed by the phytoplankton assemblage are closely linked to its composition. Key groups of organisms with their specific biogeochemical metabolism mediate the elemental fluxes in the biosphere (Falkowski et al., 2003; Le Quéré et al., 2005). The composition and succession of various phytoplankton taxa in the community are also a mirror of the ecological status of the marine environment (Devlin et al., 2007). Within this context, in the recent years the scientific interest in the comprehension of the phytoplankton assemblage structure is surging. The methods for the identification of these organisms have strongly evolved, moving from single cell counting and taxonomic identification based on the traditional microscopic techniques to most recent approach based on remote sensing investigation (IOCCG, 2014).

Proper identification of Linnaean taxonomic species that compose a natural phytoplankton assemblage requires the use of multiple combined techniques. By using the classical optical microscopy only, which remains one of the best approaches for the identification of the largest phytoplankton cells, it has been ignored for years the smallest fraction of the phytoplankton, instead detectable through specific techniques such as flow-cytometry, chemotaxonomy, epifluorescence microscopy, size-fractionation, and determination of chlorophyll *a* content with High Performance Liquid Chromatography, HPLC (Siokou-Frangou et al., 2010). A species-specific identification also requires great time demanding and needs of deep experience in the taxonomy knowledge (Reynolds, 2006). Nowadays, the systematic classification of phytoplankton at the level of phyla and of certain classes is well-established, with the agreement of microscopists and biochemicals (Reynolds, 2006). For several years, one of the most useful techniques for the algal classification at these taxonomic levels has been the HPLC. The liquid chromatography allows the separation, with resulting identification and quantification, of the main algal pigments, some of them considered markers for specific phytoplankton groups (see **Table 1**). The number of phytoplankton species is by far smaller than the terrestrial plants, but with a greater phylogenetic diversity, strictly related with the principal ecological functions (Falkowski and Raven, 1997; Falkowski

TABLE 1 | Diagnostic Pigments (DP) and their taxonomic meaning in microalgal divisions or classes (Jeffrey and Vesk, 1997; Prezelin et al., 2000; Vidussi et al., 2001; Wright and Jeffrey, 2006; Ras et al., 2008; Brunet and Mangoni, 2010).

Pigments (abbreviation)	Principal taxonomic meaning	[a]Secondary taxonomic meaning	*PSCs
Chlorophyll *a* (Chl*a*) (includes allomers and epimers)	**All divisions** unless Prochlorophyta		
Divinil-Chlorophyll *a* (Dv-Chl*a*)	**Prochlorophyta**		
Total Chlorophyll *a* (TChl*a*) (Chl*a* + Dv-Chl*a* + Chlide *a*)	**All divisions**		
Fucoxanthin (Fuco)	**Bacillariophyceae: *diatoms***	Haptophyta Chrisophyceae Raphydophyceae Dinophyta	Micro
Peridinin (Peri)	**Dinophyta: *dinoflagellates***		Micro
Alloxanthin (Allo)	**Cryptophyta**		Nano
19′-butanoiloxyfucoxanthin (But-fuco)	**Haptophyta: *coccolithophores*** Chrysophyceae		Nano
19′-hexanoiloxyfucoxanthin (Hex-fuco)	**Haptophyta: *coccolithophores***	Dinophyta	Nano/Pico
Chlorophyll *b* (Chl*b*)	**Green algae:** *- Chlorophyta* *- Euglenophyta*		Pico
Divinil Chlorophyll *b* (Dv-Chl*b*)	**Prochlorophyta**		Pico
Total Chlorophyll *b* (TChl*b*) (Chl*b* + Dv-Chl*b*)	**Green algae & Prochlorophyta**		
Zeaxanthin (Zea)	**Prokaryotes:** ***- Cyanophyta*** *- Prochlorophyta*	Chlorophyta Euglenophyta Chrisophyceae Raphydophyceae Eustigmatophyta	Pico
Chlorophyllide *a* (Chlide *a*)	Damaged centric diatoms and senescent tissue		

*Sieburth et al. (1978) Classification: Micro (>20 μm), Nano (2–20 μm), Pico (<2 μm)

[a]*DP presents in lower concentration or in some types only (Jeffrey and Vesk, 1997); *PSCs: grouping of the main taxa into size classes selected for this work (see text).*

et al., 2003). Phylogenetic studies on the oxygenic phytoplankton evolution suggested three main recognizable lineages. The first, in the prokaryotic empire, consists of all the Cyanobacteria. The other two, within the eukaryotic algae, are divided in "green lineage," characterized by chlorophyll *b* as secondary pigment and by a small quantity of several carotenoids (Phyla Chlorophyta and Euglenophyta) and in "red-lineage" including Rhodophyta, pigmented with phycobiliproteins and a number of other algal groups characterized by chlorophyll *c* and a wide variety of carotenoids. These groups involve Cryptophyta, Heterokontophyta, Haptophyta, and probably those Dinophyta pigmented with peridinin (Delwiche, 1999; Falkowski et al., 2003; Reynolds, 2006).

In order to better understand the ecological systems and monitor the ecological status of marine environment, the main target is to identify the structures and processes that can explain ecosystem dynamics, linking descriptors of state to descriptors of change. Recent trends in the comprehension of the community structure and functioning are aimed to the research of those "functional traits" species-independent able to act as non-taxonomic "descriptors of community." Two of the most relevant taxonomic-free descriptors are the body size class and the functional group (Basset et al., 2004; Mouillot et al., 2006).

The definition of "functional group" is open to different interpretations, clustering phytoplankton on the base of various ecological roles and specialized requirements. This term groups species with similar "morphological and physiological traits and ecologies" (Reynolds et al., 2002): a functional group is composed by different species that, starting from the same resource or ecological component, perform a common ecological function (Blondel, 2003). On the basis of their biogeochemical metabolism or, farther, on the "resource" shared by the organisms, main taxonomic phytoplankton groups can be assembled in four specific "functional groups" (Blondel, 2003; Falkowski et al., 2003; Litchman et al., 2007; IOCCG, 2014): nitrogen fixers (this ability is unique to the Prokaryotes), calcifiers (including the taxonomical class of Haptophyceae, generally known as coccolithophores), silicifers (represented by the class of Bacillariophyceae, tipically known as diatoms, followed by some chrysophytes, silicoflagellates, and xanthophytes, which are not very widespread in the Mediterranean Sea), and Dimethylsulfoniopropionate (DMSP) producers (referred to some marine phytoplankton organisms belonging primarily to the group of Dinoflagellates, followed by Haptophytes).

The other important "taxonomic-free" descriptor is the "size." A great number of single organism and community characteristics depend, in a known manner, on individual dimension. The "metabolic theory" of Brown et al. (2004), closely links the performance of "individuals" in terms of metabolism and energy transfer efficiency to the ecology of "population, community, and ecosystems." There is a flow of energy and matter between the various ecological systems at different hierarchical scales, depending on environmental and individual characteristics that regulate the metabolism of the single organism and consequently, the features of each hierarchical level. According to this theory, body size, together with temperature and stoichiometry, is one of three key factors that affect individual metabolism and, consequently, the community ecology.

Although, also size measurements may be affected by uncertainties, especially at ecological "individual" level, morphometric, or "body size" descriptors offer however important advantages with respect to the taxonomic ones: cell size is simpler to measure in quantitative and reproducible way and overcomes the long times and great experience required for taxonomic identification (Basset et al., 2004; Mouillot et al., 2006). In the aquatic ecosystems, the role of the individual dimension as phytoplankton community descriptor is based on the relationship between size and pigmentary content, different taxa, or stages of growth in the same taxon, photosynthetic efficiency, bio-optical phytoplankton properties, and water column dynamic (Chisholm, 1992; Raven, 1998; Organelli et al., 2007). Raven (1998), in his important work "The twelfth Tansley Lecture. Small is beautiful: the picophytoplankton", summarizes the influences of the phytoplankton cell size on its photosynthetic activity and its role in biogeochemical cycling and biodiversity. Size affects, above all, maximum specific growth rate, photon acquisition, nutrient solute, and water fluxes across the plasmalemma and loss of cells in the euphotic layer (Chisholm, 1992; Raven, 1998). On the base of the different ecological hierarchical levels of investigation, it is possible to identify several specific morphometric descriptors. For individual levels, we have bio-volume, surface area, or surface-volume ratio. Instead, for population and guild, we can consider body size-abundance distribution, body size-spectra, or biomass size fractions (Vadrucci et al., 2007). In the present work we take into account the biomass fractions of three Phytoplankton Size Classes (PSCs) related to the Sieburth et al. (1978) classification, micro- (>20 μm), nano- (2–20 μm), and pico- (<2 μm) phytoplankton and the main Phytoplankton Functional Types (PFTs).

In the recent years, several physical, biological, and ecological models have been proposed to estimate PSCs and PFTs from remote sensing data. Satellite technologies provide a great tool for a synoptic observation of the ecological state of the marine ecosystem at daily and global scale.

The most important current approaches used to detect dominant phytoplankton groups are designed for global application and are based on Brewin (2011) and IOCCG (2014): spectral-response, taking into account the specific optical signature of the different algal groups and deriving from both Apparent (AOPs) and Inherent (IOPs) Optical Properties (e.g., Ciotti et al., 2002; Sathyendranath et al., 2004; Alvain et al., 2005, 2008; Ciotti and Bricaud, 2006; Kostadinov et al., 2009; Pan et al., 2010, 2011; Roy et al., 2013; Navarro et al., 2014); phytoplankton abundance, based on the well-known inter-current relation between phytoplankton types and cell size and trophic status of environment (Chisholm, 1992; e.g., Devred et al., 2006; Uitz et al., 2006; Hirata et al., 2008, 2011; Brewin et al., 2010, 2011); ecological approach, in which additional ecological and physical information supports the ocean color data (e.g., Raitsos et al., 2008).

The applications of these approaches to the optical characteristics of Mediterranean Sea, featured by unique

optical properties in the water column, with "oligotrophic waters less blue (30%) and greener (15%) than the global ocean" (Volpe et al., 2007), are very few. The great interest in the Mediterranean Sea rises from its peculiarities of quasi enclosed sea with dimension, morphology, dynamics, and external forcing that make it a "miniature model" for a better comprehension of the global ocean complex processes, from mesoscale to basin scale (Lacombe et al., 1981; Robinson and Golnaraghi, 1995; Siokou-Frangou et al., 2010). Only recently, Navarro et al. (2014) exploited the PHYSAT method of Alvain et al. (2005), based on an empirical correlation between normalized water leaving (*nLw*) radiances (AOPs) and diagnostic pigments of an HPLC global dataset, performing its regionalization for the Mediterranean Sea. The new PHYSAT-Med has been validated mainly for nanoeukaryotes, *Prochlorococcus*, *Synechococcus,* and diatoms and provides the dominant phytoplankton group for each satellite pixel. Furthermore, Sammartino et al. (2015) exploited the capability of a global empirical model, Brewin et al. (2011), solely based on chlorophyll *a* data, to describe the phytoplankton size biomass distribution in the Mediterranean Sea.

In this work, with the intent to investigate the composition of phytoplankton assemblage and its variability, we first analyze the relationship between chlorophyll *a* content and diagnostic pigment composition of phytoplankton assemblage in the Mediterranean Sea. Afterwards, following the global abundance-based approach and selecting the Total Chlorophyll *a* (TChl*a*) as descriptor of the trophic status of the environment, we identified Mediterranean empirical relations between the concentration of TChl*a* and seven accessory pigments considered diagnostic for the main algal groups (**Table 1**). This allows us to develop new regional algorithms for satellite biomass estimates of PFTs and size classes and assess their accuracy respect to global models. Finally, we applied these new regional algorithms to the 1998–2015 TChl*a* satellite time series to compute Mediterranean PFT and PFC climatologies.

The paper is organized as follows: second section presents the *in-situ* and remote "data and methods" selected for this work, also describing the diagnostic pigment analysis performed on the Mediterranean pigment dataset; in Section "Results," we present and validate new Mediterranean regional algorithms for the identification of PFTs and PSCs and compared them with the results obtained by applying two global models; at last, Section "Discussion and Conclusions" are exposed in fourth section.

DATA AND METHODS

In-situ Pigment Data and Quality Assurance

Diagnostic pigment data for the determination of the *in-situ* PFTs and PSCs come from a Mediterranean subset of the SeaWiFS Bio-optical Archive and Storage System (SeaBASS) HPLC pigment *in-situ* dataset (Werdell and Bailey, 2005). Data were collected during different cruises and periodical activities of fixed mooring monitoring. More in details, this subset consists of data from Prosope cruise (1999, September–October), Boussole mooring data (with sampling activities nearly every month from 2001 to 2006 and only in July for the 2008) and Boum cruise (2008, July). It consists of 1,454 sets of pigments, including stations sampled in case 1 waters and in various trophic conditions. **Figure 1** shows the location of the whole SeaBASS Mediterranean *in-situ* measurements. We used all the *in-situ* data acquired in the first 50 m of the water column. Since these field samples were collected by several teams and were analyzed in different laboratories using a variety of HPLC instruments and protocols, we performed a quality assurance analysis to build up coherent combination of the data sets. At first, pigment data were visually checked in order to identify and remove suspected low quality values (for instrumental or clear stochastic errors). Then, we applied the Aiken et al. (2009) method to remove the outliers, according to Trees et al. (2000), which identified a strong log-linearity between TChl*a* and accessory pigments.

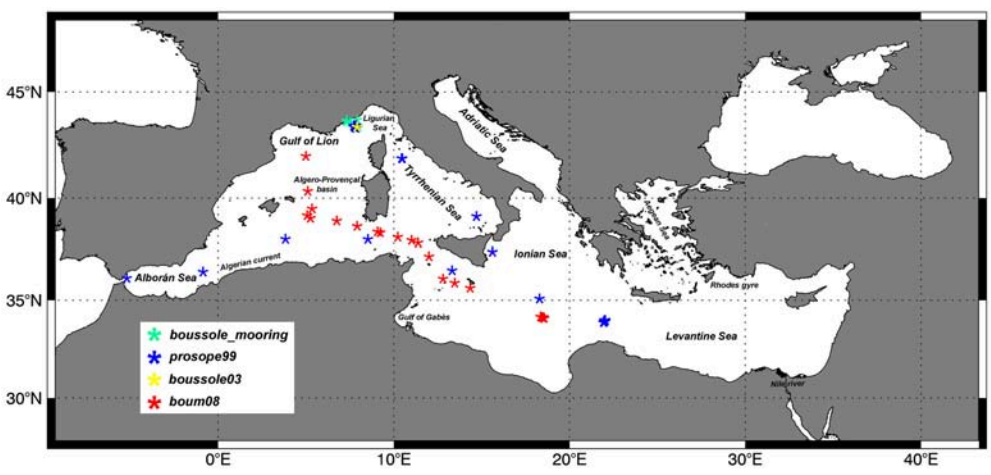


FIGURE 1 | Spatial distribution of the phytoplankton pigments in the SeaBASS Mediterranean HPLC subset (Werdell and Bailey, 2005) used for this work.

The data outside of the 95% confidence interval were eliminated. Following Hirata et al. (2011) and Brewin et al. (2010) we performed a 5-point moving average to the raw data, sorted according to increasing values of TChl*a*, to maximize the signal to noise ratio and underline the main trend of the data. The quality control reduces the useful measurements to 1,379, with values ranging from 0.02 to more than 5 mg m^{-3} (well-representative of the Mediterranean chlorophyll *a* variability). Even if the *in-situ* dataset is predominantly collected in the western and central Mediterranean Sea while the eastern Mediterranean is less sampled, it includes a significant part of samples (38% of the total) that fall in the oligotrophic chlorophyll *a* range typical of the eastern basin, therefore our dataset can be considered representative of the entire Mediterranean trophic regimes.

Determination of PFTs and PSCs from Pigment Composition: Diagnostic Pigment Analysis (DPA)

Information about the composition of phytoplankton assemblage in terms of "types" and "size classes" has been obtained from the analysis of cell's pigmentary content of *in-situ* samples, exploiting the diagnostic properties of some marker pigments.

Following Vidussi et al. (2001), according to previous works on chemotaxonomy (Wright and Jeffrey, 1987; Gieskes et al., 1988; Everitt et al., 1990; Williams and Claustre, 1991; Claustre, 1994; Jeffrey and Vesk, 1997), we take into account seven diagnostic pigments (DPs), able to detect the main phytoplankton types and to outline the size structure of the whole assemblage in the Mediterranean Sea. Some of these pigments are unambiguous markers, others typify a principal group (with a minor contribution of some other classes, see **Table 1**). In Vidussi et al. (2001) the identification of the PSCs is founded on the relation between taxonomic groups and their most common dimensions in the Mediterranean Sea. Although, the phytoplankton grouping method based on the auxiliary pigments does not exactly reflect the phytoplankton size such as the one based on the size fractionated chlorophyll (lacking for the Mediterranean Sea), nevertheless several investigations about the typical off shore composition of phytoplankton Mediterranean community have proven the validity of this approach (Vidussi et al., 2001; Siokou-Frangou et al., 2010). Therefore, it must be taken into account that, on the basis of Sieburth et al. (1978) size classification, "micro" consists of Diatoms and Dinoflagellates in general, nano includes Cryptophytes, Haptophytes, and some classes of Heterokontophytes, and pico-phytoplankton is referred to Cyanobacteria, green flagellates, and Prochlorophytes (**Table 1**). In this work we also applied the linear adjustment of Brewin et al. (2010) for the assignment of 19′ hexanoyloxyfucoxanthin (primarily marker of the Haptophytes), more traceable to pico size class rather than to nano-phytoplankton in the ultra-oligotrophic waters (Hirata et al., 2008; Ras et al., 2008).

For the quantification of each type, a now well-established method is to estimate the contribution of different phytoplankton groups to the TChl*a* of the whole assemblage on the basis of the pigment ratio of each marker to the TChl*a* (Gieskes and Kraay, 1983; Gieskes et al., 1988; Barlow et al., 1993). Following this approach, Uitz et al. (2006), carried out a multiple regression analysis between the concentrations of TChl*a* and the seven diagnostic pigments suggested by Vidussi et al. (2001), providing the best estimates of the "Total Chlorophyll *a*–Diagnostic Pigments" ratios (TChl*a*/DPs) for a global data set. Applying this method, recently Di Cicco (2014) found a regional TChl*a*–DPs relationship, based on Mediterranean data only, to evaluate the different pigment ratios of the phytoplankton assemblage that occur in this basin (Sammartino et al., 2015).

In this work, we revised this regional relationship defining new coefficients according to the new quality assurance applied to the SeaBASS data. The analysis is carried out on the 1,379 individual samples where TChl*a* and all the seven selected biomarker pigments were available at the same time. It is important to underline that in accordance with Hooker et al. (2012) we defined TChl*a* as the sum of Chlorophyll *a* with its allomers and epimers, Divinyl-Chlorophyll *a,* and Clorophyllide *a* (see **Table 1**).

Table 2 presents the best estimates resulting from the multiple regression analysis for the determination of the seven Mediterranean TChl*a*/DPs ratios. The coefficients for each DP with their standard deviation and significance level are shown. The regression is highly significant, with a determination coefficient (r^2) between the SeaBASS *in-situ* TChl*a* and TChl*a* estimated (TChl*a**, **Table 3**) equal to 0.99, and a $p < 0.001$ (based on the *t*-test).

The final estimation formulas used for the *in-situ* quantification of each PFT and PSC fractions are schematically presented in **Table 3** (each group is expressed as fraction of TChl*a**).

PSC and PFT Model Development

The *in-situ* dataset of PSC and PFT fractions resulted from the DPA has been randomly divided in two independent subsets, the first used for the model calibration (70% of the total data) and the remaining 30% for their validation. The existing co-variability founded between the accessory pigments linked to each fraction and the TChl*a* allows the use of the latter as an index of the phytoplankton assemblage structure (Chisholm, 1992; Hirata et al., 2011). For each PFT and PSC group the relative *in-situ* fractions were regressed against the corresponding $\log_{10}$-transformed *in-situ* TChl*a* concentrations

TABLE 2 | Best estimates (new coefficients) of the "TChl*a* to DP" ratios for the Mediterranean SeaBASS subset with their standard deviation and significance value (*p*-value).

Diagnostic pigments	New coefficients	Standard deviation	*p*-value
Fuco	1.60	0.01	<0.001
Peri	1.67	0.05	<0.001
Hex-fuco	1.18	0.01	<0.001
But-fuco	0.57	0.07	<0.001
Allo	2.70	0.04	<0.001
TChl*b*	0.88	0.04	<0.001
Zea	1.79	0.03	<0.001

TABLE 3 | PSCs and PFTs used in this work with their *in-situ* estimation formulas (fraction of TChl*a, ranging from 0 to 1) resulted from the Diagnostic Pigment Analysis.**

PFTs	Estimation formula	PSCs	Estimation formula
Diatoms	**1.60[Fuco]/[TChl*a**]**	**Micro**	$\frac{\mathbf{1.60[Fuco] + 1.67[Peri]}}{\mathbf{[TChl\mathit{a}^*]}}$
Dinophytes	**1.67[Peri]/[TChl*a**]**		
Cryptophytes	**2.70[Allo]/[TChl*a**]**	**Nano**	For TChl*a* $> 0.08\,mg\,m^{-3}$: $\frac{\mathbf{1.18[Hex \cdot fuco] + 0.57[But \cdot fuco] + 2.70[Allo]}}{\mathbf{[TChl\mathit{a}^*]}}$
Haptophytes[b] (Nanoflagellates)	$\frac{\mathbf{1.18[Hex \cdot fuco] + 0.57[But \cdot fuco]}}{\mathbf{[TChl\mathit{a}^*]}}$		[a]For TChl*a* $< 0.08\,mg\,m^{-3}$: $\frac{\mathbf{12.5[\mathit{TChla}]1.18[\mathit{Hex{\cdot}fuco}] + 0.57[\mathit{But{\cdot}fuco}] + 2.70[\mathit{Allo}]}}{\mathbf{[\mathit{TChla}^*]}}$
Green algae & Prochlorophytes	**0.88[TChl*b*]/[TChl*a**]**	**Pico**	For TChl*a* $> 0.08\ mg\ m^{-3}$: $\frac{\mathbf{0.88[TChl\mathit{b}] + 1.79[Zea]}}{\mathbf{[TChla^*]}}$
Prokaryotes	**1.79[Zea]/[TChl*a**]**		[a]For TChl*a* $< 0.08\ mg\ m^{-3}$: $\frac{\mathbf{(-12.5[TChl\mathit{a}] + 1)1.18[Hex \cdot fuco] + 0.88[TChl\mathit{b}] + 1.79[Zea]}}{\mathbf{[TChla^*]}}$

$\mathbf{[\mathit{TChla}^*]} = 1.60[\mathit{Fuco}] + 1.67[\mathit{Peri}] + 1.18[\mathit{Hex\text{-}fuco}] + 0.57[19'\mathit{But\text{-}fuco}] + 2.70[\mathit{Allo}] + 0.88[\mathit{TChlb}] + 1.79[\mathit{Zea}]$

[a] *Linear adjustment of Brewin et al. (2010) for the assignment of 19′ hexanoyloxyfucoxanthin to pico size class in the ultra-oligotrophic waters.*
[b] *The contribute of the But-Fuco is so low in Mediterranean data that Haptophytes can be considered the only component of the Nanoflagellates.*

(**Figure 2**), considering the log-normal distribution of this pigment (Campbell, 1995). We used the ordinary least square fit to define the functional forms that were better appropriate to represent the Mediterranean data distribution. Different functional forms were tested against our calibration dataset, starting from linear equation to more complex polynomial or exponential function, obviously also including the functions adopted by global PFT and PFC models. This allowed us to select the most appropriate functional forms, corresponding to those that better minimize the residual between the estimates and the observations. This results in six empirical relationships obtained by the regression technique, while the other three are derived as difference to maintain the mass balance. To obtain the TChl*a* concentration related to each PFT and PSC group is sufficient to just multiply the fraction for the *in-situ* TChl*a*.

Satellite Data and Processing

For the PSC and PFT determination from remote sensing we used the TChl*a* Mediterranean reprocessed product available from Copernicus Marine Environment Monitoring Service (CMEMS, see OCEANCOLOUR_MED_CHL_L3_REP_OBSERVATIONS_009_073 product). These data were produced by the CMEMS Ocean Color Thematic Assembling Centre (OCTAC) using the ESA OC-CCI (European Space Agency—Ocean Color Climate Change Initiative) processor. MERIS, MODIS-Aqua, and SeaWiFS observations were merged into a single data by applying a series of state-of-the-art algorithms, from the atmospheric correction to the band shift correction schemes (for a comprehensive overview of the ESA-CCI products see http://www.esa-cci.org). Remote Sensing Reflectance (Rrs) spectrum is used as input to compute surface TChl*a* (nominal resolution of 1 Km) via regional ocean color algorithm. The specific product used in this work, specialized for the Mediterranean Sea, is a merged Case 1—Case 2 product that takes into account the different optical properties of the offshore and inshore waters. Two different regional algorithms were applied on the reflectance: the MedOC4 algorithm (Volpe et al., 2007) for the case 1 waters, developed by the Group for Satellite Oceanography (GOS-ISAC) of the Italian National Research Council (CNR), and the AD4 (D'Alimonte and Zibordi, 2003), specialized for the case 2 ones. The exact identification of the two water types is performed by taking into account the whole light spectrum from blue to NIR bands for both two water types from *in-situ* data (D'Alimonte et al., 2003). For the waters with intermediate features, a weighted average of the two former algorithms was applied, based on the distance between the actual reflectance spectrum and the two reference reflectance spectra for case 1 and 2 waters, respectively.

For more details on the processing adopted by the data producers and the quality product assessment see Volpe et al. (2012) and http://marine.copernicus.eu/documents/QUID/CMEMS-OC-QUID-009-038to045-071-073-078-079-095-096.pdf.

In this work, we used 18 years (from 1998 to 2015) of daily TChl*a* to compute Mediterranean daily PFT and PSC maps using the new regional algorithms described in Section Empirical Algorithms for the Identification of the PFTs and PSCs: Calibration and Validation. Daily fields were then used to build-up Mediterranean PFT and PSC climatology. Taking into account the applicability range of our models (0.02–5.52 mg m^{-3}), in our processing we considered "good values" only the satellite TChl*a* data falling in this range, masking the outsider.

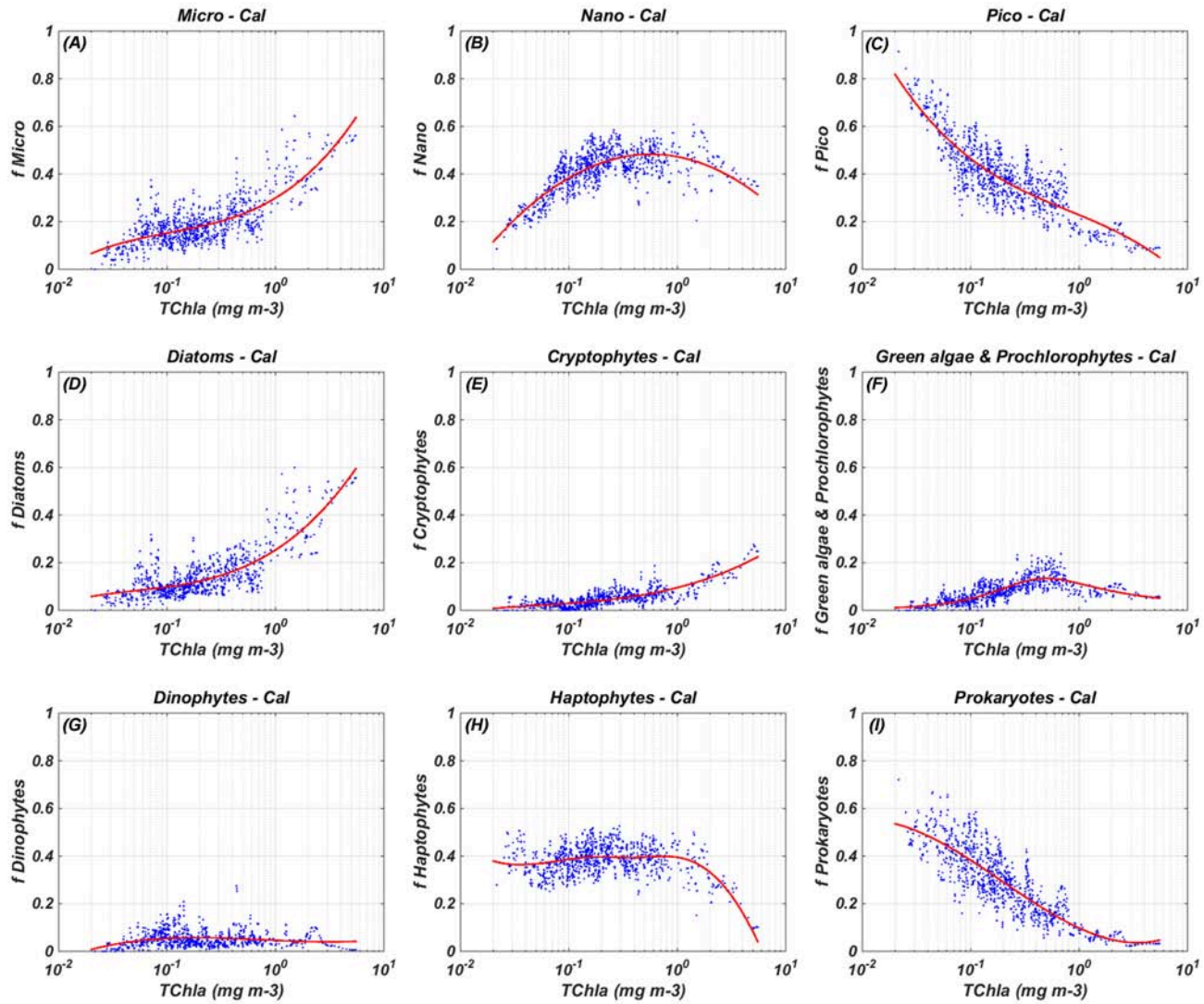


FIGURE 2 | Regional relationships between *in-situ* TChl*a* and the fraction (f) of each PSC and PFT: calibration (Cal) results. The blue dots refer to the *in-situ* PSC/PFT fraction obtained from equations in **Table 3** (966 data) and the red line indicates the best fitting curve obtained from the calibration (see equations in **Table 4**). **(A)** Micro-Cal, **(B)** Nano-Cal, **(C)** Pico-Cal, **(D)** Diatoms-Cal, **(E)** Cryptophytes-Cal, **(F)** Green algae & Prochlorophytes-Cal, **(G)** Dinophytes-Cal, **(H)** Haptophytes-Cal, and **(I)** Prokaryotes-Cal.

RESULTS

Empirical Algorithms for the Identification of the PFTs and PSCs: Calibration and Validation

The new regional algorithms with their mathematical equations and the resulting regression coefficients are showed in **Table 4**. Most of the considered phyto-groups are well outlined by simple polynomial functions (cubic for micro, Diatoms, Cryptophytes, and Prokaryotes and quadratic for nano), except for the class of "Green algae & Prochlorophytes," better represented by a different equation following the approach of Hirata et al. (2011, see **Table 4**). These functions, shown in **Table 4**, are applicable over a TChl*a* range from 0.02 to 5.52 mg m^{-3}.

Figure 2 shows the results of the algorithm calibration. Microphytoplankton function (**Figure 2A**) increases monotonically with the increase of the TChl*a*, ranging from the 8% to the 63% of the TChl*a* concentration. Pico equation (**Figure 2C**), instead, shows an opposite behavior, with minimum and maximum values for maximum and minimum TChl*a* concentrations, respectively, ranging between 5 and 81% of TChl*a*. The nano function (**Figure 2B**) presents an intermediate trend, ranging from 12 to 48% of TChl*a*, with a maximum in correspondence of about 0.57 mg m^{-3} of TChl*a*. The micro component consists almost entirely of the Diatom group contribution, represented by a cubic function (**Figure 2D**) similar to the micro one, increasing monotonically with the TChl*a* too. The contribution of the Dinophytes (**Figure 2G**) to the micro component and to the

TABLE 4 | Regional algorithms developed to estimate the PSCs and PFTs in the Mediterranean Sea (as fraction of TChl*a*, ranging from 0 to 1). For each dimensional and functional group, the equation and its relative coefficients are given.

PSCs and PFTs	Function	a	b	c	d
Micro	$ax^3 + bx^2 + cx + d$	0.0667	0.1939	0.2743	0.2994
Nano	$bx^2 + cx + d$		−0.1740	−0.0851	0.4725
Pico	$1 - micro - nano$	–	–	–	–
Diatoms *[1]*	$ax^3 + bx^2 + cx + d$	0.0482	0.1877	0.2946	0.2533
Dinophytes *[2]*	$micro - diatoms$	–	–	–	–
Cryptophytes *[3]*	$ax^3 + bx^2 + cx + d$	0.0171	0.0667	0.1153	0.0952
Green algae & Prochlorophytes *[4]*	$[\exp(ax + b) + cx]^{-1}$	−1.5780	2.1841	22.6833	
Prokaryotes *[5]*	$ax^3 + bx^2 + cx + d$	0.0664	0.1410	−0.2097	0.0979
Haptophytes *[6]*	$1 - micro - [3] - [4] - [5]$	–	–	–	–

$x = log_{10}(TChla)$

TChl*a* concentration is very low, with a small range of variation between 1 and 6% of TChl*a*, indicating that diatoms are the major constituent of the micro-phytoplankton in the Mediterranean Sea. Prokaryotes curve (**Figure 2I**) decreases monotonically from 55 to 3%, co-varying with the TChl*a* increase. At lower chlorophyll concentrations Prokaryotes represents the main component of the pico group (with the contribution of small Haptophytes in the ultra-oligotrophic water). Increasing the chlorophyll value, the non-monotonic signal of the "Green algae & Prochlorophytes" (**Figure 2F**) grows up to the maximum value (about 13%) in correspondence of the TChl*a* concentration of about 0.5 mg m^{-3}. For higher value of TChl*a* the function decreases with a weaker slope, concurring to the pico group more than to the Prokaryotes. Also the Cryptophytes (**Figure 2E**) co-vary with the TChl*a*, growing up with the increment of this pigment from a minimum of 1% to a maximum of 22% at higher TChl*a* values. At last but not least in terms of relative contribution to the TChl*a*, the Coccolithophores curve (**Figure 2H**) presents a small range of variation (35–40%) for almost the entire range of chlorophyll, decreasing up to a minimum value about 10% at maxima TChl*a* concentrations (about 5 mg m^{-3}).

The results of the application of the Mediterranean algorithms (**Table 4**) on the validation dataset are shown in **Figures 3**, **4** (right panels), for the PSCs and the PFTs, respectively. The scatter plots of the TChl*a*, estimated for each class applying the algorithms against the observed TChl*a* fractions, clearly show the goodness of the fits for all the considered groups. The data points are uniformly distributed around the 1:1 line with a very narrow scatter.

A more quantitative evaluation of the proposed algorithm performances comes from the computation of the mean absolute error (root mean squared error, RMSE) and other statistical parameters (see **Table 5** for the relative reference equations) with respect to the original PSC and PFT *in-situ* data (**Table 6**, calibration; **Table 7**, validation). A hindcast evaluation of the algorithm performances was also carried out (i.e., the same calibration data were used for fitting and testing). Furthermore, the error relative to the new regional algorithms is compared to the error associated with global abundance-based models applying them to the same validation dataset (**Table 7**). In particular, we used the empirical global relationships of Hirata et al. (2011), the only ones based on the abundance which are focused also on the PFTs, and the Brewin et al. (2010) models, applying the coefficients recalibrated in Brewin et al. (2011), developed only for the PSCs.

The new regional algorithms show good performances for most of the groups taken into account (**Table 6**, calibration; **Table 7**, validation). The results obtained applying the algorithms to the validation dataset are consistent with the hindcast evaluation. Pearson correlation coefficient, which gives an estimate of the covariance between the models and the *in-situ* validation data, shows high correlation, with values ranging from 0.75 to 0.99 both for PSCs and PFTs, excluding the group of the Dinophytes ($r = 0.60$), probably also because of this group is derived as difference. All groups show very low values of mean bias error (MBE), ranging from −0.002 to 0.003 mg m^{-3} for the validation dataset. The RMSE, which gives a measure of the spread of the estimated values around the *in-situ* observed ones, goes from 0.018 mg m^{-3} for the "Green algae & Prochlorophytes" to 0.068 mg m^{-3} for the Diatoms in the PFT group and from 0.042 mg m^{-3} for the pico- to 0.070 mg m^{-3} for the micro-phytoplankton in the PSCs.

The comparison of the scatter plot obtained by the application of the PSC regional models with respect to the global models (**Figure 3**) shows that the Mediterranean algorithms perform better than the global ones for all the three groups. This evidence is confirmed by the statistical analysis (**Tables 6, 7**). Although, the Brewin et al. (2011) models applied to the Mediterranean data show high values of the correlation coefficient (0.9 for micro- and nano-phytoplankton), the statistical results highlighted that the regionalization improves the uncertainty (MBE) and the spread (RMSE) of about one order of magnitude for all the size classes. For example, for the micro-phytoplankton the MBE decreases from 0.068 to 0 when the regional algorithm is applied. In particular, the Brewin's algorithm slightly overestimates the micro component (**Figure 3A**) in the entire dynamical range

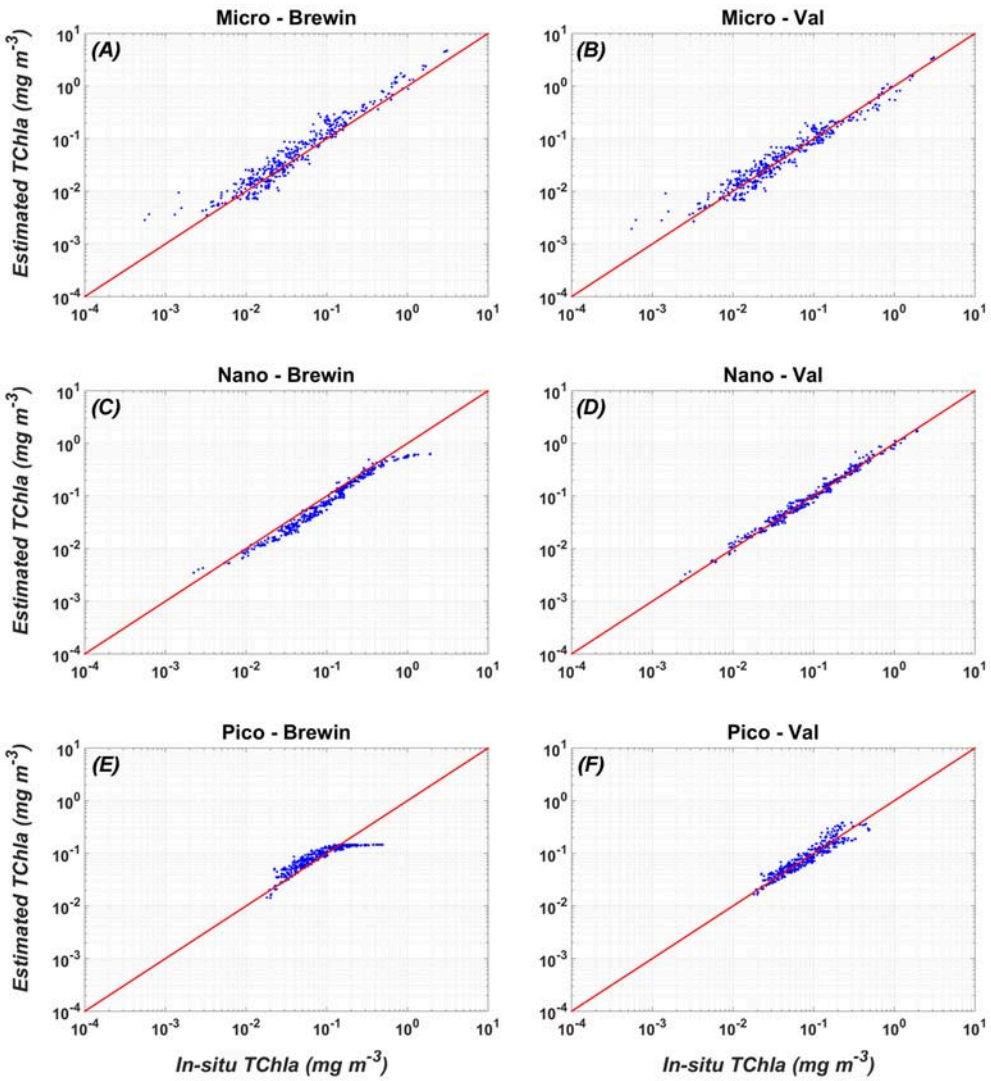


FIGURE 3 | Comparison between the validation (*in-situ* vs. estimated PSC TChl*a* concentrations) of the new PSC regional algorithms (Val, right panel: B,D,F) vs. the global PSCs model of Brewin et al. (2011) (Brewin, left panel: **A,C,E**). For the statistics, see **Table 7**.

of concentration. The behavior of this global model is exactly the opposite for the nano class, always underestimating the observed values (**Figure 3C**) resulting into a MBE of −0.056. As consequence, the main trend is the overestimation of the pico-phytoplankton component for concentrations lower than 0.1 mg m^{-3}, and the underestimation at greater values (**Figure 3E**).

Figure 4 shows that applying the global models of Hirata et al. (2011) to the Mediterranean data Prokaryotes (**Figure 4M**), Haptophytes (**Figure 4G**), and Diatoms (**Figure 4A**) would be underestimated. This underestimation results into a mean relative percentage difference, RPD, of −20, −19, and −29%, respectively. The "Green algae & Prochlorophytes" (**Figure 4I**), instead, are overestimated (RDP = 116%). The predictive power for the Dinophytes (**Figure 4C**) is negligible, as in the global validation of the model (see (Hirata et al., 2011) for more details), with an $r = 0.26$ (**Table 7**). It must be taken into account that, unlike to this work, the development of the models of Diatoms and Haptophytes in Hirata et al. (2011) is based on *in-situ* Fuco and Hex-fuco data at which a background correction was applied. The Fuco signal in oligotrophic waters (<0.25 mg m^{-3}) is assumed to be due to smaller Haptophytes rather than Diatoms. This correction is significant only at lower TChl*a* concentrations. It means that, applying this global model in this TChl*a* range, the estimates of Diatoms could be slightly improved for the Mediterranean Sea but, at the same time, the estimates of the Haptophytes would get worse. As for the PSCs, also for the PFTs (**Figure 4**) the regionalization reduces the bias of about one order of magnitude for all the types (**Tables 6, 7**). The preliminary analysis of r, MBE, and RMSE has showed a very good predictive power for all the new regional models. The best performances seem to be associated with the algorithms

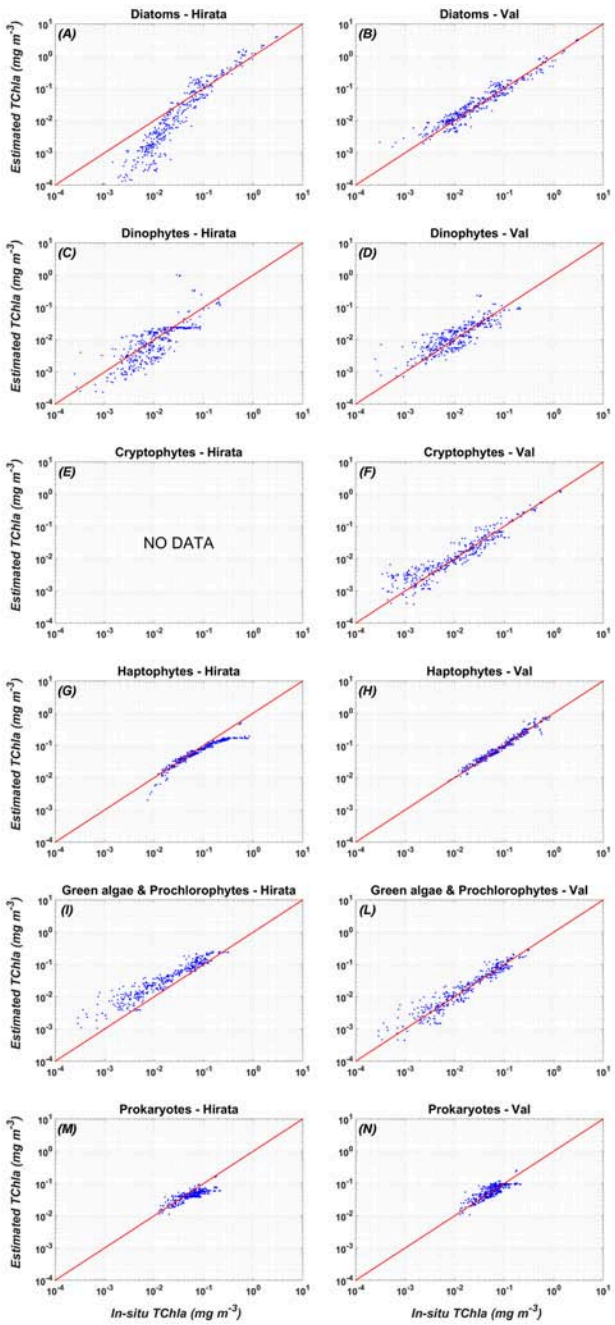


FIGURE 4 | Comparison between the validation of the new PFT regional algorithms (Val, right panel: B,D,F,H,L,N) vs. the global PFT model of Hirata et al. (2011) (Hirata, left panel: **A,C,E,G,I,M)**: *in-situ* (x axis) vs. estimated (y-axis) PFT TChla concentrations. **(A)** Diatoms-Hirata, **(B)** Diatoms-Val, **(C)** Dinophytes-Hirata, **(D)** Dinophytes-Val, **(E)** Cryptophytes-Hirata, **(F)** Cryptophytes-Val, **(G)** Haptophytes-Hirata, **(H)** Haptophytes-Val, **(I)** Green algae & Prochlorophytes-Hirata, **(L)** Green algae & Prochlorophytes - Val, **(M)** Prokaryotes - Hirata, **(N)** Prokaryotes - Val. For the statistics, see **Table 7**.

of Cryptophytes and "Green algae & Prochlorophytes" for the PFTs, followed by Haptophytes and Diatoms, and with the nano model for the PSC group. The study of the RPD and the mean absolute percentage difference, APD, has integrated this statistic information, taking also into account the different dynamical range of the TChl*a* concentration represented by each class. Weighing the uncertainty on the dynamical range of the observed concentration values, statistical data confirmed the goodness of the fits for all the phytoplankton groups (**Table 7**) and showed the best predicted power for the algorithms which estimate the nano (RDP = 3% and ADP = 12%) and Haptophytes (RDP = 4% and ADP = 13%) components, for PSCs and PFTs, respectively. These considerations are also confirmed by the validation results represented in **Figures 3**, **4**.

TABLE 5 | Mathematical equations used to compute the statistic parameters.

Mean bias error (MBE)	$MBE = \frac{1}{N}\sum_{i=1}^{N}(Model_i - Situ_i)$
Root mean squared error (RMSE)	$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Model_i - Situ_i)^2}$
Pearson's correlation coefficient (r)	$\frac{\sum_{i=1}^{N}\left(Model_i-\overline{Model_i}\right)\left(Situ_i-\overline{Situ_i}\right)}{\sqrt{\sum_{i=1}^{N}\left(Model_i-\overline{Model_i}\right)^2}\sqrt{\sum_{i=1}^{N}\left(Situ_i-\overline{Situ_i}\right)^2}}$
Mean relative percentage difference (RPD)	$RPD = \frac{1}{N}\sum_{i=1}^{N}\left(\frac{Model_i-Situ_i}{Situ_i}\right) x\ 100$
Mean absolute percentage difference (APD)	$APD = \frac{1}{N}\sum_{i=1}^{N}\left\lvert\frac{Model_i-Situ_i}{Situ_i}\right\rvert x\ 100$

TABLE 6 | Statistical results of the new regional algorithms (Med) applied to the calibration dataset (70% of the entire subset = 966 data).

PSCs and PFTs	Med (Calibration)				
	MBE	RMSE	*r*	RDP(%)	ADP(%)
Micro	0.001	0.054	0.988	11	28
Nano	−0.002	0.040	0.990	2	11
Pico	0.001	0.033	0.907	5	18
Diatoms	0	0.053	0.987	16	35
Dinophytes	0.001	0.023	0.646	47	71
Cryptophytes	−0.001	0.027	0.983	41	67
Green algae & Prochlorophytes	0	0.015	0.958	13	36
Prokaryotes	0	0.021	0.781	9	25
Haptophytes	−0.001	0.038	0.971	2	12

The statistic is computed on TChla concentration values (mg m^{-3}). MBE and RMSE are expressed in mg m^{-3}, while r, RPD (%), and APD (%) are dimensionless.

Application of the New Regional Algorithms to the Daily Mediterranean Reprocessed TChl*a* CASE1–2 Time Series: PSC and PFT Climatology (1998–2015)

The new regional algorithms (**Table 4**) are applied on an 18 years' time series of TChl*a* satellite estimates (see Section Data and Methods) to compute PSCs and PFTs. **Figures 5**, **6** show annual PSC and PFT (1998–2015) climatology, respectively.

TABLE 7 | Statistical results of the new regional algorithms (Med) validation compared with the statistics resulting by applying the global models (Brewin et al., 2011; Hirata et al., 2011) on the same validation dataset (30% of the entire subset = 413 data).

PSCs	Brewin et al. (2011)					Med (Validation)				
	MBE	RMSE	*r*	RDP (%)	ADP (%)	MBE	RMSE	*r*	RDP (%)	ADP (%)
Micro	0.068	0.219	0.989	36	45	0	0.070	0.985	11	29
Nano	−0.056	0.168	0.907	−23	24	0.001	0.045	0.989	3	12
Pico	−0.012	0.062	0.780	10	27	−0.001	0.042	0.875	4	19
PFTs	**Hirata et al. (2011)**					**Med (Validation)**				
Diatoms	0.039	0.164	0.980	−29	61	−0.001	0.068	0.983	19	38
Dinophytes	0.007	0.099	0.263	30	89	0	0.028	0.599	46	72
Cryptophytes	–	–	–	–	–	−0.002	0.027	0.990	41	66
Green algae & Prochlorophytes	0.022	0.035	0.909	116	117	0	0.018	0.946	13	35
Prokaryotes	−0.020	0.033	0.722	−20	29	−0.001	0.024	0.752	8	26
Haptophytes	−0.058	0.126	0.805	−19	24	0.003	0.048	0.956	4	13

The statistic is computed on TChla concentration values (mg m^{-3}). MBE and RMSE are expressed in mg m^{-3}, while r, RPD (%) and APD (%) are dimensionless.

In the left panel, each map shows the fractions of the TChl*a* represented by each phytoplankton component. For each pixel the percentage maps give fraction values relative to the chlorophyll concentration (**Figure 5**, top panel), whose distribution is typically characterized by a West–East decreasing gradient in the Mediterranean Sea (Siokou-Frangou et al., 2010; Estrada and Vaqué, 2014). In the right panel, instead, the maps show the relevance of each class in terms of TChl*a* estimates (mg m^{-3}).

All the three size classes reach their maxima absolute values, >3 mg m^{-3} of TChl*a* for micro (**Figure 5B**), and about 1.6 and 0.4 mg m^{-3} for nano (**Figure 5D**) and pico (**Figure 5F**), respectively, in the more productive zones of the basin (see **Figure 5**, top panel). In the eastern basin these areas are: the North Adriatic Sea and in general the whole Adriatic coast (due to the great nutrient supply from the Po river); the south-eastern area of the Levantine basin influenced by the outflow of the Nile river; the Northern Aegean Sea and the Gulf of Gabès (probably only an area of very shallow water). Otherwise, in the western basin, these more productive regions are the Gulf of Lion, the eastern cost of Spain, and the Tyrrhenian Sea coast. Very high values (greatest for the nano class) are also evident in the North–western Alborán Sea and along the Algerian and Tunisian coasts up to the Sicily channel, in the Liguro-Provençal and part of the Catalan Basin and in the cyclonic area of the North Tyrrhenian Sea. As expected, the existing co-variability between the accessory pigments linked to each fraction and the TChl*a* is highlighted by the percentage maps (**Figure 5**, left panel). In the more oligotrophic eastern Mediterranean Sea, where TChl*a* climatology shows the lower absolute concentrations (**Figure 5**, top panel), the relative dominant component is the pico class (**Figure 5E**). Here, the TChl*a* concentration of the pico-phytoplankton (**Figure 5F**) is about five times micro (**Figure 5B**) and two times the nano one (**Figure 5D**). On the contrary, in the regions where the TChl*a* reaches higher values, the percentage contribution of the nano and micro components increase (**Figures 5A,C**, respectively). Generally, in the whole basin the nano component shows intermediate values, in particular ranging from 30 to 40% of the TChl*a* in the western basin, growing up to 45% in highly productive areas (**Figure 5C**).

Moving to the PFTs, the abundance of each class in terms of TChl*a* concentration (**Figure 6**, right panel) reflects, as occurred in the size classes, the gradient of this pigment (**Figure 5**, top panel), showing for all groups higher values in the western basin and in the already mentioned high productive zones of the entire basin. This is true for all groups. In the climatological analysis, Prokaryotes constitute the principal component of the pico-phytoplankton in almost all areas, both in terms of percentage (**Figure 6M**) and concentration (**Figure 6N**). They are the absolute dominant group in the oligotrophic and ultra-oligotrophic waters of the eastern basin, but also for the western basin in the southern Tyrrhenian Sea and in some areas of the Algero-Provençal basin (**Figure 6M**). In the Levantine basin the second group in terms of TChl*a* concentration is the Haptophytes (**Figure 6H**). They represent the dominant class within the nano-phytoplankton in the whole Mediterranean Sea and constitute the main group featuring the case 1 water of the western basin (**Figure 6G**). Diatoms (**Figures 6A,B**) dominate in the micro-phytoplankton and can be considered the third group in terms of TChl*a* concentration in the open sea, followed by the Green algae (**Figure 6L**), Cryptophytes (**Figure 6F**) and, finally, Dinoflagellates (**Figure 6D**). Otherwise, in coastal areas Diatoms dominate, reaching values about of 3 mg m^{-3}. This is well-evident in the North Adriatic Sea and in general in the entire Adriatic coast, in the southeastern area of the Levantine basin influenced by the outflow of the Nile River and in the Gulf of Gabès (**Figure 6B**). In general, in terms of chlorophyll concentration, in the coastal areas we find the predominance of the Diatoms and Haptophytes, followed by the remaining classes (**Figure 6**, right panel).

At last but not least in terms of biological importance, the phytoplankton distribution in the Alborán Sea and along the Algerian-Tunisian coasts is characterized by the dominance of

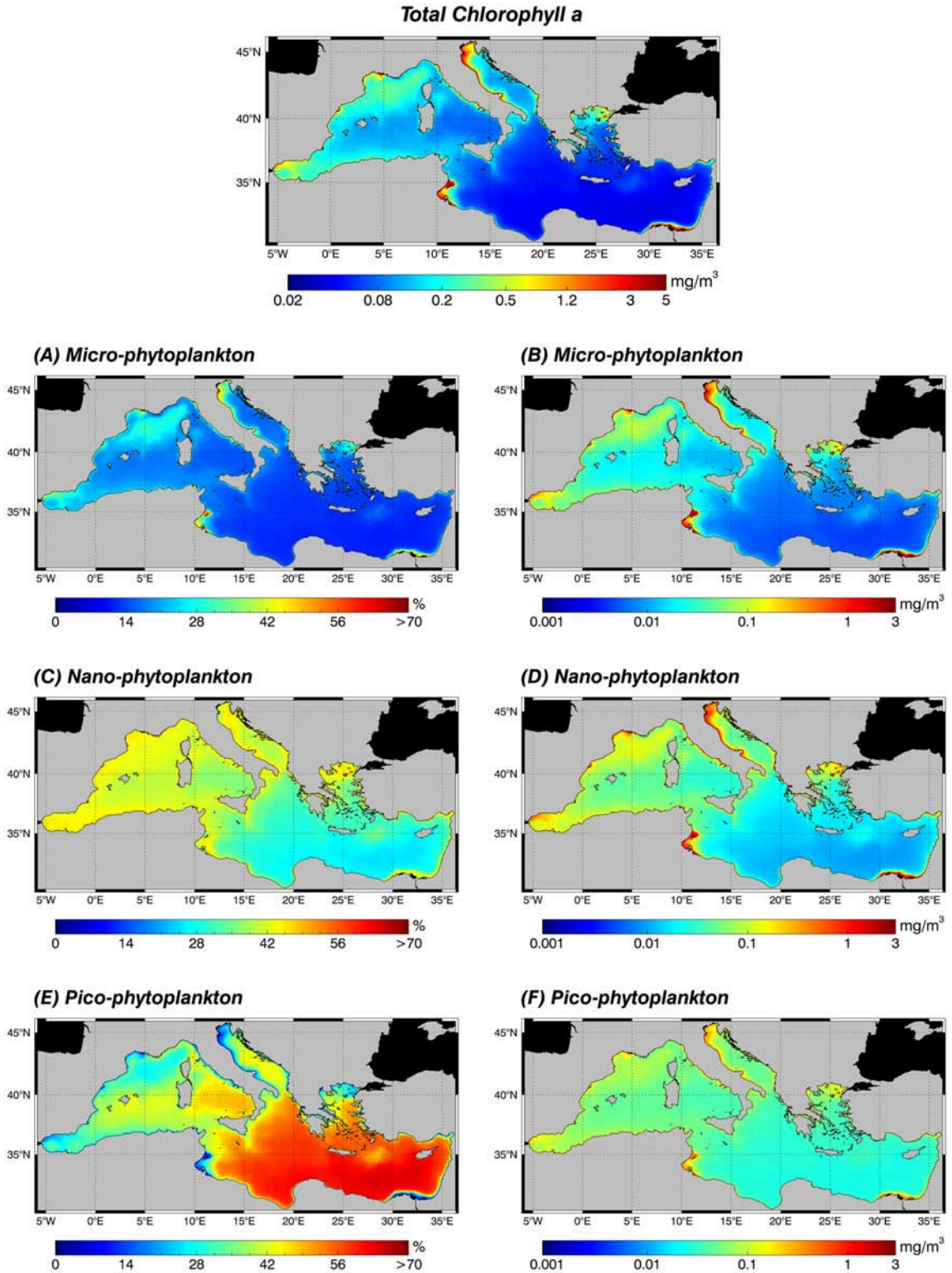


FIGURE 5 | Annual climatological maps (1998–2015) of the TChl*a* (top panel) and the PSCs: for each pixel, fractions are expressed in % with respect to the TChl*a* (left panel: A,C,E) and the contribution of each PSC group to the TChl*a* concentration in mg m^{-3} (right panel: **B,D,F**). PSCs are retrieved applying the new regional algorithms (see **Table 4**) on the daily TChl*a* times series (Mediterranean reprocessed product produced by the CMEMS-OCTAC). **(A,B)** Micro-phytoplankton, **(C,D)** Nano-phytoplankton, **(E,F)** Pico-phytoplankton.

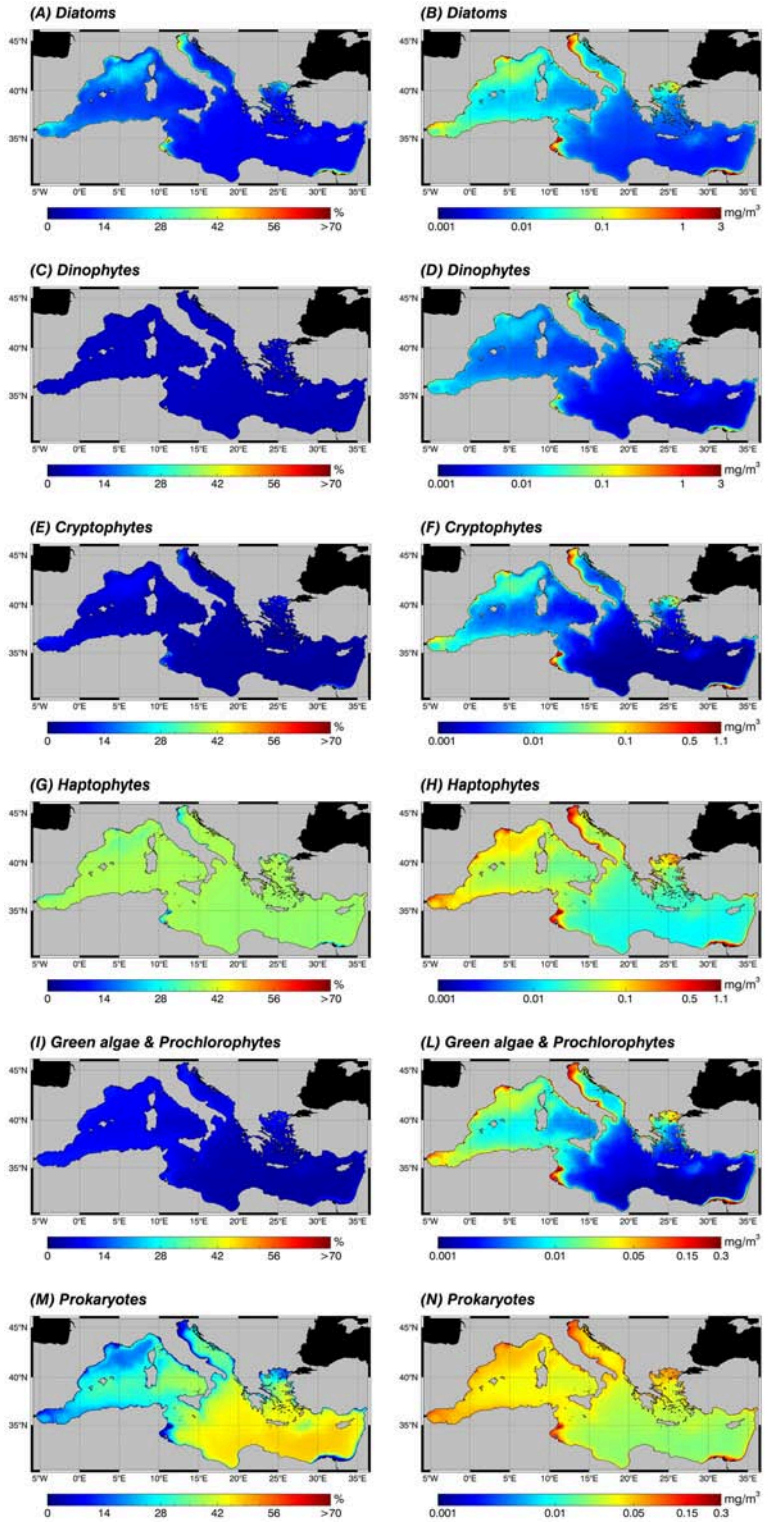


FIGURE 6 | Annual climatological maps (1998–2015) of the PFTs: for each pixel, fractions are expressed in % with respect to the TChl*a* (left panel: A,C,E,G,I,M) and the contribution of each PFT group to the TChl*a* concentration in mg m^{-3} (right panel: **B,D,F,H,L,N**). PFTs are retrieved applying the new regional algorithms (see **Table 4**) on the daily TChl*a* time series (Mediterranean reprocessed product produced by the CMEMS-OCTAC). Since the wide range of TChl*a* variability related to each PFT, note that the color scales could be different. **(A,B)** Diatoms, **(C,D)** Dinophytes, **(E,F)** Cryptophytes, **(G,H)** Haptophytes, **(I,L)** Green algae & Prochlorophytes, **(M,N)** Prokaryotes.

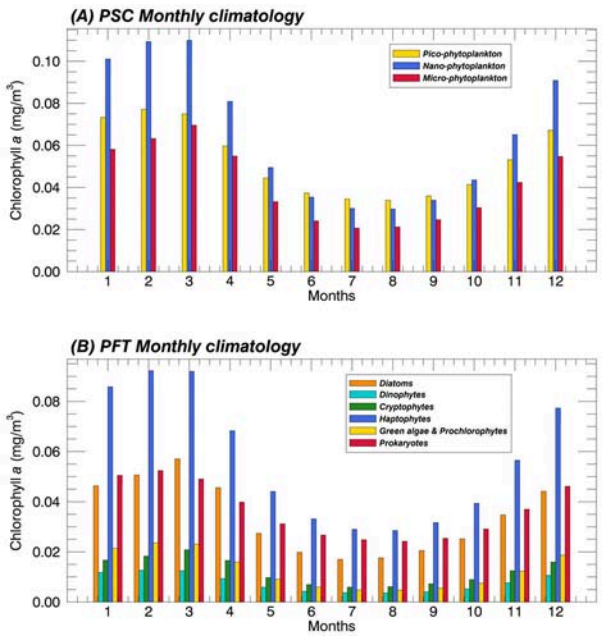


FIGURE 7 | Mean monthly climatology (1998–2015) of the PSCs (A) and the PFTs **(B)** over the Mediterranean Sea in terms of their contribution to the TChl*a* (mg m^{-3}). The PSCs and PFTs are retrieved applying the new regional algorithms (see **Table 4**) on the daily Total Chlorophyll *a* time series (Mediterranean reprocessed product produced by the CMEMS-OCTAC).

the Haptophytes, followed by Diatoms, and then Prokaryotes, Green algae, Cryptophytes, and at last Dinoflagellates.

In addition, in **Figure 7** we show the results of the monthly climatology of each group, averaged over the whole Mediterranean Sea. On basin scale, the component mostly representative of the TChl*a* seems to be the nano-phytoplankton for the PSCs (top panel), especially in the bloom periods typical of the midlatituds. Nano is followed by pico and the component with the lower contribution to the TChla is the micro one. Only in summer, pico-phytoplankton dominates on the TChl*a* concentration, exceeding the other two classes. In the same season, the micro component reaches its minimum values. The monthly mean PFT climatology (**Figure 7**, bottom panel) confirms the predominance of the Haptophytes within the nano-phytoplankton in the whole Mediterranean Sea as highlighted in the previous PFT map analysis (**Figure 6H**). This is also the predominant group all over the year. The Haptophytes are followed by Diatoms and Prokaryotes representing the main component for the micro- and the pico-phytoplankton, respectively. More in detail, the contribute of the Diatoms to the TChl*a* concentration is greater than the Prokaryotes only in the early spring. The two classes show similar concentrations in late autumn and winter season while, in the remaining part of the year, Prokaryotes dominate. Cryptophytes and "Green algae & Prochlorophytes" always reveal a same contribution, even if the concentration of the latter is slightly greater in late winter—early spring. These two functional groups represent the smallest fractions within the nano- and pico- size classes, respectively. At last, Dinophytes constitute the class with the lowest TChl*a* concentration all over the year.

DISCUSSION AND CONCLUSIONS

The Mediterranean Sea is typically characterized by peculiar optical properties that make its color different from the global ocean (Volpe et al., 2007). In addition to an abundant aerosol dominated by continental anthropogenic pollution (Moulin et al., 1997) and the presence of Saharan dust in the water column (Claustre et al., 2002), one of the main reasons that justifies its color seems to be a different phytoplankton assemblage structure typical of this basin (Volpe et al., 2007). This is also confirmed by the presence of pigment ratios different with respect to those of the global ocean (Sammartino et al., 2015). This implies the need of regional algorithms that take into account all these peculiar characteristics. In the last year, several specialized algorithms have been proposed for the detection of the chlorophyll *a* concentration (e.g., Volpe et al., 2007; Santoleri et al., 2008). Instead, PFT and PSC regional algorithms do not exist, except for the recent work of Navarro et al. (2014). They adapted the previous version of the PHYSAT method of Alvain et al. (2005, 2008), providing regional estimates of dominant PFT groups. In our work, for the first time, new regional algorithms have been advanced to identify, together, the contribution of each PSC and PFT group to the satellite estimates of TChl*a* concentration. This different approach, based on the close link existing between the abundance of each group and the trophic status of the environment (Margalef, 1967, 1978; Brewin, 2011), provides new kind of information, complementary to the results of the PHYSAT-Med.

Our assessment of the uncertainty associated to the new developed regional algorithms and the most used global models based on the same approach, highlight and confirm that a regionalization for the PSC and PFT satellite algorithms is required. As shown by our validation results (Section Empirical Algorithms for the Identification of the PFTs and PSCs: Calibration and Validation), the use of Mediterranean PSC and PFT algorithms allowed to eliminate the bias between observations and estimates and to reduce the RMSE of an order of magnitude respect the global models.

Even if the uneven distribution of the *in-situ* observations between western-central Mediterranean Sea and the eastern basin could imply that the new formulations are more appropriate for the western basin, we are confident that the derived parameterizations can be applied also in the eastern Mediterranean Sea without introducing a significant bias on satellite estimates. In fact, the *in-situ* dataset used for the algorithm calibration includes the typical values of chlorophyll *a* observed in the oligotrophic waters of eastern Mediterranean Sea (ranging from 0.02 to 0.14 mg m^{-3}). The number of the samples that fall in this chlorophyll range represent the 38% of the total number of the calibration data, 18% of which are acquired in the eastern Mediterranean Sea. This implies that the oligotrophic condition is well represented in our dataset. For

a further assessment, we made a preliminary evaluation of the new parameterizations limited to the eastern Mediterranean Sea using all available *in-situ* observation in our dataset. This results into a bias (from −0.001 to 0.001 depending on the PFT/PCS parameterization) and RMS (from 0.002 to 0.004 depending on the PFT/PCS parameterization), values comparable with the bias obtained in western Mediterranean for the same TChl*a* range and with the values resulting from the algorithm validation (see **Table 7**). Even if this result cannot be considered conclusive since has been obtained with a limited number of *in-situ* observations, the statistical results seem to indicate that our parameterization should not introduce any significant bias to satellite derived estimates.

The analysis of the phytoplankton assemblage distribution patterns resulted from the application of our new algorithms to the Mediterranean multi-sensor reprocessed dataset (1998–2015) is consistent with the main previous knowledge, both in terms of distribution and phytoplankton ecology (Siokou-Frangou et al., 2010; Uitz et al., 2012; Estrada and Vaqué, 2014; Navarro et al., 2014). Pico-phytoplankton, with Prokaryotes as the main component, is widespread throughout the whole basin and always dominant in oligotrophic and stratified waters (see Section Results) in agreement with the observation of Siokou-Frangou et al. (2010), according to which the pico component constitutes more than 50% of the total biomass in these conditions (Estrada and Vaqué, 2014). Furthermore, our results show that pico class reaches its maximum value (about 0.4 mg m^{-3}) in the more productive areas. These considerations are also in consonance with ecological behavior and strategy of this group. Size affects nutrient solute and water fluxes across the plasmalemma, favoring the smaller sized cells in the oligotrophic water. This is due to the larger surface to volume ratio of small cells with respect to the larger ones, which make the former efficient nutrient absorber in very low nutrient conditions. Moreover, in stratified environments the probability of sinking out from the euphotic zone is greater in the micro-sized cells than in the smaller ones, undergoing a lower loss of organisms (Chisholm, 1992; Raven, 1998). However, this does not imply that Prokaryotes and pico-phytoplankton in general reach the maxima values in terms of chlorophyll concentration in the oligotrophic conditions. In fact, Chisholm (1992) suggested that they usually achieve their "maximum potential biomass" (of about 0.5 mg m^{-3}) in high nutrient conditions, in accordance with our estimates. We showed that, in the more productive region of the basin, the chlorophyll concentrations of the pico and Prokaryotes classes correspond to relative small percentage of TChl*a* (Chisholm, 1992). Indeed, our climatological maps shows the dominance, in these conditions, of the micro component with values that exceeds up to the 50% of the TChl*a*, followed by nano- and, at last, by a minor contribute of the pico-phytoplankton (about 10–15%).

Within the micro group, the major contribute is clearly due to the Diatoms in the higher nutrient areas. This is justified by the ecological strategy of this functional group, physiologically better adapted to high dynamic conditions and more efficient in the nutrient absorption. Moreover, they are also able to subtract nutrients from the surrounding environment and to store them in their large vacuoles, depriving other groups and supporting their growth at the same time (Margalef, 1978; Falkowski et al., 2003; Litchman et al., 2007; Estrada and Vaqué, 2014). About the Dinoflagellates, their contribution to the micro-phytoplankton is very low with respect to the Diatoms one. This is probably due to their different ecological strategy, well-adapted to high dynamical environment but with a higher affinity for low nutrient conditions. Moreover, Estrada and Vaqué (2014) suggested that the use of peridinin as biomarker pigment for the Dinoflagellates identification could cause an underestimation of their abundance because it could be not present in some organisms of this class (Jeffrey and Vesk, 1997).

An important result of this study is the information on the Nanoflagellate distribution, mainly represented by the Haptophytes in our dataset. This can represent a precious novelty, considering the lack of knowledge on the Nanoflagellate spatial distribution, improved only in the last years thanks to the more diffused usage of chemotaxonomic and molecular techniques (Latasa et al., 2010). The widespread distribution of the nano component and its high contribution to the TChl*a* in the whole basin confirms the Uitz et al. (2012) results, according to which the primary production in the Mediterranean Sea is mainly due to the nano-phytoplankton component.

The lack of data on the phytoplankton biogeography at different spatio-temporal scales in the whole Mediterranean Sea and the well-known difficulties in the long-term acquisition of *in-situ* data at basin scale make essential the use of the remote sensing technique for a synoptic observation of the phytoplankton assemblage composition and its diversity. Our analysis revealed the importance of providing regional algorithms strictly required to suit the peculiar bio-optical properties featuring this basin. The statistical results demonstrated the goodness of the performance and the applicability of our models for the abundance estimations of PSCs and PFTs together.

Nowadays, in the context of international Climate Change Initiatives and cooperation, a synergic effort of the Space Agencies in collaboration with remote sensing scientist is conducted to identify the major gaps (both instrumental and scientific) that should be filled to improve the accuracy of satellite estimates of the phytoplankton groups and their variability (Bracher et al., 2017). In this framework, the following actions summarize our future perspectives to improve remote observations on the Mediterranean Sea: (a) to extend the validation and calibration of the PFT new regional algorithms including new *in-situ* dataset of HPLC Total Chlorophyll *a* and diagnostic pigments acquired, in recent years, by the Mediterranean scientific community; (b) to improve the accuracy of the PSC algorithms with a new calibration and validation only based on TChl*a* size-fractions; (c) to carry on the *in-situ* bio-optical measurements to cover all the un-sampled Mediterranean regions, also with the intent to exploit different approaches (e.g., spectral response-based); (d) to extend this regionalization activity to new generation sensors (e.g., OLCI for Sentinel-3) to obtain higher resolution information also

for phytoplankton dynamical studies at mesoscale; (e) to analyse the Mediterranean PFT and PSC trends, thanks to the availability of consistent long term satellite observation time series.

On time scales larger than the period we considered, climate or human induced changes in environmental conditions can produce modifications of phytoplankton pigment composition and thus the pigment ratios to the Total chlorophyll a. This implies that the simple empirical relations used to compute the PFTs and PFCs from the chlorophyll observations need to be re-evaluated and or a more sophisticated approach which links the pigment ratios, the PFT, and PSC composition and the major environmental forcing should be developed.

AUTHOR CONTRIBUTIONS

All the authors contributed to the conception and design of the work and approved the final version of the manuscript to publish. AD carried out the work, performing the Diagnostic Pigment Analysis, the algorithm development and their analysis supported by the experience of RS and SM and drafting the most of the manuscript. MS processed the satellite data, developed all the climatological maps and edited all the images of the manuscript. MS, SM, and RS revised critically the manuscript and contributed to its draft.

FUNDING

This research was supported by the European Commission in the framework of the Copernicus Marine Environmental Services—Ocean Color Thematic Assembling Center Project (Grant agreement 9836100). The research was also supported by the "Ministero dell'Istruzione, dell'Università e della Ricerca" in the framework of the Italian Flagship Project RITMARE (la Ricerca ITaliana per il MARE).

ACKNOWLEDGMENTS

The authors would like to thank the OC-CCI and CMEMS Project that generated the satellite chlorophyll data used in this paper, which are free available at http://marine.copernicus.eu/. We also acknowledge the SeaBASS archive for the *in-situ* bio-optical dataset, free available at https://seabass.gsfc.nasa.gov/. We want to thank ESA/ESRIN for the "Phytoplankton Diversity at Global and Regional Scale" session, within the "Color and Light from Earth Observation" (CLEO) workshop. We are grateful to Simone Colella for his invaluable scientific and technical advices. This work also benefited from discussions with Alessio Ansuini, Vittorio Brando, Riccardo Droghei, Marco Picone, Jaime Pitarch Portero, and Gianluca Volpe. We also would like to thank Vega Forneris and Flavio La Padula for their valuable technical support.

REFERENCES

Aiken, J., Pradhan, Y., Barlow, R., Lavender, S., Poulton, A., Holligan, P., et al. (2009). Phytoplankton pigments and functional types in the Atlantic Ocean: a decadal assessment, 1995-2005. *Deep Sea Res. II* 56, 899–917. doi: 10.1016/j.dsr2.2008.09.017

Alvain, S., Moulin, C., Dandonneau, Y., and Bréon, F. M. (2005). Remote sensing of phytoplankton groups in case 1 waters from global SeaWiFS imagery. *Deep Sea Res. I* 52, 1989–2004. doi: 10.1016/j.dsr.2005.06.015

Alvain, S., Moulin, C., Dandonneau, Y., and Loisel, H. (2008). Seasonal distribution and succession of dominant phytoplankton groups in the global ocean: a satellite view. *Global Biogeochem. Cycles* 22:GB3001. doi: 10.1029/2007GB003154

Barlow, R. G., Mantoura, R. F. C., Gough, M. A., and Fileman, T. W. (1993). Pigment signatures of the phytoplankton composition in the northeastern Atlantic during the 1990 spring bloom. *Deep Sea Res II* 40, 459–477. doi: 10.1016/0967-0645(93)90027-K

Basset, A., Sangiorgio, F., and Pinna, M. (2004). Monitoring with benthic macroinvertebrate: advantage and disadvantages of body size descriptors. *Aquat. Conserv* 14, S43–S58. doi: 10.1002/aqc.649

Blondel, J. (2003). Guilds or functional groups: does it matter? *Oikos* 100, 223–231. doi: 10.1034/j.1600-0706.2003.12152.x

Bracher, A., Bouman, H., Brewin, R. J. W., Bricaud, A., Brotas, V., Ciotti, A. M., et al. (2017). Obtaining phytoplankton diversity from ocean color: a scientific roadmap for future development. *Front. Mar. Sci.* 4:55. doi: 10.3389/fmars.2017.00055

Brewin, R. J., Sathyendranath, S., Hirata, T., Lavender, S. J., Barciela, R. M., and Hardman-Mountford, N. J. (2010). A three-component model of phytoplankton size class for the Atlantic Ocean. *Ecol. Modell.* 221, 1472–1483. doi: 10.1016/j.ecolmodel.2010.02.014

Brewin, R. J. W. (2011). *Detecting Phytoplankton size Class Using Satellite Earth Observation*. Ph.D. Thesis, University of Plymouth.

Brewin, R. J. W., Devred, E., Sathyendranath, S., Lavender, S. J., and Hardman-Mountford, N. J. (2011). Model of phytoplankton absorption based on three size classes, *Appl. Opt.* 50, 4353–4364. doi: 10.1364/AO.50.004535

Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M., and West, G. B. (2004). Toward a metabolic theory of ecology. *Ecology* 85, 1771–1789. doi: 10.1890/03-9000

Brunet, C., and Mangoni, O. (2010). "Determinazione quali-quantitativa dei pigmenti fitoplanctonici mediante HPLC," in *Metodologie di Studio del Plancton Marino*, Manuali e linee guida 56/2010, eds G. Socal, I. Buttino, M. Cabrini, O. Mangoni, A. Penna, and C. Totti (Rome: ISPRA), 379–387.

Campbell, J. W. (1995). The lognormal distribution as a model for bio-optical variability in the sea. *J. Geophys. Res.* 100, 13237–13254. doi: 10.1029/95JC00458

Chisholm, S. W. (1992). "Phytoplankton size," in *Primary Productivity and Biogeochemical Cycles in the Sea*, eds P. G. Falkowski and A. D. Woodhead (New York, NY: Plenum Press), 213–237.

Ciotti, A. M., and Bricaud, A. (2006). Retrievals of a size parameter for phytoplankton and spectral light absorption by coloured detrital matter from water-leaving radiances at SeaWiFS channels in a continental shelf off Brazil. *Limnol. Oceanogr. Methods* 4, 237–253. doi: 10.4319/lom.2006.4.237

Ciotti, A. M., Lewis, M. R., and Cullen, J. J. (2002). Assessment of the relationships between dominant cell size in natural phytoplankton communities and the spectral shape of the absorption coefficient. *Limnol. Oceanogr.* 47, 404–417. doi: 10.4319/lo.2002.47.2.0404

Claustre, H. (1994). The trophic status of various oceanic provinces as revealed bv phytoplankton pigment signatures. *Limnol. Oceanogr.* 39, 1206–1210. doi: 10.4319/lo.1994.39.5.1206

Claustre, H., Morel, A., Hooker, S. B., Babin, M., Antoine, D., Oubelkheir, K., et al. (2002). Is desert dustmaking oligotrophic waters greener? *Geophys. Res. Lett.* 29, 107.1–107.4. doi: 10.1029/2001GL014056

D'Alimonte, D., Mélin, F., Zibordi, G., and Berthon, J.-F. (2003). Use of the novelty detection technique to identify the range of applicability of the empirical

ocean color algorithms. *IEEE Trans. Geosci. Remote Sens.* 41, 2833–2843. doi: 10.1109/TGRS.2003.818020

D'Alimonte, D., and Zibordi, G. (2003). Phytoplankton determination in an optically complex coastal region using a multilayer perceptron neural network. *IEEE Trans. Geosci. Remote Sens.* 41, 2861–2868. doi: 10.1109/TGRS.2003.817682

Delwiche, C. (1999). Tracing the thread of plastid diversity through the tapestry of life. *Am. Nat.* 154, 164–177. doi: 10.1086/303291

Devlin, M., Best, M., Coates, D., Bresnan, E., O'Boyle, S., Park, R., et al. (2007). Establishing boundary classes for the classification of UK marine waters using phytoplankton communities. *Mar. Pollut. Bull.* 55, 91–103. doi: 10.1016/j.marpolbul.2006.09.018

Devred, E., Sathyendranath, S., Stuart, V., Maas, H., Ulloa, O., and Platt, T. (2006). A two-component model of phytoplankton absorption in the open ocean: theory and applications. *J. Geophys. Res.* 111:C03011. doi: 10.1029/2005JC002880

Di Cicco, A. (2014). *Spatial and Temporal Variability of Dominant Phytoplankton Size Classes in the Mediterranean Sea from Remote Sensing*. Ph.D. thesis in Ecology and Management of Biological Resources, Tuscia University.

Estrada, M., and Vaqué, D. (2014). "Microbial components," in *The Mediterranean Sea, Its History and Present Challenges,* eds S. Goffredo and Z. Dubinsky (New York, NY, London: Springer), 87–111.

Everitt, D. A., Wright, S. W., Volkman, J. K., Thomas, D. P., and Lindstrom, E. J. (1990). Phytoplankton community compositions in the western equatorial Pacific determined from chlorophyll and carotenoid pigment distributions. *Deep-Sea Res.* 37, 975–997. doi: 10.1016/0198-0149(90)90106-6

Falkowski, P. G., Laws, E. A., Barber, R. T., and Murray, J. W. (2003). "Chapter 4: Phytoplankton and their role in primary, new, and export production," in *Ocean Biogeochemistry: A Synthesis of the Joint Global Ocean Flux Study (JGOFS),* Global change–the IGBP series, ed J. R. Michael Fasham (Berlin; Heidelberg: Springer-Verlag), 1619–2435.

Falkowski, P. G., and Raven, J. A. (1997). *Aquatic Photosynthesis*. Oxford: Blackwell Scientific Publishers.

Gieskes, W. W. C., and Kraay, G. W. (1983). Dominance of Cryptophyceae during the phytoplankton spring bloom in the central North Sea detected by HPLC analysis of pigments. *Mar. Biol.* 75, 179–185. doi: 10.1007/BF00406000

Gieskes, W. W. C., Kraay, G. W., Nontji, A., and Setiapermana, D. (1988). Monsoonal alternation of a mixed and a layered structure in the phytoplankton of the euphotic zone of the Banda Sea (Indonesia): a mathematical analysis of algal pigment fingerprints. *Neth. J. Sea Res.* 22, 123–137. doi: 10.1016/0077-7579(88)90016-6

Hirata, T., Aiken, J., Hardman-Mountford, N., Smyth, T. J., and Barlow, R. G. (2008). An absorption model to determine phytoplankton size classes from satellite ocean colour. *Remote Sens. Environ.* 112, 3153–3159. doi: 10.1016/j.rse.2008.03.011

Hirata, T., Hardman-Mountford, N. J., Brewin, R. J. W., Aiken, J., Barlow, R., Suzuki, K., et al. (2011). Synoptic relationships between surface Chlorophyll-a and diagnostic pigments specific to phytoplankton functional types. *Biogeosciences* 8, 311–327. doi: 10.5194/bg-8-311-2011

Hooker, S. B., Clementson, L., Thomas, C. S., Schlüter, L., Allerup, M., Ras, J. et al. (2012). *The Fifth SeaWiFS HPLC 779 Analysis Round-Robin Experiment (SeaHARRE-5)*. Greenbelt, MD: NASA Tech. Memo 780 2012–217503, NASA Goddard Space Flight Center, 98.

IOCCG (2014). "Phytoplankton functional types from space," in *Reports of the International Ocean-Colour Coordinating Group, No. 15*, ed S. Sathyendranath (Dartmouth, NS: IOCCG).

Jeffrey, S. W., and Vesk, M. (1997). "Introduction to marine phytoplankton and their pigment signatures," in *Phytoplankton Pigments in Oceanography: Guidelines to Modern Methods,* eds S. W. Jeffrey, R. F. C. Mantoura, and S. W. Wright (Paris: UNESCO), 37–84.

Kostadinov, T. S., Siegel, D. A., and Maritorena, S. (2009). Retrieval of the particle size distribution from satellite ocean color observations. *J. Geophys. Res.* 114, C09015. doi: 10.1029/2009JC005303

Lacombe, H., Gascard, J. C., Gonella, J., and Bethoux, J. P. (1981). Response of the Mediterranean to the water and energy fluxes across its surface, on seasonal and interannual scales. *Oceanol. Acta* 4, 247–255.

Latasa, M., Scharek, R., Vidal, M., Vila-Reixach, G., Emelianov, M., Gutiérrez-Rodríguez, A., et al. (2010). Preferences of phytoplankton groups for waters of different trophic status in the northwestern Mediterranean Sea. *Mar. Ecol. Prog. Ser.* 407, 27–42. doi: 10.3354/meps08559

Le Quéré, C., Harrison, S. P., Colin Prentice, I., Buitenhuis, E. T., Aumont, O., Bopp, L., et al. (2005). Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. *Glob. Change Biol.* 11, 2016–2040. doi: 10.1111/j.1365-2486.2005.1004.x

Litchman, E., Klausmeier, C. A., Schofield, O. M., and Paul Falkowski, G. (2007). The role of functional traits and trade-offs in structuring phytoplankton communities: scaling from cellular to ecosystem level. *Ecol. Lett.* 10, 1170–1181. doi: 10.1111/j.1461-0248.2007.01117.x

Margalef, R. (1967). "Some concepts relative to the organization of plankton," in *Oceanography and Marine Biology: An Annual Review,* Vol. 5 (London: Aberdeen University Press; Allen & Unwin), 257–289.

Margalef, R. (1978). Life-forms of phytoplankton as survival alternatives in an unstable environment. *Oceanol. Acta* 1, 493–509.

Mouillot, D., Spathris, S., Reizopoulou, S., Laugier, T., Sabetta, L., Basset, A., et al. (2006). Alternatives to taxonomic-based approaches to assess changes in transitional water communities. *Aquat. Conserv.* 16, 469–482. doi: 10.1002/aqc.769

Moulin, C., Dulac, F., Lambert, C. E., Chazette, P., Jankowiak, I., Chatenet, B., et al. (1997). Long-term daily monitoring of Saharan dust load over ocean using Meteosat ISCCPB2 data.2. Accuracy of the method and validation using Sun photometer measurements. *J. Geophys. Res.* 102, 16959–16969. doi: 10.1029/96JD02598

Navarro, G., Alvain, S., Vantrepotte, V., and Huertas, I. E. (2014). Identification of dominant Phytoplankton Functional Types in the Mediterranean Sea based on a regionalized remote sensing approach. *Remote Sens. Environ.* 152, 557–575. doi: 10.1016/j.rse.2014.06.029

Organelli, E., Nuccio, N., and Massi, L. (2007). "Individuazione dei principali gruppi fitoplanctonici in base al loro contributo di assorbimento e retrodiffusione nella riflettanza," in *Ecologia Limnologia e Oceanografia*: *Quale Futuro per L'ambiente* (Ancona), 181–187.

Pan, X., Mannino, A., Marshall, H. G., Filippino, K. C., and Mulholland, M. R. (2011). Remote sensing of phytoplankton community composition along the northeast coast of the United States. *Remote Sens. Environ.* 115, 3731–3747. doi: 10.1016/j.rse.2011.09.011

Pan, X., Mannino, A., Russ, M. E., Hooker, S. B., and Harding Jr, L. W. (2010). Remote sensing of phytoplankton pigment distribution in the United States northeast coast. *Remote Sens. Environ.* 114, 2403–2416. doi: 10.1016/j.rse.2010.05.015

Prezelin, B. B., Hofmann, E. E., Mengelt, C., and Klinck, J. M. (2000). The linkage between Upper Circumpolar Deep Water (UCDW) and phytoplankton assemblages on the west Antarctic Peninsula continental shelf. *J. Mar. Res.* 58, 165–202. doi: 10.1357/002224000321511133

Raitsos, D. E., Lavender, S. J., Maravelias, C. D., Haralambous, J., Richardson, A. J., and Reid, P. C. (2008). Identifying four phytoplankton functional types from space: an ecological approach. *Limnol. Oceanogr.* 53, 605–613. doi: 10.4319/lo.2008.53.2.0605

Ras, J., Claustre, H., and Uitz, J. (2008). Spatial variability of phytoplankton pigment distributions in the Subtropical South Pacific Ocean: comparison between *in-situ* and predicted data. *Biogeosciences* 5, 353–369. doi: 10.5194/bg-5-353-2008

Raven, J. A. (1998). The twelfth Tansley Lecture. Small is beautiful: the picophytoplankton. *Funct. Ecol.* 12, 503–513. doi: 10.1046/j.1365-2435.1998.00233.x

Reynolds, C. S. (2006). *The Ecology of Phytoplankton*. New York, NY: Cambridge University Press.

Reynolds, C. S., Huszar, V., Kruk, C., Naselli-Flores, L., and Melo, S. (2002). Towards a functional classification of the freshwater phytoplankton. *J. Plankton Res.* 24, 417–428. doi: 10.1093/plankt/24.5.417

Robinson, A. R., and Golnaraghi, M. (1995). "The physical and dynamical oceanography of the Mediterranean Sea," in *Ocean Processes in Climate Dynamics: Global and Mediterranean Examples*, eds P. Malanotte-Rizzoli and A. R. Robinson (Dordrecht: Kluwer Academic Publishers), 255–306.

Roy, S., Sathyendranath, S., Bouman, H., and Platt, T. (2013). The global distribution of phytoplankton size spectrum and size classes from their light-absorption spectra derived from satellite data. *Remote Sens. Environ.* 139, 185–197. doi: 10.1016/j.rse.2013.08.004

Sammartino, M., Di Cicco, A., Marullo, S., and Santoleri, R. (2015). Spatio-temporal variability of micro-, nano- and pico-phytoplankton in the Mediterranean Sea from satellite ocean colour data of SeaWiFS. *Ocean Sci.* 11, 759–778. doi: 10.5194/os-11-759-2015

Santoleri, R., Volpe, G., Marullo, S., and Nardelli, B. B. (2008). "Open waters optical remote sensing of the Mediterranean Sea," in *Remote Sensing of the European Seas,* eds V. Barale and M. Gade (Springer Netherlands), 103–116.

Sathyendranath, S., Watts, L., Devred, E., Platt, T., Caverhill, C., and Maass, H. (2004). Discrimination of diatoms from other phytoplankton using ocean-colour data. *Mar. Ecol. Prog. Ser.* 272, 59–68. doi: 10.3354/meps272059

Sieburth, J. M., Smetacek, V., and Lenz, J. (1978). Pelagic ecosystem structure- Heterotrophic compartments of the plankton and their relationship to plankton size fractions. *Limnol. Oceanogr.* 23, 1256–1263. doi: 10.4319/lo.1978.23.6.1256

Siokou-Frangou, I., Christaki, U., Mazzocchi, M. G., Montresor, M., Ribera d'Alcalá, M., Vaqué, D., et al. (2010). Plankton in the open Mediterranean Sea: a review. *Biogeosciences* 7, 1543–1586. doi: 10.5194/bg-7-1543-2010

Trees, C. C., Clark, D. K., Bidigare, R. R., Ondrusek, M. E., and Mueller, J. L. (2000). Accessory pigments versus chlorophyll a concentrations within the euphotic zone: a ubiquitous relationship. *Limnol. Oceanogr.* 45, 130–1143. doi: 10.4319/lo.2000.45.5.1130

Uitz, J., Claustre, H., Morel, A., and Hooker, S. B. (2006). Vertical distribution of phytoplanton communities in open ocean: an assessment based on surface chlorophyll. *J. Geophys. Res.* 111:C08005. doi: 10.1029/2005JC003207

Uitz, J., Stramski, D., Gentili, B., D'Ortenzio, F., and Claustre, H. (2012). Estimates of phytoplankton class-specific and total primary production in the Mediterranean Sea from satellite ocean color observations. *Global Biogeochem. Cycles* 26:GB2024. doi: 10.1029/2011GB004055

Vadrucci, M. R., Cabrini, M., and Basset, A. (2007). Biovolume determination of phytoplankton guilds in transitional water ecosystems of Mediterranean Ecoregion. *Transit. Waters Bull.* 2, 83–102. doi: 10.1285/i1825229Xv1n2p83

Vidussi, F., Claustre, H., Manca, B. B., Luchetta, A., and Marty, J. C. (2001). Phytoplankton pigment distribution in relation to upper thermocline circulation in the eastern Mediterranean Sea during winter. *J. Geophys. Res.* 106, 19939–19956. doi: 10.1029/1999JC000308

Volpe, G., Colella, S., Forneris, V., Tronconi, C., and Santoleri, R. (2012). The mediterranean ocean colour observing system–System development and product validation. *Ocean Sci.* 8, 869–883. doi: 10.5194/os-8-869-2012

Volpe, G., Santoleri, R., Vellucci, V., Ribera d'Alcala, M., Marullo, S., and D'Ortenzio, F. (2007). The colour of the Mediterranean Sea: global versus regional bio-optical algorithms evaluation and implication for satellite chlorophyll estimates. *Remote Sens. Environ.* 107, 625–638. doi: 10.1016/j.rse.2006.10.017

Werdell, P. J., and Bailey, S. W. (2005). An improved *in-situ* bio-optical data set for ocean color algorithm development and satellite data product validation. *Remote Sens. Environ.* 98, 122–140. doi: 10.1016/j.rse.2005.07.001

Williams, R., and Claustre, H. (1991). Photosynthetic pigments as biomarkers of phytoplankton populations and processes involved in the transformation of particulate organic matter at the Biotrans site (47°N, 20°W). *Deep Sea Res.* 38, 347–355. doi: 10.1016/0198-0149(91)90072-N

Wright, S. W., and Jeffrey, S. W. (1987). Fucoxanthin pigment markers of marine phytoplankton analysed by HPLC and HPTLC. *Mar. Ecol. Progr. Ser.* 38, 259–266. doi: 10.3354/meps038259

Wright, S. W., and Jeffrey, S. W. (2006). "Pigment markers for phytoplankton production," in *Marine Organic Matter: Biomarkers, Isotopes and DNA,* ed J. K. V*olkman* (Berlin: Springer-Verlag), 71–104.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Di Cicco, Sammartino, Marullo and Santoleri. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Marine Science

ORIGINAL RESEARCH
published: 30 May 2017
doi: 10.3389/fmars.2017.00163

Extended Formulations and Analytic Solutions for Watercolumn Production Integrals

Žarko Kovač[1], Trevor Platt[2], Suzana Antunović[3], Shubha Sathyendranath[2,4], Mira Morović[1] and Charles Gallegos[5]*

[1] Physical Oceanography Laboratory, Institute of Oceanography and Fisheries, Split, Croatia, [2] Plymouth Marine Laboratory, Plymouth, United Kingdom, [3] Faculty of Civil Engineering, Architecture, and Geodesy, Split, Croatia, [4] Plymouth Marine Laboratory, National Centre for Earth Observations, Plymouth, United Kingdom, [5] Smithsonian Environmental Research Center, Edgewater, MD, United States

The effect of biomass dynamics on the estimation of watercolumn primary production is analyzed, by coupling a primary production model to a simple growth equation for phytoplankton. The production model is formulated with depth- and time-resolved biomass, and placed in the context of earlier models, with emphasis on the canonical solution for watercolumn production. A relation between the canonical solution and the general solution for the case of an arbitrary depth-dependent biomass profile was derived, together with an analytical solution for watercolumn production in case of a depth dependent biomass profile described with the shifted Gaussian function. The analysis was further extended to the case of a time-dependent, mixed-layer biomass, and two additional analytical solutions to this problem were derived, the first in case of increasing mixed-layer biomass and the second in case of declining biomass. The solutions were tested with Hawaii Ocean Time-series data. The canonical solution for mixed-layer production has proven to be a good model for this data set. The shifted Gaussian function was demonstrated to be an accurate model for the measured biomass profiles and the shifted Gaussian parameters extracted from the measured profiles were further used in the analytical solution for watercolumn production and results compared with data. The influence of time-dependent biomass on mixed-layer production was studied through analytical solutions. Re-examining the Critical Depth Hypothesis we derived an expression for the daily increase in mixed-layer biomass. Finally, the work was placed in a remote sensing context and the time-dependent model for biomass related to the remotely sensed-biomass.

Keywords: primary production, watercolumn production integrals, analytic solutions, growth models, critical depth criterion, remote sensing

OPEN ACCESS

Edited by:
Katja Fennel,
Dalhousie University, Canada

Reviewed by:
Toru Hirawake,
Hokkaido University, Japan
Alexandre Mignot,
UMR7093 Laboratoire D'océanographie de Villefranche, France

****Correspondence:***
Žarko Kovač
kovac@izor.hr

Specialty section:
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

Received: *29 October 2016*
Accepted: *12 May 2017*
Published: *30 May 2017*

Citation:
Kovač Ž, Platt T, Antunović S, Sathyendranath S, Morović M and Gallegos C (2017) Extended Formulations and Analytic Solutions for Watercolumn Production Integrals. Front. Mar. Sci. 4:163. doi: 10.3389/fmars.2017.00163

1. INTRODUCTION

In the ocean, phytoplankton form the foundation of the pelagic ecosystem and by virtue of their phototsynthesis (primary production) act as a source of organic carbon for the remainder of the ecosystem (Chavez et al., 2011). The abundance and growth of virtually all marine life on earth depend on phytoplankton. Consequently, the world's largest fisheries are concentrated around ocean areas with high primary production (Cushing, 1971; Mann and Lazier, 2006). Moreover, the

role of phytoplankton does not end with the food chain itself. Through the action of the so-called biological pump, a complex ecosystem process which starts with primary production (Volt and Hoffert, 1985), phytoplankton contribute to the transfer of carbon into the deep ocean (Longhurst and Harrison, 1989) and subsequently affect atmospheric carbon concentration on longer time scales (from decadal to millennial) (Honjo et al., 2008). With this in mind, prediction of primary production is important, not just for the open ocean, but also for coastal seas, and is also relevant to fisheries and climate change research. Given the vastness of the ocean, the basic means by which such predictions are made is through the combined use of primary production models and ocean-color data, acquired by satellites (Platt and Sathyendranath, 1991; Siegel et al., 2014).

In such applications prediction of the total amount of carbon assimilated in the water column in 1 day, i.e., daily watercolumn production, is the target (Platt et al., 1991b). Models for watercolumn production predict the amount of carbon assimilated by phytoplankton per unit area of ocean surface during the day (Platt and Sathyendranath, 1988, 1993). Typical models have chlorophyll concentration as an initial condition and use photosynthesis parameters to determine the response of phytoplankton to light (Platt et al., 1977). The light attenuation coefficient and daylength are also required to determine the depth interval in which photosynthesis takes place and the time interval during which photosynthesis occurs (Kirk, 2011). Surface photosynthetically-active radiation is the forcing variable, which integrated over daylength gives the amount of energy available for photosynthesis.

Various attempts have been made to relate watercolumn production mathematically to environmental factors. Models of various complexities have been proposed and equations have been derived for predicting the amount of primary production per unit area of ocean surface (Kirk, 2011). These equations are often referred to as estimators of watercolumn production (Platt and Sathyendranath, 1993). A straightforward application of such estimators is in converting satellite images of ocean color into primary production maps of the ocean (Platt and Sathyendranath, 1988; Campbell et al., 2002; Platt et al., 2008). For example, with such applications the global annual primary production has been estimated at $\sim 45-50$ (Longhurst et al., 1995), ~ 52 (Westberry et al., 2008), and 58 ± 7 (Buitenhuis et al., 2013) giga tonnes of carbon per year.

Some of the earliest primary production estimators date back to Ryther (1956), Ryther and Yentsch (1957), and Talling (1957), all semi empirical. These are followed by Rodhe (1965) whose model has later been used by Bannister (1974) and others (Smith and Baker, 1978; Eppley et al., 1985). Platt (1986) developed a linear model and finally in 1990 the first, and until today the only, analytical solution of the nonlinear model for daily watercolumn production was given by Platt et al. (1990). A thorough historical review of the topic is found in Platt and Sathyendranath (1993). A common feature of the models mentioned is their assumption of daily time-independent, vertically-uniform biomass (Platt and Sathyendranath, 1993). Therefore, strictly speaking, they are valid only for calculating watercolumn production occurring in the mixed layer during non bloom conditions (when net growth is low). These conditions do prevail over vast ocean areas and for prolonged periods of the year, but it is precisely when highest primary production occurs, that they do not.

Since for the open ocean the depth of the mixed layer is on the same order of magnitude as the photic depth (de Boyer Montegut et al., 2004), vertical uniformity in biomass is thought not to be a severe limitation for calculating watercolumn production. However, during stratified periods the mixed layer production (i.e., production taking place from the ocean surface up to the base of the mixed layer) may no longer constitute a major segment of watercolumn production. The reason is that when stratification is strong, the mixed layer depth is often found to be much shallower then the photic depth (Longhurst, 1998). In such conditions biomass tends to develop vertical dependency below the base of the mixed layer, and the water column below it can contribute significantly to daily watercolumn production. These conditions tend to prevail during summer periods of intense surface sunlight and in oligotrophic environments (Mignot et al., 2014). In such conditions the correctness of applying models with vertical uniformity in biomass, for the whole photic depth, may be challenged.

Also, resolving the daily time dependence of biomass when calculating watercolumn production may be advantageous in some situations. Consider for example the period of a bloom. After stratification sets in, circulation of phytoplankton to greater depths is prevented and rapid growth can occur (Sverdrup, 1953; Sathyendranath et al., 2015). Certainly, during such a period the usage of time-independent biomass in watercolumn production models can be challenged. After all, the bloom is defined as the period of rapid growth in biomass. For estimating production during such conditions a model with time-dependent biomass would be more suitable. Another case in which the time dependence of biomass may be important is during periods of sharp decline in biomass, which may be caused by intense grazing or dilution losses from deepening of the mixed layer (Zhai et al., 2010). Thanks to remote sensing technologies, time dependence in biomass is easily seen in serial satellite records of chlorophyll (Platt and Sathyendranath, 2008; Racault et al., 2012; Cabre et al., 2016).

Therefore, during periods of blooming/declining biomass or periods of strong biomass stratification, a different specification of model biomass may be more suitable. In this work, we begin with an outline of the model, followed by mathematical descriptions of the aforementioned problems and derivation of the solutions. The basic model we use is already established in the literature. It was first put forward by Platt et al. (1990) and has since received numerous applications. This primary production model was coupled to hydrodynamical models (Platt and Sathyendranath, 1991), put in the context of Sverdrup critical depth theory (Platt et al., 1991a), used in the estimation of primary production from satellite data (Platt and Sathyendranath, 1993), in studying the interaction between the mixed layer and watercolumn production (Platt et al., 1994) and finally in explaining the dynamics of high nutrient low chlorophyll zones (Platt et al., 2003). Here we extend this model, providing additional solutions for watercolumn production and analyzing its relation to biomass dynamics through simple

growth models. By doing so we broaden the range of applicability of the model and place it on a more rigorous foundation.

First we define the watercolumn production integral and proceed to discuss different ways of calculating watercolumn production based on the way biomass is specified. We then explore the general case of depth dependent biomass and give an exact solution for watercolumn production with the shifted Gaussian biomass. After that we explore the case of time dependent mixed layer biomass and provide analytical solutions for the case of growing and declining biomass. We test the new solutions on data collected at the Hawaii Ocean Time-series station located in the Pacific. Finally, we discuss the implications of this model for Sverdrup's Critical Depth Hypothesis and for remote sensing applications.

2. THEORY

Watercolumn Primary Production

Let the z axes be positive downwards with the origin at the ocean surface (**Figure 1**). Let time t equal zero at sunrise and D at sunset. At an arbitrary depth z and time t, primary production $P(z,t)$ (measured in mg C m^{-3} h^{-1}) is the product of phytoplankton biomass $B(z,t)$ (measured in mg Chl m^{-3}) and the biomass-normalized production $P^B(z,t)$ (measured in mg C (mg Chl)$^{-1}$ h^{-1}):

$$P(z,t) = B(z,t)P^B(z,t). \tag{1}$$

In analytical models of primary production, the usual definition of watercolumn primary production is that of a double integral of the product of initial biomass $B(z,0)$ and the biomass-normalized production $P^B(z,t)$:

$$P_{Z,T} = \int_0^D \int_0^\infty B(z,0)P^B(z,t)\,\mathrm{d}z\,\mathrm{d}t. \tag{2}$$

The subscript Z denotes integration over depth, whereas subscript T denotes integration over time, following the notation of Platt et al. (1990). The biomass-normalized production is a function of irradiance and is specified with the photosynthesis irradiance function $p^B(I)$ (Jassby and Platt, 1976), which is stated as:

$$P^B(z,t) = p^B(I(z,t)), \tag{3}$$

where $I(z,t)$ is the irradiance, calculated from surface irradiance $I_0(t)$ with the aid of a light penetration model (Kirk, 2011). Typically, production increases linearly with irradiance, the increase begins to decline with higher irradiance and saturation occurs, or production is reduced if irradiance reaches high enough levels, i.e., photoinhibition occurs (Platt et al., 1980). Neglecting photoinhibition leaves the photosynthesis-irradiance function determined by two parameters: the initial slope α^B and the assimilation number P_m^B (Platt and Sathyendranath, 1988). The effects of nutrients and temperature on production are assumed to be included implicitly in the magnitude of the photosynthesis parameters, following the approach of Platt and Jassby (1976). When evaluating the integral (2), various approaches and assumptions may be adopted. It can either be integrated over depth, following integration over time, or vice versa. Depth dependence of biomass can be specified, or biomass can be set constant with depth. Similarly, surface irradiance can be considered as time-dependent or constant, spectrally-resolved or spectrally integrated. Different photosynthesis-irradiance functions can be used (Jassby and Platt, 1976). How to solve integral (2) depends in large part on mathematical convenience, in the context of the particular problem under study.

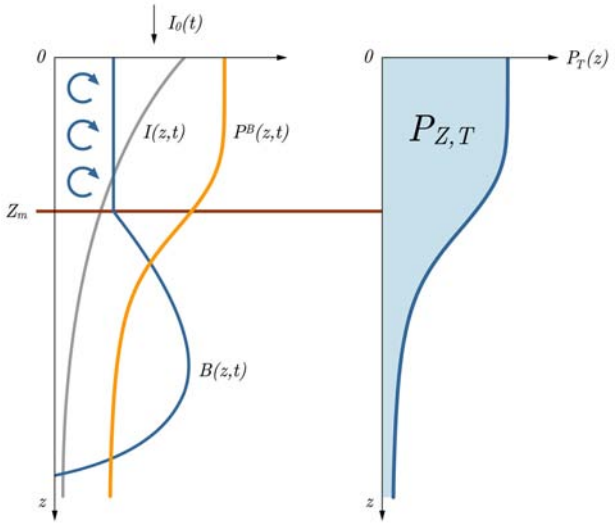


FIGURE 1 | Sketch of the basic model relations with respect to depth. With the information on surface irradiance $I_0(t)$ and the optical properties of the water column, the irradiance at depth $I(z,t)$ (gray curve) is first calculated. Using $I(z,t)$ and the photosynthesis irradiance function $p^B(I)$, along with the information on biomass $B(z,t)$ (blue curve on the left hand side), normalized instantaneous production $P^B(z,t)$ is obtained (orange curve). Further on, taking the product of $B(z,t)$ and $P^B(z,t)$ and integrating it over time we get the daily production at depth (blue curve on the right hand side). In this depiction watercolumn production $P_{Z,T}$ is the blue surface under the $P_T(z)$ curve. Blue arrows indicate mixing and Z_m marks the mixed layer depth. Daily production from the surface up to Z_m is marked with $P_{Z_m,T}$.

Using dimensional analysis Platt and Sathyendranath (1993) showed that the canonical form for the solution of integral (2), in the case of vertically uniform biomass $B(z,0) = B$, is:

$$P_{Z,T} \sim \frac{BP_m^B D}{K} f(I_*^m), \tag{4}$$

where K is the diffuse attenuation coefficient for downward irradiance (Kirk, 2011) and $I_*^m = \alpha^B I_0^m / P_m^B$ the scaled noon irradiance, with I_0^m as noon irradiance. The $f(I_*^m)$ function is determined by the formulation of the production-light relationship. In case of the Platt photosynthesis irradiance function (Platt et al., 1980):

$$p^B(I) = P_m^B\left(1 - \exp\left(-\alpha^B I / P_m^B\right)\right), \tag{5}$$

the $f(I_*^m)$ is:

$$f(I_*^m) = \sum_{n=1}^{\infty} \frac{2\left(I_*^m\right)^{2n-1}}{\pi\,(2n-1)\,(2n-1)!}\frac{(2n-2)!!}{(2n-1)!!} - \sum_{n=1}^{\infty} \frac{\left(I_*^m\right)^{2n}}{2n\,(2n)!}\frac{(2n-1)!!}{(2n)!!}. \quad (6)$$

In that case the exact solution to integral (2) is:

$$P_{Z,T} = \frac{BP_m^B D}{K} f(I_*^m). \quad (7)$$

When the water column is of a finite depth Z_m (**Figure 1**) the solution is:

$$P_{Z_m,T} = \frac{BP_m^B D}{K}\left(f(I_*^m) - f(I_*^m e^{-KZ_m})\right). \quad (8)$$

This solution was derived by Platt et al. (1990) and thus far no other analytical solution for watercolumn production has been reported in the literature. It is called the canonical solution. The assumptions of the model regarding light conditions are: optically-uniform water column with sinusoidally varying surface irradiance, such that the irradiance at depth is given by:

$$I(z,t) = I_0^m \sin(\pi t/D) e^{-Kz}. \quad (9)$$

Inserting this expressions into (5) gives the biomass normalized production as:

$$P^B(z,t) = P_m^B\left(1 - \exp\left(-\alpha^B I_0^m \sin(\pi t/D)\, e^{-Kz}/P_m^B\right)\right). \quad (10)$$

Time integral of $P^B(z,t)$ gives the daily normalized production $P_T^B(z)$. Here we shall use this same model setup but will relax the assumptions of homogeneous and constant biomass. We will allow for time- and depth-dependent biomass, seek the solution for $P_{Z,T}$ in specific situations and explore the influence that a time-, or depth-, dependent biomass has on watercolumn production. Therefore, in this paper, a more complete definition for $P_{Z,T}$ would be:

$$P_{Z,T} = \int_0^D \int_0^{\infty} B(z,t) P^B(z,t)\, \mathrm{d}z\, \mathrm{d}t. \quad (11)$$

A complete list of symbols with corresponding descriptions is given in Appendix A.

Biomass Specification

As formulated, this watercolumn production integral requires the specification of biomass as a function of depth and time. To specify biomass for integral (11) it is useful to write a differential equation for the time evolution of the biomass. In this way biomass dynamics are incorporated into the integral and its solution. The dynamics of biomass can be modeled by a simple equation of the following form:

$$\frac{\partial}{\partial t} B(z,t) = \frac{1}{\chi} P^B(z,t) B(z,t), \quad (12)$$

Sathyendranath and Platt (2007) assuming the carbon to chlorophyll ratio χ constant during time (Sathyendranath et al., 2009). This equation governs the time evolution of biomass resulting from photosynthesis and allows for the accumulation of biomass at each depth in accordance with (10).

Let $B^*(z,t)$ be the solution to the growth equation (12). A direct approach for calculating $P_{Z,T}$ would be to insert $B^*(z,t)$ into the integral (11) and solve for $P_{Z,T}$:

$$P_{Z,T} = \int_0^D \int_0^{\infty} B^*(z,t) P^B(z,t)\, \mathrm{d}z\, \mathrm{d}t. \quad (13)$$

With the usage of the solution $B^*(z,t)$ in (11), watercolumn production is coupled to the biomass dynamics expressed by (12). However, this approach can be mathematically complex, depending on the form of the solution $B^*(z,t)$. A simpler approach follows by recognizing that the process of biomass accumulation described by (12) is in fact primary production. Therefore, the difference between the final $B^*(z,D)$ and initial biomass $B^*(z,0)$, multiplied by χ, equals daily primary production at depth z. Mathematically, the solution $B^*(z,t)$ satisfies (12) and the following holds:

$$B^*(z,t) = \frac{\chi}{P^B(z,t)} \frac{\partial}{\partial t} B^*(z,t). \quad (14)$$

Inserting this expression into (13) and solving yields:

$$P_{Z,T} = \chi \int_0^{\infty} \left(B^*(z,D) - B(z,0)\right) \mathrm{d}z. \quad (15)$$

This expression gives the watercolumn production as the difference between the final and initial biomass multiplied by the carbon to chlorophyll ratio. The two integrals, (13) and (15), are equivalent. The advantage of the second approach (15) is evident in cases when the biomass dynamics are governed by the simple growth equation, such as (12), whereas the first approach (13) is more useful when the solution to equation (12) is a simple mathematical expression. Both approaches will be used later.

If in the time interval from $t = 0$ (sunrise) until $t = D$ (sunset) the following expression holds for the solution $B^*(z,t)$:

$$\frac{\partial B^*(z,t)}{\partial t} \approx 0, \quad (16)$$

it is safe to assume $B^*(z,D) \approx B^*(z,0)$ and justified to use initial biomass throughout the calculation of daily production. Biomass profiles of this type can be considered to be in, or close to, a steady state (Hodges and Rudnick, 2004). From the dynamical perspective this approach is crude, but solutions with initial biomass are valid in situations when the biomass does not change significantly during the time course of 1 day. In this sense, integral (2) is a special case of integral (13) when (16) is valid. The accumulation of biomass due to photosynthesis is either not significant compared with the initial biomass, or is balanced by the loss processes. With this approach, the functional form of the

initial biomass profile can be inferred from measurement, instead of by solving the model equations, which is an advantage. In fact, this is the standard practice in remote sensing applications (Platt and Sathyendranath, 1988; Behrenfeld and Falkowski, 1997).

Stratified, Time-Independent Biomass

In summer periods of intense heating and weak mixing, strong stratification in temperature, as well as stratification in biomass, tend to develop below the mixed layer (Mann and Lazier, 2006). In this stratified region, the biomass tends to remain virtually constant in time, with only slight fluctuations in the shape of the chlorophyll profile as the season advances. In the tropical ocean stratification can be a permanent feature, whereas in the temperate regions stratification tends to be eroded during autumn and winter (Longhurst and Harrison, 1989). Given that stratification usually persists for intervals of time longer than 1 day, it is safe to assume constant biomass when calculating daily watercolumn production. Depth variation in biomass is assumed to have a larger influence on the magnitude of daily watercolumn production than time dependence does. Mathematically, integral (2) is appropriate in these conditions and model solutions with time-independent biomass are valid, because we assume that biomass does not change significantly during the time course of 1 day, as stated by (16). This assumption is valid in regions of the ocean where a stratified biomass profile is observed to be persistent on time scales longer than that of 1 day.

General Case

When the biomass profile is held constant in time $B(z,0) = B(z)$, a link between the canonical solution and the general solution to integral (2) can be established. To demonstrate this relation, we take advantage of the daily normalized production profile $P_T^B(z)$. With it, integral (2) becomes:

$$P_{Z,T} = \int_0^\infty B(z) P_T^B(z)\, \mathrm{d}z. \tag{17}$$

In this way, watercolumn production is given as an integral of the product between the time-independent biomass $B(z)$ and normalized daily production $P_T^B(z)$. The solution for the daily normalized production profile $P_T^B(z)$ in case of the Platt et al. (1980) photosynthesis irradiance function (5) is Kovač et al. (2016a):

$$P_T^B(z) = P_m^B D f_z(I_*^m e^{-Kz}), \tag{18}$$

where the $f_z(I_*^m e^{-Kz})$ function is related to the $f(I_*^m)$ function in the following manner:

$$f_z(I_*^m e^{-Kz}) = -\frac{1}{K}\frac{\mathrm{d}}{\mathrm{d}z} f\Big(I_*^m \exp(-Kz)\Big). \tag{19}$$

Inserting this expression into (17) and solving by partial integration gives:

$$P_{Z,T} = \frac{P_m^B D}{K}\left(B(0) f\left(I_*^m\right) + \int_0^\infty \frac{\mathrm{d}B(z)}{\mathrm{d}z} f\left(I_*^m e^{-Kz}\right) dz\right), \tag{20}$$

Where we have used $B(\infty) f(I_*^m e^{-K\infty}) = 0$. The first term on the right hand side is simply the canonical solution (7) and the second term is recognized as the contribution arising from the shape of the biomass profile (stratification term).

The interpretation of the first term is simple: it gives the watercolumn production in the case where surface biomass stretches over the entire watercolumn. According to the second term, any change in biomass with depth causes a deviation from the canonical solution. If there is an increase in biomass, $\mathrm{d}B(z)/\mathrm{d}z > 0$, this contribution is positive, whereas if there is a decline, $\mathrm{d}B(z)/\mathrm{d}z < 0$, this contribution is negative. The change in biomass $\mathrm{d}B(z)/\mathrm{d}z$ is scaled by the $f(I_*^m e^{-Kz})$ function. The product $\mathrm{d}B(z) f(I_*^m e^{-Kz})$ equals the production that would occur below the depth z in case the biomass from z to ∞ were equal to $\mathrm{d}B(z)$. Total contribution from all these infinitesimal changes in $B(z)$ is accounted for by the integral on the right hand side of (20). With increase in depth, the contribution from biomass variation decreases, simply because production declines with increasing depth (10).

Expression (20) is a formal relation linking the canonical solution (7) to the solution for watercolumn production with stratified biomass (2). It is valid for an arbitrary biomass profile and clearly displays the role surface biomass $B(0)$ has on the magnitude of $P_{Z,T}$. Surface biomass appears as a leading factor in $P_{Z,T}$. The significance of this result is emphasized given that surface biomass is readily accessible to satellite measurement. Therefore, if the remotely-sensed surface biomass is precise, and assuming the remaining parameters of the model are characteristic of the ocean region in question, the error in the estimated watercolumn production arises solely as a consequence of the error in estimating the biomass profile, which is inaccessible to remote sensing and has to be assigned based on prior information (Platt and Sathyendranath, 1988).

Shifted Gaussian Biomass Profile

An important case of the time independent biomass profile is that of the shifted Gaussian superimposed on a constant background. The shifted Gaussian is a suitable function for the description of the vertical structure of biomass for diverse regions of the oceans (Platt et al., 1991b), especially for the case of the widespread deep chlorophyll maximum (Longhurst, 1998; Mignot et al., 2014). In remote sensing applications it has been widely used to model biomass profiles in algorithms for primary production calculation (Platt and Sathyendranath, 1988, 1991). It is also used in operational oceanography (Platt et al., 2008), with well-established and tested procedures (Platt et al., 1988; Longhurst et al., 1995). The role of the shifted Gaussian in ocean color remote sensing has been studied and modeled by many authors, including: Morel and Berthon (1989), Sathyendranath and Platt (1989), Andre (1992), Stramska and Stramski (2005), Uitz et al. (2006), Xiu et al. (2008), and Mignot et al. (2011).

In this case the biomass profile equals:

$$B(z) = B_0 + \frac{h}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(z-z_m)^2}{2\sigma^2}\right), \tag{21}$$

where the integrated biomass beneath the Gaussian curve is given by h, the depth of the maximum is at z_m and the width of the biomass peak is determined by σ. B_0 is the background biomass. The peak biomass above the background B_0 at z_m is $H = h/(\sigma\sqrt{2\pi})$. Upon inserting this expression into integral (17) we get:

$$P_{Z,T} = \int_0^\infty B_0 P_T^B(z)\,\mathrm{d}z + \int_0^\infty \frac{h}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(z-z_m)^2}{2\sigma^2}\right) P_T^B(z)\,\mathrm{d}z. \tag{22}$$

The first integral can be replaced with the canonical solution and the second integral can be simplified with the help of (18) to give:

$$P_{Z,T} = \frac{B_0 P_m^B D}{K} f(I_*^m) + \frac{h P_m^B D}{\sigma\sqrt{2\pi}} \int_0^\infty \exp\left(-\frac{(z-z_m)^2}{2\sigma^2}\right) f_z\left(I_*^m e^{-Kz}\right)\mathrm{d}z. \tag{23}$$

With reference to (20) we label the second term on the right hand side as $\Delta P_{Z,T}$ (stratification term). The exact mathematical derivation of the solution for $\Delta P_{Z,T}$ is given in the Appendix B. The final solution is:

$$\Delta P_{Z,T} = P_m^B D \frac{h}{2}\left[\sum_{n=1}^{\infty} \exp\left(\frac{z_{2n-1}^2 - z_m^2}{2\sigma^2}\right) \frac{2\left(I_*^m\right)^{2n-1}}{\pi\,(2n-1)!} \frac{(2n-2)!!}{(2n-1)!!} \left(1 + \Phi\left(\frac{z_{2n-1}}{\sqrt{2}\sigma}\right)\right) - \sum_{n=1}^{\infty} \exp\left(\frac{z_{2n}^2 - z_m^2}{2\sigma^2}\right) \frac{\left(I_*^m\right)^{2n}}{(2n)!} \frac{(2n-1)!!}{(2n)!!} \left(1 + \Phi\left(\frac{z_{2n}}{\sqrt{2}\sigma}\right)\right)\right], \tag{24}$$

where Φ is the error function, $z_{2n-1} = z_m - (2n-1)\sigma^2 K$ and $z_{2n} = z_m - 2n\sigma^2 K$. The complete solution to (22) is now:

$$P_{Z,T} = \frac{B_0 P_m^B D}{K} f(I_*^m) + \Delta P_{Z,T}, \tag{25}$$

where the $\Delta P_{Z,T}$ depends explicitly on the values of h, z_m, σ, α^B, P_m^B, I_0^m, and D. This expression calculates the amount of carbon assimilated during 1 day per meter squared of the ocean surface, by phytoplankton distributed vertically according to the shifted Gaussian function (21).

The shifted Gaussian is flexible enough to describe various features in the measured chlorophyll profiles and therefore this solution covers a wide range of situations encountered in the field. That flexibility is achieved by altering the parameters of the function, namely: B_0, z_m, σ, and h. The disadvantage is that in addition to the six basic quantities: α^B, P_m^B, B_0, I_0^m, D, and K, which appear in the canonical solution, the solution for the shifted Gaussian has three more: z_m, σ, and h. To apply the solution, the values of these quantities need to be specified.

Time-Dependent, Mixed-Layer Biomass

Accounting for time dependence may be advantageous when considering mixed-layer production, especially during the period of a bloom. Blooms are typically initiated by stratification, either from heating, river discharge or from sea ice melt (Mann and Lazier, 2006). After the onset of stratification, the phytoplankton become trapped in the upper, well-illuminated, nutrient-rich layer, where conditions for growth are fulfilled, and in this upper layer production is most intense, which can lead to rapid accumulation of biomass (Chiswell et al., 2015). Blooms last until nutrients are depleted, after which they crash (Levy, 2015), or are terminated by overgrazing. How rapidly the bloom develops will be determined by physical conditions, the physiological status of the phytoplankton population and by loses (Banse and English, 1994), resulting in diverse patterns of seasonal cycles of phytoplankton biomass, as evident in remotely sensed records of chlorophyll concentration (Vargas et al., 2009; Racault et al., 2012). In the model, the physiological status is described by the photosynthesis irradiance function, whereas the physical conditions are presented by the mixed layer depth, surface irradiance, and the attenuation coefficient. Next, we calculate primary production in the mixed layer and show how it is affected by rapidly-growing or declining biomass.

Increasing Mixed-Layer Biomass

Let us consider a mixed layer of depth Z_m (constant in time) with uniformly-distributed biomass at initial time $B(z,0) = B_0$, for $z = 0$ to $z = Z_m$. To simplify the equations, we introduce the following notation for the total biomass in the mixed layer:

$$B_{Z_m}(t) = \int_0^{Z_m} B(z,t)\,\mathrm{d}z. \tag{26}$$

At time $t = 0$ the total biomass is $B_{Z_m}(0) = B_0 Z_m$. Let us also assume that the newly-synthesized mixed layer biomass at time t is redistributed through the mixed layer during a time interval Δt, so that there is no stratification in biomass at $t + \Delta t$. Mixed-layer biomass at time $t + \Delta T$ is now:

$$B_{Z_m}(t + \Delta t) = B_{Z_m}(t) + \frac{1}{\chi} P_{Z_m}(t)\Delta t, \tag{27}$$

with $P_{Z_m}(t)$ given as $P_{Z_m}(t) = \int_0^{Z_m} P(z,t)\,\mathrm{d}z$. In the limit of $\Delta t \to 0$, that is instantaneous mixing of newly-synthesized biomass, this equation reduces to:

$$\frac{\partial}{\partial t} B_{Z_m}(t) = \frac{1}{\chi Z_m} P_{Z_m}^B(t) B_{Z_m}(t). \tag{28}$$

Derivation of this equation is given in the Appendix C. The solution for the mixed-layer production with time-dependent biomass described by (28) and initial biomass $B_{Z_m}(0) = B_0 Z_m$ is:

$$P_{Z_m,T} = \chi B_0 Z_m \left[\exp\left[\frac{P_m^B D}{\chi Z_m K}\left(f(I_*^m) - f(I_*^m e^{-KZ_m})\right)\right] - 1\right]. \tag{29}$$

The solution is found by the application of (15); details are given in the Appendix C. The term in the exponent is recognized as the canonical solution (8) divided by χ, Z_m, and B_0. It is clear that the canonical solution (8) will underestimate production in comparison with this solution. Due to photosynthesis, at a given moment there will be more biomass in the mixed layer, than in the case of time-independent biomass.

Declining Mixed-Layer Biomass

Given that the term $P^B(z,t)$ is always positive, we have considered so far only the case of growing biomass. Here we consider the case when this term is reduced in magnitude by a loss term, and investigate the effect biomass loss has on the magnitude of daily production. With this goal, equation (12) is modified by addition of a loss term in the most general form L^B:

$$\frac{\partial}{\partial t}B(z,t) = \left(\frac{1}{\chi}P^B(z,t) - L^B\right)B(z,t). \tag{30}$$

For there to be a negative growth (decline) in biomass the term L^B has to be greater than $P^B(z,t)$:

$$L^B \gg \frac{1}{\chi}P^B(z,t). \tag{31}$$

In this case the solution to equation (30) at time t is:

$$B(z,t) = B_0 e^{-L^B t}, \tag{32}$$

where we have used the initial condition $B(z,0) = B_0$. Time dependence of this type is often observed in satellite records of surface biomass, for example during the termination phase of a bloom (Cabre et al., 2016). Therefore, to calculate mixed layer production in this case, we insert the previous expression into (11) and obtain the following integral:

$$P_{Z_m,T} = \int_0^D \int_0^{Z_m} B_0 e^{-L^B t} P^B(z,t)\,\mathrm{d}z\;\mathrm{d}t. \tag{33}$$

This procedure is in accord with (13). The derivation of the solution to this integral is given in the Appendix D. The final solution is:

$$P_{Z_m,T} = \frac{B_0 P_m^B D}{K}\left[(e^{-L^B D}+1)\sum_{n=1}^{\infty}\frac{\left(I_*^m\right)^{2n-1} - \left(I_*^m e^{-KZ_m}\right)^{2n-1}}{\pi(2n-1)}\prod_{m=1}^{n}\frac{1}{(-L^B D/\pi)^2 + (2m-1)^2}\right.$$
$$\left. - \frac{(e^{-L^B D}-1)}{-L^B D}\sum_{n=1}^{\infty}\frac{\left(I_*^m\right)^{2n} - \left(I_*^m e^{-KZ_m}\right)^{2n}}{2n}\prod_{m=1}^{n}\frac{1}{(-L^B D/\pi)^2 + (2m)^2}\right]. \tag{34}$$

In the expression inside the square brackets, the loss term L^B and daylength D appear as a product $L^B D$, which imposes itself as a dimensionless factor for the problem, determining just how much the continual loss of initial biomass reduces watercolumn production. This type of loss can occur when the grazing pressure on phytoplankton is significant, or by sinking out of the mixed layer. Potentially, the magnitude of the loss term could be determined from time series of satellite-estimated surface chlorophyll.

Relation to the Canonical Solution

In case of increasing, mixed layer biomass, the canonical solution (8) is easily recognized in the exponential function on the right hand side of expression (29). Writing the exponential as a sum and rearranging, gives:

$$P_{Z_m,T} = \frac{B_0 P_m^B D}{K}\left(f(I_*^m) - f(I_*^m e^{-KZ_m})\right) + B_0\sum_{n=2}^{\infty}\frac{(P_m^B D)^n}{n!(\chi Z_m)^{n-1}K^n}\left(f(I_*^m) - f(I_*^m e^{-KZ_m})\right)^n. \tag{35}$$

The first term on the right hand side is the canonical solution (8) and the additional terms arise due to time dependent biomass. When the biomass is time independent, as is the case with the canonical solution, these terms will vanish. Therefore, when biomass is time dependent the canonical solution (8) can be interpreted as the first order approximation for mixed layer production, representing the lower limit on watercolumn production. Important to note is that biomass has to increase with time in accordance with (28) for this interpretation to hold.

On the other hand, when biomass is declining with time, as stated in (32), the canonical solution will be the upper limit on daily watercolumn production, given that at each time instant there is less biomass in the mixed layer, in comparison with constant biomass B_0. In this case we were unable to find an exact link between solution (34) and the canonical solution. Instead, we have found by numerical exercises, the canonical solution with noon biomass $B_0 e^{-L^B D/2}$, in place of initial biomass B_0, to be a good approximation for the full solution (34):

$$P_{Z_m,T} \approx B_0\, e^{-L^B D/2}\,\frac{P_m^B D}{K}\left(f(I_*^m) - f(I_*^m e^{-KZ_m})\right). \tag{36}$$

It is important to emphasize the initial assumption behind solution (34). For it to hold, the loss rate has to be significantly larger than normalized production (31) so that the growth due to production is insignificant in comparison with losses and (32) is valid. Therefore, solution (34) is valid when (31) holds and the approximation (36) is only then justified.

3. DATA

The models described so far and the solutions derived therefrom require parameter values to be implemented and tested against data. Oceanic time series are suitable sources of such information. For our model, the Hawaii Ocean Time-series (HOT) is an ideal testing ground. Data from it have already served for other model testing, with over 600 publications testifying to the quality of the data. The entire data set is publicly available

with full documentation of the methods and procedures used. Information on the data set can be found in Karl et al. (2001) and Karl and Church (2014).

HOT is a program of oceanic measurements started in 1988 at station Aloha, located near the Hawaii Islands at 22°45′N 158°W. The basic set of measurements relevant for this model encompasses: primary production using the standard *in-situ* implemented ^{14}C method (Steemann Nielsen, 1952), fluorimetric determination of chlorophyll concentration (Strickland and Parsons, 1972), optical measurements using bio-optical profilers, surface photosynthetically-available radiation (PAR) measurement using on deck radiometer, and finally the mixed-layer depth determination based on potential density.

There are in total 194 cruises with available data. Production and chlorophyll were measured at: 5, 25, 45, 75, 100, 125, 150, and 175 m. Surface PAR was given in μ E m^{-2} s^{-1} and the conversion to W m^{-2} was done using Smith and Morel's procedure (Morel and Smith, 1974). From it I_0^m was determined as $\widetilde{I}_0^m = \widetilde{I}_T \pi/2D$, where $\widetilde{I}_T$ is the total received irradiance throughout the day. Daylength was provided from PAR measurement. The diffuse attenuation coefficient was calculated from one percent light level which was given for 150 cruises, based on optical profiles. For the remaining cases, the average value of 0.0435 m^{-1} has been used. Mixed-layer depth Z_m was estimated by the offset of 0.125 kg m^{-3} in potential density at depth, from the surface value (de Boyer Montegut et al., 2004; Suga et al., 2004). Photosynthesis parameters were estimated from chlorophyll and production profiles using the method of Kovač et al. (2016a,b). The data on production, chlorophyll, PAR, one percent light depth and mixed layer depth are publicly available at hahana.soest.hawaii.edu/hot, whereas the data on the photosynthesis parameters are publicly available at jadran.izor.hr/∼kovac/parameters. There were no data on the carbon-chlorophyll ratio.

With the available data, we now proceed to test the solution for the mixed layer production and the shifted Gaussian solution. We further use the solutions with time dependent biomass to predict the influence that time-dependent biomass exerts on mixed layer production. For the measured/known value of a variable/parameter we use $\widetilde{\ }$, e.g., $\widetilde{x}$, whereas for a model variable/parameter we use ordinary symbols, e.g., x.

4. RESULTS

Testing the Canonical Solution for Mixed-Layer Production

A straightforward way of testing the canonical solution is simply to calculate mixed layer production with it. The mixed layer is by definition a region of uniform biomass and the assumption of uniformity in biomass required by the canonical solution is fulfilled. To test the model structure we calculate mixed-layer production with expression (8) and compare it with the measured mixed layer production, calculated with the trapezoidal rule from the measured production profile. The obtained results are displayed in **Figure 2**.

For the mixed-layer biomass in expression (8) the average value of the measured biomass from the first two measuring depths was used. As can be seen from the figure, the match between the modeled and the measured mixed-layer production is quite satisfactory. The coefficient of determination is $r^2 = 0.94$. Therefore, the canonical solution is a good model for mixed-layer production. Some discrepancy is seen at higher values of production. These errors may be caused by the way in which the mixed-layer depth was estimated. It may not always be the case that the mixed layer depth estimated from potential density corresponds well with the depth of active mixing and it is the depth of active mixing that is relevant for biomass homogenization (Franks, 2015). In some data sets we have observed the biomass not to be homogeneous from the surface all the way down to the base of the mixed layer estimated from potential density. This increase in biomass causes more production, than would otherwise occur without it and is reflected in the slight deviation in **Figure 2**.

Testing the Shifted Gaussian Solution

A prerequisite in the application of the solution (25) is to know the values for the parameters of the shifted Gaussian. These have to be estimated from HOT data on chlorophyll profiles. To obtain the biomass parameters we have fitted the shifted Gaussian to the measured chlorophyll profiles by adjusting the parameter values. The conjugate gradient method was used (Baldick, 2006; Knyazev and Lashuk, 2008). The shifted Gaussian was convergent for each chlorophyll profile. Average concentration of background biomass B_0 is 0.085 mg Chl m^{-3}, with a standard deviation of 0.025 mg Chl m^{-3}. Average depth of the deep chlorophyll maximum z_m is 104.12 m, with a standard deviation of 19.00 m.

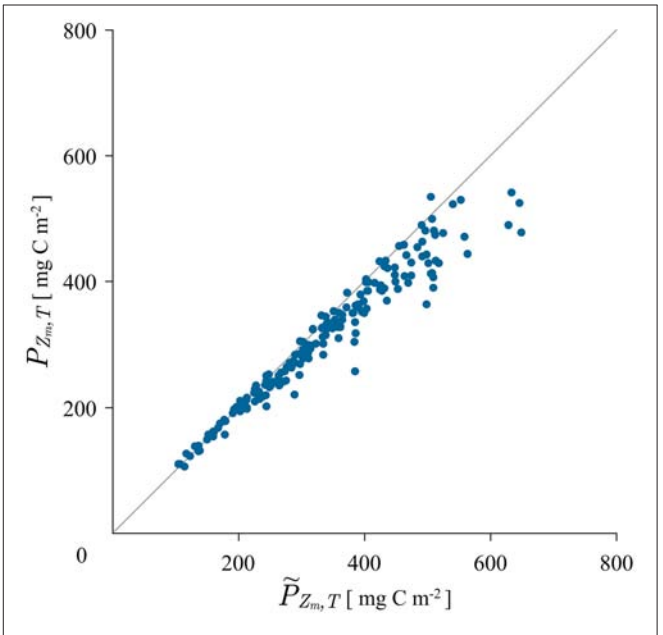


FIGURE 2 | Scatter plot of measured mixed layer production $\widetilde{P}_{Z_m,T}$ calculated with the trapezoidal rule from the measured production profile and modeled mixed layer production $P_{Z_m,T}$ calculated with the canonical solution (8) using the average measured biomass from 5 to 25 m depth.

Average width of the biomass peak σ is 21.93 m, with a standard deviation of 8.89 m, and finally the average height of the biomass peak H is 0.175 mg Chl m^{-3}, with a standard deviation of 0.075 mg Chl m^{-3}.

With the estimated parameters we have further calculated the accuracy of representing the measured biomass profiles with the shifted Gaussian. The biomass given by the shifted Gaussian was calculated at the depth of each measurement and compared with the measured value. The results are shown in **Figure 3**. There are in total 1552 measurements of chlorophyll. As can be seen, the shifted Gaussian is a good model for the biomass profile at all the measuring depths except the last two (150 and 175 m). The coefficient of determination for the data from all depths is 87.84%. However, once the last two depths are excluded the coefficient of determination jumps to a high 98.39%, signifying that the shifted Gaussian is an even better model for the biomass up to the measurement depth of 125 m. The contribution to watercolumn production from depths >125 m is expected to be minimal.

The results of applying the analytical solution (24) are displayed in **Figure 4**. The solution did not exhibit convergent behavior for all the cruise data: out of 194, it converged for 168 cruises. The reason for divergence in the remaining 26 cruises comes from the behavior of the exponential terms $\exp\left((z_{2n-1}^2 - z_m^2)/2\sigma^2\right)$ and $\exp\left((z_{2n}^2 - z_m^2)/2\sigma^2\right)$, which upon summation in solution (24) divergee when σ is high and z_m is small. This corresponds to the case of a wide chlorophyll maximum close to the surface. The solution behaves well when σ is small and z_m large, which is the case of a narrow deep maximum, i.e., a deep chlorophyll maximum. As a measure of the applicability of the given solution the 3σ rule can be applied. For the Gaussian function 99% of the biomass concentration above B_0 is located in the depth interval $(z_m - 3\sigma, z_m + 3\sigma)$. When z_m is larger than 3σ this biomass is located below the surface. That is the dominant situation at HOT station and the solution converges, as evident in **Figure 4**.

Predictions with the Time-Dependent Biomass Solutions

Application of the canonical and the shifted Gaussian solutions is straightforward, given that all the necessary parameters are available. The solutions with time-dependent biomass are more complex to apply due to the requirement for additional parameter values. However, even without knowledge of these parameter magnitudes, the solution can be used to estimate the effect of growth, or decline, in biomass on the magnitude of mixed-layer production. If the biomass is allowed to accumulate, then production is expected to increase in comparison with the case for a time-independent biomass. The opposite holds for the case of a decline in biomass. Just how strong these effects are can easily be calculated for hypothetical cases, but to be as close as possible to real scenarios we apply the solutions with the parameter values typical of HOT.

To illustrate our point, we calculate mixed-layer production using the new solutions with the median values for the mixed-layer depth, assimilation number, mixed-layer biomass, and the attenuation coefficient from HOT. We plot the solutions as a function of the dimensionless noon irradiance to demonstrate

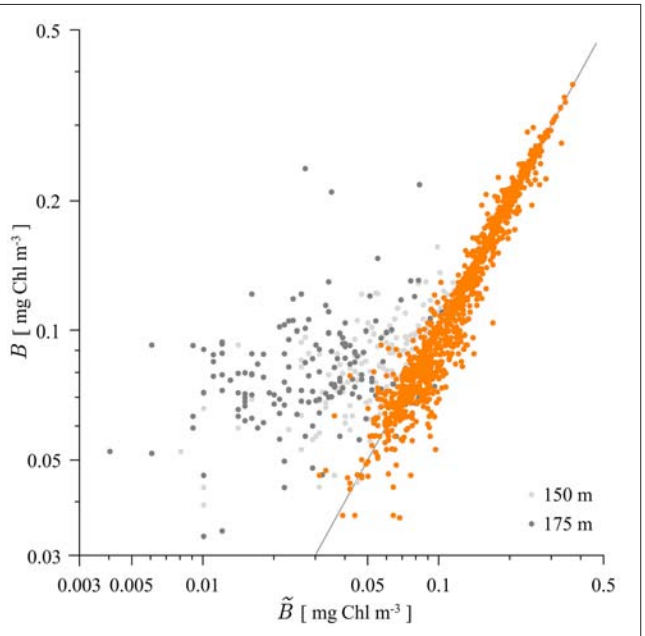


FIGURE 3 | Scatter plot of measured $\widetilde{B}$ and modeled biomass B with the shifted Gaussian (21). The first six measurement depths are given in orange, whereas the data from 150 m measurement depth is given in light gray and the data from 175 m measurement depth is given in gray. The gray line represents the 1:1 model vs. data ratio. For the points above/below the gray line the shifted Gaussian overestimates/underestimates the measured biomass.

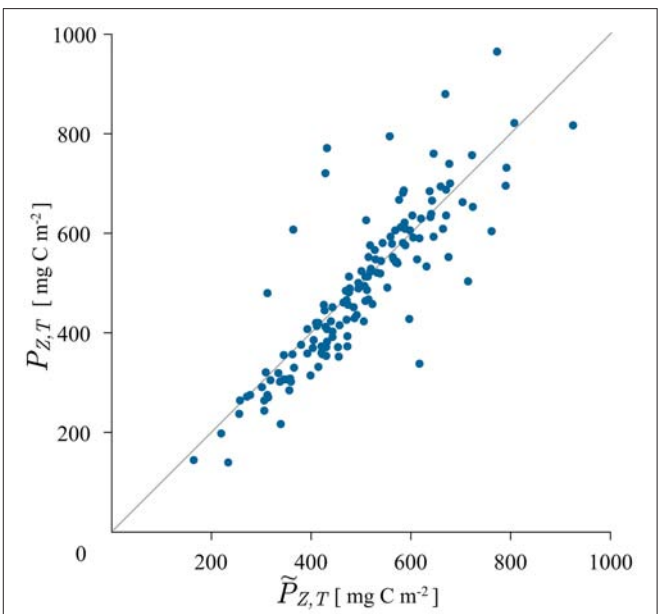


FIGURE 4 | Scatter plot of measured watercolumn production $\widetilde{P}_{Z,T}$, calculated with the trapezoidal rule from the measured production profile, and modeled watercolumn production $P_{Z,T}$, calculated with the analytical solution for the shifted Gaussian biomass (24). The solution did not converge for 26, out of 194 tested cruises.

the effect light has on the magnitude of mixed-layer production (**Figure 5**). In this exercise daylength D is set to 10 h. The blue curve on the figure gives the canonical solution for mixed layer production with the median parameter values: $Z_m = 54.75$ m, $P_m^B = 7.85$ mg C (mg Chl)$^{-1}$ h^{-1}, $B = 0.072$ mg Chl m^{-3}, and $K = 0.043$ m^{-1}. The orange curve is the solution (29) with the carbon-chlorophyll ratio equal to 150 mg C (mg Chl)$^{-1}$ and the light orange with the carbon-chlorophyll ratio equal to to 100 mg C (mg Chl)$^{-1}$. As expected, the production is higher when increase in biomass is accounted for. The red curve corresponds to the solution (34) with the dimensionless factor $L^B D = 1$. In this example $D = 10$ h, which gives $L^B = 0.1$ h^{-1}, i.e., a 10% decrease of biomass per hour. At this rate the biomass at sunset declines to its e-folding value $B(D) = B_0/e$. Finally the pink curve is the approximation (36), of the solution (34), with the same parameter values.

In both cases the effect on mixed layer production is significant. The effect of growth is more pronounced for phytoplankton with lower values of carbon-chlorophyll ratio. This is simply understood by considering that for higher value of χ more carbon is required for a unit increase of chlorophyll. Mathematically, χ appears in the denominator in equation (28) signifying that the change in biomass per unit time is inversely proportional to χ. A straightforward conclusion is that the growth effect on primary production will be more pronounced for phytoplankton with lower values of χ and vice versa. The effect of loss on primary production is also straightforward: the decline of biomass during the day results in lower production. The greater the loss, the greater the diminution in production in comparison with the canonical solution. The exact magnitude of the reduction is now easily calculated using solution (34). The agreement between the approximation (36) with the exact solution (34) is remarkable.

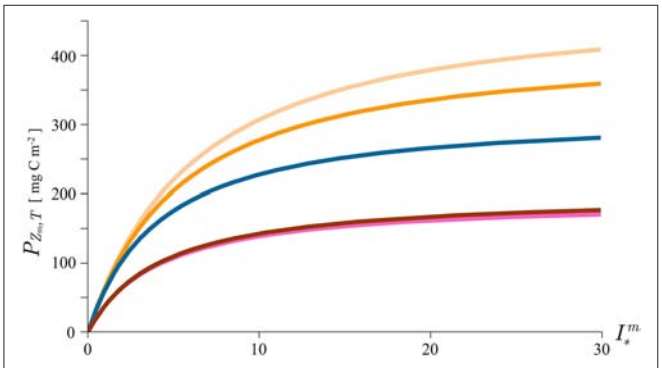


FIGURE 5 | Mixed layer production calculated with the canonical solution (8, blue curve), solution for growing biomass (29, two orange curves with different carbon-chlorophyll ratio), the solution with declining biomass (34, red curve) and the approximation to the solution with declining biomass (36, pink curve). The parameters used in the solutions are typical of the Hawaii area and are given by the median values of: $Z_m = 54.75$ m, $P_m^B = 7.85$ mg C (mg Chl)$^{-1}$ h^{-1}, $B = 0.072$ mg Chl m^{-3}, and $K = 0.043$ m^{-1}. Daylength D is set to 10 h, the dimensionless factor $L^B D$ to 1, carbon to chlorophyll ratio χ to 150 mg C (mg Chl)$^{-1}$ (orange curve) and to 100 mg C (mg Chl)$^{-1}$ (light orange curve).

5. DISCUSSION

Taken together, the solutions presented cover a wide range of different mixing and growth conditions (**Figure 6**), such as might be encountered in the field: from intense mixing and low growth—canonical solution (8), intense mixing and high growth—increasing mixed-layer biomass solution (29), intense mixing and negative growth—declining mixed-layer biomass solution (34), and finally low mixing and low growth—shifted Gaussian solution (24). As all these conditions are indeed observable at times in the ocean, so too are the assumptions of the outlined solutions fulfilled to some extent, and the application of the solutions justified. The only case not solved here is of low mixing and high growth in biomass, which corresponds to the case of time-dependent stratified biomass. Analytical solution to this problem has yet to be found and is a potential topic for future theoretical work. The problem can be treated as formulated in this work, by first solving the growth equation for biomass (12) and inserting the solution directly into the integral for watercolumn production with time-dependent biomass (11).

Irrespective of the model, there are in essence three possible outcomes concerning temporal evolution of biomass. It can be constant, or accumulating or declining. Time dependence in biomass is easily seen in satellite records of chlorophyll (Racault et al., 2012; Cabre et al., 2016) which can potentially serve for assessing whether or not time dependence in biomass should be accounted for when calculating primary production.

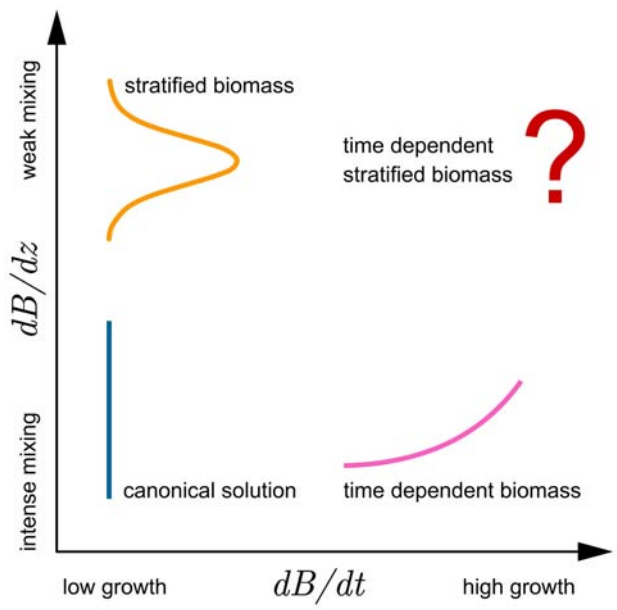


FIGURE 6 | Conceptual diagram highlighting the relations of the model assumptions and analytical solutions with respect to growth (abscissa) and mixing (ordinate). Growth is presented as change in biomass over time dB/ dt, whereas the effects of mixing are presented as dB/ dz. Lines sketch the biomass profiles: orange for stratified biomass with week mixing and low growth, blue for uniform biomass with intense mixing and low growth, and pink for time dependent biomass with intense mixing. Solution for watercolumn production with time dependent stratified biomass has yet to be found.

If the chlorophyll time series displays quiet periods with no temporal dependence in biomass, the canonical solution is then a valid model for mixed layer production. The model with increasing time-dependent biomass is adequate for primary production calculation during a blooming period. This period is also easy to diagnose from satellite chlorophyll time series. After the bloom crashes and the decline of chlorophyll begins to show in the record, the assumption behind the model with declining mixed-layer biomass becomes valid. The value of the loss rate can potentially be determined empirically from the chlorophyll record. Therefore, with respect to the annual cycle of biomass, the solutions presented in this work are each appropriate for a particular period of the year when their respective assumptions are met. Calculating annual watercolumn production in this manner can potentially be a topic for future research.

In earlier literature, integrals for watercolumn production were usually formulated with time-independent biomass (Platt and Sathyendranath, 1993) and could therefore be applied at any stage during the course of the annual cycle. Calculating production in this manner proved suitable for ocean areas where field biomass was known not to change considerably during the interval over which production was calculated, and was in a sense mandatory from the practical standpoint, given that observations of biomass were predominantly performed once per day. However, this approach resulted in biomass accumulation due to primary production having no effect on daily production. Formulating production integrals with biomass constant in time presumes incremental production due to newly-synthesized biomass as insignificant in comparison with production arising from initial biomass. The approach presented here alleviates this limitation and allows a positive feedback between biomass accumulation and primary production, with newly-synthesized biomass contributing to primary production.

The approach advocated here treats the temporal evolution of biomass as being governed by the growth equation, with the growth term dictated by instantaneous production. Coupling this equation to the watercolumn production integral was achieved through the reformulation of the integral via a time-dependent biomass term multiplying the normalized production term. Due to mathematical complexities of this formulation, different approaches for handling the problem were proposed. First, the problem of depth-dependent biomass was analyzed, with an important note that in this context it was viewed as a special case of a steady state solution for biomass distribution. Assuming steady state is reasonable for non-bloom conditions in the oligotrophic ocean, during summer periods at higher latitudes and below the mixed layer (Fennel and Boss, 2003). Biomass profiles of this sort are assumed to be solutions to a more complex equation for biomass, which accounts for other processes besides growth, such as losses by grazing, mixing, and sinking (Beckman and Hense, 2007). These processes are easily included in vertically resolved models for biomass, but the price paid for their inclusion is the increase in model complexity, which makes them difficult to solve analytically, with numerical procedures stepping in as the method of choice to obtain solutions (Huisman et al., 2002; Taylor and Ferrari, 2011).

To circumvent the problem of finding an analytical, steady-state solution to a more general equation, involving not only the nonlinear, time-dependent production term, but also sinking and mixing, we have employed the shifted Gaussian as an approximation to the solution for the biomass profile at steady state. The types of profiles described by the shifted Gaussian are indeed in close agreement with the numerical simulations of the biomass profiles (Beckman and Hense, 2007; Liccardo et al., 2013). More importantly, they are often obtained as results of measurements, and biomass profiles observed in the open ocean are a prototype example for using the shifted Gaussian in the integral for watercolumn production (Platt and Sathyendranath, 1988; Platt et al., 1991b). For vast regions of the open ocean biomass is vertically structured (Longhurst and Harrison, 1989). The most common structure is that of the deep chlorophyll maximum (Navaro and Ruiz, 2013), which is a perfect match for the shifted Gaussian. The deep chlorophyll maximum is often observed to be a quasi permanent feature, existing for months, if not for the duration of the whole annual cycle (Platt and Sathyendranath, 1988). In this case the processes that act to create the deep chlorophyll maximum, and sustain it, are in equilibrium on time scales longer than that of 1 day (Chiswell et al., 2015), justifying the assumption of time independent biomass in daily watercolumn production calculations. Deep chlorophyll maxima are also observed on majority of HOT cruises, with the shifted Gaussian demonstrated here to be a good model for HOT data (**Figure 3**). From this alone stems the legitimacy of using the shifted Gaussian as a model for the biomass profile in the watercolumn production integral for HOT.

In addition, when biomass remains stratified for prolonged periods of time it is safe to assume that mixing is not vigorous enough to distort the established stratification (Liccardo et al., 2013). Mixing itself is caused by various physical agents such as wind, waves, and convection (Franks, 2015), which increase turbulent kinetic energy in the mixed layer. In numerical models this process is parameterized with the mixing coefficient through turbulence closure schemes (Cushman-Roisin and Beckers, 2011). In the presented model however, mixing is not expressed explicitly, but rather the consequences of mixing are assumed implicit through the effect it had on shaping the initial biomass profile. In the field, uniform biomass profiles are most likely associated with strong mixing, whereas stratified biomass profiles are surely associated with less intense mixing. For if mixing were intense, stratification in biomass would not come about, because any incipient stratification would be quickly eroded. This process is well represented in numerical models of the biomass profile (Beckman and Hense, 2007). When assuming a uniform biomass profile we are in fact assuming that mixing was strong enough to cause homogenization of biomass.

This is of course valid for the mixed layer, in which biomass is homogeneous by definition, but it is important to make a distinction between the mixed layer and a layer of active mixing, as highlighted by Franks (2015). The solutions for mixed layer production presented here are strictly valid for a layer of active mixing, since we assume that mixing of newly-synthesized biomass is occurring instantaneously. However, it is often the case that mixed layer depth is determined based on density, or

temperature, offset from the surface value (de Boyer Montegut et al., 2004), even though the mixed layer depth determined in this way may not always correspond well with the depth of active mixing (Franks, 2015). We suspect this to be the cause for the slight bias evident in **Figure 2**. For the model presented here, the assumption of active mixing is required to redistribute the newly-synthesized biomass so that growth occurs uniformly throughout the mixed layer.

Another relevant consequence of the assumption of active mixing is that it enabled the loss rate to be assumed vertically uniform in the mixed layer. In addition to vertical uniformity, not stating the loss rate explicitly left a certain flexibility in the presented model. The loss rate of the phytoplankton population is known to be a complex mixture of various processes such as mixing, sinking, predation, and mortality (Platt et al., 1991a; Zhai et al., 2010), and in the ocean these processes can combine to give a more complex pattern than simply a vertically-uniform loss rate. However, since the work of Sverdrup (1953), it is commonly assumed in theoretical considerations of mixed layer production that losses are in fact uniform due to mixing itself, and here presented model follows this basic approach.

Contrary to the loss rate, a basic feature of the model, shared also by all the previous models of watercolumn production (Platt and Sathyendranath, 1993), is depth resolved instantaneous production, caused by the decline of light intensity with depth. This allows for the accumulation of biomass in the growth equation to proceed in accord with the well established response of primary production to light, stated in (10). Therefore, production is given a more realistic treatment, than losses are. The lack of detailed treatment of the loss terms is not a serious drawback for this model, because parameterizations for respiration, excretion, grazing by micro- and macro-zooplankton, and sedimentation were already studied by Platt et al. (1991a) and Zhai et al. (2010). Inclusion of parametrization for losses given in Zhai et al. (2010) is straightforward here. Losses became as important as production in considerations of bloom dynamics and the model used here can shed some light on this topic, specifically on the Critical Depth Hypothesis.

Implications for the Critical Depth Hypothesis

The new formulation of the mixed-layer production model also has consequences for the Critical Depth Hypothesis. To demonstrate, let us again return to the case of the mixed layer with uniform biomass and the mixing depth given by Z_m. According to Sverdrup (1953), if the mixing depth is shallower than the critical depth Z_{cr} conditions for the initiation of a bloom are fulfilled (Siegel et al., 2002; Fischer et al., 2014). Critical depth is defined as the depth at which the mixed-layer production is balanced by mixed layer losses (Platt et al., 1991a; Sathyendranath et al., 2015) and is derivable from a conservation of biomass equation (Levy, 2015; Mignot et al., 2016). Mathematical formulations of the critical depth criterion are well established in the literature (Sverdrup, 1953; Platt et al., 1991a) and Sverdrup's criterion is usually acknowledged to be a necessary and a sufficient condition for bloom initiation, but not a sufficient condition for determining the amplitude of the bloom (Platt et al., 1994). It specifies whether the bloom can occur, but leaves the bloom amplitude unspecified. With the dynamic approach for primary production calculation, we now demonstrate that Sverdrup's criterion can also be used to calculate the daily increase in mixed layer biomass.

To elaborate, let us augment the production equation (28) with the loss term, so that the time evolution of mixed layer biomass becomes:

$$\frac{\partial}{\partial t}B_{Z_m}(t) = \frac{1}{\chi Z_m}\left(P^B_{Z_m}(t) - L^B_{Z_m}\right)B_{Z_m}(t), \tag{37}$$

where $L^B_{Z_m}$ represents total losses in the broadest sense, arising from respiration, excretion, grazing, sinking, and so on (Zhai et al., 2010). Let the growth and loss of the mixed layer biomass be on the same order of magnitude. We assume the loss rate to be vertically uniform and time independent, a justifiable assumption given the mixing. The solution to this equation at time $D + N$, where N marks the night interval, is then:

$$B_{Z_m}(D+N) = B_{Z_m}(0)\exp\left(\frac{1}{\chi Z_m}\left(P^B_{Z_m,T} - L^B_{Z_m,T}\right)\right). \tag{38}$$

If there is to be an increase in biomass during a 24 h period, that is:

$$B_{Z_m}(D+N) > B_{Z_m}(0), \tag{39}$$

the term in the exponential function has to be greater than zero. It will be greater than zero when the mixed layer production surpasses the losses, that is $P^B_{Z_m,T} > L^B_{Z_m,T}$. Since the net production decreases with depth, because light intensity decreases and losses remain constant, there will be a depth at which the vertically-integrated production $P^B_{Z_m,T}$ equals vertically-integrated losses $L^B_{Z_m}$. The depth at which the two terms exactly balance $P^B_{Z_{cr},T} = L^B_{Z_{cr},T}$ is recognized as the critical depth Z_{cr}, defined by Sverdrup (1953). If the mixing depth is shalower than the critical depth the term in the exponential function is positive.

Invoking the canonical solution for $P^B_{Z_{cr},T}$ we get the implicit expression for the critical depth:

$$\frac{P^B_m D}{K}\left(f(I^m_*) - f(I^m_* e^{-KZ_{cr}})\right) = L^B(D+N)Z_{cr}, \tag{40}$$

where we have used $L^B_{Z_{cr},T} = L^B(D+N)Z_{cr}$. Dividing both sides by Z_{cr} gives:

$$\frac{P^B_m D}{KZ_{cr}}\left(f(I^m_*) - f(I^m_* e^{-KZ_{cr}})\right) = L^B(D+N). \tag{41}$$

This expression states that the average production in the mixed layer equals the average loss when the mixed layer depth is equal to the critical depth, which is precisely the critical depth criterion of Sverdrup (1953). Therefore, when the mixed layer extends to the critical depth there is no accumulation of biomass. Mixing beyond the critical depth leads to losses in the mixed-layer

biomass, and finally mixing not extending to the critical depth leads to accumulation of mixed-layer biomass. The mathematical condition expressing this latter statement is simply:

$$Z_m < Z_{cr}, \tag{42}$$

with Z_{cr} given as a solution of (41). If this condition is met so is condition (39) and accumulation of mixed-layer biomass occurs. Whenever condition (42) is met, there will be a positive increase in mixed layer biomass over the course of 1 day of the following magnitude:

$$\Delta B_{Z_m} = B_{Z_m}(0)\left[\exp\left(\frac{1}{\chi Z_m}\left(P^B_{Z_m,T} - L^B_{Z_m,T}\right)\right) - 1\right]. \tag{43}$$

The critical-depth criterion can now be restated as: when the mixed layer depth is shallower/deeper than the critical depth, there is an increase/decrease in mixed layer biomass of magnitude ΔB_{Z_m}. If the critical depth and the mixed-layer depth are equal, the biomass remains constant.

Implications for Remote Sensing

Applying the models presented in this work as estimators of watercolumn production, based on remotely-sensed data on ocean color, is straightforward. The solutions can be used as a part of existing remote sensing algorithms for watercolumn production, requiring only alterations to be made on the module that calculates watercolumn production, with the analytical solutions taking place of the commonly employed numerical ones. The models are formulated in a similar fashion to our previous ones and when implemented require information on the same parameters and variables (Platt and Sathyendranath, 1993), of which the following are accessible by remote sensing: chlorophyll concentration, surface irradiance and the attenuation coefficient. A relevant distinction from the previous models concerns the assumption of time dependence in biomass. This has implications for remote sensing applications, given that all prior models assumed biomass constant in time, implying that sampling biomass at any time of the day was sufficient for these estimators. However, in the newly presented models temporal evolution of biomass is accounted for and we can ask how the biomass sampled at a specific time of day relates to the initial biomass required by the models.

Acknowledging that ocean color satellites have access to surface chlorophyll (approximately up to the first photic depth $1/K$) and in line with the models presented so far, we write the equation for the time evolution of surface chlorophyll $B(0,t)$ as:

$$\frac{\partial}{\partial t}B(0,t) = \left(\frac{1}{\chi}P^B(0,t) - L^B\right)B(0,t), \tag{44}$$

ignoring advection and mixing, which makes it valid for laterally uniform fields, or for time scales short enough so that neither advection, nor mixing, cause significant changes in biomass over the course of integration. Let us assume that the satellite samples surface biomass at time t_s and label it $\widetilde{B}(0,t_s)$. From the previous equation surface biomass at time t_s is:

$$B^*(0,t_s) = B(0,0)\exp\left(\int_0^{t_s}\left(\frac{1}{\chi}P^B(0,t) - L^B\right)\mathrm{d}t\right). \tag{45}$$

Equating the remotely sensed biomass $\widetilde{B}(0,t_s)$ with $B^*(0,t_s)$ enables as to express the initial surface biomass as:

$$B(0,0) = \widetilde{B}(0,t_s)\exp\left(-\int_0^{t_s}\left(\frac{1}{\chi}P^B(0,t) - L^B\right)\mathrm{d}t\right). \tag{46}$$

This expression takes the remotely sensed biomass $\widetilde{B}(0,t_s)$ and transforms it into the initial biomass $B(0,0)$, under the assumption that biomass evolves according to equation (44). It corrects the remotely sensed surface biomass for dynamical processes of growth and loss, to yield the initial biomass. Taking the satellite overpass time to be at local noon $t_s = D/2$, further enables us to express initial biomass explicitly as:

$$B(0,0) = \widetilde{B}(0,D/2)\exp\left(-\frac{D}{2}\left(\frac{P^B_m}{\chi}f_z(I^m_*) - L^B\right)\right). \tag{47}$$

According to these expressions, having dynamically evolving biomass in the model, affects not only the magnitude of watercolumn production, but also the way in which initial biomass should be calculated from remotely sensed biomass, to compensate for growth and loss. It is important to note that the correction is not linear, but exponential, with respect to production and loss.

6. CONCLUSIONS

The work presented here extends on the standard formulation of daily watercolumn production by allowing for depth- and time-resolved biomass. In the standard formulation, biomass was specified in advance and treated as unrelated to primary production (Platt and Sathyendranath, 1993), leaving prior models without proper dynamics in this regard. To avoid having this problem, we proposed an alternative approach and stated a growth equation for biomass, thus allowing for a time-dependent solution in biomass. Coupling this equation to the watercolumn production integral was achieved by reformulating the integral with time-dependent biomass. Therefore, biomass was related to growth, and as such, subsequently used in the watercolumn production integral.

Depth-resolved biomass was set via an initial condition and we distinguished two possibilities regarding depth dependence in biomass: stratified water column and a mixed layer. For the mixed layer, we further distinguished between growing and declining biomass, providing analytical solutions for watercolumn production in both cases. For the stratified water column we used the shifted Gaussian function to represent biomass profiles and derived an exact analytical solution for daily watercolumn production in this case. No analytical solutions to these problems have been reported in the literature until now.

The new solutions were tested with data from the HOT program. The shifted Gaussian was used to model biomass profiles and it was demonstrated to be a good model. The canonical solution for mixed layer production and the solution with the shifted Gaussian were applied as models of watercolumn production. Both analytical solutions proved to be good models for this open ocean station. The solutions for growing and declining biomass were used to predict mixed layer production in the cases where biomass was increasing/decreasing in accordance with the assumptions of the models.

Sverdrup's critical depth criterion was explored further and an exact expression for mixed layer biomass increment during 1 day was derived. The final statement of the critical depth criterion remained unaltered, although it was based on the argument of growth, whereas prior statements of the criterion were based on the balance between watercolumn production and losses (Sverdrup, 1953; Platt et al., 1991a). The two approaches are now seen equivalent with respect to the final outcome, that being the critical depth criterion.

It was further discussed how to merge the temporally-evolving surface biomass with the remotely-sensed surface biomass, to get to the initial condition on biomass. We expect the processes of growth and loss to affect surface biomass significantly during the initiation or termination phases of a bloom, when the biomass is changing rapidly on time scales shorter than 1 day. A potential course for future research would be to implement this approach in producing maps of chlorophyll from remotely-sensed data.

AUTHOR CONTRIBUTIONS

CG and TP had the original ideas. TP, SS, and ŽK formulated the mathematical models. ŽK, TP, and SA solved the problems analytically. SA did the mathematical analyses of the derivations. CG found the approximate solution. ŽK implemented the models. ŽK and TP wrote the original draft. SS, SA, MM, and CG provided critical reviews and commentary on the draft, and wrote the final manuscript.

FUNDING

This work has been supported in part by Croatian Science Foundation under the projects: MARIPLAN (IP-2014-09-3606) and SCOOL (IP-2014-09-5747). This work is a contribution to the European Space Agency Projects "STSE Marine Primary Production: Model Parameters from Space" and "SEOM Photosynthetically Active Radiation for Primary Production." Additional support from the National Centre for Earth Observation of the Natural Environment Research Council (UK) is also acknowledged. TP acknowledges the support provided by a Jawaharlal Nehru Science Fellowship (Government of India).

ACKNOWLEDGMENTS

We acknowledge the Hawaii Ocean Time Series for the publicly available database (hahana.soest.hawaii.edu/hot). We thank their scientists and staff, and in particular David Karl for his interest.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmars.2017.00163/full#supplementary-material

REFERENCES

Andre, J.-M. (1992). Ocean color remote-sensing and the subsurface vertical structure of phytoplankton pigments. *Deep-Sea Res.* 39, 763–779. doi: 10.1016/0198-0149(92)90119-E

Banse, K., and English, D. C. (1994). Seasonality of coastal zone color scanner phytoplankton pigment in the offshore oceans. *J. Geophys. Res.* 99, 7323–7345. doi: 10.1029/93JC02155

Baldick, R. (2006). *Applied Optimization: Formulation and Algorithms for Engineering Systems, 1st Edn.* Cambridge: Cambridge University Press.

Bannister, T. T. (1974). Production equations in terms of chlorophyll concentration, quantum yield, and upper limit to production. *Limnol. Oceanogr.* 19, 1–12. doi: 10.4319/lo.1974.19.1.0001

Beckman, A., and Hense, I. (2007). Beneath the surface: characteristics of oceanic ecosystems under weak mixing conditions - A theoretical investigation. *Progr. Oceanogr.* 75, 771–796. doi: 10.1016/j.pocean.2007.09.002

Behrenfeld, M. J., and Falkowski, P. G. (1997). Photosynthetic rates derived from satellite-based chlorophyll concentration. *Limnol. Oceanogr.* 42, 1–20. doi: 10.4319/lo.1997.42.1.0001

Buitenhuis, E. T., Hashioka, T., and Quéré, L. (2013). Combined constraints on global ocean primary production using observations and models. *Glob. Biogeochem. Cycles* 27, 847–858. doi: 10.1002/gbc.20074

Cabre, A., Shields, D., Marinov, I., and Kostadinov, T. S. (2016). Phenology of size-partitioned phytoplankton carbon-biomass from ocean color remote sensing and cmip5 models. *Front. Mar. Sci.* 3:39. doi: 10.3389/fmars.2016.00039

Campbell, J., Antoine, D., Armstrong, R., Arrigo, K., Balch, W., Barber, R., et. al. (2002). Comparison of algorithms for estimating ocean primary production from surface chlorophyll, temperature, and irradiance. *Glob. Biogeochem. Cycles* 16, 9.1–9.15. doi: 10.1029/2001GB001444

Chavez, F. P., Messié, M., and Pennington, J. T. (2011). Marine primary production in relation to climate variability and change. *Annu. Rev. Mar. Sci.* 3, 227–260. doi: 10.1146/annurev.marine.010908.163917

Chiswell, S. M., Calil, P. H. R., and Boyd, P. W. (2015). Spring blooms and annual cycles of phytoplankton: a unified perspective. *J. Plankton Res.* 37, 500–508. doi: 10.1093/plankt/fbv021

Cushing, D. H. (1971). Upwelling and the production of fish. *Adv. Mar. Biol.* 9, 255–334. doi: 10.1016/S0065-2881(08)60344-2

Cushman-Roisin, B., and Beckers, J. M. (2011). *Introduction to Geophysical Fluid Dynamics: Physical and Numerical Aspects, 2nd Edn.* San Diego, CA: Academic Press.

de Boyer Montegut, C. B., Madex, G., Fischer, A. S., Lazar, A., and Iudicone, D. (2004). Mixed layer depth over the global ocean: an examination of profile data and a profile-based climatology. *J. Geophys. Res.* 109, C12003. doi: 10.1029/2004JC002378

Eppley, R. W. E., Stewart, E., Abbott, M. R., and Heyman, U. (1985). Estimating ocean primary production from satellite chlorophyll. Introduction to regional differences and statistics for the southern California Bight. *J. Plankton Res.* 7, 57–70. doi: 10.1093/plankt/7.1.57

Fennel, K., and Boss, E. (2003). Subsurface maxima of phytoplankton and chlorophyll: steady-state solutions from a simple model. *Limnol. Oceanogr.* 48, 1521–1534. doi: 10.4319/lo.2003.48.4.1521

Fischer, A., Moberg, E., Alexander, H., Brownlee, E., Hunter-Cevera, K., Pitz, K., et al. (2014). Sixty years of Sverdrup: a retrospective of progress in the study of phytoplankton blooms. *Oceanography* 27, 222–235. doi: 10.5670/oceanog.2014.26

Franks, P. J. S. (2015). Has Sverdrup's critical depth hypothesis been tested? Mixed layers vs. turbulent layers. *ICES J. Mar. Sci.* 72, 1897–1907. doi: 10.1093/icesjms/fsu175

Hodges, B. A., and Rudnick, D. L. (2004). Simple models of steady deep maxima in chlorophyll and biomass. *Deep Sea Res. Part I Oceanogr. Res. Pap.* 51, 999–1015. doi: 10.1016/j.dsr.2004.02.009

Honjo, S., Manganini, S. J., Krishfield, R. A., and Francois, R. (2008). Particulate organic carbon fluxes to the ocean interior and factors controlling the biological pump: a synthesis of global sediment trap programs since 1983. *Progr. Oceanogr.* 76, 217–285. doi: 10.1016/j.pocean.2007.11.003

Huisman, J., Arrayas, M., Ebert, U., and Sommeijer, B. (2002). How do sinking phytoplankton species manage to persist? *Am. Nat.* 159, 245–254. doi: 10.1086/338511

Jassby, A. D., and Platt, T. (1976). Mathematical formulation of the relationship between photosynthesis and light for phytoplankton. *Limnol. Oceanogr.* 21, 540–547. doi: 10.4319/lo.1976.21.4.0540

Karl, D. M., and Church, M. J. (2014). Microbial oceanography and the Hawaii Ocean Time-series programme. *Nat. Rev. Microbiol.* 12, 1–15. doi: 10.1038/nrmicro3333

Karl, D. M., Dore, J. E., Lukas, R., Michaels, A. F., Bates, R. N., and Knap, A. (2001). Building the long-term picture: the U. S. JGOFS time-series programs. *Oceanography* 14, 6–17. doi: 10.5670/oceanog.2001.02

Kirk, J. T. O. (2011). *Light and Photosynthesis in Aquatic Ecosystems, 3rd Edn.* New York, NY: Cambridge University Press.

Knyazev, A. V., and Lashuk, I. (2008). Steepest descent and conjugate gradient methods with variable preconditioning. *SIAM J. Matrix Anal. Appl.* 29, 1267–1281. doi: 10.1137/060675290

Kovač, Ž., Platt, T., Sathyendranath, S., and Morović, M. (2016a). Analytical solution for the vertical profile of daily production in the ocean. *J. Geophys. Res.* 121, 3532–3548. doi: 10.1002/2015JC011293

Kovač, Ž., Platt, T., Sathyendranath, S., Morović, M., and Jackson, T. (2016b). Recovery of photosynthesis parameters from *in situ* profiles of phytoplankton production. *ICES J. Mar. Sci.* 73, 275–285. doi: 10.1093/icesjms/fsv204

Levy, M. (2015). Exploration of the critical depth hypothesis with a simple NPZ model. *ICES J. Mar. Sci.* 72, 1916–1925. doi: 10.1093/icesjms/fsv016

Liccardo, A., Fierro, A., Iudicone, D., Bouruet-Aubertot, P., and Dubroca, L. (2013). Response of the deep chlorophyll maximum to fluctuations in vertical mixing intensity. *Progr. Oceanogr.* 109, 33–46. doi: 10.1016/j.pocean.2012.09.004

Longhurst, A., Sathyendranath, S., and Platt, T. (1995). An estimate of global primary production in the ocean from satellite radiometer data. *J. Plankton Res.*, 17, 1245–1271. doi: 10.1093/plankt/17.6.1245

Longhurst, A. R. (1998). *Ecological Geography of the Sea, 2nd Edn.* New York, NY: Academic Press.

Longhurst, A. R., and Harrison, W. G. (1989). The biological pump: profiles of plankton production and consumption in the upper ocean. *Progr. Oceanogr.* 22, 47–123. doi: 10.1016/0079-6611(89)90010-4

Mann, K. H., and Lazier, J. R. N. (2006). *Dynamics of Marine Ecosystems: Biological-Physical Interactions in the Oceans, 3rd Edn.* Oxford: Blackwell Publishing.

Mignot, A., Claustre, H., D´Ortenzio, F., Xing, X., D´Poteau, A., and Ras, J. (2011). From the shape of the vertical profile of *in vivo* fluorescence to Chlorophyll-*a* concentration. *Biogeosciences* 8, 2391–2406. doi: 10.5194/bg-8-2391-2011

Mignot, A., Claustre, H., Uitz, J., Poteau, A., D´Ortenzio, F., and Xing, X. (2014). Understanding the seasonal dynamics of phytoplankton biomass and the deep chlorophyll maximum in oligotrophic environments: a Bio-Argo float investigation. *Glob. Biogeochem. Cycles* 28, 856–876. doi: 10.1002/2013GB004781

Mignot, A., Ferrari, R., and Mork, K. A. (2016). Spring bloom onset in the Nordic Seas. *Biogeosciences* 13, 3485–3502. doi: 10.5194/bg-13-3485-2016

Morel, A., and Berthon, J.-F. (1989). Surface pigments, algal biomass profiles, and potential production of the euphotic layer: relationships reinvestigated in view of remote-sensing applications. *Limnol. Oceanogr.* 34, 1545–1562. doi: 10.4319/lo.1989.34.8.1545

Morel, A., and Smith, R. C. (1974). Relation between total quanta and total energy for aquatic photosynthesis. *Limnol. Oceanogr.* 19, 591–600. doi: 10.4319/lo.1974.19.4.0591

Navaro, G., and Ruiz, J. (2013). Hysteresis conditions the vertical position of deep chlorophyll maximum in the temperate ocean. *Glob. Biogeochem. Cycles* 27, 1013–1022. doi: 10.1002/gbc.20093, 2013

Platt, T. (1986). Primary production of the ocean water column as a function of surface light intensity: algorithms for remote sensing. *Deep-Sea Res.* 33, 149–163. doi: 10.1016/0198-0149(86)90115-9

Platt, T., and Jassby, A. D. (1976). The relationship between photosynthesis and light for natural assemblages of coastal marine phytoplankton. *J. Phycol.* 12, 421–430. doi: 10.1111/j.1529-8817.1976.tb02866.x

Platt, T., Bird, D. F., and Sathyendranath, S. (1991a). Critical depth and marine primary production. *Proc. R. Soc. B* 246, 205–217.

Platt, T., Broomhead, D. S., Sathyendranath, S., Edwards, A. M., and Murphy, E. J. (2003). Phytoplankton biomass and residual nitrate in the pelagic ecosystem. *Proc. R. Soc. A* 459, 1063–1073. doi: 10.1098/rspa.2002.1079

Platt, T., Caverhill, C., and Sathyendranath, S. (1991b). Basin-scale estimates of oceanic primary production by remote sensing: the North Atlantic. *J. Geophys. Res.* 96, 15147–15159. doi: 10.1029/91JC01118

Platt, T., Denman, K. L., and Jassby, A. D. (1977). "Modelling the productivity of phytoplankton," in *The Sea: Ideas and Observations on Progress in the Study of the Seas*, eds E. D. Goldberg, I. N. McCave, J. J. O'Brien, and J. H. Steele (New York, NY: Wiley), 807–856.

Platt, T., Gallegos, C. L., and Harrison, W. G. (1980). Photoinhibition of photosynthesis in natural assemblages of marine phytoplankton. *J. Mar. Res.* 38, 687–701.

Platt, T., and Sathyendranath, S. (1988). Oceanic primary production: estimation by remote sensing at local and regional scales. *Science* 241, 1613–1620. doi: 10.1126/science.241.4873.1613

Platt, T., and Sathyendranath, S. (1991). Biological production models as elements of coupled, atmosphere-ocean models for climate research. *J. Geophys. Res.* 96, 2585–2592. doi: 10.1029/90JC02305

Platt, T., and Sathyendranath, S. (1993). Estimators of primary production for interpretation of remotely sensed data on ocean color. *J. Geophys. Res.* 98, 14561–14576. doi: 10.1029/93JC01001

Platt, T., and Sathyendranath, S. (2008). Ecological indicators for the pelagic zone of the ocean from remote sensing. *Remote Sens. Environ.* 112, 3426–3436. doi: 10.1016/j.rse.2007.10.016

Platt, T., Sathyendranath, S., Caverhill, C. M., and Lewis, M. R. (1988). Ocean primary production and available light: further algorithms for remote sensing. *Deep-Sea Res.* 35, 855–879. doi: 10.1016/0198-0149(88)90064-7

Platt, T., Sathyendranath, S., Forget, M., White, G. N., Caverhill, C., Bouman, H., et al. (2008). Operational estimation of primary production at large geographical scales. *Remote Sens. Environ.* 112, 3437–3448. doi: 10.1016/j.rse.2007.11.018

Platt, T., Sathyendranath, S., and Ravindran, P. (1990). Primary production by phytoplankton: analytic solutions for daily rates per unit area of water surface. *Proc. R. Soc. B* 214, 101–111. doi: 10.1098/rspb.1990.0072

Platt, T., Woods, J. D., Sathyendranath, and Barkmann, W. (1994). Net primary production and stratification in the ocean. *Geophys. Monogr.* 85, 247–254.

Racault, M.-F., Quéré, L., Buitenhuis, E., Sathyendranath, S., and Platt, T. (2012). Phytoplankton phenology in the global ocean. *Ecol. Indic.* 14, 152–163. doi: 10.1016/j.ecolind.2011.07.010

Rodhe, W. (1965). "Standard correlations between pelagic photosynthesis and light," in *Primary Productivity in Aquatic Environment*, ed C. R. Goldman (Berkeley: University of California Press), 365–381.

Ryther, J. H. (1956). Photosynthesis in the ocean as a function of light intensity. *Limnol. Oceanogr.* 1, 61–70. doi: 10.4319/lo.1956.1.1.0061

Ryther, J. H., and Yentsch, C. S. (1957). The estimation of phytoplankton production in the ocean from chlorophyll and light data. *Limnol. Oceanogr.* 11, 281–286. doi: 10.1002/lno.1957.2.3.0281

Sathyendranath, S., Ji, R., and Browman, H. I. (2015). Revisiting Sverdrup's critical depth hypothesis. *ICES J. Mar. Sci.* 72, 1892–1896. doi: 10.1093/icesjms/fsv110

Sathyendranath, S., and Platt, T. (1989). Remote sensing of ocean chlorophyll: consequence of nonuniform pigment profile. *Appl. Opt.* 28, 490–495. doi: 10.1364/AO.28.000490

Sathyendranath, S., and Platt, T. (2007). Spectral effects in bio-optical control on the ocean system. *Oceanologia* 49, 5–39.

Sathyendranath, S., Stuart, V., Nair, A., Oka, K., Nakane, T., Bouman, H., et al. (2009). Carbon-to-chlorophyll ratio and growth rate of phytoplankton in the sea. *Mar. Ecol. Progr. Ser.* 383, 73–84. doi: 10.3354/meps07998

Siegel, D. A., Buesseler, K. O., Doney, S. C., Sailley, S. F., Behrenfeld, M. J., and Boyd, P. W. (2014). Global assessment of ocean carbon export by combining satellite observations and food-web models. *Glob. Biogeochem. Cycles* 28, 181–196. doi: 10.1002/2013GB004743

Siegel, D. A., Doney, S. C., and Yoder, J. A. (2002). The North Atlantic spring phytoplankton bloom and sverdrup's critical depth hypothesis. *Science* 296, 730–732. doi: 10.1126/science.1069174

Smith, R. C., and Baker, K. S. (1978). The bio-optical state of ocean waters and remote sensing. *Limnol. Oceanogr.* 23, 247–259. doi: 10.4319/lo.1978.23.2.0247

Steemann Nielsen, E. (1952). The use of radioactive carbon (^{14}C) for measuring organic production in the sea. *J. Cons. Int. Pour l'Explor. Mer*, 18, 117–140. doi: 10.1093/icesjms/18.2.117

Stramska, M., and Stramski, D. (2005). Effects of a nonuniform vertical profile of chlorophyll concentration on remote-sensing reflectance of the ocean. *Appl. Opt.* 44, 1735–1747. doi: 10.1364/AO.44.001735

Strickland, J. D. H., and Parsons, T. R. (1972). *A Practical Handbook on Seawater Analysis*. Otawa, ON: Fisheries research Board of Canada.

Suga, T., Motoki, K., Aoki, Y., and MacDonald, A. M. (2004). The North Pacific climatology of winter mixed layer and mode waters. *J. Phys. Oceanogr.* 34, 3–22. doi: 10.1175/1520-0485(2004)034<0003:TNPCOW>2.0.CO;2

Sverdrup, H. U. (1953). On conditions for the vernal blooming of phytoplankton. *J. Cons. Int. Pour l Explor. Mer* 18, 287–295. doi: 10.1093/icesjms/18.3.287.

Talling, J. F. (1957). The phytoplankton population as a compound photosynthetic system. *New Phytol.* 56, 133–149. doi: 10.1111/j.1469-8137.1957.tb06962.x

Taylor, J. R., and Ferrari, R. (2011). Shutdown of turbulent convection as a new criterion for the onset of spring phytoplankton blooms. *Limnol. Oceanogr.* 56, 2293–2307. doi: 10.4319/lo.2011.56.6.2293

Uitz, J., Claustare, H., Morel, A., and Hooker, S. B. (2009). Vertical distribution of phytoplankton communities in open ocean: an assessment based on surface chlorophyll. *J. Geophys. Res.* 111, C08005. doi: 10.1029/2005JC 003207

Vargas, M., Brown, C. W., and Sapiano, M. R. P. (2009). Phenology of marine phytoplankton from satellite ocean color measurements. *Geophys. Res. Lett.* 36, L01608. doi: 10.1029/2008GL036006

Volt, T., and Hoffert, M. I. "Ocean carbon pumps: analysis of relative strength and efficiencies of in ocean-driven circulation atmospheric CO2 changes," in *The Carbon Cycle and Atmospheric CO2: Natural Variation Archean to Present*, eds E. T. Sundquist and W. S. Broecker (Washington, DC: American Geophysical Union), 99–110.

Westberry, T., Behrenfeld, M. J., Siegel, D. A., and Boss, E. (2008). Carbon-based primary productivity modeling with vertically resolved photoacclimation. *Glob. Biogeochem. Cycles* 22, GB2024. doi: 10.1029/2007 GB003078

Xiu, P., Liu, Y., and Tang, J. (2008). Variations of ocean color parameters with nonuniform vertical profiles of chlorophyll concentration. *Int. J. Remote Sens.* 29, 831–849. doi: 10.1080/01431160701280991

Zhai, L., Platt, T., Tang, C., Sythyendranath, S., Fuentes-Yaco, C., Devred, E. et al. (2010). Seasonal and geographic variations in phytoplankton losses from the mixed layer on the Northwest Atlantic Shelf. *J. Mar. Syst.* 80, 36–46. doi: 10.1016/j.jmarsys.2009.09.005

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Kovač, Platt, Antunović, Sathyendranath, Morović and Gallegos. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Marine Science

ORIGINAL RESEARCH
published: 05 July 2017
doi: 10.3389/fmars.2017.00203

Synergistic Exploitation of Hyper- and Multi-Spectral Precursor Sentinel Measurements to Determine Phytoplankton Functional Types (SynSenPFT)

Svetlana N. Losa[1], Mariana A. Soppa[1], Tilman Dinter[1,2], Aleksandra Wolanin[2], Robert J. W. Brewin[3,4], Annick Bricaud[5,6], Julia Oelker[2], Ilka Peeken[1], Bernard Gentili[5,6], Vladimir Rozanov[2] and Astrid Bracher[1,2]*

[1] Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany, [2] Institute of Environmental Physics, University of Bremen, Bremen, Germany, [3] Plymouth Marine Laboratory, Plymouth, United Kingdom, [4] Plymouth Marine Laboratory, National Centre for Earth Observation, Plymouth, United Kingdom, [5] Sorbonne Universités, UPMC-Université Paris-VI, UMR 7093, Laboratoire d'Océanographie de Villefranche, Observatoire Océanologique, Villefranche-sur-Mer, France, [6] Centre National de la Recherche Scientifique, UMR 7093, LOV, Observatoire océanologique, Villefranche-sur-Mer, France

OPEN ACCESS

Edited by:
Laura Lorenzoni, University of South Florida, United States

Reviewed by:
Lionel Arteaga, Princeton University, United States
Raphael M. Kudela, University of California, Santa Cruz, United States
Cecile S. Rousseaux, USRA, United States

***Correspondence:**
Svetlana N. Losa
svetlana.losa@AWI.de

Specialty section:
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

Received: *08 March 2017*
Accepted: *13 June 2017*
Published: *05 July 2017*

Citation:
Losa SN, Soppa MA, Dinter T, Wolanin A, Brewin RJW, Bricaud A, Oelker J, Peeken I, Gentili B, Rozanov V and Bracher A (2017) Synergistic Exploitation of Hyper- and Multi-Spectral Precursor Sentinel Measurements to Determine Phytoplankton Functional Types (SynSenPFT). Front. Mar. Sci. 4:203. doi: 10.3389/fmars.2017.00203

We derive the chlorophyll *a* concentration (Chla) for three main phytoplankton functional types (PFTs) – diatoms, coccolithophores and cyanobacteria – by combining satellite multispectral-based information, being of a high spatial and temporal resolution, with retrievals based on high resolution of PFT absorption properties derived from hyperspectral satellite measurements. The multispectral-based PFT Chla retrievals are based on a revised version of the empirical OC-PFT algorithm applied to the Ocean Color Climate Change Initiative (OC-CCI) total Chla product. The PhytoDOAS analytical algorithm is used with some modifications to derive PFT Chla from SCIAMACHY hyperspectral measurements. To combine synergistically these two PFT products (OC-PFT and PhytoDOAS), an optimal interpolation is performed for each PFT in every OC-PFT sub-pixel within a PhytoDOAS pixel, given its Chla and its *a priori* error statistics. The synergistic product (SynSenPFT) is presented for the period of August 2002 March 2012 and evaluated against PFT Chla data obtained from *in situ* marker pigment data and the NASA Ocean Biogeochemical Model simulations and satellite information on phytoplankton size. The most challenging aspects of the SynSenPFT algorithm implementation are discussed. Perspectives on SynSenPFT product improvements and prolongation of the time series over the next decades by adaptation to Sentinel multi- and hyperspectral instruments are highlighted.

Keywords: synergistic, Sentinel, satellite retrievals, phytoplankton functional type

1. INTRODUCTION

Phytoplankton supplies over 90% of the nutrition consumed by the higher trophic levels of the marine ecosystem and contributes to 50% of the global primary production (Field et al., 1998). Therefore, it is very important for global biogeochemical fluxes (e.g., carbon) since it fixes atmospheric carbon, CO_2, and produces organic carbon compounds. In combination with

physical factors, this process helps to determine which part of the ocean is a sink or source of CO_2 (Laufkötter et al., 2016). Furthermore, global biogeochemical fluxes can be impacted by the size and composition of phytoplankton, in addition to the structure of the trophic community. Ocean color remote sensing has revolutionized our understanding of these processes in the past decades by providing globally continuous data on surface chlorophyll *a* concentration (Chla, mg m^{-3}). Chla, however, is an index of total phytoplankton biomass within which each phytoplankton group has a specific morphology and photophysiology and plays a particular role in biogeochemical cycling. For instance, diatoms are the phytoplankton silicifiers, which contribute to most of the primary production and biomass during the spring bloom in temperate and polar regions (Buesseler, 1998). Their importance is related to the efficiency of carbon export through the direct sinking of single cells, key grazing pathways and through mass sedimentation events at the end of the spring blooms when nutrients are depleted (Le Qur et al., 2005). Coccolithophores are the main planktonic calcifiers in the ocean: through building and releasing calcium carbonate plates, coccoliths, coccolithophores make a major contribution to the total content of particulate inorganic carbon in the open oceans (Milliman, 1993; Ackleson et al., 1988). Cyanobacteria regenerate nutrients and, therefore, influence the marine recycled production (Waterbury et al., 1986; Morán et al., 2004).

The ability to observe the phenology and variability of different phytoplankton groups simultaneously is "a scientific priority for understanding the marine food web, and ultimately predicting the oceans role in regulating climate and responding to climate change on various time scales" (Bracher et al., 2017b).

Abbreviations: AWI, Alfred Wegener Institute; Chla, chlorophyll *"a"* concentration; c-micro, microphytoplankton Chla; CNRS, Centre National de la Recherche Scientifique; c-nano, nanophytoplankton Chla; c-pico, picophytoplankton Chla; coc, coccolithophores Chla; cya, cyanobacteria Chla; dia, diatom Chla; DOAS, Differential Optical Absorption Spectroscopy; DPA, Diagnostic Pigment Analysis; ESA, European Space Agency; f-micro, fraction of microplankton Chla to total Chla; f-nano, fraction of nanoplankton Chla to TChla; f-pico, fraction of picoplankton Chla to TChla; f-PSC, fraction of phytoplankton size class Chla to TChla; IOCCG, International Ocean-Color Coordinating Group; IOP, Inherent Optical Properties; IUP, Institute of Environmental Physics; LOV, Laboratoire d'Oceanographie de Villefranche; LUT, Look Up Table; MAE, Mean Absolute Error; MAD, Mean Absolute Differences; micro, microplankton; nano, nanoplankton; NASA, National Aeronautics and Space Administration; NOBM, NASA Ocean Biogeochemical Model; OCI, Ocean Color Instrument; OC-PFT, Algorithm of Hirata et al. (2011) to retrieve phytoplankton functional types; OC-CCI, Ocean Color Climate Change Initiative; OI, Optimal Interpolation; OLCI, Ocean Land Color Instrument; OMI, Ozone Monitoring Instrument; PACE, Plankton, Aerosol, Cloud, ocean Ecosystem; PCA, Principal Component Analysis; PFT, Phytoplankton Functional Type; PhytoDOAS, DOAS applied for retrieval of phytoplankton and PFT biomass; pico, picoplankton; PSC, Phytoplankton Size Class; PML, Plymouth Marine Laboratory; PVR, product validation report; QAA Quasi-Analytical Algorithm; RMSD, root mean square difference; Rrs, Rotational Raman Scattering; RTM, Radiative Transfer Model; SCIAMACHY, Scanning Imaging Absorption Spectrometers for Atmospheric Chartography; SCIATRAN, Radiative Transfer Model and Retrieval Algorithm; SEOM, Scientific Exploitation of Operational Missions; Sf, phytoplankton size factor; SynSenPFT, Synergistic hyper- and multispectral satellite PFT product; SZA, Solar zenith angle; TChla, total chlorophyll *"a"* concentration; TC, Triple collocation; TOA, Top of Atmosphere; TROPOMI, TROPOspheric Monitoring Instrument; UV-A, Ultraviolet A; VRS, Vibrational Raman Scattering; WF, Weighting-function.

Highly resolved information about phytoplankton diversity is also essentially required for a variety of socio-economic applications dealing with predicting harmful algal blooms, eutrophication, hypoxia and other events affecting water quality (see IOCCG 2009). Nearly all global ocean color products presently available are retrieved from multispectral ocean color sensors (IOCCG, 2014; Mouw et al., 2017) providing information on water leaving radiance at number of spectral bands (up to 5 used for ocean color) with the bands width varying from 7.5 nm to 20 nm. Compared to available sensors allowing for more detailed spectral information (hyperspectral), multi-spectral sensors provide data with relatively high spatial (∼1 to 4.6 km) resolution and temporal (within a few days) coverage. This spatial and temporal resolution as well as the legacy of the data make them especially attractive for exploiting in phytoplankton diversity and phenology.

Some of the well-known multispectral-based algorithms differentiating phytoplankton groups (IOCCG 2014) are applied to the total Chla (TChla) product that due to strong absorption of Chla (a pigment produced by all phytoplankton species) at 443 nm can be retrieved quite accurately given the spectral resolution and radiometric sensitivity of the ocean color sensors (Mouw et al., 2017). To derive either the fraction or Chla of phytoplankton size classes (PSC) (Uitz et al., 2006; Brewin et al., 2010, 2015; Hirata et al., 2011) or phytoplankton functional types (PFT) (Hirata et al., 2011), these so-called abundance-based approaches use empirical relationships between TChla and *in situ* marker pigments determined using high precision liquid chromatography (HPLC). The HPLC method itself is based on spectral absorption properties of the specific phytoplankton marker pigments, which allows to identify particular phytoplankton groups. However, the empirical nature of their relation to TChla (usually estimated in a global context) leads to some limitations of the abundance-based approaches: they cannot predict atypical associations, the relationships derived may differ regionally (Soppa et al., 2014), may vary with environmental conditions and thus the model parameters may change in a future ocean state (Brewin et al., 2010; Ward, 2015; Brewin et al., 2017) or if shifts in phytoplankton composition occur without any change in TChla. For instance, the shift from diatom-dominated composition to haptophyte-dominated, as observed in the Arctic Ocean under resent warming conditions and indicated by (Nöthig et al., 2015).

Other so-called spectral-based approaches, besides being different in underlying physical principles and input information (radiance, absorption, backscattering), rely explicitly on spectral signatures of specific PFT or PSC retrieved directly from the input information (or satellite radiometric measurements) (for a comprehensive overview see Mouw et al. 2017) and, therefore, have the potential to discriminate blooms of phytoplankton groups which are not correlated with TChla. These methods also use to some extent empirical relationships by parameterizing their spectral models based on, for example, choosing certain (mean) phytoplankton absorption spectra measured in the field as representative of specific PSC or PFT. Current discretization of multispectral-based inputs (up to five wavebands) limits the ability to differentiate among the optical imprints of different

water constituents. Although phytoplankton types have different marker pigments, the differences in the spectral absorption structures can be small, since they also have many pigments in common. The small number of wavelength bands and the broadband resolution of multispectral sensors provide only limited information on the difference of the phytoplankton absorption structures (Bricaud et al., 2004; Organelli et al., 2011). In the study by Organelli et al. (2013) it was highlighted that the use of hyperspectral data (radiometric information with a spectral resolution higher than those coming from the current ocean color sensors) would allow to improve the accuracy of spectral-based phytoplankton composition retrievals. This was also supported by Wolanin et al. (2016b) who applied PFT retrievals to a large synthetically simulated data set testing different band settings for multispectral data and various resolutions of hyperspectral information. The authors showed that hyperspectral data are the most beneficial for various PFT retrievals.

Former and current satellite instruments with a very high spectral resolution (1 nm and higher) provide the opportunity for distinguishing more accurately multiple PFTs using spectral approaches. The capability to retrieve quantitatively major PFTs (even several PFTs simultaneously) based on their absorption properties has been demonstrated with the Phytoplankton Differential Optical Absorption Spectroscopy (PhytoDOAS) method (Bracher et al., 2009; Sadeghi et al., 2012) in the open ocean using hyperspectral satellite data from the sensor "SCanning Imaging Absorption Spectrometer for Atmospheric CHartographY" (SCIAMACHY). Being originally developed for atmospheric applications, hyperspectral sensors like SCIAMACHY do not provide operational water-leaving radiance products as do ocean color sensors. Hence, the PhytoDOAS algorithm was designed to retrieve three PFTs directly from top-of-atmosphere radiances, which requires handling properly strong atmospheric absorbers. Due to the high spectral resolution of SCIAMACHY (<0.5 nm) it is possible to separate high frequency absorptions features of each particular PFT from optical signature of the relevant atmospheric absorbers (broad band effects are accounted for by a low order polynomial). However, the global exploitation of hyperspectral satellite data for ocean color applications has been so far very limited, since the pixel size of these data is very large (30 km by 60 km per pixel) and global coverage by these measurements is reached only within six days. This circumstance constrains any assessment of the retrievals accuracy with *in situ* point measurements. It also limits the application of hyperspectral-based PFT data products.

The aim of this study was to overcome the aforementioned short-comings of current multispectral PFT products (supplying either dominant groups only, Bracher et al. 2017a, or data products with strong linkage to *a priori* information) and of current PhytoDOAS-data products (low temporal and spatial coverage). This was done based on the synergistic use of low spatial resolution hyperspectral data with higher spatial and temporal resolution multispectral data. In this study a quantitative estimation of the abundance (given as Chla in mg/m^3) of the same (diatoms and cyanobacteria/i.e., prokaryotic phytoplankton) or similar (coccolithophores vs. haptophytes) PFTs are obtained by the PhytoDOAS (Bracher et al., 2009; Sadeghi et al., 2012) and by the abundance-based OC-PFT algorithm (Hirata et al., 2011). Comparison to other satellite products bears the limitation that different aspects of diversity of phytoplankton groupings are compared (see also Bracher et al. 2017a for mismatch between satellite products). **Figure 1** summarizes the different classifications of phytoplankton used as PFTs within the study (PhytoDOAS: Bracher et al. 2009, OC-PFT: Hirata et al. 2011). Note that OC-PFT retrieves haptophytes while PhytoDOAS coccolithophores, a sub-group of haptophytes (often dominating). **Figure 1** also illustrates how the considered PFTs relate to main PSCs (Brewin et al., 2010, 2015) and size factor (Sf, Ciotti and Bricaud 2006).

This paper highlights the first development of a synergistic algorithm (SynSenPFT) and shows its potential via evaluation against *in situ* PFT data (derived from HPLC) and intercomparison with the satellite phytoplankton composition products focusing on the variation in phytoplankton size (Ciotti and Bricaud, 2006; Brewin et al., 2010, 2015) and the PFT products from the NASA ocean biogeochemical model (NOBM, Gregg and Casey 2007). Given the evaluation results, we discuss the potential of this method for future applications to other recent and upcoming sensors data, such as the Ocean Land Color Instrument (OLCI) on Sentinel-3 (in operation since 2016), the hyperspectral sensor Ozone Monitoring Instrument (OMI) on Aura (in operation since 2004) and the TROPOspheric Monitoring Instrument (TROPOMI) on Sentinel-5P to be launched in 2017. Such a synergistic approach shall enable a PFT time series from 2002 onwards which can then be extended by exploiting the data of the hyperspectral global Ocean Color

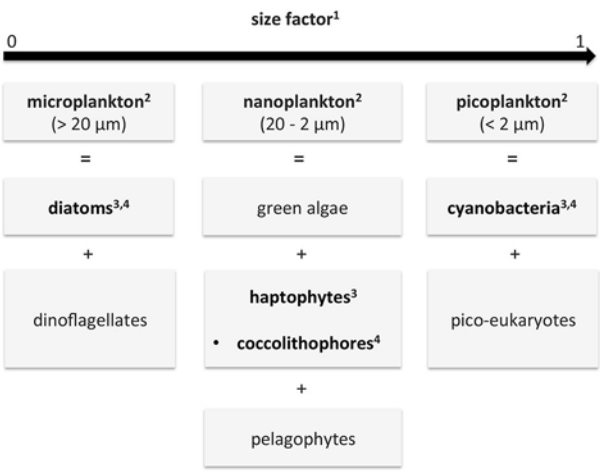


FIGURE 1 | Diagram describing the different phytoplankton classifications given by the different data products of this study and how they relate to each other: functional types (PFTs), size classes (PSCs) and size factor (Sf). In bold are the data used in this study: in the SynSenPFT algorithm (PFTs) and for validation (PSC and Sf). The numbers in the boxes indicate the corresponding references of the products used in this study: [1]Ciotti and Bricaud (2006), [2]Brewin et al. (2010, 2015), [3]Hirata et al. (2011), [4]Bracher et al. (2009) and Sadeghi et al. (2012). Note that different PFTs can vary among size classes, e.g., diatoms also can spread into the nanoplankton fraction.

Instrument (OCI) planned by NASA for launching within the "Plankton, Aerosol, Cloud, ocean Ecosystem" (PACE) mission in the early 2020 (see Gregg and Rousseaux 2017).

The structure of the paper is as follows. Section 2 introduces the materials and methods used: the input PhytoDOAS and OC-PFT products with the theoretical description and updates to the algorithms provided; the independent PFT *in situ* observations and satellite retrievals; a method to evaluate the initial products; details on the synergistic algorithm SynSenPFT and its implementation using the OC-PFT and PhytoDOAS data products. Section 3 provides results and discussions on intercomparison of the initial OC-PFT and PhytoDOAS Chla products, introduces an example of the SynSenPFT product and discusses the SynSenPFT product evaluation. Summary and Outlook are presented in Section 4.

2. MATERIALS AND METHODS

2.1. Initial PFT Products

2.1.1. Multispectral Retrieval: OC-PFT

The OC-PFT, an abundance based approach developed by Hirata et al. (2011), was applied to multispectral-based satellite-derived TChla (OC-CCI version 2, https://rsg.pml.ac.uk/thredds/catalog-cci.html, OC-CCI 2015) with revised empirical relationships between TChla and PFTs or PSCs to retrieve diatom, haptophyte and prokaryotic phytoplankton Chla. Being an empirical approach, the OC-PFT requires refinement when additional data become available to improve the retrievals for both global and under-sampled oceans, as shown by Soppa et al. (2014). Thus, the OC-PFT algorithm of Hirata et al. (2011) was revised using a larger and more evenly spatially distributed data set of *in situ* phytoplankton pigment data (Section 2.2.1) compared to Hirata et al. (2011).

Following Hirata et al. (2011) the Diagnostic Pigment Analysis (DPA) developed by Vidussi et al. (2001) and modified by Uitz et al. (2006) was used to derive PFTs and PSCs (microplankton, nanoplankton and picoplankton) from a large global *in situ* HPLC phytoplankton pigment data set (**Table 1** in 2.2.1, detailed description is provided in Supplementary Section 1). Based on the fractions of PFTs (*f*-PFT) and TChla in each *in situ* sample, a statistical model of the relationship between TChla and *f*-PFT was built. Such a statistical model when applied to satellite TChla data allows to retrieve the *f*-PFT on a global scale. The revised statistical models and the new ones proposed were found to represent well the relationship between the *f*-PFTs (diatom, haptophyte and prokaryote, in particular) and TChla (see Supplementary Section 2). Based on the results, the model of Soppa et al. (2014) for *f*-Diatoms with here revised parameters and the new models for *f*-Haptophytes and *f*-Prokaryotes based on Hirata et al. (2011), also with revised parameters, were chosen to retrieve the *f*-PFTs.

Given the models, their statistical parameters and OC-CCI TChla, the diatom, haptophyte and cyanobacteria were derived in a fractional form (*f*-PFT) and then in terms of abundance (mgChla/m^3). To retrieve the PFT abundance, the *f*-PFT is multiplied by the TChla value of each sample/pixel. This OC-PFT Chla data product is provided globally over the ocean with the

TABLE 1 | List of databases, two French campaigns (Atalante-3, KEOPS, and Bonus Good Hope) and German research vessels (RV) Maria S. Merian, Meteor, Polarstern, Poseidon and Sonne.

Database/cruise	References
SeaWiFS Bio-optical Archive and Storage System (SeaBASS)	Werdell et al., 2003
Marine Ecosystem Data (MAREDAT)	Peloquin et al., 2013
LTer long term research station	https://portal.lternet.edu/nis/home.jsp
Bermuda Atlantic Time-Series Study	http://bats.bios.edu/
AESOP-CSIRO database	http://aesop.csiro.au/
KEOPS	Uitz et al., 2009
Bonus Good Hope	Sauzde et al., 2015
ANT-XVIII/2 (EisenEx)	Peeken and Nachtigall, 2014a
	Soppa et al., 2014
ANTXXI/3 (EIFEX)	Peeken and Hoffmann, 2014c
	Peeken and Hoffmann, 2014b
	Soppa et al., 2014
ANTXXIII/1	Bracher et al., 2015b
	Bracher et al., 2015a
ANTXXIV/1	Bracher, 2015b
	Bracher et al., 2015a
ANTXXIV/2	Peeken et al., 2017c
ANTXXIV/4	Bracher, 2015c
	Bracher et al., 2015a
ANTXXV/1	Taylor et al., 2011a
	Taylor et al., 2011b
ANTXXV/3 (LOHAFEX)	R. Roy (CSIR, National Institute of Oceanography (NIO),
	Goa, India, pers. communication)
ANT XXVI/3	Peeken and Nachtigall, 2014b
	Soppa et al., 2014
ANT XXVI/4	Bracher, 2015d
	Bracher et al., 2015a
ANT XXVII/2	Bracher, 2015e
	Trimborn et al., 2015
ANT XXVIII/3	Bracher, 2014a
	Soppa et al., 2014
ARK XXIV/1	Nöthig et al., 2015
ARK XXIV/2	Nöthig et al., 2015
ARK XXV/1	Nöthig et al., 2015
ARK XXV/2	Nöthig et al., 2015
ARK XXVI/3	I. Peeken, AWI (pers. communication)
MSM 9/1	Bracher and Taylor, 2017
MSM 18/3	Bracher, 2015a
	Bracher et al., 2015a
Sonne SO202/2	Taylor and Bracher, 2012
	Zindler et al., 2012a
Sonne SO218	Bracher, 2014b
	Cheah et al., 2013
POS 320-1	Peeken and Quack, 2017
	Quack et al., 2007
POS 348	Peeken and Walter, 2017
Atalante 3	Zindler et al., 2012b

(Continued)

TABLE 1 | Continued

Database/cruise	References
Meteor 68	Franklin et al., 2009
Indonesia 2005	Peeken et al., 2017d
Marion Dufresne 7/1	Peeken et al., 2017a
Marion Dufresne 7/2	Peeken et al., 2017b
Meteor 55	Peeken and Hoffmann, 2014a
	Soppa et al., 2014
Meteor 60	Peeken and Nachtigall, 2014c
	Soppa et al., 2014

Cruises with RV Polarstern to the Southern hemisphere are labeled as ANT, to the Northern hemisphere as ARK, with RV Poseidon as POS, with RV Maria S. Merian with MSM, with RV Meteor as M and with RV Sonne as SO. A campaign in Indonesian waters was conducted from a small fisher boat.

∼4 km (sinusoidal) spatial resolution over the period of August 2002–March 2012 on a daily time scale.

2.1.2. Hyperspectral-Based Retrieval: PhytoDOAS

The other PFT Chla data products considered in this study are based on the Differential Optical Absorption Spectroscopy (DOAS) method (Perner and Platt, 1979) exploited with respect to phytoplankton absorption properties (PhytoDOAS), as introduced by Bracher et al. (2009). The current PhytoDOAS PFT retrieval of Chla for cyanobacteria (that includes all prokaryotic phytoplankton), diatoms and coccolithophores was obtained with the algorithm PhytoDOAS version 3.3 applied to SCIAMACHY hyperspectral measurements of the radiance at the top of atmosphere (TOA) for the sensor's operation period of August 2002–March 2012. The cyanobacteria and diatom PhytoDOAS retrievals are based on the algorithm by Bracher et al. (2009). For coccolithophores the algorithm by Sadeghi et al. (2012) has been used. Both algorithms, however, have been slightly modified to obtain optimal results for the whole time series. In particular, the changes include the following:

- Data aquired during SCIAMACHY instrument decontamination are excluded in the analysis.
- SCIAMACHY level-1b input data for PhytoDOAS are now version 7.04 data instead of version 6.0.
- The wavelength window for all three PFTs fit factor starts at 427.5 nm instead of 429 nm.
- Coccolithophores fit factors are obtained when fitting simultaneously only diatoms and coccolithophores instead of a triple fit as in Sadeghi et al. (2012) where dinoflagellates were additionally fitted.
- Vibrational Raman Scattering (VRS) is now fitted directly in the blue spectrum following Dinter et al. (2015), instead of in the UV-A region as in Vountas et al. (2007). However, here a shortened wavelength window is used: 450–495 nm instead of 450–521 nm as in Dinter et al. (2015). Moreover, the correction for some instrumental effects varying over time is done with a VRS pseudo absorption spectrum calculated following Vountas et al. (2007) based on a daily solar background spectrum measured by SCIAMACHY instead of using a radiative transfer model (RTM) simulated background spectrum as done in Dinter et al. (2015).
- The PFT Chla are derived from the ratio of the PFT fit factor to the VRS fit factor applying a LUT (Look Up Table) approach. The LUT is derived from the SCIATRAN RTM simulations (Rozanov et al., 2014) and accounts also for changes in solar zenith angle (SZA).

Within the current study we used seven days composites of the PhytoDOAS PFT Chla retrievals interpolated onto a 0.5° × 0.5° grid covering the global ocean on a daily basis. Data are available at http://doi.pangea.de/10.1594/PANGAEA.870486 (Bracher et al., 2017b).

2.2. Independent PFT Estimates Used for Evaluation

2.2.1. *In situ* Observations

The *in situ* data set used for the SynSenPFT satellite validation and improvement of OC-PFT approach was built with HPLC pigment data compiled from several databases and individual cruises (**Table 1**). The spatial distribution of the *in situ* data set is depicted in Figure S5 (Supplementary Section 3). Chla of PFTs from *in situ* phytoplankton pigment data was derived using the Diagnostic Pigment Analysis of Vidussi et al. (2001) and Uitz et al. (2006) modified as in Hirata et al. (2011) and Brewin et al. (2015). The weights were revised in accordance to this enlarged *in situ* pigment data set (**Table 1**). We refer the reader to the Supplementary Material (Section 1) for more details. The final version of the global data set is available at https://doi.pangaea.de/10.1594/PANGAEA.875879 (Soppa et al., 2017).

2.2.2. Size Factor

The estimate of the dimensionless size factor (Sf), following Ciotti et al. (2002) and Ciotti and Bricaud (2006), is based on the fact that the shape of the phytoplankton absorption spectrum flattens with increasing cell size ("packaging effect"). Sf varies between 0 and 1, which represent the extreme situations (100 % microphytoplankton and 100 % picophytoplankton, respectively). The Sf value thus provides information on the dominant size of the population: values close to 0 (1) indicate dominance of microplankton (picoplankton). Note that the contribution of nanophytoplankton is not explicitly taken into account. Also, for a given population, the shape of the absorption spectrum (and therefore the Sf value) can be affected by photoacclimation to ambient light. On the other hand, as the method is spectral-based, Sf variations are not constrained by Chla variations. The Sf maps have been derived from remote sensing reflectances data from daily OC-CCI (version 2) products covering the same time period as the OC-PFT data set. First, the total absorption coefficients were computed using the Quasi-Analytical Algorithm (QAA) (Lee et al., 2002) version 6. This step was followed by an optimization procedure (revised from Bricaud et al. 2012) which retrieves three independent variables: Sf and two parameters related to Colored Detrital Matter absorption properties. The Sf data were obtained at daily temporal resolution binned globally on a 4 km sinusoidal grid.

2.2.3. Phytoplankton Size Classes

The model of Brewin et al. (2010), using global parameters from Brewin et al. (2015), was used to compute the fractions of total chlorophyll for three PSC (pico- < 2 μm, nano- 2-20 μm and micro-phytoplankton > 20 μm). The model is designed to estimate the fractions of the three phytoplankton size classes as a continuous function of TChla. Based on the work of Sathyendranath et al. (2001), the model assumes small cells are incapable of growing beyond a particular Chla, with an upper limit imposed possibly from a combination of bottom-up (e.g., nutrient control) and top-down (e.g., grazing) processes, and that, beyond this value, chlorophyll is added to a system solely by the addition of larger size classes of phytoplankton (see Raimbault et al. 1988; Chisholm 1992). For further details on the model, the reader is referred to Brewin et al. (2010, 2015). The model was applied to the same daily OC-CCI TChla data set as used for the OC-PFT product, for the period of August 2002 – March 2012, producing global data on the fractions of the three size classes. The PSC data (*f*-micro, *f*-nano and *f*-pico) and TChla were obtained with the same spatial and temporal resolution as used for OC-PFT and Sf. TChla data where then multiplied by *f*-PSC to obtain the respective Chla of PSC (c-PSC).

2.2.4. Phytoplankton Functional Types from the NASA Ocean Biogeochemical Model

The NOBM (Gregg and Casey, 2007) is a global biogeochemical model with coupled circulation and radiative models (Gregg, 2002). NOBM simulations combine assimilated global SeaWiFS and MODIS Chla data with global data sets on nutrient distributions, sea surface temperature and current conditions to calculate four PFTs: diatoms, coccolithophores, cyanobacteria (as defined in PhytoDOAS) and chlorophytes. NOBM PFT data span from 84° S to 72° N on a 1.25° by 2/3° of longitude/latitude grid. Data were obtained at daily and monthly temporal resolution at http://giovanni.gsfc.nasa.gov/giovanni/ for the same time period as other aforementioned products.

2.3. Initial (Input) Product Evaluation

Given the PFT Chla information based on two distinct algorithms, different not only with respect to temporal and spatial resolution but also with respect to underlying physical and/or statistical principles, it is essential to compare these OC-PFT and PhytoDOAS PFT data against each other to evaluate whether/where and to what extent one could expect the retrieval algorithms to provide similar (different) results. For this purpose, we consider the triple collocation (TC) analysis (Stoffelen, 1998) known as a powerful tool for global scale product evaluation and intercomparisons (Gruber et al., 2016).

The triple collocation method requires analysing three independent products (but, following Zwieback et al. 2012, can be extended to a larger amount of products) describing exactly the same variable state: Chla of PFT in our case. To complete the required triplet, the PFT Chla estimates of the NOBM (Gregg and Casey, 2007) were considered in addition to those of PhytoDOAS and OC-PFT products. Opposed to the PSC and Sf algorithms, the NOBM provides estimations of diatom, coccolithophores and cyanobacteria abundance in terms of Chla as the PhytoDOAS and OC-PFT do (except that OC-PFT provides haptophytes Chla instead of coccolithophores Chla). The TC method is briefly described in the following subsection.

2.3.1. Triple Collocation Analysis

The TC allows to estimate the absolute error variances ($\sigma^2_{\varepsilon_i}$), or the root-mean-square deviation (RMSD, σ_{ε_i}) of three collocated data sets ($i = \{1, 2, 3\}$, in our case $i = \{$PhytoDOAS, OC-PFT, NOBM$\}$) with unknown uncertainties and assuming uncorrelated errors. Among other assumptions underlying the method are the linearity, stationarity of the signal and errors, and independence of the error from the variability of the measured signal itself (orthogonality). Under the aforementioned assumptions, the σ_{ε_i} can be estimated from the unique terms covariance matrix (McColl et al., 2014) (Q_{11}, Q_{12}, Q_{13}, Q_{22}, Q_{23}, Q_{33}):

$$\sigma_{\varepsilon_i} = \begin{bmatrix} \sqrt{Q_{11} - \frac{Q_{12}Q_{13}}{Q_{23}}} \\ \sqrt{Q_{22} - \frac{Q_{12}Q_{23}}{Q_{13}}} \\ \sqrt{Q_{33} - \frac{Q_{13}Q_{23}}{Q_{12}}} \end{bmatrix} \quad (1)$$

It is worth mentioning that the TC analysis is usually used for calibration purposes (Stoffelen, 1998; Vogelzang et al., 2011). The estimates provided by the analysis are normally so-called "unscaled" and include also any biases, if available, of the particular data product with respect to the "truth". The "truth" is determined by the "joint covariance" σ^2_{Θ} observed by all three analyzed data products. Here we consider the "unscaled" uncertainties σ_{ε_i} as a merit of difference between the distinct features of PFT Chla temporal variability observed and described by the satellite retrievals (PhytoDOAS and OC-PFT) and NOBM numerical model over a particular time period.

2.4. Synergistic Product

The synergistic (SynSenPFT) product is obtained as Chla for diatoms, coccolithophores and cyanobacteria presented globally on a 4 km sinusoidal grid on a daily basis over the period of August 2002 – March 2012.

2.4.1. SynSenPFT Algorithm

The SynSenPFT combines OC-PFT and PhytoDOAS level-3 Chla products with an optimal interpolation (OI, Gandin and Hardin, 1965). In a generalized form, the OI method is formulated as following:

$$x^a = x^b + W(y - Hx^b) \quad (2)$$

x^a denotes state analysis, in our case SynSenPFT product; x^b is a background, for the particular application it is the OC-PFT product. y refers to observations – PhytoDOAS. H is the observation operator projecting OC-PFT into PhytoDOAS space. W is a weight matrix reflecting data error statistics.

In terms of Kalman-type filtering, the synergistic estimates of x or so called state vector analysis x^a is expressed as:

$$x(t_n)^a = x(t_n)^b + K_n(y_n - Hx(t_n)^b) \quad (3)$$

where $x(t_n)^a$ and $x(t_n)^b$ denote the analysis (SynSenPFT) and OC-PFT, respectively, at certain time t_n and OC-PFT grid points, y_n is the PhytoDOAS observations available at t_n, and K is the so-called Kalman gain:

$$K_n = P_n^b H^T (H P_n^b H^T + R)^{-1} \quad (4)$$

Here P_n^b and R are the OC-PFT and PhytoDOAS error covariance matrices, respectively.

As seen, the SynSenPFT is an update of OC-PFT with PhytoDOAS values weighted in accordance to our degree of belief to both initial-input data products. Note that within the current version of SynSenPFT algorithm the update is done for every subpixel of OC-PFT within a PhytoDOAS pixel (**Figure 2**). Thus, SynSenPFT in every OC-PFT sub-pixel on average is nudged toward PhytoDOAS values as close as allowed by the prescribed R matrix (considered diagonal for simplicity). The SynSenPFT spatial distribution within the PhytoDOAS pixel is based on the information carried by P_n^b that reflects OC-PFT spatial structure if the P_n^b matrix is estimated based on the OC-PFTs covariances within the PhytoDOAS grid cell.

As an illustration, **Figure 3** shows examples of SynSenPFT diatom Chla (on the OC-PFT sub-grid) within three different PhytoDOAS pixels. As seen, on average within a PhytoDOAS pixel, SynSenPFT Chla is closer to PhytoDOAS retrievals, subject to assumed PhytoDOAS error statistics (an absolute error value of 0.4 mgChla m^{-3}, given on **Figures 3a–c**, or the relative error of 40% (0.15 mgChla m^{-3}), given on **Figure 3d**. Comparing **Figure 3c** and **Figure 3d** one can conclude about the sensitivity of SynSenPFT product to the assumed PhytoDOAS errors: the more accurate PhytoDOAS PFT product is assumed, the more it influences the SynSenPFT Chla estimates. In comparison with OC-PFT, the spatial variability of the SynSenPFT Chla will be smoother, however. In this study as absolute PhytoDOAS errors approximating the R matrix, we assume 0.4, 0.3, and 0.1 mgChla m^{-3} for diatom, coccolithophores and cyanobacteria, respectively.

3. RESULTS AND DISCUSSIONS

3.1. Initial Product TC Analysis

Here we present the results of the TC analysis carried out with respect to OC-PFT, PhytoDOAS, and NOBM PFT data products. **Figure 4** shows an example of the spatial distribution of PhytoDOAS, OC-PFT, and NOBM unscaled uncertainties for diatom, coccolithophore (haptophytes for OC-PFT) and cyanobacteria Chla, calculated with the TC analysis for each 0.5° by 0.5° box based on PhytoDOAS, OC-PFT and NOBM PFT information over the period of 2003–2009. When comparing σ_ε of PhytoDOAS diatom (**Figure 4A**) against those of OC-PFT (**Figure 4B**) and NOBM (**Figure 4C**) one can notice much higher PhytoDOAS diatom Chla temporal variability all over the World Ocean, except for Northern Hemisphere mid- and high latitudes and a tiny belt around the Antarctica, while the OC-PFT diatom Chla product reveals higher σ_ε in the upwelling regions and in the Arabian Sea (**Figure 4B**). The σ_ε of PhytoDOAS coccolithophores Chla (**Figure 4D**) also exceeds similar statistical estimates of NOBM coccolithophores (**Figure 4F**) and OC-PFT haptophytes products (**Figure 4E**), now even also in the high Northern latitudes, and the OC-PFT shows higher haptophytes Chla ranges in the Arabian Sea and upwelling regions (**Figure 4E**). However, NOBM has very low coccolithophores Chla σ_ε (also compared to NOBM diatom Chla σ_ε) all over low- and mid latitudes. The spatial distribution of

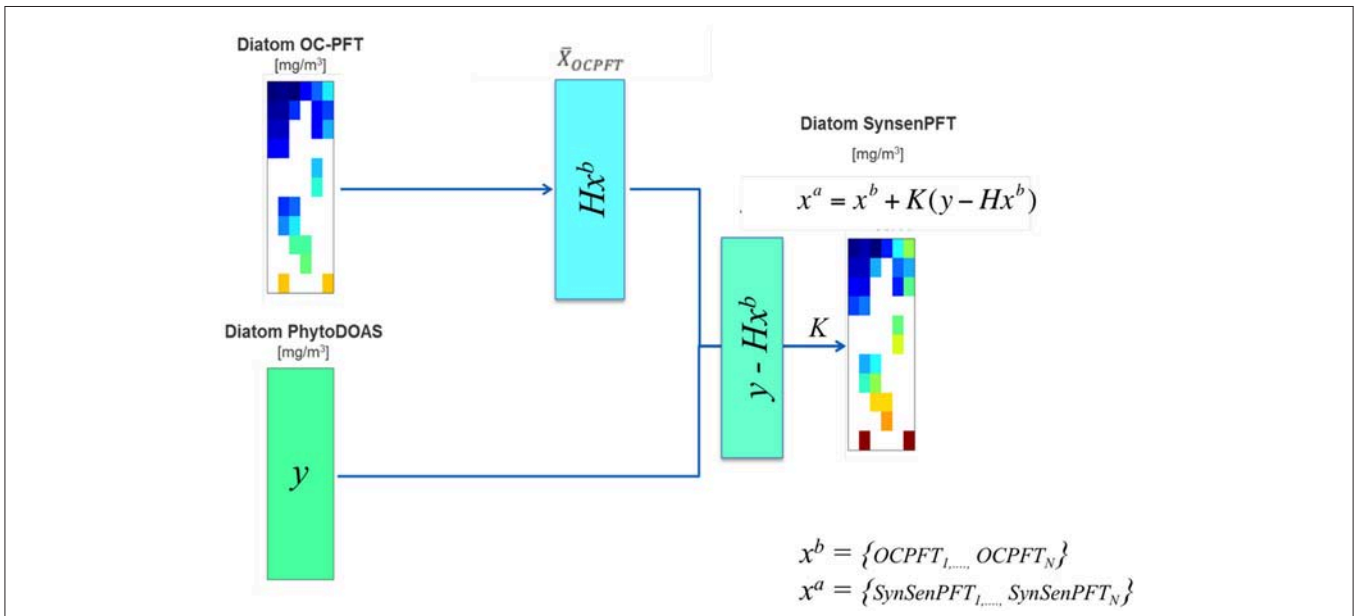


FIGURE 2 | Diagram exemplifying the SynSenPFT algorithm (Equations 2–4) that updates the PFT information from each OC-PFT sub-pixel (x^b) within a PhytoDOAS pixel given the PhytoDOAS Chla (y) and *a priori* error statistics introduced in K, which results in SynSenPFT Chla (x^a) at each OC-PFT sub-pixel.

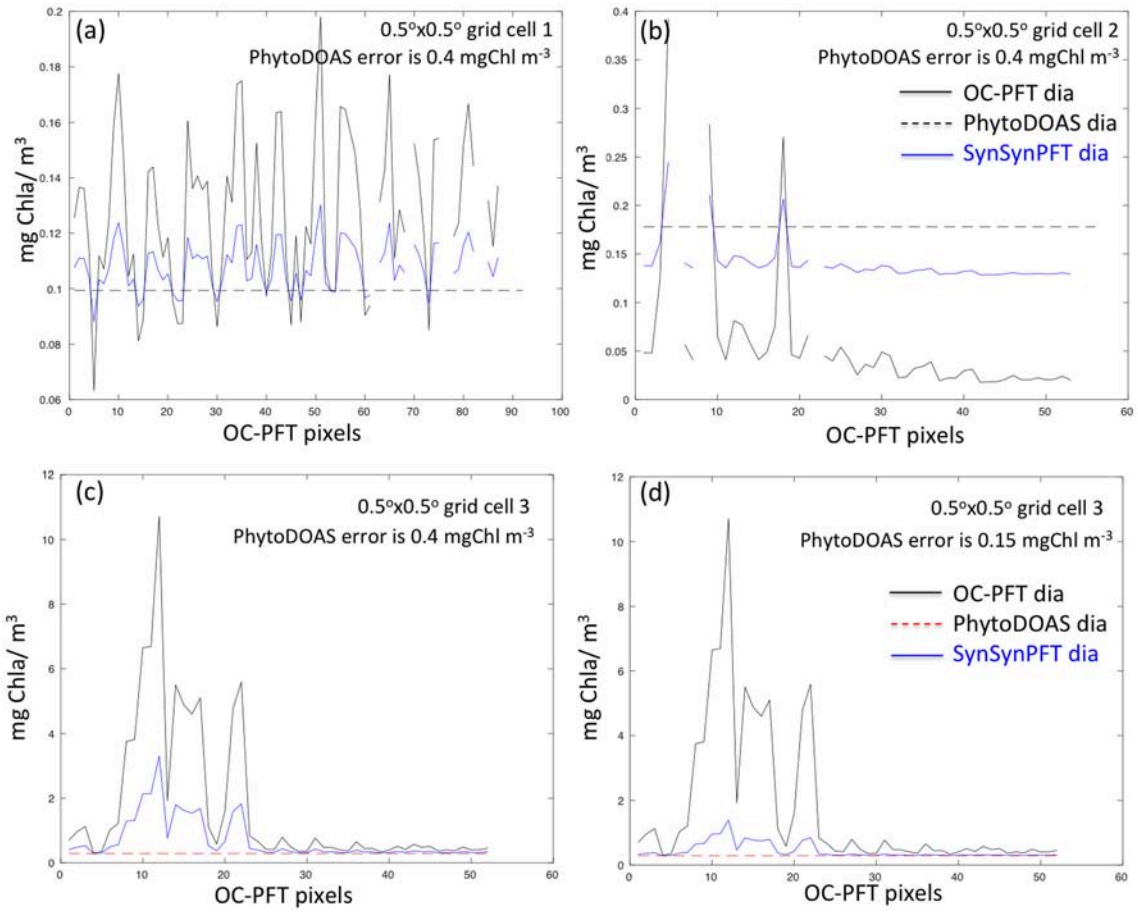


FIGURE 3 | Examples illustrating variability of OC-PFT and updated (nudged toward PhytoDOAS) SynSenPFT diatom Chla within three PhytoDOAS pixels with different OC-PFT Chla variability, PhytoDOAS Chla on average either lower **(a,c)** or higher **(b)** than OC-PFT Chla and given PhytoDOAS error statistics (0.4 mgChla m^{-3} for **a–c**). **(c,d)** depict SynSenPFT diatom Chla obtained for the same PhytoDOAS grid cell but assuming different PhytoDOAS errors: 0.4 mgChla m^{-3} **(c)** and 0.15 mgChla m^{-3} **(d)**. Thus, lower panels illustrate the sensitivity of SynSenPFT diatom Chla product to the assumed PhytoDOAS error statistics.

cyanobacteria σ_{ε_i} is depicted on **Figures 4G–I**. Opposed to OC-PFT, NOBM, and PhytoDOAS show larger cyanobacteria Chla deviation in the oligotrophic and equatorial regions with higher σ_ε values in the PhytoDOAS data product (**Figure 4G**). These are the regions where we can expect most of the differences between the OC-PFT and synergistic cyanobacteria Chla.

3.1.1. Approximation of *Prior* Error Statistics

As concluded from subsection 2.3.1, the performance of the SynSenPFT strongly depends on the plausibility of the assumed quality of the input data products approximated by the R and P_n^b matrices. As mentioned above, P_n^b matrix can be (was) presented by the OC-PFT Chla covariances within a PhytoDOAS pixel. However, *per se*, we can say little about the quantitative estimates of PhytoDOAS errors to properly approximate the R matrix. The following is dedicated to the opportunity and challenges of the R matrix specification.

The PhytoDOAS error covariance matrix can be provided by using the TC analysis (Crow and van den Berg, 2010), since the TC analysis in general provides quantitative information about quality of the data product considered (Stoffelen, 1998; Gruber et al., 2016). Nevertheless, the estimates presented in **Figure 4** and discussed in 2.3.1 can be considered only as the merit of differences between the collocated products unless the quality of at least one data product is well-known. Not as an attempt to find a possible scaling for our TC estimates, but rather for an additional evaluation of our TC analysis, the OC-PFT TC unscaled uncertainties were compared against statistics of the OC-PFT match-up with *in situ* observations. **Figure 5** illustrates the mean absolute error (MAE) of OC-PFT diatom (**Figure 5A**), haptophytes (**Figure 5C**), and cyanobacteria (**Figure 5E**) Chla relative to *in situ* Chla calculated for the biomes determined following Hardman-Mountford et al. (2008). In addition, it presents the ratio of the MAE to the OC-PFT TC-based uncertainties (**Figures 5B,D,F**). The ratio values less than 1 indicate that the uncertainties obtained with the TC overestimate the product error statistics while the values larger than 1 would correspond to the situation with the TC uncertainties being underestimated. For instance, the TC-based OC-PFT unscaled uncertainties for diatoms (haptophytes) are

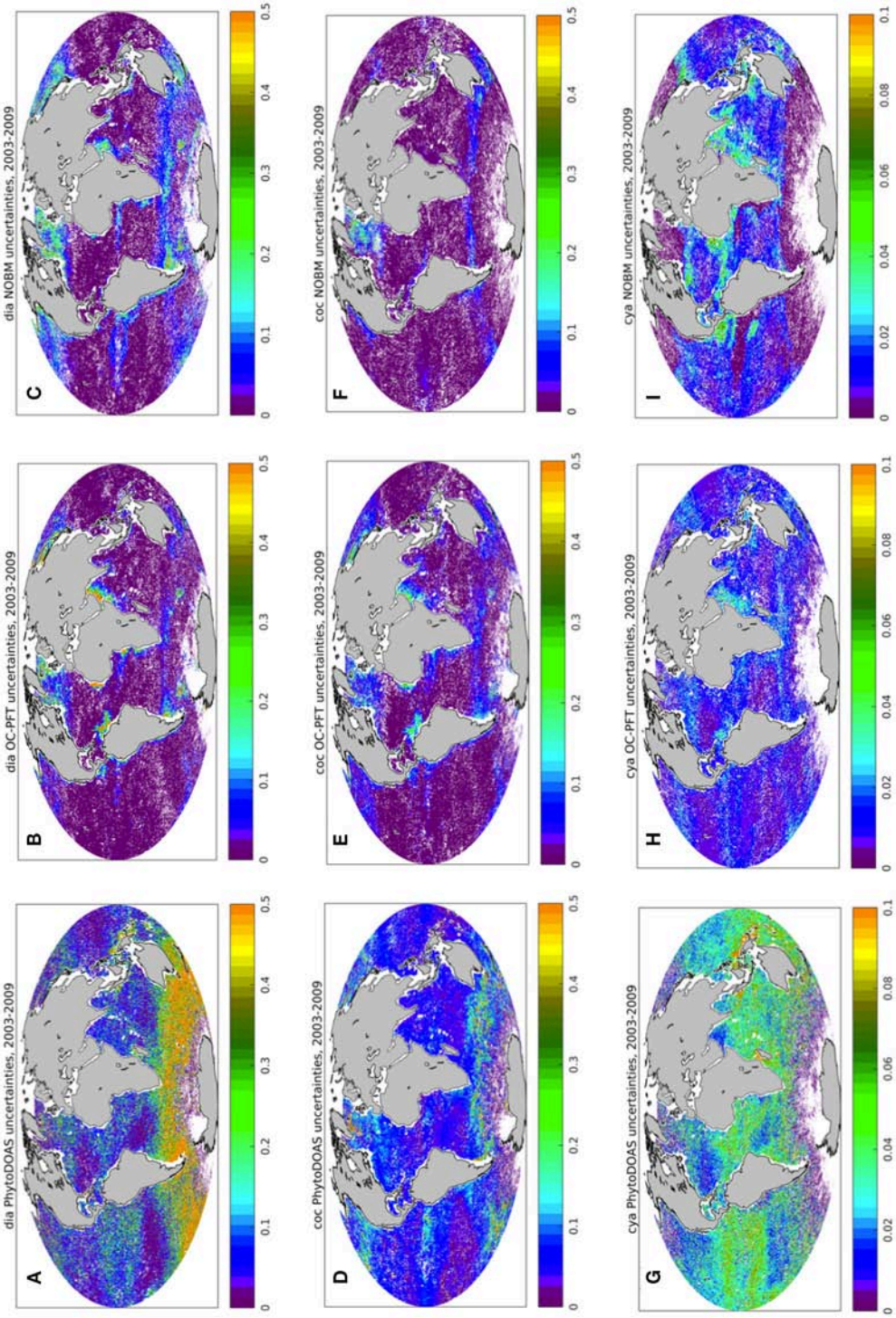


FIGURE 4 | Triple collocation analysis: spatial distribution of PhytoDOAS **(A,D,G)**, OC-PFT **(B,E,H)**, and NOBM **(C,F,I)** unscaled uncertainties σ_ε (mg m^{-3}) for diatom Chla **(A–C)**, coccolithophore Chla **(D–F)**, and cyanobacteria Chla **(G–I)** based on the time period of 2003–2009. White dots are the grid points with negative $\sigma^2_{\varepsilon_i}$ and/or with no satellite information available.

overestimated (underestimated) in the "high Chla" biome (Hardman-Mountford et al., 2008) covering, for instance, southwest mid-to-high northern latitudes, southwest Southern Ocean, the edge of equatorial upwelling regions, the Arabian Sea and the shelf seas of South-East Asia (**Figure 5**). In the "low-intermediate Chla" biome, the TC σ_ε for diatoms and

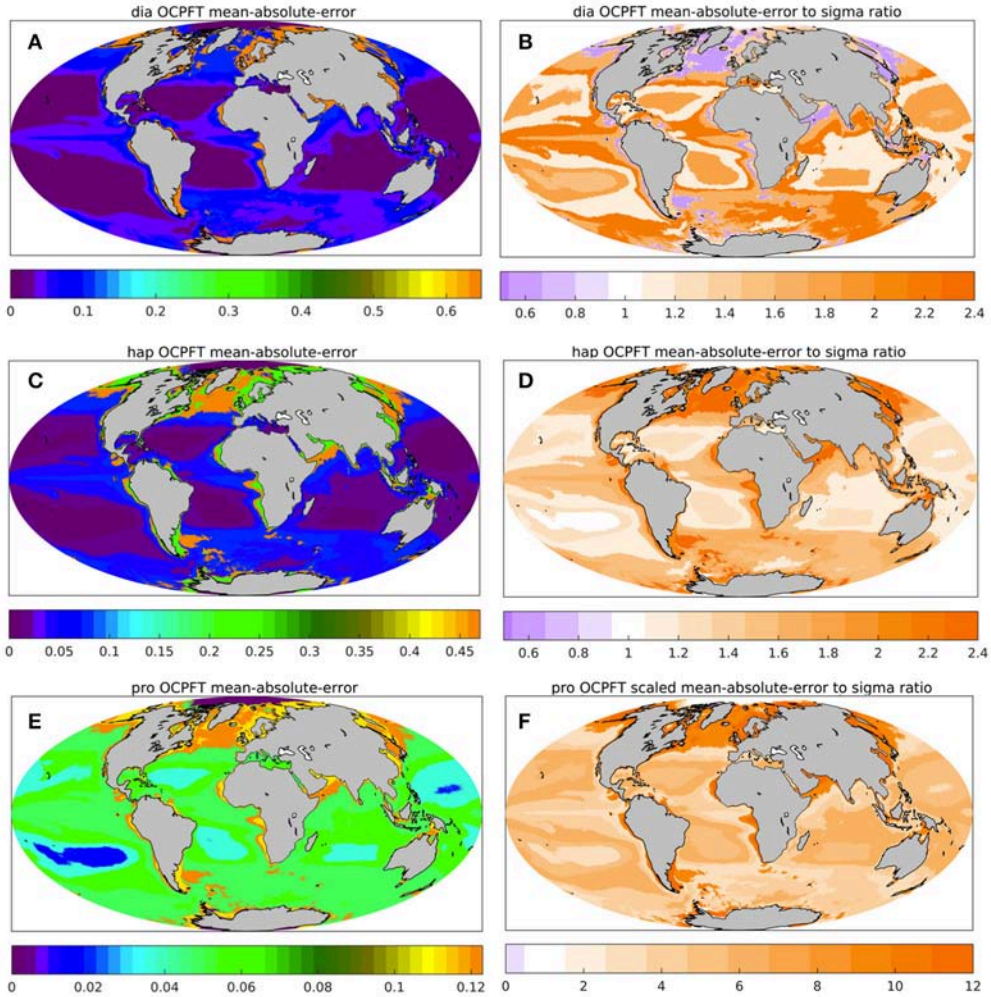


FIGURE 5 | Additional statistics assessing OC-PFT against *in situ* on the biomes grid by Hardman-Mountford et al. (2008): the mean-absolute-error (MAE, mgChla m^{-3}, **A,C,E**) and the ratio of MAE to the OC-PFT unscaled uncertainties obtained based on the TC analysis **(B,D,F)** for for diatom **(A,B)**, haptophytes **(C,D)** and prokaryotes **(E,F)**.

haptophytes are in agreement with the MAE estimates. The MAE of OC-PFT diatom Chla is higher than TC σ_ε in the "high-intermediate Chla" biome (for example, in equatorial areas outside the upwelling regions and in the Pacific, easten Atlantic and Indian Ocean sectors of the Southern Ocean). As seen from **Figure 5F**, the TC estimates for OC-PFT cyanobacteria are highly underestimated all over the ocean. That might indicate that PhytoDOAS overestimates cyanobacteria Chla. Following (Gruber et al., 2016) an additional evaluation of the TC estimates was done (not shown) with the fractional mean-squared-error ($fMSE_i$) that can be also calculated within the frame of the triple collocation analysis. The $fMSE_i$ criterion explains the existence of white dots in **Figure 4**, correspondent to negative $\sigma^2_{\varepsilon_i}$, by too low values of the product noise to the true signal ratio, which shows that the joint variability of all three products does not always exist (Supplementary Section 4, Figure S6). Or due to underestimation of the variability of one collocated triple or because of some violations of the assumptions essential for TC, the joint variability might not represent the truth. Before treating the TC estimates as product uncertainties, a more detailed evaluation of the plausibility of our TC-based PFTs uncertainties as well as the product evaluation with respect to violating the assumptions underlying the method might be needed. For instance, errors of NOBM PFT estimates and OC-PFT can be correlated since both are based on multispectral satellite data from the SeaWiFS and MODIS sensors: OC-PFT algorithm input OC-CCI Chla product includes in addition to MERIS, also SeaWiFS and MODIS information; NOBM assimilates SeaWiFS and MODIS TChla.

The more detailed assessment of the TC analysis of SynSenPFT initial input product would allow in future to use the TC estimates for approximating R matrix. In this study, the values of the R diagonal are assumed constant and partly based on the current TC results. It is worth commenting

on representation/representativeness errors due to 'physical' mismatch between the initial algorithms (Bracher et al., 2017a), and up-(down-)scaling the products when regridding. These errors also impact the TC estimates. In data assimilation applications (Losa et al., 2012, 2014; Yang et al., 2016) to account for such representation errors as well as for any possible errors in the P_n^b approximation, the assumed σ_{ε_i} is normally enlarged. So did we in our study: the assumed errors exceed the PhytoDOAS unscaled uncertainties discussed in 2.3.1.

3.2. Example of SynSenPFT Product

Figure 6 illustrates the monthly mean SynSenPFT Chla product for diatom, coccolithophores and cyanobacteria in September 2006. **Figure 7** depicts absolute differences between SynSenPFT and OC-PFT averaged over the September 2006 and shows where the synergistically combined PFT information are influenced by PhytoDOAS product and over which regions OC-PFT dominates. For instance, by directly comparing **Figures 6A,C**, **7A,C** one can conclude that at high latitudes SynSenPFT diatom and cyanobacteria estimates contain mostly the OC-PFT signal (because of low dia-(cya-)PhytoDOAS coverage in these regions). Note that the spatial distribution of the mean absolute differences (MAD) between SynSenPFT and OC-PFT Chla corresponds to either PhytoDOAS or OC-PFT TC uncertainties patterns (**Figure 4**). We remind that the TC uncertainties are considered as a merit of differences between OC-PFT and PhytoDOAS (and NOBM) products, therefore they indicate *per se* the areas where the largest updates of OC-PFT Chla by values of PhytoDOAS are expected. Consequently, the dia-SySenPFT Chla differ from OC-PFT estimates most of all in the Southern Ocean (except for the very high latitudes), in the Arabian Sea and north of Australia. The mean differences reach 0.5 mgChla m^{-3}. The differences between coc-PhytoDOAS and OC-PFT haptophytes Chla are also in agreement with the TC estimates, but hardly exceed 0.2 mgChla m^{-3}, and are most apparent in the Equatorial Pacific, southwest of Africa, in the Arabian Sea and north as well as southeast of Australia. For the considered time period of September 2006 relatively high MAD values for cyanobacteria are distributed over the tropical areas with a maximum of 0.05 mgChla m^{-3} in the equatorial upwelling systems, in the northwest part of the Indian Ocean and north of Australia.

3.3. SynSenPFT Product Evaluation

The SynSenPFT Chla products were evaluated by comparison with *in situ* observations, other satellite PFT/PSC products and the NOBM model simulations over the period of August 2002–March 2012.

3.3.1. Match-Ups with *In situ* Observations

Daily SynSenPFT products at ∼4 km spatial resolution were matched with samples from the global *in situ* validation dataset. The *in situ* dataset was matched with satellite data from OC-CCI Chla product (but restricted to the life time of SCIAMACHY) at daily temporal resolution and satellite values were retained when located at the same date within a distance of 4 km from the *in situ* measurement. Note: for OC-PFT model development

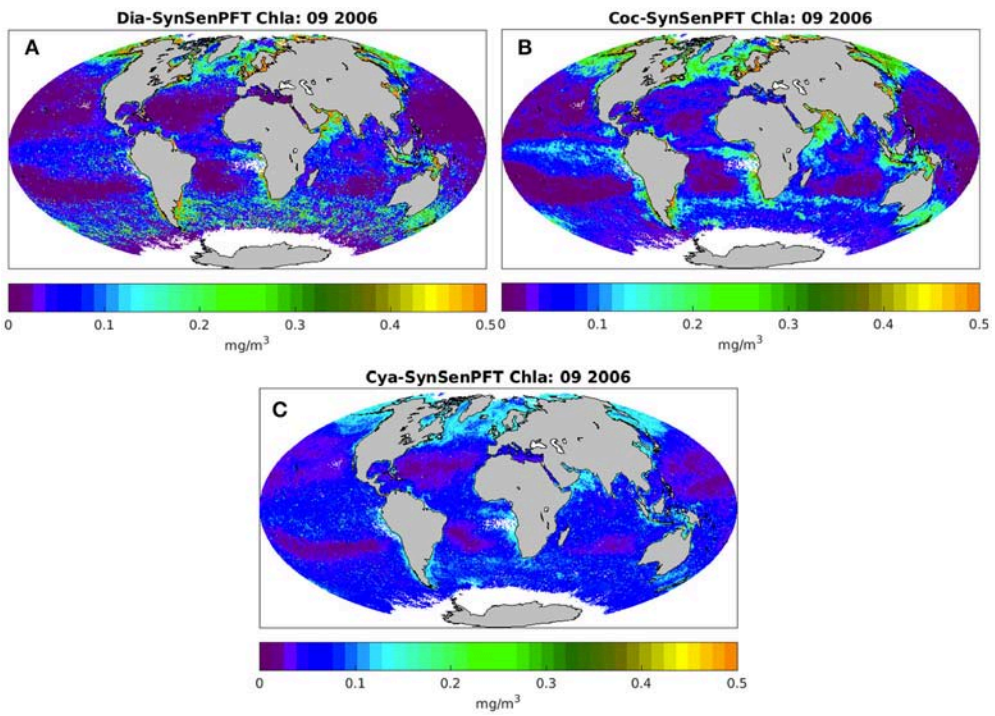


FIGURE 6 | Monthly mean SynSenPFT Chla product of diatoms **(A)**, coccolithophores **(B)**, and cyanobacteria **(C)** for September 2006.

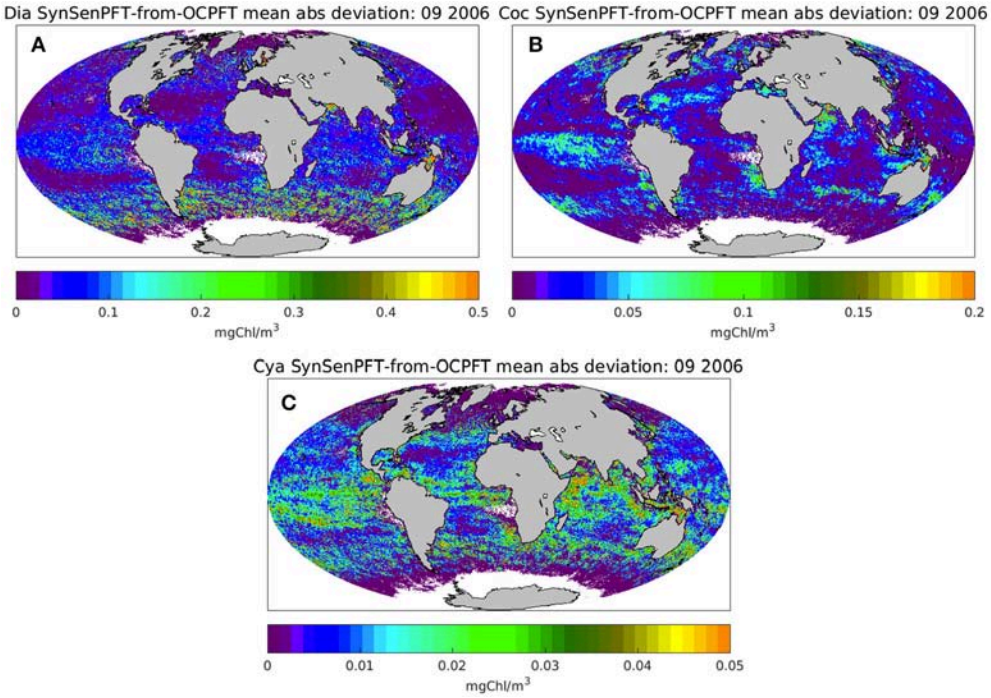


FIGURE 7 | Monthly mean absolute deviation of SynSenPFT Chla from OC-PFT Chla for diatoms **(A)**, coccolithophores **(B)**, and cyanobacteria **(C)** in September 2006.

the *in situ* pigment data not matching the OC-CCI TChla were used, such that the validation data are independent of the data used to train the model. In the case of more than one *in situ* sample within one satellite pixel, respective match-ups were treated independently. To finally derive high quality match-ups, those were considered valid where the coefficient of variation of the 3 x 3 pixels of satellite TChla (OC-CCI) around each *in situ* was lower than 0.15. This is a similar quality control as applied by Werdell et al. (2007), but here we applied it to TChla instead of Rrs data. In addition, only match-ups above a threshold of 0.01 mgChla m^{-3} were selected. The rationale for this threshold is that the surface Chla values encountered in the clearest ocean waters (south Pacific gyre) were found to be in the range 0.01–0.02 mg m^{-3} (Morel et al., 2007). Therefore, values below 0.01 mg m^{-3} may be considered as questionable.

To compare the satellite SynSenPFT products and *in situ* Chla of the PFTs the mean absolute error (MAE), root mean squared difference (RMSD), un-biased RMSD and bias (as formulated in Sá et al. 2015) were used. The determination coefficient, slope and intercept (type II regression based on $\log_{10}$ data) were also computed. **Figure 8** and **Table 2** present the results of the validation of SynSenPFT Chla against *in situ* PFT Chla considering the data all over the global ocean. High R-square values around 0.5 were achieved for coc-SynSenPFT and dia-SynSenPFT with a slightly better slope, intercept, MAE, and RMSD for coc-SynSenPFT, but lower bias for dia-SynSenPFT. The dia-SynSenPFT RMSD and bias are comparable to the statistics presented by Brewin et al. (2017) for different optical water types in the North Atlantic (RMSD varying from 0.29 to 0.60, bias varying from −0.30 to 0.23). For all three PFTs, the estimated SynSenPFT bias is less than reported in the study by Gregg and Rousseaux (2017).

Looking at the grouping of diatom match-ups according to latitude (**Figure 8A**), we can note the underestimation of diatom Chla at high southern latitudes, which could be related to the general underestimation of TChla when applying standard algorithms to Southern polar areas (Johnson et al., 2013). It is worth keeping in mind that south of 65° most of the dia-SynSenPFT values are derived from OC-PFT since PhytoDOAS pixel based information was sparse at these latitudes and low SZA.

Figure 8C shows that correlation is not significant for cya-SynSenPFT to *in situ* PFT data which may be due to the limited spread of the *in situ* PFT as compared to the two other PFT products evaluated. Cya-SynSenPFT Chla reaches a maximum concentration at −0.8 $\log_{10}$Chla (0.18 mg Chla/m^3). This fact implies a high weighting toward OC-PFT, since this feature is an artifact of the OC-PFT approach. Nevertheless, the amount of *in situ* cya Chla exceeding 0.18 mg Chla/m^3 is quite small (see Supplementary, Figure S7). **Figure 8D** depicts the same as **Figure 8C** but with the density of cya-SynSenPFT match-ups as a background. From this figure panel one can see that the most frequent or dense match-ups are located close to the 1:1 curve. As a result, the MAE and RMSD are lower for cya-SynSenPFT than for dia-SynSenPFT.

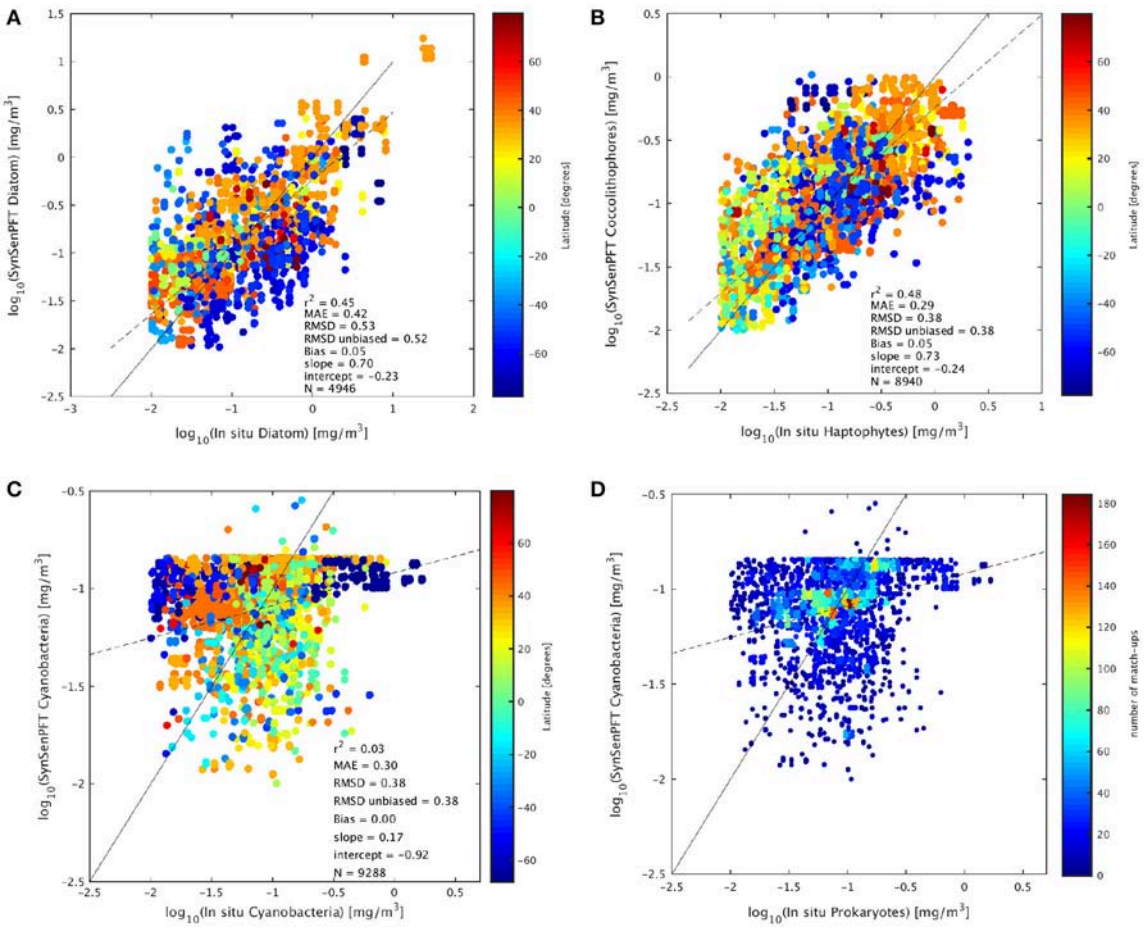


FIGURE 8 | Comparison of *in situ* Chla and SynSenPFT Chla for diatom **(A)**, haptophytes against coccolithophores **(B)**, and prokaryotes vs. cyanobacteria **(C)**. The continuous line represents the 1:1 line and the dashed line the type II linear regression. **(D)** Depicts the same as **(C)**, but colored in accordance with the matchups density. The statistical analysis was performed in $\log_{10}$ space.

TABLE 2 | SynSenPFT match-ups statistics.

PFT	r^2	MAE	RMSD	Bias	Slope	Number of match-ups
diatom	0.45	0.42	0.53	0.05 (5.6%)	0.70	4946
coccolithophores	0.48	0.29	0.38	0.05 (4%)	0.73	8940
cyanobacteria	0.03	0.30	0.38	0.00	0.17	9288

3.3.2. Comparison Against Satellite Sf and PSC and Model PFT

The comparison with other satellite-derived products and model simulations was performed by investigating time-latitude Hovmöller diagrams (longitudinal averaged from 180°W to 180°E). All analyses were based on monthly averages. Monthly averages of SynSenPFT were calculated from daily binned data averaged in $\log_{10}$ space as Chla data are typically log-normally distributed (Campbell, 1995). Monthly averages of Sf and c-PSC were calculated from daily data onto a 0.25° spatial grid. Monthly NOBM PFT Chla were resampled to 0.25° spatial grid in $\log_{10}$ space. Those intercomparison data representing Chla concentration, as the NOBM PFT and c-PSC products, were averaged in $\log_{10}$ space, while Sf were averaged in linear space.

The Hovmöller diagrams are presented in **Figures 9–12**, on the left, together with the climatological annual cycle, on the right. Generally the NOBM model results data are available year round at every latitude between 72°N to 80°S as opposed to the satellite products (SynSenPFT Chla, c-PSC, and Sf) which are limited by light availability and cloud-, ice-, glint, foam free conditions. While NOBM and SynSenPFT provide Chla for exactly the same PFT, c-PSC shows the Chla of the respective size class (micro, nano, pico), and Sf and 1-Sf show the contribution of pico- and microplankton, respectively, to total Chla.

When comparing c-micro and dia-NOBM (**Figures 9B,C**) directly, c-micro has in general higher Chla between 80°N and 40°S as expected since c-micro include not only diatoms but also dinoflagellates. However, south of 45°S dia-NOBM is higher than c-micro. Dia-SynSenPFT (**Figure 9D**) follow rather strictly the patterns of c-micro but values are higher in low and mid latitudes

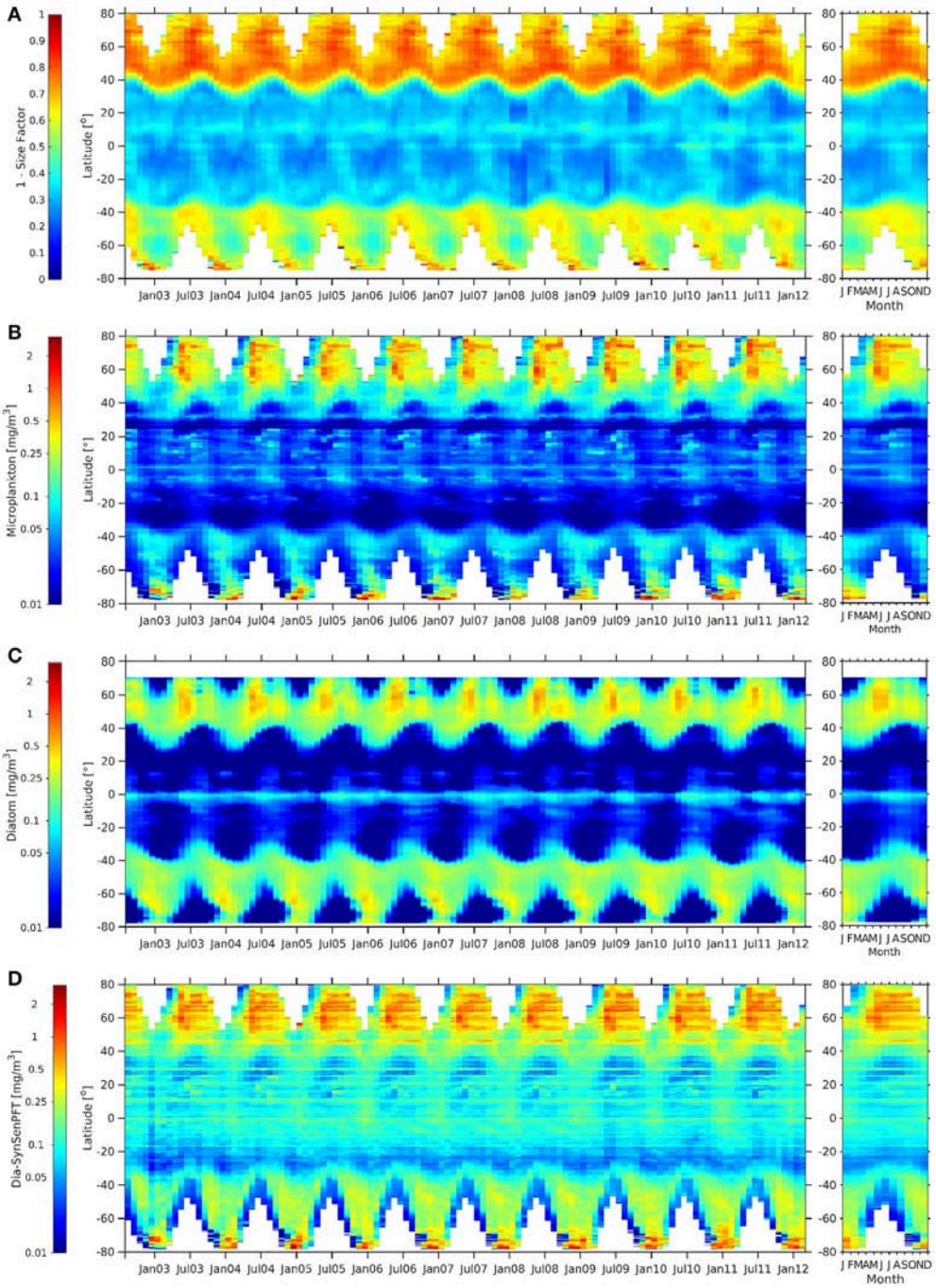


FIGURE 9 | Hovmöller diagrams of 1-Sf (Ciotti and Bricaud, 2006) **(A)**, c-micro [mg m $^{-3}$] **(B)**, dia-NOBM [mg m $^{-3}$] **(C)**, and dia-SynSenPFT [mg m $^{-3}$] **(D)** from August 2002 to March 2012 (left panel) together with the corresponding climatological annual cycle (right panels).

and the gyres are not as pronounced as for the c-micro and dia-NOBM. At the polar regions c-micro and dia-SynSenPFT show similar ranges. However, in the Arctic Ocean dia-SynSenPFT spread for longer in the year (Apr-Oct), while higher values are rather limited to Apr-Jun for c-micro. In the Southern Ocean dia-SynSenPFT and dia-NOBM have consistently higher Chla than c-micro. The 1-Sf (**Figure 9A**) shows a weaker seasonality and indicates high contribution of microplankton throughout the observed seasons for the Arctic which seems to agree more with the dia-SynSenPFT results than the c-micro. For the Southen Ocean region 1-Sf shows rather an inverse seasonality as opposed to Chla given by the other three products. The former peaks

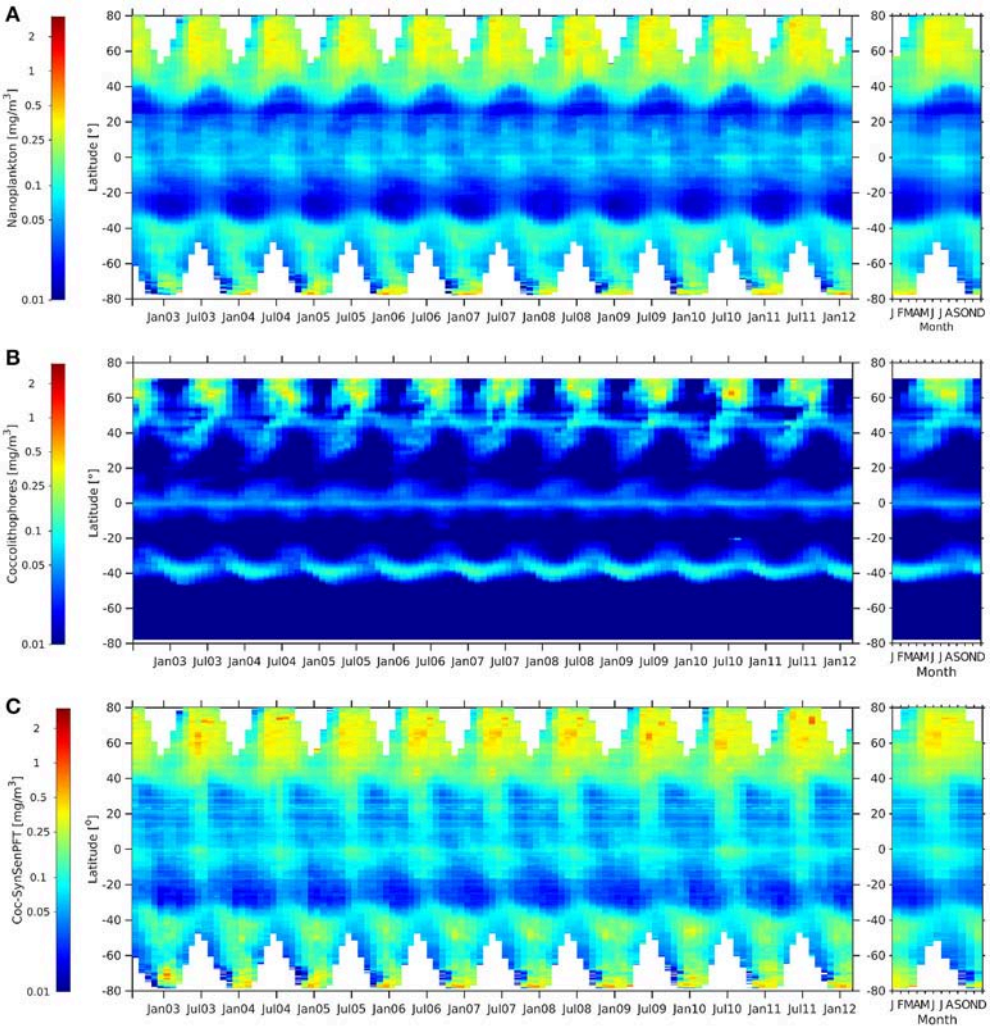


FIGURE 10 | Hovmöller diagrams of c-nano [mg m $^{-3}$] **(A)**, coc-NOBM [mg m $^{-3}$] **(B)**, and coc-SynSenPFT [mg m $^{-3}$] **(C)** from Aug. 2002 to Mar. 2012 (left panel) together with the corresponding climatological annual cycle (right panels).

at 40°S to 60°S in late winter to spring (Aug to Oct), while the others show highest Chla in summer (Dec-Feb) which is , e.g., in agreement with the diatom phenology studied by Soppa et al. (2016). It is recalled here that (1-Sf) represents a relative contribution of microplankton to total biomass, therefore its seasonality can be different from that observed for the other products which consider the magnitude of biomass in terms of Chla.

Similar distributions can be found between coc-SynSenPFT and c-nano (**Figure 10**). Values are lower in most parts for coc-SynSenPFT, which is probably related to the fact that coccolithophores are only part of c-nano. C-nano might also include green algae, *Phaeocystis* sp. and small diatoms for instance. However, coc-SynSenPFT values are a bit higher at 20°N to 40°N. Opposed to that, large differences are observed between c-nano and coc-SynSenPFT when compared with coc-NOBM (**Figure 10**). Especially south of 45°S coc-NOBM shows absents of coccolithophores which is in disagreement not only to coc-SynsenPFT, but also to findings by O'Brien et al. (2013). They show that coccolithophores have been sampled during many research campaigns up to 80°S. While coc-SynSenPFT and c-nano show elevated values at equator with similar seasonality, coc-NOBM is low, but enhanced north and south of the equator as opposed to the two other products. However, it is worth pointing out that the NOBM is the product that is not purely satellite-based (though it assimilates some satellite data) so it may not be surprising that it shows the most different patterns. All three products (c-nano, coc-SynSenPFT and coc-NOBM) show similar seasonal cycles at around 60°N.

Generally c-pico and cya-SynSenPFT (**Figure 11**) show very similar patterns and a low chl-a conc. range. Both products show clear maxima in the Arctic and nearly as high values in the Southern Ocean and at the equator, while values in the subtropics are low, especially in the gyres. They show a

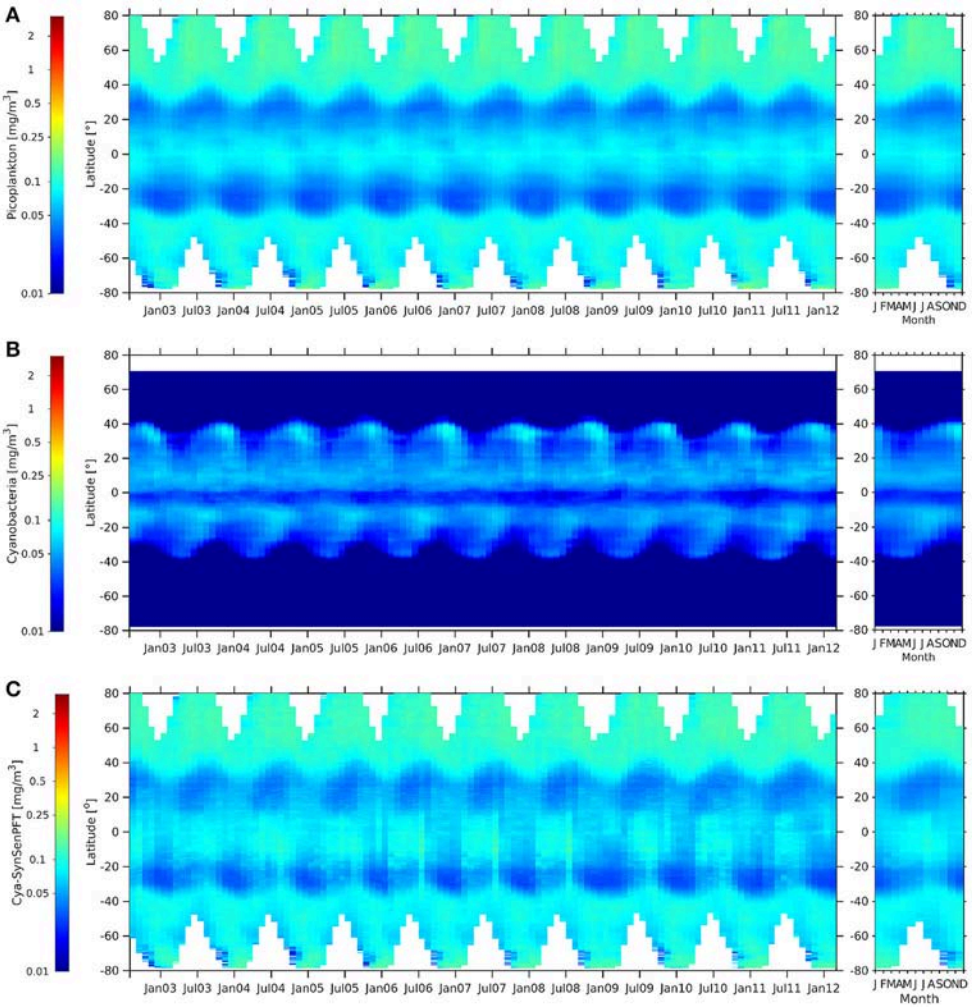


FIGURE 11 | Hovmöller diagrams of c-pico [mg m $^{-3}$] **(A)**, cya-NOBM [mg m $^{-3}$] **(B)**, and cya-SynSenPFT [mg m $^{-3}$] **(C)** from August 2002 to March 2012 (left panel) together with the corresponding climatological annual cycle (right panels).

similar seasonality at the mid- and high latitudes as for the c-nano and coc-SynSenPFT products. Opposed to that, the cya-NOBM shows quite different patterns: values are very low at latitudes higher than the subtropics and show also a minimum at the equator. Oligothrophic areas at around 20° to 35° are seen by low values in all three products. However, there, cya-NOBM Chla increases from January till May, while for c-pico and cya-SynSenPFT an increase in Chla is observed from June till September (**Figure 11**, right panels). Again, concentrations of PSC, in this case c-pico are higher than cya-NOBM, mainly north of 40°S, south of 40°S and just south of 0° (**Figure 11**). Between 0° and 40°N and 5°S and 40°S, cya-NOBM has larger concentrations. Cya-NOBM is nearly absent in latitudes higher than 40° as a result of the used cyanobacteria growth parametrisation with the growth rate decreased in cold waters (Gregg and Casey, 2007).

Sf (**Figure 12**) indicates clearly the dominance of picoplankton for the tropical waters which is reflecting what can be concluded from comparing globally the three c-PSC and the three SynSenPFT products among each other. The Hovmöller diagram for Sf (**Figure 12A**) also shows that phytopicoplankton contribution is higher (especially during summer) in the Antarctic than in the Arctic. Similar could be concluded by looking at the Hovmöller plots of *f*-Cya-SynSenPFT (**Figure 12B**) calculated as a ratio of cya-SynSenPFT product to the OC-CCI TChla. Though, for SynSenPFT it should be noted that the product is designed to identify three different PFTs which are only part of the whole phytoplankton community. Although cyanobacteria make up the largest fraction in the tropics, high diversity of cyanobacteria and other phytopicoplankton has been reported also in the Atlantic Arctic region (e.g., Fram Strait and the Greenland Sea, Díez et al. 2012).

When comparing *f*-Cya-SynSenPFT against Sf, we see that the observed dominance of the cyanobacteria in north tropical regions is weaker (not in Autumn however) and reveals more

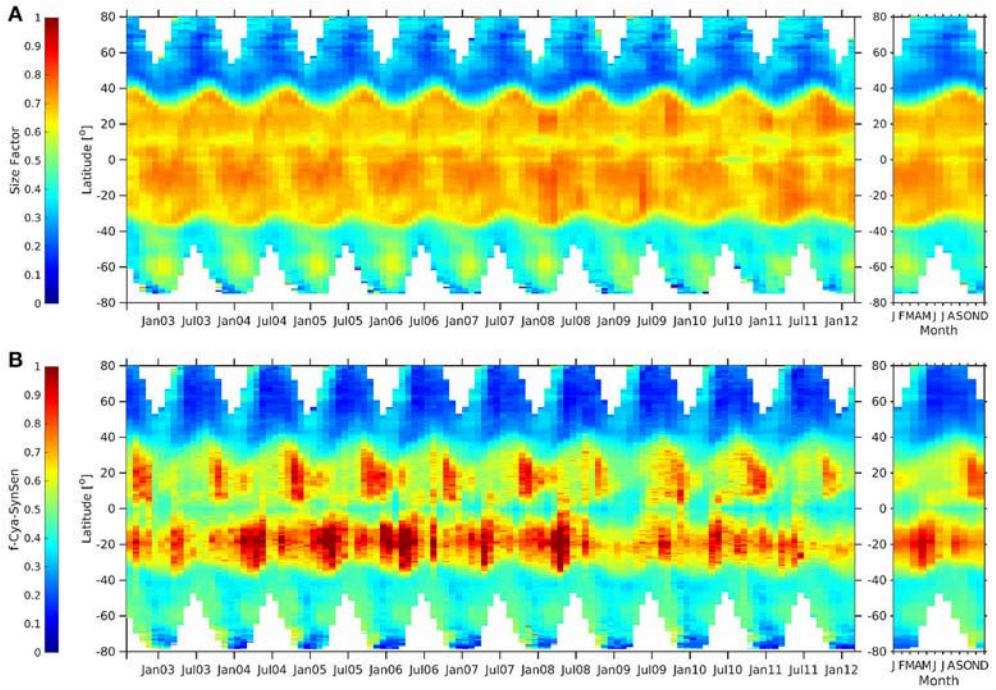


FIGURE 12 | Hovmöller diagrams of Sf (Ciotti and Bricaud, 2006) **(A)**, fraction of cya-SynSenPFT **(B)** with respect to OC-PFT TChla for the period of August 2002–March 2012 (left panel) together with the corresponding climatological annual cycle (right panel).

pronounced seasonality for the SynSenPFT product. The *f*-Cya-SynSenPFT is lower than Sf also in the equatorial region, while *f*-Cya-SynSenPFT estimates exceed Sf in southern tropical areas. One can also notice differences in temporal variability of pico-plankton dominance in the Southern Ocean. It is worth, however, emphasizing one more time that SynSenPFTs represent only a part of the whole pico-phytoplankton community.

Overall, all three SynSenPFT global products seem to reproduce the seasonality in their respective dominant phytoplankton size class (the c-PSC products), except that dia-SynSenPFT seems to be generally higher. Patterns of 1-Sf and Sf indicating the contribution of pico- and microplankton to the phytoplankton community seem also in agreement with the two former products, in terms of the spatial distribution. The temporal pattern is often different which may be due to differences between contribution of a size class and abundance. However, NOBM products, which identify the same PFTs as SynSenPFT, show for most regions quite different patterns and Chla ranges in the distibution of cyanobacteria and coccolithophores, while diatom results are closer to dia-SynSenPFT (and to c-micro). The model has parametrizations limiting the latitudinal extension of the two former groups (Gregg and Casey, 2007). Distribution of global HPLC marker pigments for PFTs (Peloquin et al., 2013), PFT abundance based on HPLC pigment data sets (Swan et al., 2016) and coccolithophore counts (O'Brien et al., 2013) clearly show presence of cyanobacteria and coccolithophores also in temperate and high latitudes. In this respect, these studies as well as the satellite-based study by Alvain et al. (2008) support the global patterns of the SynSenPFT, Sf and PSC.

4. SUMMARY AND OUTLOOK

A first version of a synergetic hyper- and multispectral-based satellite product for diatom, coccolithophore, and cyanobacteria (i.e., prokaryotic phytoplankton) Chla was developed globally (being binned at 4 × 4 km resolution) on a daily basis from August 2002 to April 2012. As input data OC-PFT Chla level-3 at the same resolution and daily 7-days composites of PhytoDOAS PFT Chla on a 0.5° by 0.5° resolution grid were used.

The SynSenPFT products were evaluated globally through validation against *in situ* PFT Chla data and intercomparison with satellite retrievals of Phytoplankton Size Classes Chla (c-PSC) of Brewin et al. (2015), Size Factor (Sf) of Ciotti and Bricaud (2006) and PFT Chla from the NASA Ocean Biogeochemical Model (NOBM) of Gregg and Casey (2007). When compared to *in situ* PFT data, good performance of SynSenPFT was achieved for all three products. Although cya-SynSenPFT does not correlate with global *in situ* match-ups (maybe due to the limited Chla range), nevertheless its RMSD and MAE were lower than for dia-SynSenPFT in the *in situ* validation. Also cya-SynSenPFT shows reasonable spatial and temporal patterns as detected by the other satellite-derived (especially c-pico) estimates and with results shown in literature based on *in situ* measurements.

Overall we can conclude that this first version of SynSenPFT products compare reasonably well with *in situ* data and

also with other satellite products. Nevertheless, the current version of the synergistic cya-SynSenPFT as well as dia-SynSenPFT in the higher than 65° regions is mostly influenced by OC-PFT data because of large gaps in the PhytoDOAS version 3.3 diatoms and cyanobacteria products at high latitudes.

It is worth emphasizing that we do not introduce the SynSenPFT only as a PFT product, but also as an established dynamic system (see **Figure 13**) including several components within a network allowing the system to further educate and develop. Any possible future improvements of the SynSenPFT system's components would lead to improved quality of the synergistic PFT product. In the following we consider each (1–4) component in a perspective of the future development.

(1) *Input initial product.* To reduce the deficiencies of the PhytoDOAS diatom Chla product, the PhytoDOAS algorithm may be optimized by including the packaging effect when converting its DOAS fit factor to its Chla via extending the LUT. Accounting for the packaging effect is probably not appropriate for the coc-SynSenPFT and cya-SynSenPFT products, since a very low pigment packaging effect in coccolithophores and if any in cyanobacteria is expected. This may reduce the observed overestimation of dia-SynSenPFT as compared to other products.

The input OC-PFT and PhytoDOAS algorithms can be applied to various similar multi- and hyperspectral-based information from other satellite missions. For instance, OC-PFT is worth applying to OLCI (Sentinel-3), while PhytoDOAS product might be based on the information from OMI (Aura) or TROPOMI- and UVN-instruments onboard Sentinel-5P and Sentinel-4 and -5. Respectively, OMI products, and especially the hyperspectral Sentinel-PFT information might allow PhytoDOAS global coverage to be improved by a factor of 2-3 temporally with spatial resolution also being improved by a factor of 4 for OMI and 40 for the upcoming hyperspectral Sentinel instruments.

PhytoDOAS is currently sensitive to the number of retrieved PFTs as shown by the PhytoDOAS sensitivity study using radiative transfer simulations with SCIATRAN (Wolanin et al., 2016a) and the model study assessing the potential of the retrievals on remote sensing reflectance data (Wolanin et al., 2016b). The SCIAMACHY retrievals were limited to channel 3 cluster 15 of the sensor which represents the wavelength range from 425 to 529 nm. Above and below this range the pixel resolution is worse (240 by 30 km). The OMI and TROPOMI sensors allow for a continuous exploitation from the UV to 500 nm. Therefore, the algorithms may be further revised and extended to enable accurate estimates of enlarged number of PFTs simultaneously retrieved.

(2) *In situ and independent information.* Continuous enlargement of the *in situ* pigment data set would enable OC-PFT algorithm and product to be revised, evaluation of the satellite retrievals and model estimates to be more accurate, and the error statistics to be more precise.

(3) *Metrics.* A more detailed TC-based analysis is desired to prescribe *a priori* error statistics spatially variable. Accounting for a seasonal variability of the used error statistics approximation would be also preferable in the context of stationarity (one of the assumptions underlying the TC).

More precise information on data quality – measurement errors traced through particular retrieval algorithms as well as estimates of the any possible representation errors – would allow for better SynSenPFT algorithm performance due to more plausible approximation of the input product *a priori* error statistics.

There is also a need in additional metrics for *a posteriori* assessment of the SynSenPFT product as well as for initial input products evaluation: the PFT/PSC/Sf phenology

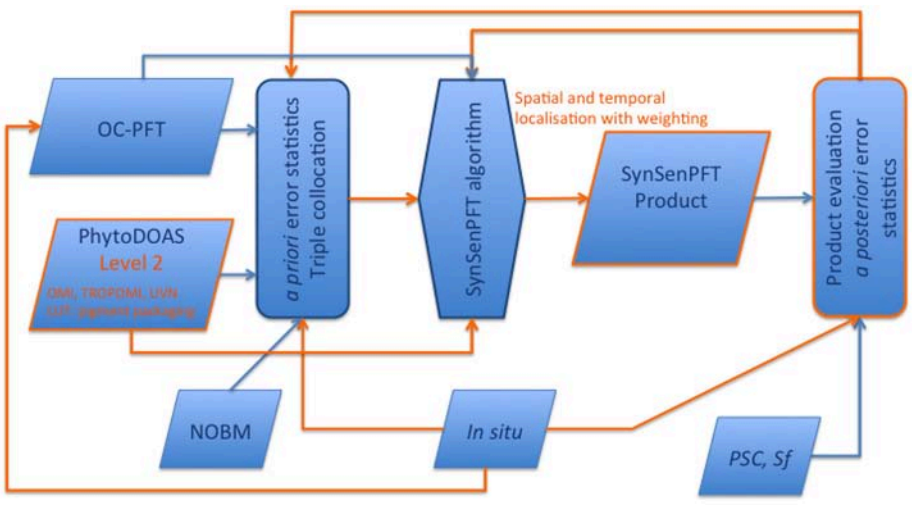


FIGURE 13 | Schematic presentation of the SynSenPFT system and development flowchart. In orange color possibilities for further improvements and/or augments are emphasized.

intercomparison on a weekly basis (for example) for various biogeochemical provinces (e.g., according to Longhurst 1998).

(4) *SynSenPFT algorithm.* The synergistic algorithm is to be further explored to include the PhytoDOAS level-2 information within a certain time window. The level-2 information can then be weighted relatively to the time interval between the date of the analysis and the date of the PhytoDOAS information used. A similar approach can be used to introduce a spatial radius of data influence. That would, however, require a calibration of the algorithm, since such a radius is spatially variable being dependent on the system dynamics which vary a lot between different regions. The synergistic algorithm can be extended with respect to augmenting input products (including, for instance, biogeochemical modeling). Details of the implementation and the algorithm design will depend on the model resolution, the output, nevertheless, could even represent a vertical structure of PFTs and cover the areas and time periods where (when) no satellite data are available.

AUTHOR CONTRIBUTIONS

ABra is the initiator and leader of the SynSenPFT project, she contributed significantly to the writing process. SL implemented the SynSenPFT algorithm, compiled the first draft and updated the manuscript according the revisions by all the co-authors: ABra, MS, RB, ABri, TD, AW, JO, VR, and IP. MS produced the OC-PFT PFT, suggested the intercomparison design and wrote the most of the Supplementary Material. TD, AW, JO, VR, BG, and ABra were involved in the PhytoDOAS version 3.3 development. ABri and BG contributed with the Sf data product, and RB provided PSC data. ABri and RB as well as IP substantially participated in the PFT products intercomparison and discussion.

FUNDING

This work was supported by the ESA SEOM SY-4Sci Synergy project (No 400112410/14/I-NB_SEOM_SY-4SciSynergy). Funding to SL was also supported by the SFB/TR 172 (AC)^3 "Arctic Amplification" subproject C03; DFG-Priority Program SPP 1158 "Antarktis" PhySyn BU2913/3-1, and by the Helmholtz Climate Initiative REKLIM (Regional Climate Change), a joint research project of the Helmholtz Association of German Research Centres (HGF).

ACKNOWLEDGMENTS

We are thankful to Steven Delwart (ESA-ESRIN) for valuable and inspiring discussion throughout the SynSenPFT project. We thank the reviewers and Editor for their nice comments and constructive suggestions leading to the improved manuscript. We thank ESA for OC-CCI data, MERIS and SCIAMACHY data; NASA Global Modeling and Assimilation Office (GMAO) for NOBM PFT output, NASA Goddard Space Flight Center's Ocean Data Processing System (ODPS) for the MODIS and SeaWiFs data. We thank Dr. Rajdeep Roy (CSIR National Institute of Oceanography (NIO), Goa, India) and all the scientists and crews involved in HPLC data collection and analyses for providing their pigment data. PFT Chla data sets used and produced within our SynSenPFT study are available at https://doi.org/10.1594/PANGAEA.873210.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmars.2017.00203/full#supplementary-material

REFERENCES

Ackleson, S., Balch, W., and Holligan, P. (1988). White waters of the gulf of maine. *Oceanography* 1, 18–22.

Alvain, S., Moulin, C., Dandonneau, Y., and Loisel, H. (2008). Seasonal distribution and succession of dominant phytoplankton groups in the global ocean: a satellite view. *Global Biogeochem. Cycles* 22:GB3001. doi: 10.1029/2007GB003154

Bracher, A. (2014a). *Phytoplankton Pigment Concentrations during POLARSTERN Cruise ANT-XXVIII/3.* Alfred Wegener Institute, Helmholtz Center for Polar and Marine Research, Bremerhaven. doi: 10.1594/PANGAEA.848588

Bracher, A. (2014b). *Phytoplankton Pigments Measured on Water Bottle Samples during SONNE Cruise SO218.* Alfred Wegener Institute, Helmholtz Center for Polar and Marine Research, Bremerhaven. doi: 10.1594/PANGAEA.848589

Bracher, A. (2015a). *Phytoplakton Pigment Concentrations during Maria S. Merian Cruise MSM18/3.* Alfred Wegener Institute, Helmholtz Center for Polar and Marine Research, Bremerhaven. doi: 10.1594/PANGAEA.848586

Bracher, A. (2015b). *Phytoplakton Pigment Concentrations during POLARSTERN Cruise ANT-XXIV/1.* Alfred Wegener Institute, Helmholtz Center for Polar and Marine Research, Bremerhaven. doi: 10.1594/PANGAEA.848583

Bracher, A. (2015c). *Phytoplakton Pigment Concentrations during POLARSTERN Cruise ANT-XXIV/4.* Alfred Wegener Institute, Helmholtz Center for Polar and Marine Research, Bremerhaven. doi: 10.1594/PANGAEA.848584

Bracher, A. (2015d). *Phytoplakton Pigment Concentrations during POLARSTERN Cruise ANT-XXVI/4.* Alfred Wegener Institute, Helmholtz Center for Polar and Marine Research, Bremerhaven. doi: 10.1594/PANGAEA.848585

Bracher, A. (2015e). *Phytoplankton Pigment Concentrations during POLARSTERN Cruise ANT-XXVII/2.* Alfred Wegener Institute, Helmholtz Center for Polar and Marine Research, Bremerhaven. doi: 10.1594/PANGAEA.848590

Bracher, A., Bouman, H., Brewin, R., Bricaud, A., Brotas, V., Ciotti, A., et al. (2017a). Obtaining phytoplankton diversity from ocean color: A scientific roadmap for future development. *Front. Mar. Sci.* 4:55. doi: 10.3389/fmars.2017.00055

Bracher, A., Dinter, T., Wolanin, A., Rozanov, V. V., Losa, S., and Soppa, M. A. (2017b). *Global Monthly Mean Chlorophyll a Surface Concentrations from August 2002 to April 2012 for Diatoms, Coccolithophores and Cyanobacteria from PhytoDOAS Algorithm Version 3.3 Applied to SCIAMACHY Data, Link to NetCDF Files in ZIP Archive.* PANGAEA. doi: 10.1594/PANGAEA.870486

Bracher, A., and Taylor, B. B. (2017). *Phytoplankton Pigment Concentrations Measured by HPLC during Maria S. Merian cruise MSM9/1.* Alfred Wegener Institute, Helmholtz Center for Polar and Marine Research, Bremerhaven. doi: 10.1594/PANGAEA.873070

Bracher, A., Taylor, M., Taylor, B., Dinter, T., Roettgers, R., and Steinmetz, F. (2015a). Using empirical orthogonal functions derived from remote sensing reflectance for the prediction of phytoplankton pigments concentrations. *Ocean Sci.* 11, 139–158. doi: 10.5194/os-11-139-2015

Bracher, A., Taylor, M. H., Taylor, B. B., Dinter, T., Röttgers, R., and Steinmetz, F. (2015b). *Phytoplakton Pigment Concentrations during POLARSTERN Cruise ANT-XXIII/1*. Bremerhaven: PANGEA. doi:10.1594/PANGAEA.871713.

Bracher, A., Vountas, M., Dinter, T., Burrows, J. P., Röttgers, R., and Peeken, I. (2009). Quantitative observation of cyanobacteria and diatoms from space using PhytoDOAS on SCIAMACHY data. *Biogeosciences* 6, 751–764. doi: 10.5194/bg-6-751-2009

Brewin, R., Ciavatta, S., Sathyendranath, S., Jackson, T., Tilstone, G., Curran, K., et al. (2017). Uncertainty in ocean-color estimates of chlorophyll for phytoplankton groups. *Front. Mar. Sci.* 4:104. doi: 10.3389/fmars.2017.00104

Brewin, R. J., Sathyendranath, S., Hirata, T., Lavender, S. J., Barciela, R. M., and Hardman-Mountford, N. J. (2010). A three-component model of phytoplankton size class for the atlantic ocean. *Ecol. Model.* 221, 1472–1483. doi: 10.1016/j.ecolmodel.2010.02.014

Brewin, R. J., Sathyendranath, S., Jackson, T., Barlow, R., Brotas, V., Airs, R., et al. (2015). Influence of light in the mixed-layer on the parameters of a three-component model of phytoplankton size class. *Remote Sensing Environ.* 168, 437–450. doi: 10.1016/j.rse.2015.07.004

Bricaud, A., Ciotti, A., and Gentili, B. (2012). Spatial-temporal variations in phytoplankton size and colored detrital matter absorption at global and regional scales, as derived from twelve years of seawifs data (1998–2009). *Global Biogeochem. Cycles* 26:GB1010. doi: 10.1029/2010GB003952

Bricaud, A., Claustre, H., Ras, J., and Oubelkheir, K. (2004). Natural variability of phytoplanktonic absorption in oceanic waters: Influence of the size structure of algal populations. *J. Geophys. Res. Oceans* 109:C11010. doi: 10.1029/2004jc002419

Buesseler, K. O. (1998). The decoupling of production and particulate export in the surface ocean. *Global Biogeochem. Cy.* 12, 297–310. doi: 10.1029/97GB03366

Campbell, J. W. (1995). The lognormal distribution as a model for bio-optical variability in the sea. *J. Geophys. Res. Oceans* 100, 13237–13254. doi: 10.1029/95JC00458

Cheah, W., Taylor, B. B., Wiegmann, S., Raimund, S., Krahmann, G., Quack, B., et al. (2013). Photophysiological state of natural phytoplankton communities in the south china sea and sulu sea. *Biogeosci. Discuss.* 10, 12115–12153. doi: 10.5194/bgd-10-12115-2013

Chisholm, S. W. (1992). "Phytoplankton size," in *Primary Productivity and Biogeochemical Cycles in the Sea* (Springer), 213–237.

Ciotti, A. M., and Bricaud, A. (2006). Retrievals of a size parameter for phytoplankton and spectral light absorption by colored detrital matter from water-leaving radiances at seawifs channels in a continental shelf region off brazil. *Limnol. Oceanogr. Methods* 4, 237–253. doi: 10.4319/lom.2006.4.237

Ciotti, A. M., Lewis, M. R., and Cullen, J. J. (2002). Assessment of the relationships between dominant cell size in natural phytoplankton communities and the spectral shape of the absorption coefficient. *Limnol. Oceanogr.* 47, 404–417. doi: 10.4319/lo.2002.47.2.0404

Crow, W. T., and van den Berg, M. J. (2010). An improved approach for estimating observation and model error parameters in soil moisture data assimilation. *Water Res. Res.* 46:W12519. doi: 10.1029/2010WR009402

Díez, B., Bergman, B., Pedrós-Alió, C., Antó, M., and Snoeijs, P. (2012). High cyanobacterial nifh gene diversity in arctic seawater and sea ice brine. *Environ. Microbiol. Reports* 4, 360–366. doi: 10.1111/j.1758-2229.2012.00343.x

Dinter, T., Rozanov, V., Burrows, J., and Bracher, A. (2015). Retrieval of light availability in ocean waters utilizing signatures of vibrational raman scattering in hyper-spectral satellite measurements. *Ocean Sci.* 11, 373–389. doi: 10.5194/os-11-373-2015

Field, C. B., Behrenfeld, M. J., Randerson, J. T., and Falkowski, P. (1998). Primary production of the biosphere: Integrating terrestrial and oceanic components. *Science* 281, 237–240. doi: 10.1126/science.281.5374.237

Franklin, D. J., Poulton, A. J., Steinke, M., Young, J., Peeken, I., and Malin, G. (2009). Dimethylsulphide, dmsp-lyase activity and microplankton community structure inside and outside of the mauritanian upwelling. *Progr. Oceanogr.* 83, 134–142. doi: 10.1016/j.pocean.2009.07.011

Gandin, L. S., and Hardin, R. (1965). *Objective Analysis of Meteorological Fields*, Vol. 242. Jerusalem: Israel Program for Scientific Translations Jerusalem.

Gregg, W. W. (2002). "A coupled ocean-atmosphere radiative model for global ocean biogeochemical models," in *NASA Global Modeling and Assimilation Series*, Vol. 22, ed M. Suarez (Greenbelt, MD: NASA Technical Memorandum 2002-104606), 33.

Gregg, W. W., and Casey, N. W. (2007). Modeling coccolithophores in the global oceans. *Deep Sea Research II: Topical Studies in Oceanography*, 54, 447–477. doi: 10.1016/j.dsr2.2006.12.007

Gregg, W. W., and Rousseaux, C. S. (2017). Simulating pace global ocean radiances. *Front. Marine Sci.* 4:60. doi: 10.3389/fmars.2017.00060

Gruber, A., Su, C.-H., Zwieback, S., Crow, W., Dorigo, W., and Wagner, W. (2016). Recent advances in (soil moisture) triple collocation analysis. *Int. J. Appl. Earth Observ. Geoinform.* 45 (Pt B), 200–211. doi: 10.1016/j.jag.2015.09.002

Hardman-Mountford, N. J., Hirata, T., Richardson, K. A., and Aiken, J. (2008). An objective methodology for the classification of ecological pattern into biomes and provinces for the pelagic ocean. *Remote Sensing Environ.* 112, 3341–3352. doi: 10.1016/j.rse.2008.02.016

Hirata, T., Hardman-Mountford, N., Brewin, R., Aiken, J., Barlow, R., Suzuki, K., et al. (2011). Synoptic relationships between surface chlorophyll-a and diagnostic pigments specific to phytoplankton functional types. *Biogeosciences* 8, 311–327. doi: 10.5194/bg-8-311-2011

IOCCG (2009). "Remote sensing fisheries and aquaculture: the societal benefits," in *in Reports of the International Ocean Color Coordinating Group No. 8*, eds M.-H. Forget, V. Stuart, and T. Platt (Dartmouth, NS: IOCCG).

IOCCG (2014). "Phytoplankton functional types from space," in *in Reports of the International Ocean Color Coordinating Group No. 15*, eds S. Sathyendranath and V. Stuart (Dartmouth, NS: IOCCG).

Johnson, R., Strutton, P. G., Wright, S. W., McMinn, A., and Meiners, K. M. (2013). Three improved satellite chlorophyll algorithms for the southern ocean. *J. Geophys. Res. Oceans* 118, 3694–3703. doi: 10.1002/jgrc.20270

Laufkötter, C., Vogt, M., Gruber, N., Aumont, O., Bopp, L., Doney, S. C., (2016). Projected decreases in future marine export production: the role of the carbon flux through the upper ocean ecosystem. *Biogeosciences* 13, 4023–4047. doi: 10.5194/bg-13-4023-2016

Lee, Z., Carder, K. L., and Arnone, R. A. (2002). Deriving inherent optical properties from water color: a multiband quasi-analytical algorithm for optically deep waters. *Appl. Opt.* 41, 5755–5772. doi: 10.1364/AO.41.005755

Le Qur, C., Harrison, S. P., Colin Prentice, I., Buitenhuis, E. T., Aumont, O., Bopp, L., et al. (2005). Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. *Global Change Biol.* 11, 2016–2040. doi: 10.1111/j.1365-2486.2005.1004.x

Longhurst, A. R. (1998). *Ecological Geography of the Sea*. San Diego, CA; London: Academic Press.

Losa, S. N., Danilov, S., Schröter, J., Janjić, T., Nerger, L., and Janssen, F. (2014). Assimilating NOAA SST data into BSH operational circulation model for the north and baltic seas: part 2. sensitivity of the forecast's skill to the prior model error statistics. *J. Marine Sys.* 129, 259–270. doi: 10.1016/j.jmarsys.2013.06.011

Losa, S. N., Danilov, S., Schröter, J., Nerger, L., Maßmann, S., and Janssen, F. (2012). Assimilating NOAA SST data into the BSH operational circulation model for the north and baltic seas: Inference about the data. *J. Marine Sys.* 105–108, 152–162. doi: 10.1016/j.jmarsys.2012.07.008

McColl, K. A., Vogelzang, J., Konings, A. G., Entekhabi, D., Piles, M., and Stoffelen, A. (2014). Extended triple collocation: Estimating errors and correlation coefficients with respect to an unknown target. *Geophys. Res. Lett.* 41, 6229–6236. doi: 10.1002/2014GL061322

Milliman, J. D. (1993). Production and accumulation of calcium carbonate in the ocean: Budget of a nonsteady state. *Global Biogeochem. Cycles* 7, 927–957. doi: 10.1029/93GB02524

Morán, X. A. G., Fernández, E., and Pérez, V. (2004). Size-fractionated primary production, bacterial production and net community production in subtropical and tropical domains of the oligotrophic ne atlantic in autumn. *Marine Ecol. Progr. Ser.* 274, 17–29. doi: 10.3354/meps274017

Morel, A., Gentili, B., Claustre, H., Babin, M., Bricaud, A., Ras, J., et al. (2007). Optical properties of the "clearest" natural waters. *Limnol. Oceanogr.* 52, 217–229. doi: 10.4319/lo.2007.52.1.0217

Mouw, C. B., Hardman-Mountford, N. J., Alvain, S., Bracher, A., Brewin, R. J. W., Bricaud, A., et al. (2017). A consumer's guide to satellite remote sensing of multiple phytoplankton groups in the global ocean. *Front. Marine Sci.* 4:41. doi: 10.3389/fmars.2017.00041

Nöthig, E.-M., Bracher, A., Engel, A., Metfies, K., Niehoff, B., Peeken, I., et al. (2015). Summertime plankton ecology in the Fram Sztrait: a compilation of short- and long-term observations. *Polar Res.* 34:23349. doi: 10.3402/polar.v34.23349

O'Brien, C., Peloquin, J., Vogt, M., Heinle, M., Gruber, N., et al. (2013). Global plankton functional type biomass distributions: Coccolithophores. *Earth Sys. Sci. Data* 5, 259–276. doi: 10.5194/essd-5-259-2013

OC-CCI (2015). *Ocean Color Climate Change Initiative Product User Guide Version 2*. Plymouth: Plymouth Marine Laboratory.

Organelli, E., Bricaud, A., Antoine, D., and Uitz, J. (2013). Multivariate approach for the retrieval of phytoplankton size structure from measured light absorption spectra in the mediterranean sea (boussole site). *Appl. Opt.* 52, 2257–2273. doi: 10.1364/AO.52.002257

Organelli, E., Nuccio, C., Melillo, C., and Massi, L. (2011). Relationships between phytoplankton light absorption, pigment composition and size structure in offshore areas of the mediterranean sea. *Adv. Oceanogr. Limnol.* 2, 107–123. doi: 10.4081/aiol.2011.5320

Peeken, I., Bluhm, K., and Zöllner, E. (2017a). *Phytoplankton pigments measured from underway and water bottle samples during Marion Dufresne cruise MD158 (OOMPH MD07 leg 1).* PANGAEA. doi: 10.1594/PANGAEA.873203

Peeken, I., Bluhm, K., and Zo ́llner, E. (2017b). *Phytoplankton Pigments Measured from Underway and Water Bottle Samples during Marion Dufresne Cruise MD160 (OOMPH MD07 leg 2).* PANGAEA. doi: 10.1594/PANGAEA.873208

Peeken, I., Bracher, A., and Murawski, S. (2017c). *Phytoplankton Pigments Measured on Water Bottle Samples during POLARSTERN Cruise ANT-XXIV/2.* PANGAEA. doi: 10.1594/PANGAEA.873199

Peeken, I., and Hoffmann, L. (2014a). *Phytoplankton Pigments and Nutrients Measured on Water Bottle Samples during METEOR Cruise M55.* PANGAEA. doi: 10.1594/PANGAEA.869828

Peeken, I., and Hoffmann, L. (2014b). Phytoplankton pigments in *Surface Water during POLARSTERN Cruise ANT-XXI/3 (EIFEX).* PANGAEA. doi: 10.1594/PANGAEA.869825

Peeken, I. and Hoffmann, L. (2014c). *Phytoplankton Pigments Measured on Water Bottle Samples during POLARSTERN Cruise ANT-XXI/3 (EIFEX).* PANGAEA. doi: 10.1594/PANGAEA.869824

Peeken, I., and Nachtigall, K. (2014a). *Phytoplankton Pigments Measured on Water Bottle Samples during POLARSTERN Cruise ANT-XVIII/2 (EisenEx).* PANGAEA. doi: 10.1594/PANGAEA.869823

Peeken, I., and Nachtigall, K. (2014b). *Pigments in Surface Water during POLARSTERN Cruise ANT-XXVI/3.* PANGAEA. doi: 10.1594/PANGAEA.869827

Peeken, I., and Nachtigall, K. (2014c). *Pigments Measured on Water Bottle Samples during METEOR Cruise M60.* PANGAEA. doi: 10.1594/PANGAEA.869830

Peeken, I., and Quack, B. (2017). *Phytoplankton Pigments Measured on Underway and Water Bottle Samples during POSEIDON Cruise POS320.* PANGAEA. doi: 10.1594/PANGAEA.873198

Peeken, I., and Walter, S. (2017). *Phytoplankton Pigments Measured from Underway and Water Bottle Samples during POSEIDON Cruise POS348.* PANGAEA. doi: 10.1594/PANGAEA.873200

Peeken, I., Wilken, S., Bremer, K., and Schnack-Schiel, S. B. (2017d). *Phytoplankton Pigments Measured from Seawater Samples from the Spermonde Archipelago in Southwest Sulawesi, Indonesia.* Alfred Wegener Institute, Helmholtz Center for Polar and Marine Research, Bremerhaven. doi: 10.1594/PANGAEA.873209

Peloquin, J., Swan, C., Gruber, N., Vogt, M., Claustre, H., Ras, J., et al. (2013). The maredat global database of high performance liquid chromatography marine pigment measurements. *Earth Sys. Sci. Data* 5, 109–123. doi: 10.5194/essd-5-109-2013

Perner, D., and Platt, U. (1979). Detection of nitrous acid in the atmosphere by differential optical absorption. *Geophys. Res. Lett.* 6, 917–920. doi: 10.1029/GL006i012p00917

Quack, B., Peeken, I., Petrick, G., and Nachtigall, K. (2007). Oceanic distribution and sources of bromoform and dibromomethane in the mauritanian upwelling. *J. Geophys. Res. Oceans* 112, 1–13. doi: 10.1029/2006jc003803

Raimbault, P., Rodier, M., and Taupier-Letage, I. (1988). Size fraction of phytoplankton in the ligurian sea and the algerian basin (mediterranean sea): size distribution versus total concentration. *Marine Microb. Food Webs* 3, 1–7.

Rozanov, V., Rozanov, A., Kokhanovsky, A., and Burrows, J. (2014). Radiative transfer through terrestrial atmosphere and ocean: Software package SCIATRAN. *J. Quant. Spectr. Radiat. Transfer* 133, 13–71. doi: 10.1016/j.jqsrt.2013.07.004

Sá, C., D'Alimonte, D., Brito, A., Kajiyama, T., Mendes, C., Vitorino, J., et al. (2015). Validation of standard and alternative satellite ocean-color chlorophyll products off western iberia. *Remote Sensing Environ.* 168, 403–419. doi: 10.1016/j.rse.2015.07.018

Sadeghi, A., Dinter, T., Vountas, M., Taylor, B., Altenburg-Soppa, M., and Bracher, A. (2012). Remote sensing of coccolithophore blooms in selected oceanic regions using the PhytoDOAS method applied to hyper-spectral satellite data. *Biogeosciences* 9, 2127–2143. doi: 10.5194/bg-9-2127-2012

Sathyendranath, S., Cota, G., Stuart, V., Maass, H., and Platt, T. (2001). Remote sensing of phytoplankton pigments: a comparison of empirical and theoretical approaches. *Int. J. Remote Sensing* 22, 249–273. doi: 10.1080/014311601449925

Sauzde, R., Claustre, H., Jamet, C., Uitz, J., Ras, J., Mignot, A., et al. (2015). Retrieving the vertical distribution of chlorophyll a concentration and phytoplankton community composition from in situ fluorescence profiles: a method based on a neural network with potential for global-scale applications. *J. Geophys. Res. Oceans* 120, 451–470. doi: 10.1002/2014JC010355

Soppa, M. A., Hirata, T., Silva, B., Dinter, T., Peeken, I., Wiegmann, S., et al. (2014). Global retrieval of diatom abundance based on phytoplankton pigments and satellite data. *Remote Sensing* 6, 10089–10106. doi: 10.3390/rs61010089

Soppa, M. A., Peeken, I., and Bracher, A. (2017). *Global Chlorophyll "a" Concentrations for Diatoms, Haptophytes and Prokaryotes Obtained with the Diagnostic Pigment Analysis of HPLC Data Compiled from Several Databases and Individual Cruises.* PANGAEA. doi: 10.1594/PANGAEA.875879

Soppa, M. A., Voelker, C., and Bracher, A. (2016). Diatom Phenology in the Southern Ocean: Mean Patterns, Trends and the Role of Climate Oscillations. *Remote Sensing* 8:420. doi: 10.3390/rs8050420

Stoffelen, A. (1998). Toward the true near-surface wind speed: error modeling and calibration using triple collocation. *J. Geophys. Res. Oceans (1978–2012)* 103, 7755–7766. doi: 10.1029/97JC03180

Swan, C., Vogt, M., Gruber, N., and Laufkötter, C. (2016). A global seasonal surface ocean climatology of phytoplankton types based on CHEMTAX analysis of HPLC pigments. *Deep Sea Res. Part I* 109, 137–156. doi: 10.1016/j.dsr.2015.12.002

Taylor, B. B., and Bracher, A. (2012). *Pigment Concentrations Measured in Surface Water during SONNE Cruise SO202/2 (TRANSBROM).* PANGAEA. doi: 10.1594/pangaea.804525

Taylor, B. B., Torrecilla, E., Bernhardt, A., Taylor, M. H., Peeken, I., Röttgers, R., et al. (2011a). Bio-optical provinces in the eastern atlantic ocean and their biogeographical relevance. *Biogeosciences* 8, 3609–3629. doi: 10.5194/bg-8-3609-2011

Taylor, B. B., Torrecilla, E., Bernhardt, A., Taylor, M. H., Peeken, I., Röttgers, R., et al. (2011b). *Pigments of Phytoplankton during POLARSTERN Cruise ANT-XXV/1.* Alfred Wegener Institute, Helmholtz Center for Polar and Marine Research, Bremerhaven. doi: 10.1594/PANGAEA.819070

Trimborn, S., Hoppe, C. J., Taylor, B. B., Bracher, A., and Hassler, C. (2015). Physiological characteristics of open ocean and coastal phytoplankton communities of western antarctic peninsula and drake passage waters. *Deep Sea Res. I Oceanogr. Res. Papers* 98, 115–124. doi: 10.1016/j.dsr.2014.12.010

Uitz, J., Claustre, H., Griffiths, F. B., Ras, J., Garcia, N., and Sandroni, V. (2009). A phytoplankton class-specific primary production model applied to the kerguelen islands region (southern ocean). *Deep Sea Res. I Oceanogr. Res. Papers* 56, 541–560. doi: 10.1016/j.dsr.2008.11.006

Uitz, J., Claustre, H., Morel, A., and Hooker, S. B. (2006). Vertical distribution of phytoplankton communities in open ocean: an assessment based on surface chlorophyll. *J. Geophys. Res.* 111:C08005. doi: 10.1029/2005jc003207

Vidussi, F., Claustre, H., Manca, B. B., Luchetta, A., and Marty, J.-C. (2001). Phytoplankton pigment distribution in relation to upper thermocline circulation in the eastern mediterranean sea during winter. *J. Geophys. Res. Oceans (1978–2012)* 106, 19939–19956. doi: 10.1029/1999JC000308

Vogelzang, J., Stoffelen, A., Verhoef, A., and Figa-Saldaa, J. (2011). On the quality of high-resolution scatterometer winds. *J. Geophys. Res. Oceans* 116:C10033. doi: 10.1029/2010jc006640

Vountas, M., Dinter, T., Bracher, A., Burrows, J. P., and Sierk, B. (2007). Spectral studies of ocean water with space-borne sensor SCIAMACHY using

Differential Optical Absorption Spectroscopy (doas). *Ocean Sci.* 3, 429–440. doi: 10.5194/os-3-429-2007

Ward, B. A. (2015). Temperature-correlated changes in phytoplankton community structure are restricted to polar waters. *PLoS ONE* 10:e0135581. doi: 10.1371/journal.pone.0135581

Waterbury, J. B., Watson, S. W., Valois, F. W., and Franks, D. G. (1986). Biological and ecological characterization of the marine unicellular cyanobacterium synechococcus. *Can. Bull. Fish Aquat. Sci.* 214:120.

Werdell, P. J., Bailey, S., Fargion, G., Pietras, C., Knobelspiesse, K., Feldman, G., et al. (2003). Unique data repository facilitates ocean color satellite validation. *EOS Trans. Am. Geophys. Union* 84, 377–387. doi: 10.1029/2003EO380001

Werdell, P. J., Franz, B. A., Bailey, S. W., Harding, L. W. Jr., and Feldman, G. C. (2007). Approach for the long-term spatial and temporal evaluation of ocean color satellite data products in a coastal environment. *Proc. SPIE* 6680:66800G. doi: 10.1117/12.732489

Wolanin, A., Dinter, T., Soppa, M., and Bracher, A. (2016a). *D2.1 Report on Using Radiative Transfer Modeling to Develop a Correction Scheme and Investigate the Sensitivity of the Improved Phytodoas (version3.0) Algorithm.* SY-4Sci Synergy R and D Study 4: Phytoplankton Functional Types (SynSenPFT).

Wolanin, A., Soppa, M. A., and Bracher, A. (2016b). Investigation of spectral band requirements for improving retrievals of phytoplankton functional types. *Remote Sensing Environ.* 8:871. doi: 10.3390/rs8100871

Yang, Q., Losch, M., Losa, S. N., Jung, T., Nerger, L., and Lavergne, T. (2016). Brief communication: The challenge and benefit of using sea ice concentration satellite data products with uncertainty estimates in summer sea ice data assimilation. *Cryosphere* 10, 761–774. doi: 10.5194/tc-10-761-2016

Zindler, C., Bracher, A., Marandino, C. A., Taylor, B., Torrecilla, E., Kock, A., et al. (2012a). Sulphur compounds, methane, and phytoplankton: interactions along a north-south transit in the western Pacific Ocean. *Biogeosci. Discuss.* 9, 15011–15049. doi: 10.5194/bgd-9-15011-2012

Zindler, C., Peeken, I., Marandino, C., and Bange, H. W. (2012b). Environmental control on the variability of DMS and DMSP in the Mauritanian upwelling region. *Biogeosciences* 9, 1041–1051. doi: 10.5194/bg-9-1041-2012

Zwieback, S., Scipal, K., Dorigo, W., and Wagner, W. (2012). Structural and statistical properties of the collocation technique for error characterization. *Nonlinear Process. Geophys.* 19, 69–80. doi: 10.5194/npg-19-69-2012

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Losa, Soppa, Dinter, Wolanin, Brewin, Bricaud, Oelker, Peeken, Gentili, Rozanov and Bracher. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Marine Science

ORIGINAL RESEARCH
published: 02 August 2017
doi: 10.3389/fmars.2017.00246

Reproduction of Spatio-Temporal Patterns of Major Mediterranean Phytoplankton Groups from Remote Sensing OC-CCI Data

Gabriel Navarro[1]*, *Pablo Almaraz*[1], *Isabel Caballero*[1], *Águeda Vázquez*[2] and *Isabel E. Huertas*[1]

[1] Department of Ecology and Coastal Management, Institute of Marine Sciences of Andalusia (CSIC), Puerto Real, Spain,
[2] Department of Applied Physics, Higher Engineering School, University of Cadiz, Puerto Real, Spain

During the last two decades, several satellite algorithms have been proposed to retrieve information about phytoplankton groups using ocean color data. One of these algorithms, the so-called PHYSAT-Med, was developed specifically for the Mediterranean Sea due to the optical peculiarities of this basin. The method allows the detection from ocean color images of the dominant Mediterranean phytoplankton groups, namely nanoeukaryotes, *Prochlorococcus, Synechococcus*, diatoms, coccolithophorids, and *Phaeocystis*-like phytoplankton. Here, we present a new version of PHYSAT-Med applied to the Ocean Colour—Climate Change Initiative (OC-CCI) database. The OC-CCI database consists of a multi-sensor, global ocean-color product that merges observations from four different sensors. This retuned version presents improvements with respect to the previous version, as it increases the temporal range (since 1998), decreases the cloud cover, improves the bias correction and a validation exercise was performed in the NW Mediterranean Sea. In particular, the PHYSAT-Med version has been used here to analyse the annual cycles of the major phytoplankton groups in the Mediterranean Sea. Wavelet analyses were used to explore the spatial variability in dominance both in the time and frequency domains in several Mediterranean sub-regions, such as the Alboran Sea, Ligurian Sea, Northern Adriatic Sea, and Levantine basin. Results extended the interpretation of previously detected patterns, indicating the dominance of *Synechococcus*-like vs. prochlorophytes throughout the year at the basin level, and the predominance of nanoeukaryotes during the winter months. The method successfully reproduced the diatom blooms normally detected in the basin during the spring season (March to April), especially in the Adriatic Sea. According to our results, the PHYSAT-Med OC-CCI algorithm represents a useful tool for the spatio-temporal monitoring of dominant phytoplankton groups in Mediterranean surface waters. The successful applications of other regional ocean color algorithms to the OC-CCI database will give rise to extended time series of phytoplankton functional types, with promising applications to the study of long-term oceanographic trends in a global change context.

Keywords: PHYSAT-Med algorithm, OC-CCI database, phytoplankton functional types, Mediterranean Sea, wavelet analysis

OPEN ACCESS

Edited by:
Astrid Bracher, Alfred-Wegener-Institute Helmholtz Center for Polar and Marine Research, Germany

Reviewed by:
Alison Palmer Chase, University of Maine, United States
Takafumi Hirata, Hokkaido University, Japan

***Correspondence:**
Gabriel Navarro
gabriel.navarro@icman.csic.es

Specialty section:
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

Received: *15 March 2017*
Accepted: *19 July 2017*
Published: *02 August 2017*

Citation:
Navarro G, Almaraz P, Caballero I, Vázquez Á and Huertas IE (2017) Reproduction of Spatio-Temporal Patterns of Major Mediterranean Phytoplankton Groups from Remote Sensing OC-CCI Data. Front. Mar. Sci. 4:246. doi: 10.3389/fmars.2017.00246

INTRODUCTION

Since the launch of the Coastal Zone Color Scanner (CZCS) in the late 1970s, ocean color remote sensing has deeply improved our understanding of the ocean system by providing global estimations of the surface chlorophyll concentration (Chl*a*), a parameter known to be a good proxy of phytoplankton biomass (e.g., McClain, 2009). Marine phytoplankton are located at the base of the marine food web (Chassot et al., 2010; and references therein), play a major role in the global biogeochemical cycles (Field et al., 1998) and participate actively in the regulation of the global climate (Sabine et al., 2004). During the last 40 years, observations of regional-to-global Chl*a* data have been acquired by different ocean color sensors (IOCCG, 2012), such as Sea-viewing Wide Field-of-view Sensor (SeaWiFS), Moderate Resolution Imaging Spectroradiometer (MODIS), Medium-Resolution Imaging Spectrometer (MERIS) and Visible Infrared Imager Radiometer Suite (VIIRS). In order to extend the existing time series beyond that provided by a single satellite sensor, the European Space Agency (ESA) has recently generated the Ocean Colour—Climate Change Initiative (OC-CCI), a multi-sensor, global, ocean-color product mainly devoted to climate research (Storm et al., 2013) that merges observations from four different sensors: SeaWiFS, MODIS, MERIS, and VIIRS. As an ESA-funded CCI project, the OC-CCI focuses specifically in creating a consistent, error-characterized time-series of ocean-color products, with a strong focus in climate-change studies (Brewin et al., 2015). Remote-sensing reflectance (Rrs) data from MODIS-Aqua and MERIS are then band-shifted to match the wavelengths of SeaWiFS by using an in-water bio-optical model (e.g., see Mélin and Sclep, 2015). The main reason behind this choice is that SeaWiFS is widely considered as the highest quality sensor with the best match to *in situ* observations, and is commonly used in peer literature (Couto et al., 2016). This dataset improves the bias correction, thus reducing the sensitivity to medium-term changes and extending the method applicability beyond the lifetime of SeaWiFS. As a result, the current OC-CCI database allows for the examination of the spatial and temporal variability of surface Chl*a* since September 1997 (Couto et al., 2016).

Even though remote sensing derived phytoplankton types does not provide a full description of the marine ecosystem, its spatio-temporal distribution (including phenology, Kostadinov et al., 2017), and identification of key groups give powerful insights on the dynamics of the marine food web and the ocean's role in climate regulation in the context of the global change (Bracher et al., 2017). This relevance was early recognized by Platt et al. (2006), who concluded that detection of phytoplankton from remote sensing images was a major challenge in ocean optics. Therefore, over the last decade, several remote sensing algorithms have been developed to characterize the global distribution patterns of phytoplankton functional types (PFT) or size classes (PSC; e.g., Sathyendranath et al., 2004; Alvain et al., 2005, 2008; Ciotti and Bricaud, 2006; Raitsos et al., 2008; Aiken et al., 2009; Bracher et al., 2009; Brewin et al., 2010; Kostadinov et al., 2010; Hirata et al., 2011; Uitz et al., 2012; see recent summary in Table 2 in (Bracher et al., 2017 and Table 3 in Mouw et al., 2017). A complete guide of the available approaches can be found in Mouw et al. (2017). Some of these algorithms are based on various spectral features, such as backscattering (e.g., Kostadinov et al., 2010), absorption (e.g., Ciotti and Bricaud, 2006; Bracher et al., 2009; Mouw and Yoder, 2010; Roy et al., 2013) or a hybrid between absorption and backscattering (Fujiwara et al., 2011). Other algorithms exploit second-order anomalies of reflectance spectra (Alvain et al., 2005, 2008), which is the case of the so-called PHYSAT that was first developed at a global scale by Alvain et al. (2005, 2008). The PHYSAT approach relies on the identification of specific signatures in the normalized water leaving radiance (nLw) spectra measured by an ocean color sensor (Alvain et al., 2005, 2008), thereby enabling the identification of nanoeukaryotes, haptophytes (a major component of the nanoflagellates), *Synechococcus*-like cyanobacteria, diatoms, *Prochlorococcus, Phaeocystis*-like phytoplankton, and coccolithophorids. The PHYSAT method has been successfully validated with phytoplankton *in situ* data and extensively used by many authors (e.g., Bopp et al., 2005; Arnold et al., 2010; D'Ovidio et al., 2010; Gorgues et al., 2010; Masotti et al., 2010, 2011; Alvain et al., 2012, 2013; Belviso et al., 2012; Demarcq et al., 2012; De Monte et al., 2013; Hashioka et al., 2013; Ben Mustapha et al., 2014; Thyssen et al., 2015).

Navarro et al. (2014) later proposed a regionalized version of the algorithm for the Mediterranean Sea (**Figure 1**), the PHYSAT-Med, using the MODIS era (2002–2013) for identification of nanoeukaryotes, *Prochlorococcus*, *Synechococcus*-like cyanobacteria and diatoms, which was compared with more than 3,000 high-performance liquid chromatography (HPLC) *in situ* measurements (see Table 3 in Navarro et al., 2014). The main utility of the PHYSAT-Med is that it allows for the tracking of specific features of phytoplankton community structure occurring in the basin, along with their associated bio-optical relationships that are heavily affected by continental inputs, such as desert dust events and rivers discharge (Bricaud et al., 2002; Claustre et al., 2002; Alvain et al., 2006; Loisel et al., 2011). Volpe et al. (2007) early suggested that the unique phytoplankton assemblages of the basin could alter its spectral signature, therefore being responsible for the peculiar color of the Mediterranean. Due to these characteristics, standard remote sensing approaches tend to either overestimate or underestimate Chl*a* levels in the Mediterranean. In fact, Volpe et al. (2007) also showed that NASA SeaWiFS standard chlorophyll products are affected by an uncertainty in the order of 100%. Specific algorithms have been thus developed to retrieve Chl*a* in the region, namely DORMA-SeaWiFS (D'Ortenzio et al., 2002), BRIC-SeaWiFS (Bricaud et al., 2002), MedOC4-SeaWiFS (Volpe et al., 2007), MedOC3-MODIS (Santoleri et al., 2008), and MedOC4ME-MERIS (Santoleri et al., 2008).

Furthermore, these bio-optical characteristics of the basin described above clearly indicate the necessity to use customized algorithms to detect PFT or PSC in Mediterranean Sea. Recently, Sammartino et al. (2015) described the temporal variability of PSC in Mediterranean Sea using the model proposed by Brewin et al. (2011). Di Cicco et al. (2017) presented a new regional algorithm to identify simultaneously the contribution of each

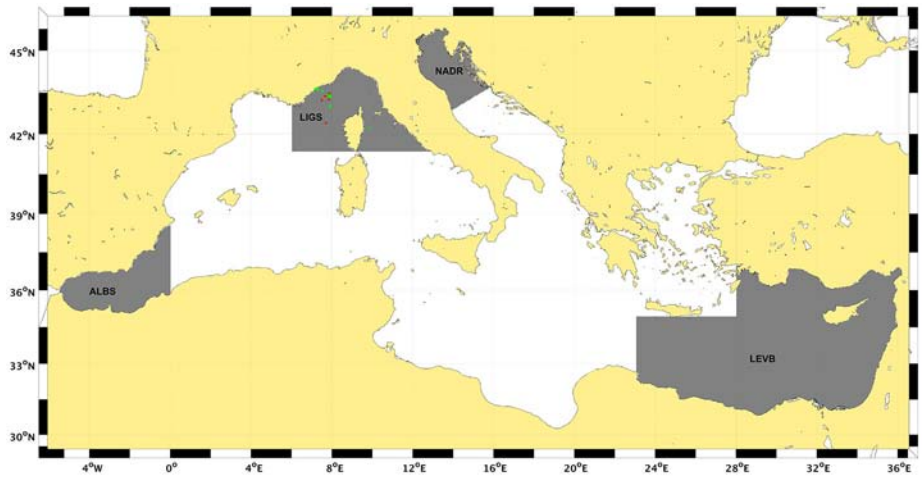


FIGURE 1 | Map of the Mediterranean Sea. Gray areas show the sub-regions selected in this study for a regional assessment: Alboran Sea (ALBS), Ligurian Sea (LIGS), Northern Adriatic Sea (NADS), and Levantine basin (LEVB).

PSC and PFT group to the satellite estimates of total Chl*a* concentration in the basin.

The Mediterranean (**Figure 1**) is the largest inland ocean basin on Earth, only connected to the rest of the world's oceans by the Strait of Gibraltar. It exhibits an oligotrophic regime (Krom et al., 1991), notwithstanding relatively external high inputs of essential nutrients (Ludwig et al., 2009; Huertas et al., 2012; Powley et al., 2016). Nevertheless, local physical structures generate convergences zones, which are reflected in the distinct biogeochemical properties of the two Mediterranean sub-regions, the Western and the Levantine. Thus, a notable decreasing gradient in Chla concentration is detected from the west to the east, which causes a significant longitudinal variation in primary production (Turley et al., 2000; Uitz et al., 2012). This gradient in oligotrophy is evidenced both by *in situ* measurements (Tanhua et al., 2013) and satellite data (D'Ortenzio and Ribera d'Alcalá, 2009). However, the seasonal evolution of Chla distribution still follows the typical succession of temperate regions, characterized by a phytoplankton biomass increase in late winter/early spring, a decrease during the summer season and a second smaller phytoplankton bloom in autumn (Siokou-Frangou et al., 2010; Sammartino et al., 2015).

Phytoplankton community structure in oligotrophic areas throughout the world's ocean is mainly composed by picoplankton and ultraplankton (Li et al., 1983; Brunet et al., 2006; Dandonneau et al., 2006). Nevertheless, the Mediterranean phytoplankton communities structures reveals a considerable variability over both temporal and spatial scales, and large dissimilarities in phytoplankton assemblage composition along with other microorganisms across the basin have been also highlighted (Siokou-Frangou et al., 2010). Many studies have pointed to the dominance of picoplankton as the fingerprint of the Mediterranean Sea and its overriding oligotrophy, but the occurrence of regional phytoplankton blooms cause the coexistence of numerous microalgal groups (Siokou-Frangou et al., 2010).

The satellite empirical model applied by Sammartino et al. (2015) encompassed this unusual and complex community structure in the Mediterranean Sea and allowed assessment of the spatio-temporal variability of the three phytoplankton size classes (micro-, nano-, and pico-plankton) during the entire SeaWiFS era (1998–2010). Previously, Navarro et al. (2014) had redefined the PHYSAT algorithm (Alvain et al., 2005, 2008) to the Mediterranean Sea's bio-optical characteristics to estimate the dominant functional phytoplankton types (*Prochlorococcus, Synechococcus*, diatoms, nanoeukaryotes, coccolithophorids, and *Phaeocystis*-like) from the MODIS sensor. More recently, Di Cicco et al. (2017) developed a new regional algorithm for satellite biomass estimates of PSC and PFT in Mediterranean Sea and assessed their accuracy with respect to global models, improving the uncertainty and the spread of about one order of magnitude for all phytoplankton classes.

Regarding the distribution of chlorophyll, low values (less 0.2 mg/m^3) are found over vast areas of the basin, with the exception of large blooms observed in late winter and early spring in the North Western Mediterranean (Siokou-Frangou et al., 2010). Mesoscale activity also increases the chlorophyll concentration mainly in the Alboran Sea, Balearic-Catalan Sea, Adriatic Sea and the South Eastern Levantine Sea, by about one order of magnitude for all phytoplankton classes. In other coastal areas close to major rives, such us the Po in the North Adriatic Sea, the Rhone in Gulf of Lions and the Nile in the Levantine Sea, and river discharge generates a large increase in chlorophyll levels (Siokou-Frangou et al., 2010). In the eastern basin, the Chl*a* rarely exceeds 0.5 mg/m^3, with minima as low as 0.003 mg/m^3 (Siokou-Frangou et al., 2010). Low biomass values are generally associated with the dominance of cyanobacteria, prochlorophytes and picoplankton-sized flagellates (Siokou-Frangou et al., 2010 and references therein), and represents 59% of the total Chl*a* and 65% of the primary production. However, nanoflagellates are the dominant group in terms of cell numbers throughout most of the year in the Mediterranean Sea. Finally, observed increases in

Chl*a* correlate with decreases in the contribution of picoplankton and nanoplankton, and increases in diatom concentration during February and March (Siokou-Frangou et al., 2010).

Bracher et al. (2017) recently highlighted the limited applicability of global satellite algorithms to determine the composition of phytoplankton at a regional scale as one of the major gaps in satellite research. Accordingly, these authors suggest a roadmap for future developments in regionally adapted algorithms. The main goal of this paper is thus to bridge the gap diagnosed by Bracher et al. (2017) by updating and improving the original version of the PHYSAT-Med method (Navarro et al., 2014) with the new OC-CCI database. Firstly, the advantages of this new version are: (a) an increase in the temporal range (1997–2015; (b) a decrease of the cloud cover due to the use of several ocean color sensors; (c) an improvement in the bias correction, thus reducing sensitivity to medium-term changes; and (d) the validation of the temporal range in the NW Mediterranean Sea using diagnostic pigments analysis (DPA, Vidussi et al., 2001). Secondly, wavelet analysis is applied to the retuned version in order to analyse the contributions of different temporal cycles of dominance variability of the major phytoplankton groups in the Mediterranean Sea.

MATERIALS AND METHODS

PHYSAT-Med OC-CCI Algorithm

The OC-CCI is a long-term, consistent and error-characterized dataset generated from merged normalized remote-sensing reflectance derived from four satellite sensors: SeaWiFS, MODIS, MERIS, and VIIRS (Storm et al., 2013; Jackson et al., in press). In this work, we have used OC-CCI v3.0, where more data have been included (VIIRS and SeaWiFS LAC). In addition, the bias correction has been improved, reducing sensitivity to medium-term changes and extending the method to work beyond the lifetime of SeaWiFS. Daily level 3 remote sensing reflectance data (Rrs) at 412, 443, 490, 510, 555, and 670 nm and diffuse attenuation coefficient (Kd_{490}) were downloaded from the OC-CCI website covering the period from January 1998 to December 2015 (**Figure 2**, step 1). These products were displayed on a regular 4 km grid, with an equi-rectangular projection with constant longitude and latitude steps. Error specification (RMSE and bias) is based on comparison with match-up *in situ* data and extrapolation to global scale ocean.

In a second step (**Figure 2**, step 2), the Chl*a* concentration in the Mediterranean Sea was calculated using a regional algorithm (MedOC4, Mediterranean ocean color four-bands, Volpe et al., 2007) developed for the basin for SeaWiFS bands (or CCI),

$$MedOC4 - Chla = 10^{(0.4424 - 3.686R + 1.076R^2 + 1.684R^3 - 1.437R^4)} \quad (1)$$

where

$$R = log_{10}\left[MAX\left(Rrs_{555}^{443}, Rrs_{555}^{490}, Rrs_{555}^{510}\right)\right] \quad (2)$$

This bio-optical algorithm is based on a fourth-power polynomial regression between log-transformed Chl*a* and log-transformed maximum band ratio (MBR). It is known that using multiple Rrs ratios decreases the noise-to-signal ratio, thereby enhancing the algorithm's performance (O'Reilly et al., 1998). The MedOC4 algorithm was calibrated on a representative open-water bio-optical dataset collected in the Mediterranean Sea, and is the best algorithm matching the requirement of unbiased satellite orophyll estimates (Volpe et al., 2007; Santoleri et al., 2008). At a global scale, the SeaWiFS algorithms have shown errors in the range of <5% for radiances and <35% for chlorophyll (Mueller and Austin, 1995; Gregg and Casey, 2004). The accuracy limit for chlorophyll using these standard algorithms has been shown to be unrealistic in Mediterranean Sea, yielding a severe overestimation (>70% for chlorophyll <0.2 mg/m^3; Volpe et al., 2007, 2012).

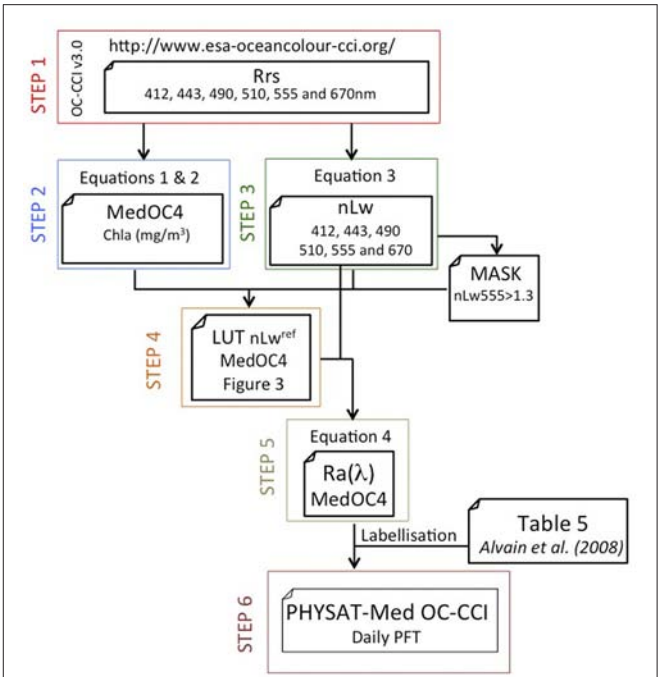


FIGURE 2 | Schematic view of steps of the PHYSAT-Med OC-CCI algorithm.

At the third step (**Figure 2**), the Rrs was converted to nLw using the nominal band solar irradiance (Fo, in mW cm^{-2} μm^{-1}) for any specific spectral band (λ) of the SeaWiFS sensor (Gregg et al., 1993; Thuillier et al., 2003).

$$nLw_{(\lambda)} = Rrs_{(\lambda)} * F_{o(\lambda)} \quad (3)$$

During step 4 (**Figure 2**), a new Look-Up-Table (LUT, **Figure 3** and Table 1 in Supplementary Material) of $nLw^{ref}(\lambda, Chla)$ was empirically generated for the Mediterranean Sea from a large dataset of OC-CCI Chl*a* and nLw pixels for all daily images contained within the study period (January 1998 to December 2015). Turbid pixels (defined as nLw555>1.3 mW cm^{-2} mm^{-1} sr^{-1}, Nezlin and DiGiacomo, 2005) were excluded in order to minimize the impact of high-suspended matter loads. Briefly, nLw^{ref} is calculated from nLw data, and the associated Chl*a* computed from the MedOC4 algorithm within the concentration

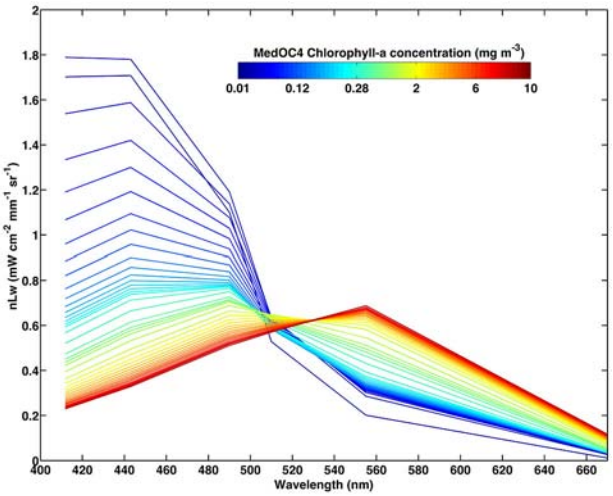


FIGURE 3 | Look-Up-Table for PHYSAT-Med OC-CCI algorithm. Normalized water-leaving radiance nLw as a function of wavelength for various MedOC4-Chla concentration (color scale, in mg/m^3) for the Mediterranean Sea (excluding Black Sea) during the study period (January 1998–December 2015).

range between 0.01 and 10 mg/m^3 (41 narrow intervals). This figure is similar to the one used in the development of the PHYSAT-Med algorithm (Navarro et al., 2014).

Once the new LUT (**Figure 3**, Table 1 in Supplementary Material) for the Mediterranean Sea was calculated using the regional MedOC4-Chl*a* algorithm, the radiance anomalies [Ra(λ), see **Figure 3**, step 5] were computed for all daily OC-CCI wavelengths analyzed using Equation (4) for all available wavelengths (412, 443, 490, 510, 555, and 670 nm). Ra(λ) is an adimensional parameter independent of the Chl*a* level, and hence also independent of the biomass. Ra(λ) thus represents the second order variation in nLw(λ) after removal of the first order effect of the Chl*a* variation (Alvain et al., 2005):

$$Ra_{(\lambda)} = {}^{nLw_{(\lambda)}}\!/_{nLw^{ref}_{(\lambda)}} \quad (4)$$

The analyses by Alvain et al. (2005) showed that for a given Chl*a* concentration, the particle scattering variability explains the largest fraction of the remotely sensed Ra spectral variability, especially when focusing on Ra magnitude changes. The labellization step was performed using the thresholds of Ra for each of the six phytoplankton groups examined in PHYSAT-v2008 (see Table 5 in Alvain et al., 2008), which is specifically set up for SeaWiFS channels. These thresholds were used to process daily images to calculate daily PFTs map (**Figure 3**, step 6). For a spectrum to be associated with one group, all criteria must be fulfilled. Thresholds (Table 5 in Alvain et al., 2008) were fixed in order to avoid any overlapping. Pixels with nLw values that were not classified for any phytoplankton groups were cataloged as "unidentified (unid.)," and this can sum up a significant fraction (Navarro et al., 2014). PHYSAT-Med retrieves the dominant group for a given satellite image pixel (4 km) for Mediterranean Sea, where a given phytoplankton group is the major contributor to the radiance anomaly. From this database (near to 6,600 daily images), 10-day and monthly maps of dominant phytoplankton groups were obtained by calculating the phytoplankton group that was present more days during the integration period (10-day or monthly, respectively) at each geographical pixel, not including "unidentified" pixels. To estimate the proportion of each phytoplankton group in the entire basin and several Mediterranean sub-regions (Alboran Sea, Ligurian Sea, Northern Adriatic Sea, and Levantine basin in **Figure 1**, Bricaud et al., 2002), the number of the pixels of each PFT during 10-day or 1 month was calculated for each area in proportion to all the identified pixels, excluding the unidentified pixels. The box plot figures were created using Matlab® software (boxplot.m function), where the central mark corresponds to the median, the edges of the box are the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers, and outliers are plotted individually. These statistics were calculated for each climatological month by considering all the values obtained during a particular month for the whole data series (from 1998 to 2015).

In situ Validation

In order to validate the PHYSAT-Med output, we have compared the temporal variability of PFT obtained from remote sensing algorithm with diagnostic pigments analysis (DPA, Vidussi et al., 2001) obtained from HPLC in North-Western Mediterranean Sea. A total of 5,400 samples were collected from the basin and analyzed by HPLC (**Figure 1**). This dataset consisted of samples from the DYFAMED (Dynamics of Atmospheric Fluxes in the Mediterranean Sea) time series included in the MAREDAT global database of HPLC (Peloquin et al., 2013) and the BOUSSOLE (Buoy for the acquisition of a Long-Term Optical Time Series) program (i.e., Antoine et al., 2006). Details of HPLC methods used can be found in the aforementioned references. Here, we only considered samples limited to the first optical depth (Z_{90}), which reduces the number of available pigment inventories to 1,615 samples and comprising the temporal range analyzed (1998–2015). The first optical depth was calculated using daily Kd_{490} images from OC-CCI data [$Z_{90} = 1/Kd_{490}$], which is about 15–35 m on average in the Mediterranean Sea (D'Ortenzio and Ribera d'Alcalá, 2009). The OC-CCI Kd_{490} product is computed from the inherent optical properties (IOPs) at 490 nm (Lee et al., 2005; Grant et al., 2016).

This comparison is based on many pigments specific to individual phytoplankton taxa or groups (i.e., Gieskes et al., 1988; Goericke and Repeta, 1993; Claustre and Marty, 1995; Jeffrey and Vesk, 1997). For instance, Divinyl Chlorophyll-a (dChl*a*) is a typical marker of prochlorophytes (Goericke and Repeta, 1992; Claustre and Marty, 1995; Vidussi et al., 2001), whereas zeaxanthin (Zeax) is associated with cyanobacteria (Guillard et al., 1985). Fucoxanthin (Fuco) pigment is the principal marker of diatoms (Jeffrey, 1980). For nanoplankton cuantification, Vidussi et al. (2001) used three diagnostic pigments: alloxanthin (Allox), that is a pigment typical of the cryptomonads (Gieskes and Kraay, 1983); 19′-hexanoyloxyfucoxanthin (HexFuco), whose concentration is related to prymnesiophytes and

chromophytes nanoflagellates (Wright and Jeffrey, 1987); and 19′-butanoyloxyfucoxanthin (ButFuco), a typical marker of chromophytes nanoflagellates (Wright and Jeffrey, 1987). Other pigments used in this method are total chlorophyll-b (TChlb, chlorophyll b + Divinyl-Chlorophyll b) and peridinin (Perid), which appears in small dinoflagellates (Jeffrey and Hallegraeff, 1987). This approach has been used at global scale (Uitz et al., 2006) and particularly in the Mediterranean Sea (i.e., Vidussi et al., 2001; Marty et al., 2002; Sammartino et al., 2015; Di Cicco et al., 2017; Mayot et al., 2017).

In this study, we have compared the *in situ* chlorophyll concentration of nanoplankton and diatoms using the method recently applied to the Mediterranean Sea by Di Cicco et al. (2017). Following the DPA procedure, originally proposed by Vidussi et al. (2001) and later refined by Uitz et al. (2006) to scale diagnostic pigments to Chl*a*, it is possible to apply DPA-based approaches to satellite-derived Chl*a*:

$$Chla_{diatoms} = \left(\frac{1.60\ [Fuco]}{\Sigma DPW}\right)\ x\ Chla \tag{5}$$

$$Chla_{nano} = \left(\frac{1.18\,[HexFuco] + 0.57\ [ButFuco] + 2.70\ \left[Allo\right]}{\Sigma DPW}\right) x\ \ Chla \text{ if TChla} > 0.08\ mg/m^3 \tag{6}$$

$$Chla_{nano} = \left(\frac{\begin{array}{c}12.5\,[TChla]\\ +1.18\,[HexFuco] + 0.57\ [ButFuco] + 2.70\ \left[Allo\right]\end{array}}{\Sigma DPW}\right) x\ \ Chla \text{ if TChla} < 0.08\ mg/m^3 \tag{7}$$

where

$$\Sigma DPW = 1.18\,[HexFuco] + 0.57\ [ButFuco] + 2.70\ \left[Allo\right] + 1.67\left[Perid\right] + 1.60\,[Fuco] + 0.88\ \left[TChlb\right] + 1.79\,[Zeax] \tag{8}$$

Alternatively, the method described by Alvain et al. (2005) was used to compare *Prochlorococcus* and *Synechococcus* time series, as this method uses the pigment ratio of divinyl-Chl*a* and zeaxanthin, respectively, and has been applied previously for the Mediterranean Sea (Navarro et al., 2014):

$$P_{rel} = {}^{P}\!\!\diagup_{(Chla\,+\,dChla)} \tag{9}$$

where P is the measured pigment concentration (dChl*a* or zeaxhantin), and Chl*a* and dChl*a* are the concentrations of chlorophyll-*a* and divinyl Chlorophyll-*a*, respectively.

As HPLC and PHYSAT-Med OC-CCI output data are measured in different units, we used Spearman's rank-order (non-parametric) correlations to assess the strength of the temporal association between both variables for each of the phytoplankton functional types. We used 10,000 bootstrap samples to construct 90% empirical confidence intervals for the correlations (Efron and Tibshirani, 1993). In order to check for possible time lags between the timing of seasonal blooms within the Ligurian Sea, as measured with both methods, we inspected the effect of different time lags in both variables on the strength of the association.

Wavelet Analysis

Wavelet analysis has emerged as a tool for characterizing periodicities in non-stationary time series, as it decomposes a time series both in the frequency and time domains (Percival and Walden, 2000). In this study, wavelet analysis has been used to characterize the different periodic components of the variability in dominance of the major phytoplankton groups in the Mediterranean Sea across time. Wavelet analysis performs a time-scale decomposition of the signal by estimating its frequency characteristics as a function of time (Torrence and Compo, 1998; Grinsted et al., 2004; Winder and Cloern, 2010). In order to normalize time series data and obtain the wavelet power spectrum of the different phytoplankton groups, the continuous Morlet wavelet transform was applied by using the Matlab® toolbox provided by Torrence and Compo (1998) and Grinsted et al. (2004) (http://atoc.Colorado.edu/research/wavelets/). The wavelet power spectrum identifies the periods that are the most important sources of variability across time. Additionally, it is possible to define a global wavelet spectrum, which identifies the variance associated to each period for a given time series, and is similar to Fourier spectra (Percival and Walden, 2000). Wavelet analysis was performed over the 10-day times series of nanoeukaryotes, *Prochlorococcus, Synechococcus*-like cyanobacteria and diatoms for the entire Mediterranean Sea and for four selected sub-regions (Alboran Sea, Liguarian Sea, Northern Adriatic Sea and Levantine sea).

RESULTS AND DISCUSSION

Validation

Figure 4 shows the temporal variability of several diagnostic pigments and PFT in the Ligurian Sea, where DYFAMED and BOUSSOLE sampling stations are located. The comparison exercise covers all range of the PFT analyzed, except for zeaxanthin pigment for which there were no values during 1998. *Prochlorococcus* showed maximum values in autumn over several years, at the end of the stratification period (**Figure 4A**). The maxima found by the PHYSAT-Med OC-CCI algorithm were in close agreement with the maxima in the concentrations of the pigment ratio for dChl*a* measured by HPLC method, which is indicative of prochlorophytes (Goericke and Repeta, 1992; Claustre and Marty, 1995; Vidussi et al., 2001). These results agree with the pattern reported by Vaulot et al. (1990) and Marty et al. (2002) estimated by flow cytometry and HPLC analysis, respectively. During summer, also coinciding with the stratification period, the dominant group is *Synechococcus*, and a maximum in zeaxanthin concentration is observed across the basin all years (**Figure 4B**). This pigment is associated with cyanobacteria (Guillard et al., 1985) and has been widely used to estimate *Synechococcus* concentration in the Ligurian Sea (Vidussi et al., 2001; Marty et al., 2002). However, during the spring bloom period, when the mixed layer depth is at its maximum (Marty et al., 2002), the diatom group and diatom Chl*a* concentration also reached its highest value (**Figure 4C**). Finally, the nanoeukaryotes distribution presented maxima during winter, normally around January, coinciding with the maximum of nanoplankton chlorophyll concentration estimated

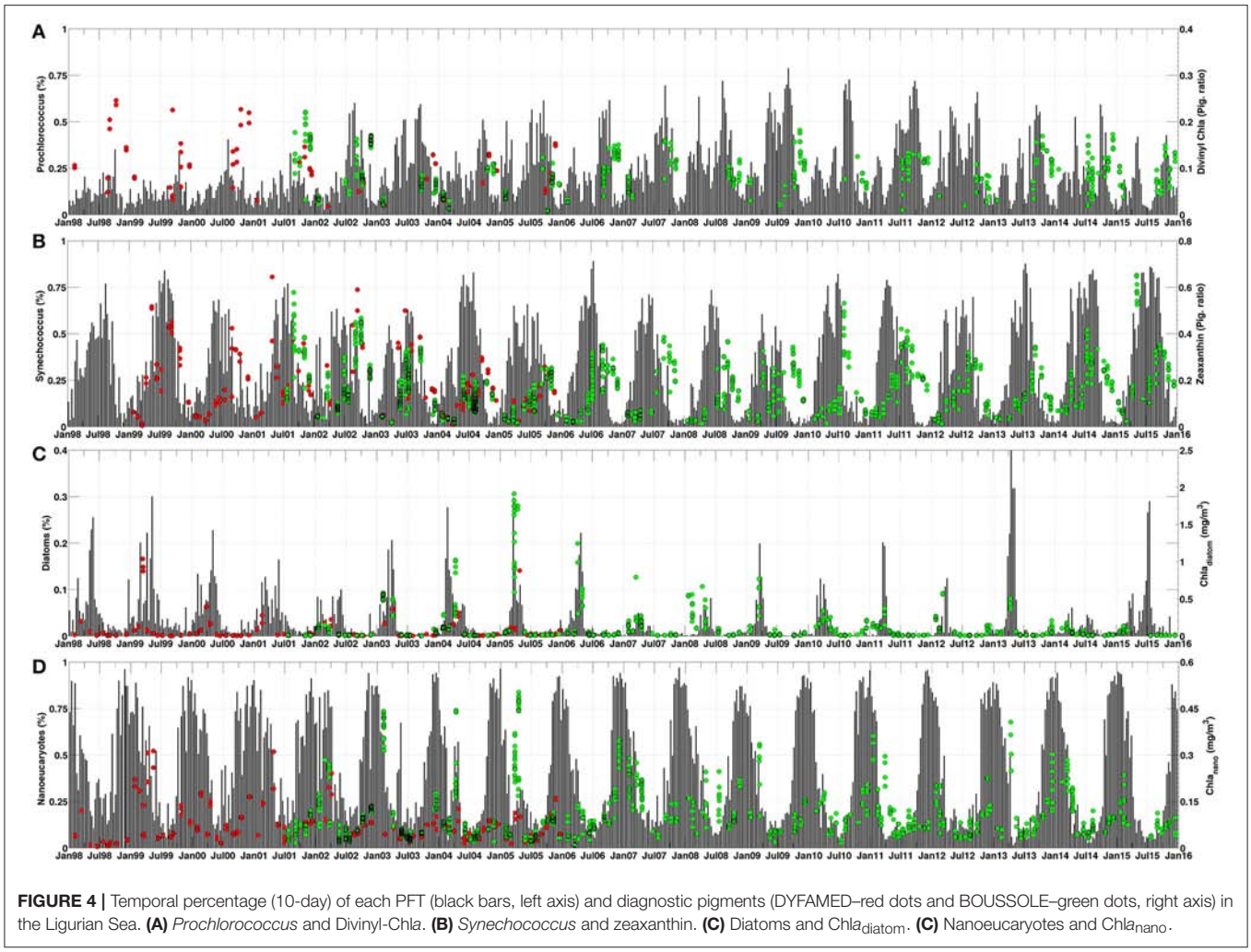


FIGURE 4 | Temporal percentage (10-day) of each PFT (black bars, left axis) and diagnostic pigments (DYFAMED–red dots and BOUSSOLE–green dots, right axis) in the Ligurian Sea. **(A)** *Prochlorococcus* and Divinyl-Chl*a*. **(B)** *Synechococcus* and zeaxanthin. **(C)** Diatoms and Chla_{diatom}. **(C)** Nanoeucaryotes and Chla_{nano}.

using Equations 6 and 7 (**Figure 4D**; Di Cicco et al., 2017). Overall, the broad coincidence between PHYSAT-Med outputs and HPLC pigments in the temporal pattern suggests that the new version of PHYSAT-Med algorithm using OC-CCI v3.0 database is in agreement with the results obtained through long-term monitoring programs for phytoplankton distribution, at least in the Ligurian Sea area.

Even though the OC-CCI database provided per-pixel errors (RMSE and bias) for all OC-CCI products, this approach is not so common in retrieving phytoplankton functional types, except for the recent works published by Brewin et al. (2017) and Di Cicco et al. (2017) for the Mediterranean Sea. In fact, Di Cicco et al. (2017) showed the improvements obtained from the use of regional models with respect to the global models, with the reduction of bias being of about one order of magnitude. As we described above, the PHYSAT algorithm allows for the detection of dominant PFT. This approach is based on the analysis of the second order variation in nLw measurements after removal of the impact of Chl*a* variation. Alvain et al. (2012) found acceptable results for diatoms (73%) and nanoeucaryotes (82%), but relatively low for *Prochlorococcus* and cyanobacteria (61 and 57% of successful identification, respectively). For PHYSAT-Med, Navarro et al. (2014) found similar results for *Synechococcus* and nanoeucaryotes (61 and 74%, respectively).

Table 1 shows the results of the validation exercise. For both nanoeucaryotes and diatoms the correlation between the PHYSAT-Med OC-CCI and the HPLC data is relatively large, with narrow bootstrapped confidence intervals non-overlapping 0. For *Synechococcus* the correlation is weaker but, again, the confidence interval does not contain 0. In contrast, for *Prochlorococcus* the association between PHYSAT-Med and HPLC data is weaker, and now the 0 is included within the confidence interval. In this later case, the sample size is clearly lower. The reason might be that the largest *Prochlorococcus* abundance is located near the deep chlorophyll maximum, deeper than the first optical depth (Siokou-Frangou et al., 2010).

Spatio-Temporal Patterns at the Basin Scale

Figure 5 shows the monthly climatology (1998–2015) of the dominant phytoplankton groups in the Mediterranean Sea

TABLE 1 | Validation results of the PHYSAT-Med dataset with HPLC data from the Ligurian Sea.

Functional type	Sample size	r_S	90% B.C.I.
Nanoeucaryotes	231	0.426	0.339 – 0.509
Diatoms	228	0.397	0.299 – 0.491
Synechococcus	234	0.112	0.019 – 0.204
Prochlorococcus	115	−0.133	−0.272 – 0.008

Shown are the sample size of the time series used, the Spearman's rank-order correlation between both time series (r_s) and the 90% bootstrapped confidence interval. 10,000 bootstrapped samples of the original series were used to construct the intervals.

obtained with the OC-CCI database. These patterns are similar to those obtained by Navarro et al. (2014) using the PHYSAT-Med and the MODIS imagery for the period comprised from July 2002 to May 2013. In addition, the analysis of the PFT is consistent with the previous knowledge of this area (Siokou-Frangou et al., 2010; Uitz et al., 2012; Sammartino et al., 2015; Di Cicco et al., 2017). It is evident that *Synechococcus* is the most abundant group detected at the basin scale and particularly during spring-summer months, whereas nanoeukaryotes seem to dominate during autumn-winter months. *Prochlorococcus* is preferentially distinguished during February and October in offshore waters, in opposition to diatoms that prevail in coastal areas, such as the Gulf of Lions, the Ligurian Sea and the northern Adriatic Sea, and mostly during the spring season. This last finding agrees with the microplankton distribution provided by Sammartino et al. (2015) and Di Cicco et al. (2017) who concluded that the fraction of microplankton significantly increases in the Northwestern Mediterranean Sea, reaching values from 30 to 57% (Sammartino et al., 2015). The presence of coccolithophorids in the basin is particularly evident along the Mediterranean coastline and particularly in the surroundings of the large river mouths (Ebro, Rhone and Nile) and in the Adriatic Sea (**Figure 5**). However, it is worthy to highlight that fluvial inputs of terrestrial matter or suspended solids may slightly mask the signals and affect phytoplankton distribution (Navarro et al., 2014). Even though PHYSAT-Med is also appropriate for detecting *Phaeocystis*-like phytoplankton, no signal of this group was found in the current study. These spatio-temporal patterns were subsequently corroborated by the time series of monthly climatology (**Figures 6B**, **7B**, **8B**, and **9B**) although only nanoeukaryotes, *Prochlorococcus, Synechococcus*-like cyanobacteria and diatoms were considered because these groups were compared with *in situ* pigment markers (**Figure 4**).

The abundance of nanoeukaryotes in the Mediterranean Sea (**Figure 6A**) follows recurrent 12-months cycles across time, as suggested by the power of the wavelet spectrum at this cycle (**Figure 6C**). Some weaker 6-months cyclic components can be observed during certain particular years (2003, 2006–2014). The mean annual cycle or the monthly climatology for this group (**Figure 6B**) showed a maximum percentage of abundance during the winter months, mainly November, December and January. The global wavelet spectrum (**Figure 6D**) demonstrated that the 12-months periodicity was highly significant, with a minor peak at 6-months also contributing to the variance. The amplitude of the seasonal variations of nanoeukaryotes at a basin scale is similar to that described by Sammartino et al. (2015) and Di Cicco et al. (2017), who found minimum values of abundance for nanoplankton during summer and maximum during winter, when the mixed layer depth (MLD) is also deeper (Siokou-Frangou et al., 2010). In the Levantine basin this group was the second most abundant in terms of Chl*a*, but it was the main group in the western basin (Di Cicco et al., 2017). Nanoplankton make a dominant contribution (up to 43–50%) to total primary production throughout the year at the basin scale (Uitz et al., 2012).

For *Synechococcus* the wavelet power spectrum also revealed a persistent 12-months periodicity (**Figures 7C,D**), with virtually no secondary 6-months cycles. The higher percentages of *Synechococcus* abundance were effectively observed during the summer season (**Figures 4**, **5**, and **7B**), particularly in June and July, coinciding with the stratification period (Siokou-Frangou et al., 2010).

Interestingly, the temporal patterns of *Prochlorococcus* (**Figure 8A**) exhibited a less periodic fluctuation in dominance at the basin scale as compared to those of *Synechococcus* (**Figure 7A**).The continuous wavelet spectrum (**Figures 8B,C**) suggests that prior to 2002, no cyclic component dominated (the time series conformed to a white-noise process). From this year onwards, a 6-months periodicity pattern became apparent, particularly from 2007. An annual cycle also appeared during this period, but the power was smaller as suggested by the global wavelet spectrum (**Figure 8D**). This indicates that most of the temporal variance in the dominance of *Prochlorococcus* in the Mediterranean Sea occurs at different periodicities, perhaps dominated by a seasonal period. *Synechococcus* tends to be more abundant at the surface waters, whereas *Prochlorococcus* thrives mainly in the deep-chlorophyll maximum (Marty and Chiavérini, 2002; Casotti et al., 2003).

The temporal patterns of diatoms were characterized by a robust periodicity of 12 months across time (**Figure 9**). In this case, the largest dominance was observed during spring, in agreement with the diatom blooms reported in the basin over this season (Marty and Chiavérini, 2002). In contrast, the minima occurred in September, coinciding with nutrient exhaustion.

Spatio-Temporal Patterns at the Sub-Regional Scale

The PHYSAT-MedOC-CCI approach was also applied in four selected sub-regions in order to track the temporal evolution of phytoplankton groups at smaller spatial scales. To allow for meaningful comparisons, the chosen areas resemble those considered by Sammartino et al. (2015): the Alboran Sea, the Ligurian Sea, the North Adriatic Sea and the Levantine basin (**Figure 1**). **Figure 10** shows the monthly climatology in the percentage of dominance of the four phytoplankton groups for each sub-region. Nanoeukaryotes occurrence exhibited a marked

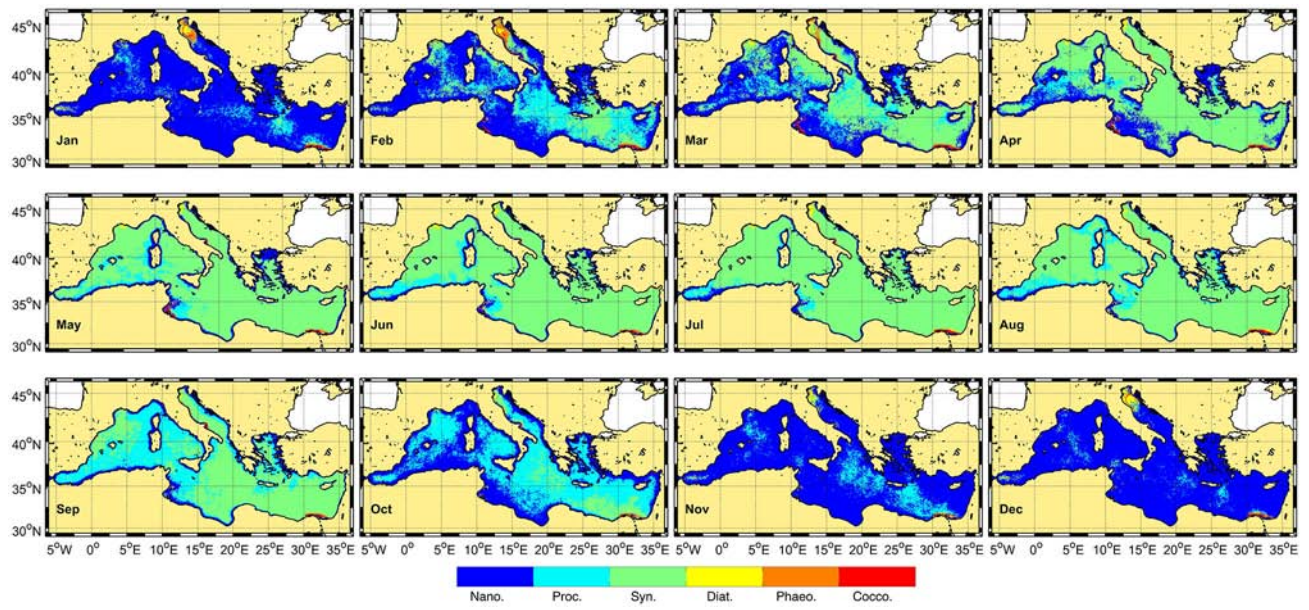


FIGURE 5 | Monthly climatology of the dominant phytoplankton groups detected by PHYSAT-Med OC-CCI during the study period (January 1998–December 2015).

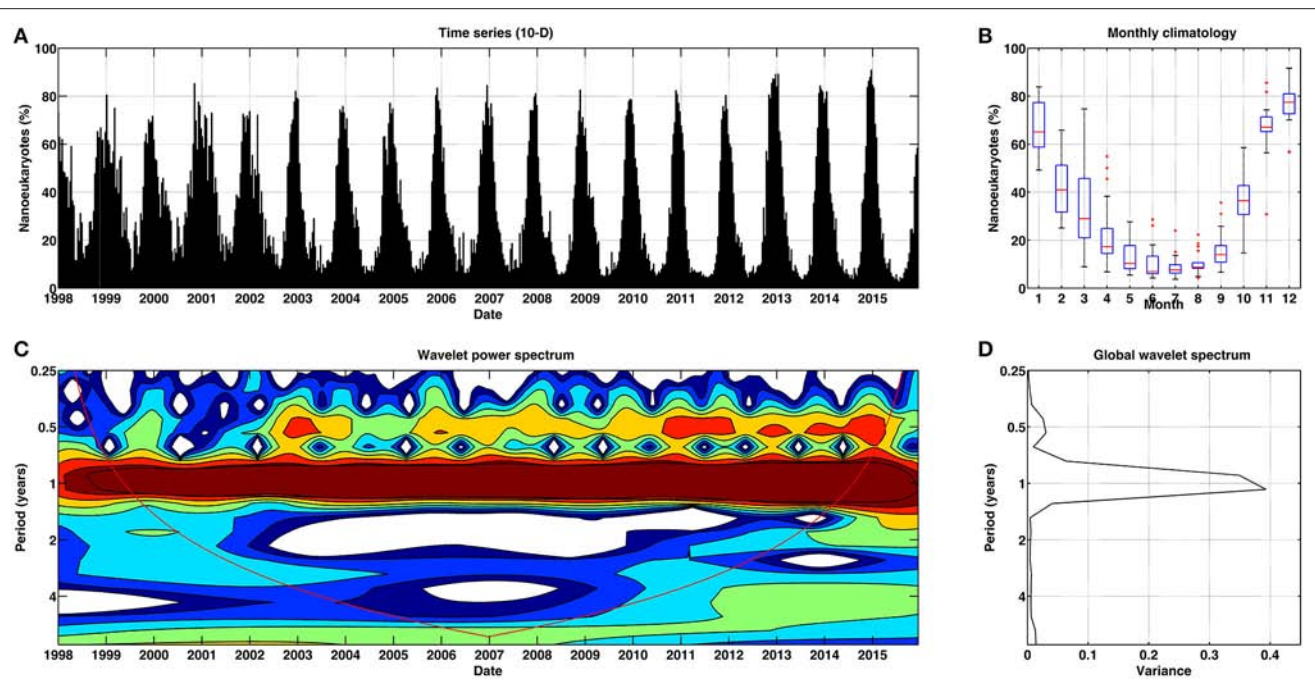


FIGURE 6 | Temporal patterns of variability for nanoeukaryotes for Mediterranean Sea. **(A)** Temporal percentage (10-day) of nanoeukaryotes. **(B)** Box plot (red lines stand for the median, blue box spans from the first to the second quartiles, and black lines represent 5th and 95th percentiles, respectively) of monthly climatology of nanoeukaryotes. **(C)** Continuous wavelet power spectrum for the 10-day time series. Red line indicates the cone of influence, and is the region affected by the edges of the data and should not be considered. **(D)** Global wavelet spectrum for the 10-day time series.

longitudinal gradient, with a higher abundance in the western basin (Alboran Sea and Ligurian Sea) in relation to the Levantine basin. Nonetheless, a marked annual cycle is evident in all sub-regions, which is characterized by the presence of maxima over the winter months and a minimum in summer. This was confirmed by a regional wavelet analysis that clearly revealed a consistent 12-month periodicity (**Figure 11**). Interestingly however, a longitudinal increase in the importance of 6-months

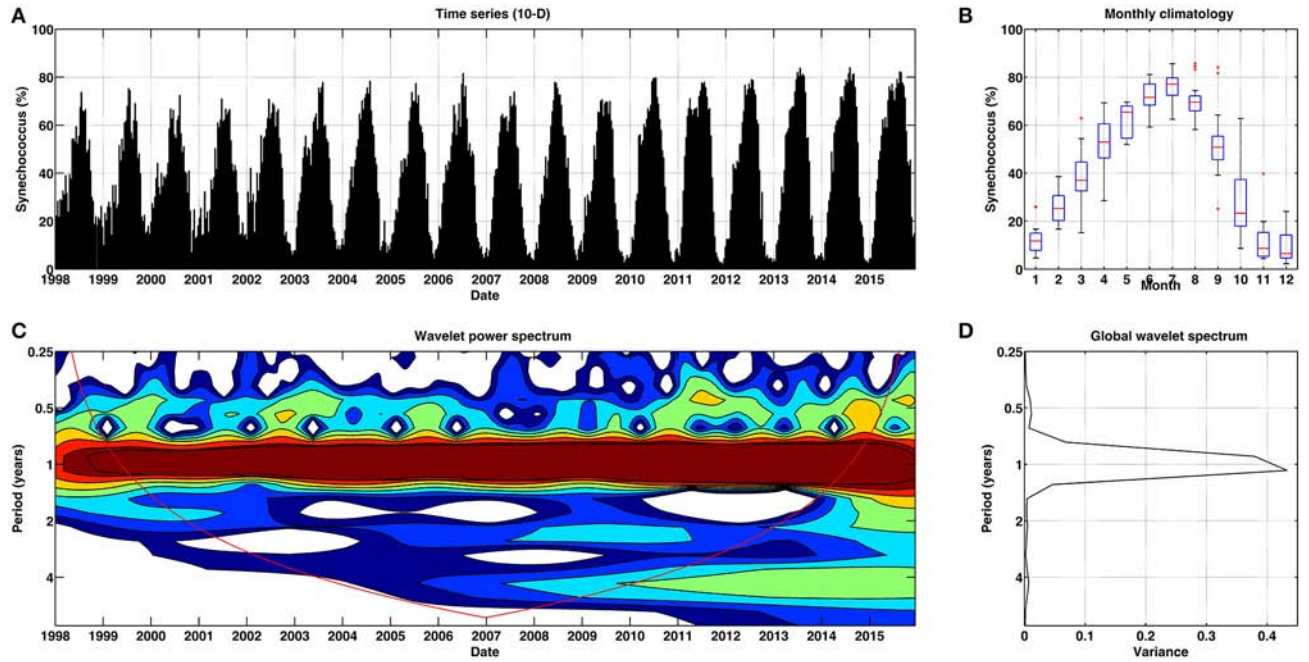


FIGURE 7 | Temporal patterns of variability for *Synechococcus* for Mediterranean Sea. **(A)** Temporal percentage (10-day) of *Synechococcus*. **(B)** Box plot (red lines stand for the median, blue box spans from the first to the second quartiles, and black lines represent 5th and 95th percentiles, respectively) of monthly climatology of *Synechococcus*. **(C)** Continuous wavelet power spectrum for the 10-day time series. Red line indicates the cone of influence, and is the region affected by the edges of the data and should not be considered. **(D)** Global wavelet spectrum for the 10-day time series.

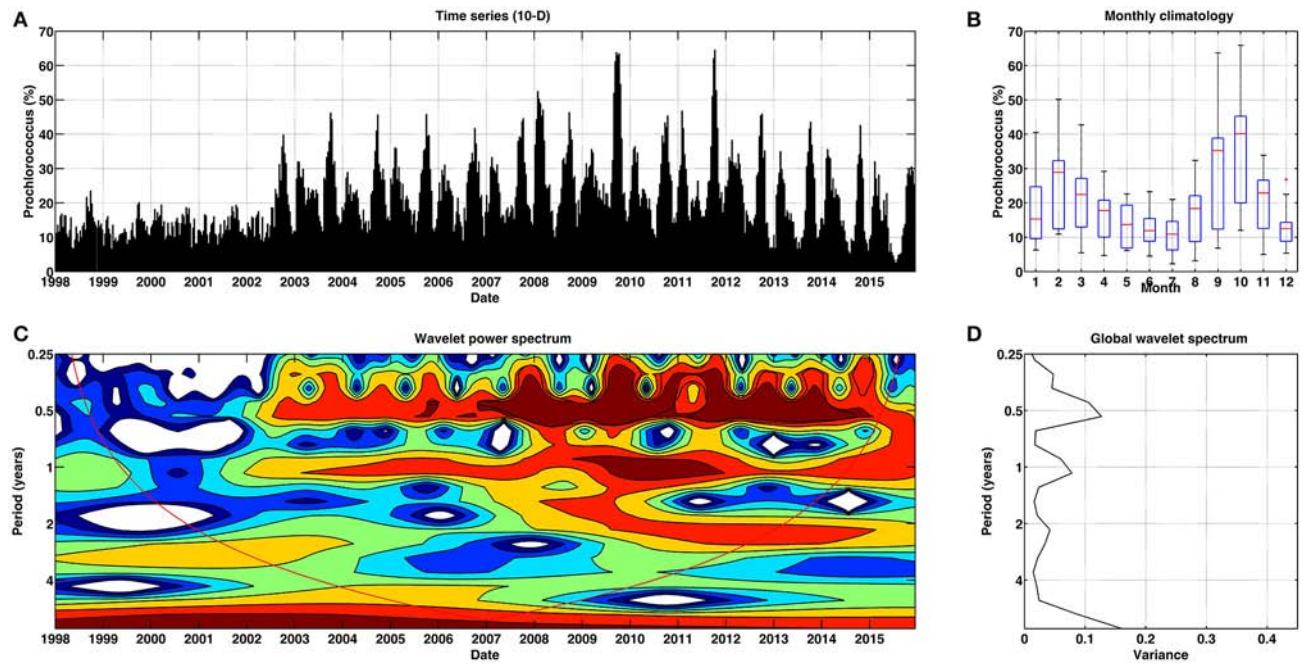


FIGURE 8 | Temporal patterns of variability for *Prochlorococcus* for Mediterranean Sea. **(A)** Temporal percentage (10-day) of *Prochlorococcus*. **(B)** Box plot (red lines stand for the median, blue box spans from the first to the second quartiles, and black lines represent 5th and 95th percentiles, respectively) of monthly climatology of *Prochlorococcus*. **(C)** Continuous wavelet power spectrum for the 10-day time series. Red line indicates the cone of influence, and is the region affected by the edges of the data and should not be considered. **(D)** Global wavelet spectrum for the 10-day time series.

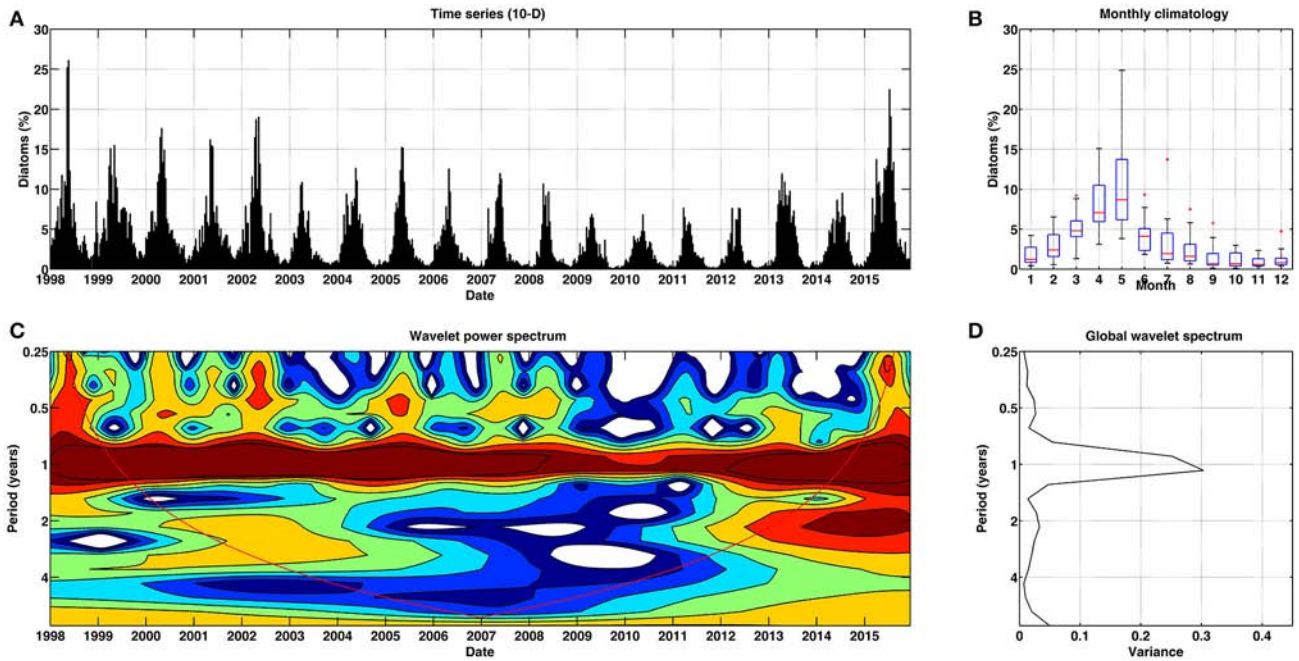


FIGURE 9 | Temporal patterns of variability for diatoms for Mediterranean Sea. **(A)** Temporal percentage (10-day) of diatoms. **(B)** Box plot (red lines stand for the median, blue box spans from the first to the second quartiles, and black lines represent 5th and 95th percentiles, respectively) of monthly climatology of diatoms. **(C)** Continuous wavelet power spectrum for the 10-day time series. Red line indicates the cone of influence, and is the region affected by the edges of the data and should not be considered. **(D)** Global wavelet spectrum for the 10-day time series.

periodicities also became apparent: both in the North Adriatic Sea and in the Levantine basin, recurrent seasonal periods contribute to the overall variability (**Figure 11**). Nevertheless, the overall abundance of this group kept values above 20%, with the exception of the Levantine basin over summer months. This temporal pattern resembles that of nanoplankton found in the selected areas by Sammartino et al. (2015), and reflects the constant contribution of this group to primary production, as previously reported (Vidussi et al., 2000, 2001; Uitz et al., 2012).

Prochlorococcus abundance in the four regions was higher over the late summer months and similar in terms of percentage between the Western and Eastern basins. This phytoplankton group is, however, less represented in the Northern Adriatic Sea. It is worthy to note that an additional winter peak (February) of *Prochlorochoccus* can be identified in the Levantine basin. This pattern of abundance at a sub-regional scale coincides with the wavelet analyses (**Figure 11**). The annual cycle of *Synechococcus* is also evidenced by the maximum of abundance in the four sub-regions during the summer months coinciding with the stratification period and when Chl*a* concentration in the basin is low (Volpe et al., 2007). During this season, the temporal climatology of *Synechococcus* reached values close to 100% in the Levantine basin, in agreement with the values given by Sammartino et al. (2015) for picoplankton. These authors indicated that this size class (closely corresponding to *Synechococcus*) seems to cover homogenously the entire Mediterranean Sea, with percentages of abundance close to 70%, although a decreasing concentration gradient from west to east can be still observed, which was also revealed by our analysis particularly for *Prochlorococcus*. According to our analysis (**Figure 10**), *Synechococcus* dominated in the Eastern basin, where ultraoligotrophic conditions are present and particularly during summer (Siokou-Frangou et al., 2010). During this season, primary production by the picoplankton exhibits a maximum (Uitz et al., 2012. It is well known that due to their high surface/volume ratio, *Synechococcus* (and also *Prochlorococcus*) can cope optimally with nutrients-impoverished environments (Le Quéré et al., 2005). The presence of *Synechococcus* in the Levantine basin has been widely reported (Uitz et al., 2012 and references therein), and the PHYSAT-Med OC-CCI clearly revealed its presence in the ultraoligotrophic Levantine basin and depicted a realistic annual cycle (**Figures 10**, **11**).

As expected, diatoms were the least abundant of the four phytoplankton groups analyzed in the Mediterranean sub-regions. In fact, with the exception of the Northern Adriatic Sea, the percentage of abundance of diatoms fell within the range of 10–20% in the Western sub-regions and in the Levantine basin during the whole year. A moderate spring maximum could be still detected, coinciding with the seasonal blooms normally described for this phytoplankton group along the Mediterranean (Marty and Chiavérini, 2002). As suggested by the wavelet analyses, the annual cycle for diatoms is rather less robust compared to other groups (**Figure 11**). These findings agree with previous studies, where higher

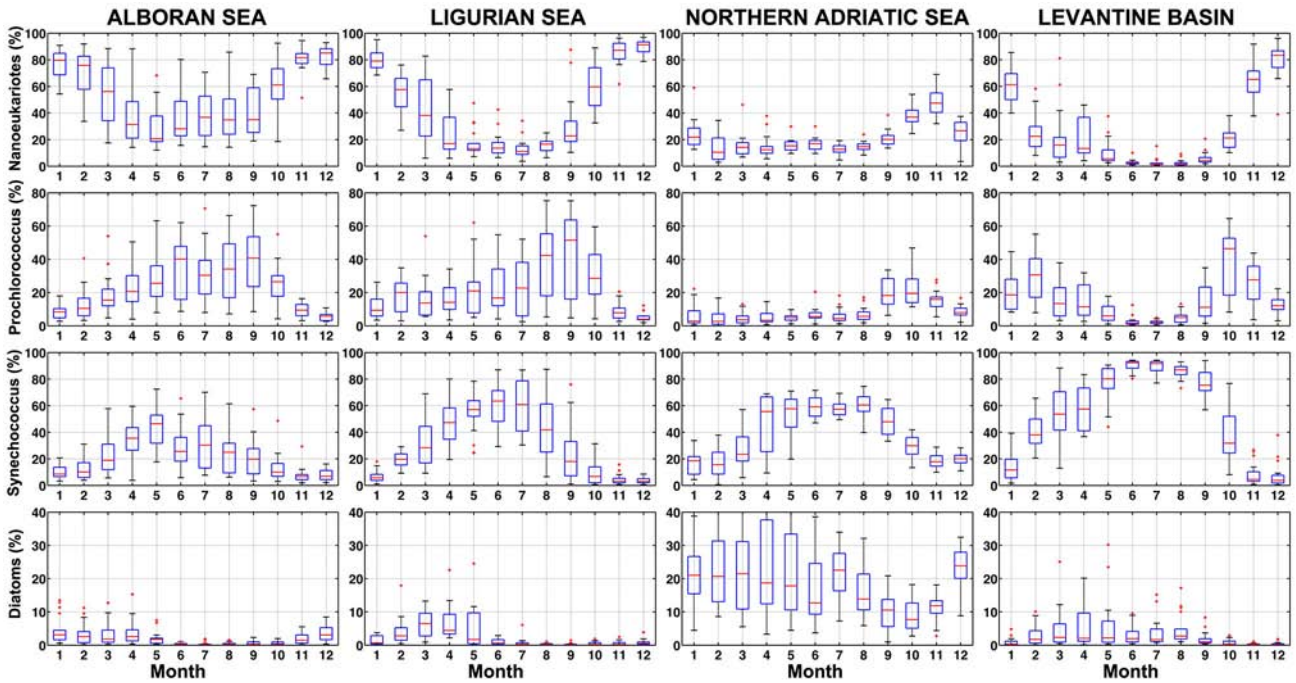


FIGURE 10 | Box plot for monthly climatology for each group in Mediterranean sub-regions. Each row represents each phytoplankton group (nanoeukaryotes, *Prochlorococcus*, *Synechococcus,* and diatoms) and each column corresponds to the same sub-region (Alboran Sea, Ligurian Sea, Northern Adriatic Sea, and Levantine Basin). In the box plot figure red lines stand for the median, blue box spans from the first to the second quartiles, and black lines represent 5th and 95th percentiles, respectively, of monthly climatology for each group.

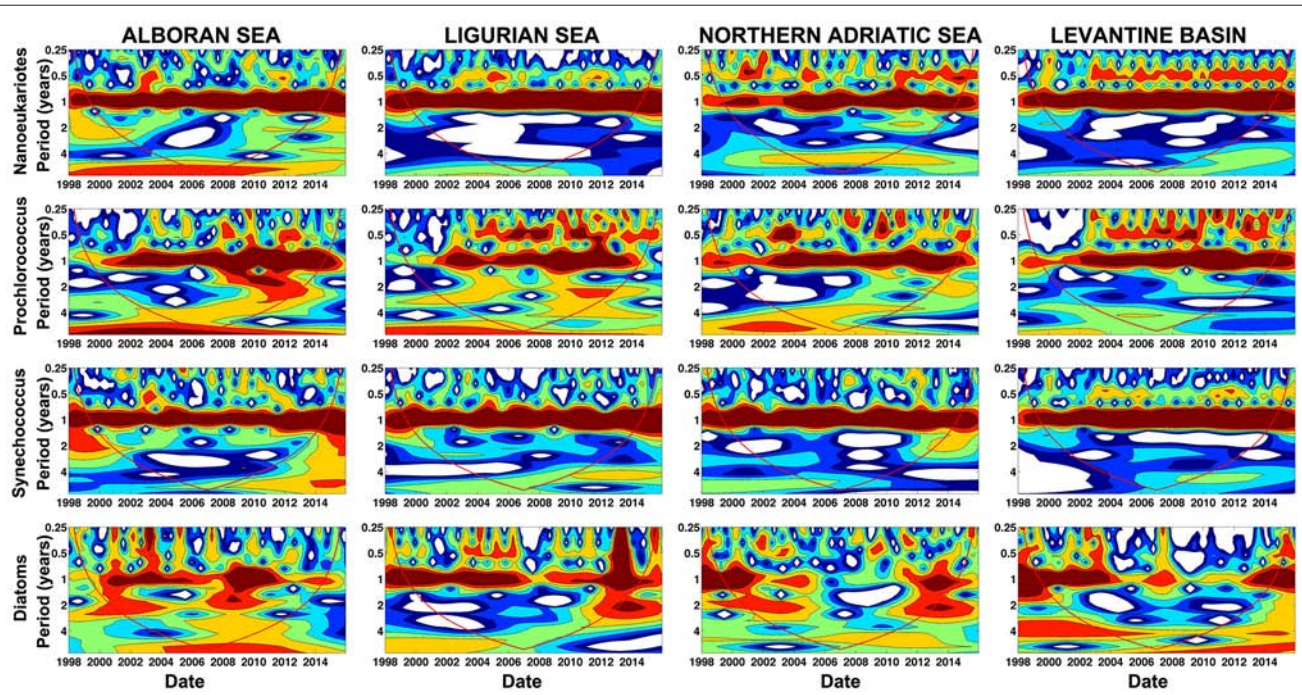


FIGURE 11 | Continuous wavelet power spectrum for the 10-day time series for each group and four sub-regions. Each row represents each phytoplankton group (nanoeukaryotes, *Prochlorococcus*, *Synechococcus,* and diatoms) and each column corresponds to the same sub-region (Alboran Sea, Ligurian Sea, Northern Adriatic Sea, and Levantine Basin). Red line indicates the region of time and frequency affected by the edges of the data and should not be considered.

phytoplankton biomass, particularly diatoms, was found in the Adriatic Sea (Socal et al., 1999; Casotti et al., 2003). This group represented an important faction (14%) of phytoplankton only in winter (Socal et al., 1999). Nevertheless, the contribution of picoplankton typically exceeds that of microplankton most of the year except during the winter-spring bloom (Uitz et al., 2012).

According to our assessment using the new OC-CCI database, the most abundant phytoplankton group during the winter months in the Mediterranean Sea was the nanoeukaryotes, but particularly in the sub-regions of the Western Mediterranean, the Alboran and the Ligurian Seas. Regardless of a specific area, the convective mixing in winter over the basin that uplifts deep nutrients to the upper layer triggers the proliferation of the bigger phytoplankton groups (Marty and Chiavérini, 2010), which is evidenced here by the major presence of nanoeukaryotes in winter months (**Figures 5**, **10**). Our data also reproduce the basin-wide diatom peaks of abundance in spring (**Figure 9**), following the Mediterranean phytoplankton succession previously reported (Marty et al., 2002) and with a consistent pattern every year (**Figures 9–11**). It should be indicated, however, that our approach presents certain limitations in diatoms detection, as acknowledged by Navarro et al. (2014). Hence, it is likely that diatoms abundance may have been slightly underestimated across the basin although studies on extensive distribution of this group along the Mediterranean are scarce to allow for an accurate comparison. Conversely, *Synechococcus* and *Prochlorococcus*, which are more favored by stratification conditions during summer due to their better efficiency under nutrient depleted conditions, were successfully identified both at the basin (**Figure 5**) and sub-regional scales (**Figure 10**). Moreover, the well-known dominance of *Synechococcus* with respect to *Prochlorococcus* (Schauer et al., 2003), particularly in the ultraoligotrophic Eastern basin, was neatly reproduced in our study.

Overall, the spatio-temporal patterns obtained by applying the PHYSAT-Med to the satellite OC-CCI database are consistent with the previous distributions of the major phytoplankton groups observed in the Mediterranean Sea (Vidussi et al., 2000, 2001; Marty and Chiavérini, 2002, 2010; Marty et al., 2002; Siokou-Frangou et al., 2010; Navarro et al., 2014; Sammartino et al., 2015; Di Cicco et al., 2017). These results suggest that our approach is highly suitable at basin scale and in selected sub-regions. This new dataset for PFT could be an efficient tool for recording and understanding the response of the marine ecosystem to human pressures and thus for detecting eutrophication in the Mediterranean Sea (Vantrepotte and Mélin, 2010; Colella et al., 2016).

CONCLUSIONS

This work presents an updated version of the PHYSAT-Med algorithm that has been specifically developed using the OC-CCI database. This ESA initiative aims at gathering ocean color measurements from four sensors since 1997. The distribution of the major phytoplankton groups in the Mediterranean basin during a 18 years period was consistent with the previous knowledge on the distribution patterns of phytoplankton in the basin. In addition, the modeled distributions are in concordance with the distribution of HPLC pigments analyzed in NW Mediterranean Sea for the whole temporal range. The utility of the updated approach was confirmed by the temporal analysis using the wavelet spectrum, which allowed for the identification of shifting patterns of periodicities across time for the dominant phytoplankton groups. Therefore, the new version of the PHYSAT-Med is appropriate for assessing the shifting spatio-temporal patterns of the most abundant phytoplankton groups in the Mediterranean Sea.

AUTHOR CONTRIBUTIONS

All authors contributed to the final manuscript: GN and IH were responsible for writing and organizing the manuscript. IC and GN implemented the PHYSAT-Med algorithm to new OC-CCI database. PA and AV performed the wavelet analysis. All authors discussed the results and drew the conclusions.

FUNDING

This work was financially supported by the Junta de Andalucía Projects PR11-RNM-7722, PIE 201530I012 and the National Project CTM2014-58181-R.

ACKNOWLEDGMENTS

The authors acknowledge the Ocean Colour Climate Change Initiative dataset, Version 3.0, European Space Agency, available online at http://www.esa-oceancolour-cci.org/, for providing access to remote sensing reflectance products. *In situ* HPLC dataset was obtained from BOUSSOLE Project and MAREDAT database. We are grateful to two reviewers for their invaluable scientific improvements.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmars.2017.00246/full#supplementary-material

REFERENCES

Aiken, J., Pradhan, Y., Barlow, R., Lavender, S., Poulton, A., Holligan, P., et al. (2009). Phytoplankton pigments and functional types in the Atlantic Ocean: a decadal assessment, 1995–2005. *AMT Spec. Issue. Deep-Sea Res. II* 56, 899–917. doi: 10.1016/j.dsr2.2008.09.017

Alvain, S., Le Quéré, C., Bopp, L., Racault, M. F., Beaugrand, G., Dessailly, D., et al. (2013). Rapid climatic driven shifts of diatoms at high latitudes. *Remote Sens. Environ.* 132, 195–201. doi: 10.1016/j.rse.2013.01.014

Alvain, S., Loisel, H., and Dessailly, D. (2012). Theoretical analysis of ocean color radiances anomalies and implications for phytoplankton groups detection in case 1 waters, *Opt. Express* 20, 1070–1083. doi: 10.1364/OE.20.001070

Alvain, S., Moulin, C., and Dandonneau, Y. (2008). Seasonal distribution and succession of dominant phytoplankton groups in the global ocean: a satellite view (SeaWiFS 1998–2006). *Global Biogeochem. Cycles* 22:GB3001. doi: 10.1029/2007GB003154

Alvain, S., Moulin, C., Dandonneau, Y., and Breon, F. M. (2005). Remote sensing of phytoplankton groups in case 1 waters from global SeaWiFS imagery. *Deep Sea Res. I* 1, 1989–2004. doi: 10.1016/j.dsr.2005.06.015

Alvain, S., Moulin, C., Dandonneau, Y., Loisel, H., and Bréon, F. M. (2006). A species-dependent bio-optical model of case I waters for global ocean color processing. *Deep Sea Res.* 53*(Pt. I*), 917–925. doi: 10.1016/j.dsr.2006.01.011

Antoine, D., Chami, M., Claustre, H., d'Ortenzio, F., Morel, A., Becu, G., et al. (2006). *BOUSSOLE: A Joint CNRS-INSU, ESA, CNES, and NASA Ocean Color Calibration and Validation Activity*. NASA Tech. Memo., 2006–21414, 59.

Arnold, S. R., Spracklen, D. V., Gebhardt, S., Custer, T., Williams, J., Peeken, I., et al. (2010). Relationships between atmospheric organic compounds and air-mass exposure to marine biology. *Environ. Chem.* 7, 232–241. doi: 10.1071/EN09144

Belviso, S., Masotti, I., Tagliabue, A., Bopp, L., Brockmann, P., Fichot, C., et al. (2012). DMS dynamics in the most oligotrophic subtropical zones of the global ocean. *Biogeochemistry* 110, 215–241. doi: 10.1007/s10533-011-9648-1

Ben Mustapha, Z., Alvain, S., Jamet, C., Loisel, H., and Dessailly, D. (2014). Automatic classification of water-leaving radiance anomalies from global SeaWiFS imagery: application to the detection of phytoplankton groups in open ocean waters. *Remote Sens. Environ.* 146, 97–12. doi: 10.1016/j.rse.2013.08.046

Bopp, L., Aumont, O., Cadule, P., Alvain, S., and Gehlen, M. (2005). Response of diatoms distribution to global warming and potential implications: a global model study. *Geophys. Res. Lett.* 32:L19606. doi: 10.1029/2005GL023653

Bracher, A., Bouman, H. A., Brewin, R. J. W., Bricaud, A., Brotas, V., Ciotti, A. M., et al. (2017). Obtaining Phytoplankton diversity from ocean color: a scientific roadmap for future development. *Front. Mar. Sci.* 4:55. doi: 10.3389/fmars.2017.00055

Bracher, A., Vountas, M., Dinter, T., Burrows, J. P., Röttgers, R., and Peeken, I. (2009). Quantitative observation of cyanobacteria and diatoms from space using PhytoDOAS on SCIAMACHY data. *Biogeosciences* 6, 751–764. doi: 10.5194/bg-6-751-2009

Brewin, R. J. W., Ciavatta, S., Sathyendranath, S., Jackson, T., Tilstone, G., Curran, K., et al. (2017). Uncertainty in ocean-color estimates of chlorophyll for phytoplankton groups. *Front. Mar. Sci*. 4:104. doi: 10.3389/fmars.2017.00104

Brewin, R. J. W., Devred, E., Sathyendranath, S., Lavender, S. J., and Hardman-Mountford, N. J. (2011). Model of phytoplankton absorption based on three size classes. *Appl. Opt.* 50, 4353–4364. doi: 10.1364/AO.50.004535

Brewin, R. J. W., Sathyendranath, S., Müller, D., Brockmann, C., Deschamps, P. Y., Devred, E., et al. (2015). The ocean colour climate change initiative: III. a round-robin comparison on in-water bio-optical algorithms. Remote Sens. Environ. 162, 271–294. doi: 10.1016/j.rse.2013.09.016

Brewin, R. J., Sathyendranath, S., Hirata, T., Lavender, S. J., Barciela, R. M., and Hardman-Mountford, N. J. (2010). A three-component model of phytoplankton size class for the Atlantic Ocean. *Ecol. Modell.* 221, 1472–1483. doi: 10.1016/j.ecolmodel.2010.02.014

Bricaud, A., Bosc, E., and Antoine, D. (2002). Algal biomass and sea surface temperature in the Mediterranean basin. intercomparison of data from various satellite sensors, and implications for primary production estimates. Remote Sens. Environ. 81, 163–178. doi: 10.1016/S0034-4257(01)00335-2

Brunet, C., Casotti, R., Vantrepotte, V., Corato, F., and Conversano, F. (2006). Picophytoplankton diversity and photoacclimation in the Strait of Sicily (Mediterranean Sea) in summer. *I. Mesoscale variations. Aqua. Microbiol. Ecol.* 44, 127–141. doi: 10.3354/ame044127

Casotti, R., Landolfi, A., Brunet, C., D'Ortenzio, F., Mangoni, O., Ribera d'Alcalà, M., et al. (2003). Composition and dynamics of the phytoplankton of the Ionian Sea (eastern Mediterranean). *J. Geophys. Res.* 108:8116. doi: 10.1029/2002JC001541

Chassot, E., Bonhommeau, S., Dulvy, N. K., Mélin, F., Watson, R., Gascuel, D., et al. (2010). Global marine primary production constrains fisheries catches. *Ecol. Lett.* 13, 495–505. doi: 10.1111/j.1461-0248.2010.01443.x

Ciotti, A., and Bricaud, A. (2006). Retrievals of a size parameter for phytoplankton and spectral light absorption by colored detrital matter from water-leaving radiances at SeaWiFS channels in a continental shelf region off Brazil. *Limnol. Oceanogr. Methods.* 4, 237–253. doi: 10.4319/lom.2006.4.237

Claustre, H., and Marty, J. C. (1995). Specific phytoplankton biomasses and their relation to primary production in the Tropical North Atlantic. *Deep Sea Res.* 42*(Pt. I*), 1475–1493. doi: 10.1016/0967-0637(95)00053-9

Claustre, H., Morel, A., Hooker, S. B., Babin, M., Antoine, D., Oubelkheir, K., et al. (2002). Is desert dust making oligotrophic waters greener? *Geophys. Res. Lett.* 29:1469. doi: 10.1029/2001GL014056

Colella, S., Falcini, F., Rinaldi, E., Sammartino, M., and Santoleri, R. (2016). Mediterranean ocean colour chlorophyll trends. *PLoS ONE* 11:e0155756. doi: 10.1371/journal.pone.0155756

Couto, A. B., Brotas, V., Mélin, F., Groom, S., and Sathyendranath, S. (2016). Inter-comparison of OC-CCI chlorophyll-a estimates with pre-cursor datasets. *Int. J. Remote Sens.* 37, 4337–4355. doi: 10.1080/01431161.2016.1209313

D'Ortenzio, F., and Ribera d'Alcalá, M. (2009). On the trophic regimes of the Mediterranean Sea: a satellite analysis. *Biogeosciences* 6, 139–148. doi: 10.5194/bg-6-139-2009

D'Ovidio, F., De Monte, S., Alvain, S., Dandonneau, Y., and Levy, M. (2010). Fluid dynamical niches of phytoplankton types. *Proc. Natl. Acad. Sci. U.S.A.* 107, 18366–18370. doi: 10.1073/pnas.1004620107

Dandonneau, Y., Montel, Y., Blanchot, J., Giraudeau, J., and Neveux, J. (2006). Temporal variability in phytoplankton pigments, picoplankton and coccolithophores along a transect through the North Atlantic and tropical southwestern Pacific. *Deep Sea Res. Oceanogr. Res. Pap.* 53*(Pt. I),* 689–712. doi: 10.1016/j.dsr.2006.01.002

De Monte, S., Soccodato, A., Alvain, S., and D'Ovidio, F. (2013). Can we detect oceanic biodiversity hotspots from space? *ISME J.* 7, 2054–2056. doi: 10.1038/ismej.2013.72

Demarcq, H., Reygondeau, G., Alvain, S., and Vantrepotte, V. (2012). Monitoring marine phytoplankton seasonality from space. *Remote Sens. Environ.* 117, 211–222. doi: 10.1016/j.rse.2011.09.019

Di Cicco, A., Sammartino, M., Marullo, S., and Santoleri, R. (2017). Regional empirical algorithms for an improved identification of Phytoplankton Functional Types and Size Classes in the Mediterranean Sea using satellite data. *Front. Mar. Sci.* 4:126 doi: 10.3389/fmars.2017.00126

D'Ortenzio, F., Marullo, S., Ragni, M., Ribera d'Alcala, M., and Santoleri, R. (2002). Validation of empirical SeaWiFS algorithms for chlorophyll-a retrieval in the Mediterranean Sea: a case study for oligotrophic seas. *Remote Sens. Environ.* 82, 79–94. doi: 10.1016/S0034-4257(02)00026-3

Efron, B., and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap. (New York, NY: Chapman & Hall/CRC).*

Field, C. B., Behrenfeld, M. J., Randerson, J. T., and Falkowski, P. (1998). Primary production of the biosphere: Integrating terrestrial and oceanic components. *Science* 281, 237–240.

Fujiwara, A., Hirawake, T., Suzuki, K., and Saitoh, S. I. (2011). Remote sensing of size structure of phytoplankton communities using optical properties of the Chukchi and Bering Sea shelf region. *Biogeosciences* 8, 3567–3580. doi: 10.5194/bg-8-3567-2011

Gieskes, W. W. C., and Kraay, G. W. (1983). Dominance of Cryptophyceae during the phytoplankton spring bloom in the central North Sea detected by HPLC analysis of pigments. *Mar. Biol.* 75, 179–185. doi: 10.1007/BF00406000

Gieskes, W. W. C., Kraay, G. W., Nontji, A., Setiapermana, D., and Sutomo, D. (1988). Monsoonal alternation of a mixed and a layered structure in the phytoplankton of the euphotic zone of the banda sea (Indonesia): a mathematical analysis of algal pigment fingerprints. *Neth. J. Sea Res.* 22, 123–137. doi: 10.1016/0077-7579(88)90016-6

Goericke, R., and Repeta, D. (1992). The pigments of *Prochlorococcus marinus*: the presence of divinyl-chlorophylls a and b in a marine prokaryote. *Limnol. Oceanogr.* 37, 425–433. doi: 10.4319/lo.1992.37.2.0425

Goericke, R., and Repeta, D. (1993). Chlorophylls a and b and divinyl chlorophylls a and b in the open subtropical North Atlantic Ocean. *Mar. Ecol. Prog. Ser.* 101, 307–313. doi: 10.3354/meps101307

Gorgues, T., Menkes, C., Slemons, L., Aumont, O., Dandonneau, Y., Radenac, M. H., et al. (2010). Revisiting the La Niña 1998 phytoplankton blooms in the equatorial Pacific. *Deep Sea Res. Oceanogr. Res. Pap.* 57(*Pt. I*), 567–576. doi: 10.1016/j.dsr.2009.12.008

Grant, M., Jackson, T., Chuprin, A., Sathyendranath, S., Zühlke, M Groom, S., et al. (2016). *Ocean Colour Climate Change Initiative (OC-CCI)—Phase Two. Product User Guide.* 2015. Available Online at: ftp://ftp.rsg.pml.ac.uk/occci-v2.0/documentation/OC-CCI-PUG-v2.0.5.pdf (accessed on 4 May 2017).

Gregg, W. W., and Casey, N. W. (2004). Global and regional evaluation of the SeaWiFS chlorophyll data set. *Remote Sens. Environ.* 93, 463–479. doi: 10.1016/j.rse.2003.12.012

Gregg, W. W., Chen, F. C., Mezaache, A. L., Chen, J. D., and Whiting, J. A. (1993). *The Simulated SeaWiFS Data Set. Version 1*. SeaWiFS Technical Report Series #9, edited by S.B. Hooker, NASA Technical Memorandum 104566. Greenbelt. Vol.09.

Grinsted, A., Moore, J. C., and Jevrejeva, S. (2004). Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Process. Geophys.* 11, 561–566. doi: 10.5194/npg-11-561-2004

Guillard, R. R. L., Murphy, L. S., Foss, P., and Liaaen-Jensen, S. (1985). *Synechococcus* spp. as likely zeaxanthin-dominant ultraphytoplankton in the North Atlantic. Limnol. Oceanogr. 30, 412–414. doi: 10.4319/lo.1985.30.2.0412

Hashioka, T., Vogt, M., Yamanaka, Y., Le Quéré, C., Buitenhuis, E. T., Aita, M. N., et al. (2013). Phytoplankton competition during the spring bloom in four plankton functional type models. *Biogeosciences* 10, 6833–6850. doi: 10.5194/bg-10-6833-2013

Hirata, T., Hardman-Mountford, N. J., Brewin, R. J. W., Aiken, J., Barlow, R., Suzuki, K., et al. (2011). Synoptic relationships between surface Chlorophyll-a and diagnostic pigments specific to phytoplankton functional types, *Biogeosciences* 8, 311–327. doi: 10.5194/bg-8-311-2011

Huertas, I. E., Ríos, A. F., García-Lafuente, J., Navarro, G., Makaoui, A., Sánchez-Román, A., et al. (2012). Atlantic forcing of the Mediterranean oligotrophy, *Glob. Biogeochem. Cycles* 26, *GB*2022. doi: 10.1029/2011GB004167

IOCCG (2012). *Mission Requirements for Future Ocean-Colour Sensors*. Reports of the International Ocean-Colour Coordinating Group, No. 13, IOCCG, eds C. R McClain and G. Meister, Dartmouth.

Jackson, T., Sathyendranath, S., and Mélin, F. (in press). An improved optical classification scheme for the Ocean Colour Essential Climate Variable and its applications. *Remote Sens. Environ.* doi: 10.1016/j.rse.2017.03.036

Jeffrey, S. W. (1980). "Algal pigment systems," in *Primary Productivity in the Sea*. ed *P. Falkowski* (New York, NY: Plenum Press), 33–58.

Jeffrey, S. W., and Hallegraeff, G. M. (1987). Phytoplankton pigments, species and light climate in a complex warm-core eddy of the East Australian Current. *Deep Sea Res. Part A Oceanogr. Res. Pap.* 34, 649–673. doi: 10.1016/0198-0149(87)90029-X

Jeffrey, S. W., and Vesk, M. (1997). "Introduction to marine phytoplankton and their pigment signatures," in *Phytoplankton Pigments in Oceanography*, eds S. W. Jeffrey, R. F. C. Mantoura, and S. W. Wright (Paris: U. N. Educ. Sci. and Cult. Org.), 407–428.

Kostadinov, T. S., Siegel, D. A., and Maritorena, S. (2010). Global variability of phytoplankton functional types from space: assessment via the particle size distribution. *Biogeosciences* 7, 3239–3257. doi: 10.5194/bg-7-3239-2010

Kostadinov, T., Cabré, A., Vedantham, H., Marinov, I., Bracher, A., Brewin, R., et al. (2017). Intercomparison of phytoplankton functional types derived from ocean color algorithms and earth system models: phenology. *Remote Sens. Environ.* 190, 162–177.doi: 10.1016/j.rse.2016.11.014

Krom, M. D., Kress, N., Brenner, S., and Gordon, L. (1991). Phosphorous limitation of primary productivity in the Eastern Mediterranean Sea. *Limnol. Oceanogr.* 36, 424–432. doi: 10.4319/lo.1991.36.3.0424

Le Quéré, C., Harrison, S. P., Prentice, I. C., Buitenhuis, E. T., Aumont, O., Bopp, L., et al. (2005). Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. *Glob. Chang. Biol.* 11, 2016–2040. doi: 10.1111/j.1365-2486.2005.1004.x

Lee, Z. P., Darecki, M., Carder, K. L., Davis, C. O., Stramski, D., and Rhea, W. J. (2005). Diffuse attenuation coefficientof downwelling irradiance: an evaluation of remote sensing methods. *J. Geophys. Res. Oceans* 110:C02017. doi: 10.1029/2004JC002573

Li, W. K. W., SubbaRao, D. V., Harrison, W. G., Smith, J. C., Cullen, J. J., and Irwin, T. (1983). Autotrophic picoplankton in the tropical ocean. *Science* 219, 292–295. doi: 10.1126/science.219.4582.292

Loisel, H., Vantrepotte, V., Norkvist, K., Mériaux, X., Kheireddine, M., Ras, J., et al. (2011). Characterization of the bio-optical anomaly and diurnal variability of particulate matter, as seen from scattering and backscattering coefficients, in ultra-oligotrophic eddies of the Mediterranean Sea. *Biogeosciences* 8, 3295–3317. doi: 10.5194/bg-8-3295-2011

Ludwig, W., Dumont, E., Meybeck, M., and Heussner, S. (2009), River discharges of water and nutrients tothe Mediterranean and Black Sea: Major drivers for ecosystem changes during past and future decades? *Prog. Oceanogr.* 80, 199–217. doi: 10.1016/j.pocean.2009.02.001

Marty, J. C., and Chiavérini, J. (2002). Seasonal and interannual variations in phytoplankton production at DYFAMED time-series station, northwestern Mediterranean Sea. *Deep Sea Res. II* 49, 2017–2030. doi: 10.1016/S0967-0645(02)00025-5

Marty, J. C., and Chiavérini, J. (2010). Hydrological changes in the Ligurian Sea (NW Mediterranean, DYFAMED site) during 1995–2007 and biogeochemical consequences, *Biogeosciences* 7, 2117–2128. doi: 10.5194/bg-7-2117-2010

Marty, J. C., Chiavérini, J., Pizay, M. D., and Avril, B. (2002). Seasonal and interannual dynamics of nutrients and phytoplankton pigments in the western Mediterranean Sea at the DYFAMED time-series station (1991–1999). *Deep Sea Res. Part II Top. Stud. Oceanogr.* 49, 1965–1985. doi: 10.1016/S0967-0645(02)00022-X

Masotti, I., Belviso, S., Alvain, S., Johnson, J. E., Bates, T. S., Tortell, P. D., et al. (2010). Spatial and temporal variability of the dimethylsulfide to chlorophyll ratio in the surface ocean: an assessment based on phytoplankton group dominance determined from space. *Biogeosciences* 7, 3215–3237. doi: 10.5194/bg-7-3215-2010

Masotti, I., Moulin, C., Alvain, S., Bopp, L., Tagliabue, A., and Antoine, D. (2011). Large-scale shifts in phytoplankton groups in the Equatorial Pacific during ENSO cycles, *Biogeosciences* 8, 539–550, doi: 10.5194/bg-8-539-2011

Mayot, N., D'Ortenzio, F., Uitz, J., Gentili, B., Ras, J., Vellucci, V., et al. (2017). Influence of the phytoplankton community structure on the spring and annual primary production in the North-Western Mediterranean Sea. *J. Geophys. Res. Oceans.* doi: 10.1002/2016JC012668. [Epub ahead of print].

McClain, C. R. (2009). A decade of satellite ocean color observations. *Annu. Rev. Mar. Sci.* 1, 19–42. doi: 10.1146/annurev.marine.010908.163650

Mélin, F., and Sclep, G. (2015). Band shifting for ocean color multi-spectral reflectance data. *Opt. Express* 23, 2262–2279. doi: 10.1364/OE.23.002262

Mouw, C. B., and Yoder, J. A. (2010). Optical determination of phytoplankton size composition from global SeaWiFS imagery. *J. Geophys. Res.* 115, C12018. doi: 10.1029/2010JC006337

Mouw, C. B., Hardman-Montford, N., Alvain, S., Bracher, A., Brewin, R., Bricaud, A., et al. (2017). A Consumer's guide to satellite remote sensing of multiple phytoplankton groups in the global ocean. *Front. Mar. Sci.* 4:41. doi: 10.3389/fmars.2017.00041

Mueller, J. L., and Austin, R. W. (1995). *Ocean Optics Protocols for SeaWiFS Validation, Revision 1*. NASA Tech. Memo. 104566, Vol. 25, edited by: Hooker, S. B., Firestone, E. R., and Acker, J. G., NASA Goddard Space Flight Center, Greenbelt, Maryland, 67.

Navarro, G., Alvain, S., Vantrepotte, V., and Huertas, I. E. (2014). Identification of dominant phytoplankton functional types in the Mediterranean Sea based on a regionalized remote sensing approach. *Remote Sens. Environ.* 152, 557–575. doi: 10.1016/j.rse.2014.06.029

Nezlin, N. P., and DiGiacomo, P. M. (2005). Satellite ocean color observations of stormwater runoff plumes alongthe San Pedro Shelf (southern California) during1997–2003. *Cont. Shelf Res.* 25, 1692–1711. doi: 10.1016/j.csr.2005.05.001

O'Reilly, J. E., Maritorena, S., Mitchell, B. G., Siegel, D. A., Carder, K. L., Garver, S. A., et al. (1998). Ocean color chlorophyll algorithms for SeaWiFS. *J. Geophys. Res.* 103, 24937–24953. doi: 10.1029/98JC02160

Peloquin, J., Swan, C., Gruber, N., Vogt, M., Claustre, H., and Ras, J. (2013). The MAREDAT global database of high performance liquid chromatography marine pigment measurements. *Earth System Science Data* 5, 109–123. doi: 10.5194/essd-5-109-2013

Percival, D. B., and Walden, A. T. (2000). *Wavelet Methods for Time Series Analysis*. Cambridge: Cambridge University Press.

Platt, T., Sathyendranath, S., and Stuart, V. (2006). Why study biological oceanography? *Aquabiology* 28, 542−557.

Powley, H. R., Dürr, H. H., Lima, A. T., Krom, M. D., and Van Cappellen, P. (2016). Direct Discharges of Domestic Wastewater are a Major Source of Phosphorus and Nitrogen to the Mediterranean Sea. *Environ. Sci. Technol.* 50, 8722–8730. doi: 10.1021/acs.est.6b01742

Raitsos, D., Lavender, S. J., Maravelias, C. D., Haralambous, J., Richardson, A. J., and Reid, P. C. (2008). Identifying four phytoplankton functional types from space: an ecological approach. *Limnol. Oceanogr.* 53, 605–613. doi: 10.4319/lo.2008.53.2.0605

Roy, S., Sathyendranath, S., Bouman, H., and Platt, T. (2013). The global distribution of phytoplankton size spectrum and size classes from their light-absorption spectra derived from satellite data. *Remote Sens. Environ.* 139, 185–197. doi: 10.1016/j.rse.2013.08.004

Sabine, C. L., Feely, R. A., Gruber, N., Key, R. M., Lee, K., Bullister, J. L., et al. (2004). The oceanic sink for anthropogenic CO2, *Science* 305, 367. doi: 10.1126/science.1097403

Sammartino, M., Di Cicco, A., Marullo, S., and Santoleri, R. (2015). Spatio-temporal variability of micro-, nano- and pico-phytoplankton in the Mediterranean Sea from satellite ocean colour data of SeaWiFS. *Ocean Sci.* 11, 759–778. doi: 10.5194/os-11-759-2015

Santoleri, R., Volpe, G., Marullo, S., and Nardelli, B. B. (2008). "Open waters optical remote sensing of the Mediterranean Sea," in *Remote Sensing of the European Seas,* eds V. Barale and M. Gade (Dordrecht: Springer Netherlands), 103–116.

Sathyendranath, S., Watts, L., Devred, E., Platt, T., Caverhill, C., and Maass, H. (2004). Discrimination of diatoms from other phytoplankton using ocean-colour data. *Mar. Ecol. Prog. Ser.* 272, 59–68. doi: 10.3354/meps272059

Schauer, M., Balagué, V., Pedrós-Alió, C., and Massana,. R. (2003). Seasonal changes in the taxonomic composition of bacterioplankton in a coastal oligotrophic system. *Aquat. Microb. Ecol.* 31, 163–174. doi: 10.3354/ame031163

Siokou-Frangou, I., Christaki, U., Mazzocchi, M. G., Montresor, M., Ribera d'Alcalá, M., Vaqué, D., et al. (2010). Plankton in the open Mediterranean Sea: a review. *Biogeosciences* 7, 1543–1586. doi: 10.5194/bg-7-1543-2010

Socal, G., Boldrin, A., Bianchi, F., Civitarese, G., De Lazzari, A., Rabitti, S., et al. (1999). Nutrient, particulate matter and phytoplankton variability in the photic layer of the Otranto strait. *J. Mar. Syst.* 20, 381–398. doi: 10.1016/S0924-7963(98)00075-X

Storm, T., Boettcher, M., Grant, M., Zühlke, M., Fomferra, N., Jackson, T., et al. (2013). *Product User Guide, Ocean Colour Climate Change Initiative.* Plymouth:Plymouth Marine Laboratory.

Tanhua, T., Hainbucher, D., Schröder, K., Cardin, V., and Álvarez, M., and Civitarese, G. (2013). The Mediterranean Sea system: a review and an introduction to the special issue. *Ocean Sci. Discuss.* 10, 581–617. doi: 10.5194/osd-10-581-2013

Thuillier, G., Herse, M., Labs, D., Foujols, T., Peetermans, W., Gillotay, D., et al. (2003). The solar spectral irradiance from 200 to 2400 nm as measured by the solspec spectrometer from the ATLAS and EURECA missions. *Solar Phys.* 214, 1–22. doi: 10.1023/A:1024048429145

Thyssen, M.,Alvain, S., Lefebvre, A., Dessailly, D., Rijkeboer, M., Guiselin, N., et al. (2015). High-resolution analysis of a North Sea phytoplankton community structure based on *in situ* flow cytometry observations and potential implication for remote sensing, *Biogeosciences* 12, 4051–4066. doi: 10.5194/bg-12-4051-2015

Torrence, C., and Compo, G. P. (1998). A practical guide to wavelet analysis. *Bull. Am. Meteorol. Soc.* 79, 61–78.

Turley, C. M., Bianchi, M., Christaki, U., Conan, P., Harris, J. R. W., Psarra, S., et al. (2000). Relationship between primary producers and bacteriain an oligotrophic sea - the Mediterranean and biogeochemical implications. *Mar. Ecol. Prog. Ser.* 193, 11–18. doi: 10.3354/meps193011

Uitz, J., Claustre, H., Morel, A., and Hooker, S. B. (2006). Vertical distribution of phytoplankton communities in Open Ocean: an assessment based on surface chlorophyll. *J. Geophys. Res.* 111, CO8005. doi: 10.1029/2005JC 003207

Uitz, J., Stramski, D., Gentili, B., D'Ortenzio, F., and Claustre, H. (2012). Estimates of phytoplankton class-specific and total primary production in the Mediterranean Sea from satellite ocean color observations. *Global Biogeochem. Cycles* 26, GB2024. doi: 10.1029/2011GB004055

Vantrepotte, V., and Mélin, F. (2010). Temporal variability in SeaWiFS derived apparent optical properties in European seas. *Continent. Shelf Res.* 30, 319–334. doi: 10.1016/j.csr.2009.11.012

Vaulot, D., Partensky, F., Neveux, J., Mantoura, F. C., and Llevellyn, C. A. (1990). Winter presence of prochlorophytes in surface waters of the northwestern Mediterranean Sea. *Limnol. Oceanogr.* 35, 1156–1164. doi: 10.4319/lo.1990.35.5.1156

Vidussi, F., Claustre, H., Manca, B., Luchetta, A., and Marty, J. C. (2001). Phytoplankton pigment distribution in relation to the upper thermocline circulation in the Eastern Mediterranean Sea during winter. *J. Geophys. Res.* 106, 19939–19956. doi: 10.1029/1999JC000308

Vidussi, F., Marty, J. C., and Chiaverini, J. (2000). Phytoplankton pigment variations during the transition from spring bloom to oligotrophy in the northwestern Mediterranean Sea. *Deep-Sea Res.* 47, 423–445. doi: 10.1016/S0967-0637(99)00097-7

Volpe, G., Colella, S., Forneris, V., Tronconi, C., and Santoleri, R. (2012). The Mediterranean Ocean Colour Observing System - System development and product validation. *Ocean Sci.* 8, 869–883. doi: 10.5194/os-8-86 9-2012

Volpe, G., Santoleri, R., Vellucci, V., Ribera d'Alcalà, M., Marullo, S., and D'Ortenzio, F. (2007). The colour of the Mediterranean Sea: global versus regional bio-optical algorithms evaluation and implication for satellite chlorophyll estimates. *Remote Sens. Environ.* 107, 625–638. doi: 10.1016/j.rse.2006.10.017

Winder, M., and Cloern, J. E. (2010). The annual cycles of phytoplankton biomass. *Philos. Trans. R. Soc. B* 365:3215. doi: 10.1098/rstb.2010.0125

Wright, S. W., and Jeffrey, S. W. (1987). Fucoxanthin pigment markers of marine phytoplankton analysed by HPLC and HPTLC. *Mar. Ecol. Prog. Ser.* 38, 259–266. doi: 10.3354/meps038259

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Navarro, Almaraz, Caballero, Vázquez and Huertas. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Marine Science

ORIGINAL RESEARCH
published: 09 August 2017
doi: 10.3389/fmars.2017.00251

Validation and Intercomparison of Ocean Color Algorithms for Estimating Particulate Organic Carbon in the Oceans

Hayley Evers-King[1], Victor Martinez-Vicente[1], Robert J. W. Brewin[1], Giorgio Dall'Olmo[1,2], Anna E. Hickman[3], Thomas Jackson[1], Tihomir S. Kostadinov[4†], Hajo Krasemann[5], Hubert Loisel[6], Rüdiger Röttgers[5], Shovonlal Roy[7], Dariusz Stramski[8], Sandy Thomalla[9,10], Trevor Platt[1] and Shubha Sathyendranath[1,2]*

[1] Plymouth Marine Laboratory, Plymouth, United Kingdom, [2] National Centre for Earth Observation, Plymouth, United Kingdom, [3] Ocean and Earth Science, National Oceanography Centre Southampton, University of Southampton, Southampton, United Kingdom, [4] Department of Geography and the Environment, University of Richmond, Richmond, VA, United States, [5] Helmholtz-Zentrum Geesthacht, Center for Materials and Coastal Research, Geesthacht, Germany, [6] LOG, Laboratoire d'Océanologie et de Géosciences, Centre National de la Recherche Scientifique, University of Littoral Cote d'Opale, University Lille, UMR 8187, Wimereux, France, [7] Department of Geography and Environmental Sciences, and School of Agriculture, Policy and Development, University of Reading, Reading, United Kingdom, [8] Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA, United States, [9] Southern Ocean Carbon and Climate Observatory, Council for Scientific and Industrial Research, Stellenbosch, South Africa, [10] Department of Oceanography, Marine Research Institute, University of Cape Town, Cape Town, South Africa

OPEN ACCESS

Edited by:
Hervé Claustre, Centre National de la Recherche Scientifique (CNRS), France

Reviewed by:
Emmanuel Boss, University of Maine, United States
David Antoine, Curtin University, Australia

***Correspondence:**
Hayley Evers-King
hek@pml.ac.uk

†Present Address:
Tihomir S. Kostadinov, Division of Hydrologic Sciences, Desert Research Institute, Reno, NV, United States

Specialty section:
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

Received: *15 March 2017*
Accepted: *21 July 2017*
Published: *09 August 2017*

Citation:
Evers-King H, Martinez-Vicente V, Brewin RJW, Dall'Olmo G, Hickman AE, Jackson T, Kostadinov TS, Krasemann H, Loisel H, Röttgers R, Roy S, Stramski D, Thomalla S, Platt T and Sathyendranath S (2017) Validation and Intercomparison of Ocean Color Algorithms for Estimating Particulate Organic Carbon in the Oceans. Front. Mar. Sci. 4:251. doi: 10.3389/fmars.2017.00251

Particulate Organic Carbon (POC) plays a vital role in the ocean carbon cycle. Though relatively small compared with other carbon pools, the POC pool is responsible for large fluxes and is linked to many important ocean biogeochemical processes. The satellite ocean-color signal is influenced by particle composition, size, and concentration and provides a way to observe variability in the POC pool at a range of temporal and spatial scales. To provide accurate estimates of POC concentration from satellite ocean color data requires algorithms that are well validated, with uncertainties characterized. Here, a number of algorithms to derive POC using different optical variables are applied to merged satellite ocean color data provided by the Ocean Color Climate Change Initiative (OC-CCI) and validated against the largest database of *in situ* POC measurements currently available. The results of this validation exercise indicate satisfactory levels of performance from several algorithms (highest performance was observed from the algorithms of Loisel et al., 2002; Stramski et al., 2008) and uncertainties that are within the requirements of the user community. Estimates of the standing stock of the POC can be made by applying these algorithms, and yield an estimated mixed-layer integrated global stock of POC between 0.77 and 1.3 Pg C of carbon. Performance of the algorithms vary regionally, suggesting that blending of region-specific algorithms may provide the best way forward for generating global POC products.

Keywords: satellite ocean color, particulate organic carbon, algorithms, validation, essential climate variables

1. INTRODUCTION

Total particulate organic carbon (POC or C_o) in the ocean is a key currency used in studies of both the biological export of carbon from the surface to the deep ocean, and the availability of food for marine organisms. The pool of POC in the ocean is relatively small (estimates include: 0.43 Pg C in the first light-attenuation depth—Gardner et al., 2006; 2.28 Pg C over a 200 m surface layer—Stramska, 2009). Despite the relative small size of the POC compartment, its components (phytoplankton, bacteria, zooplankton, and organic detritus) are responsible for large fluxes in the ocean, because of their high turnover rates. The organic tissue generated by photosynthesis in the sunlit ocean is either exported from the surface via the "biological pump" (Volk and Hoffert, 1985; Ducklow et al., 2001), transferred to higher trophic levels through the food chain, transformed into detritus, or recycled via the microbial loop, with some of it going into the pool of dissolved organic (DOC) and inorganic carbon (DIC). Organic particles are therefore involved in two important carbon fluxes in the ocean, primary production and export to either the deep ocean or the DOC and DIC pools, in addition to being an integral part of the marine food web. In addition to the components of POC arising from local sources, POC may be transported to a particular location from distant sources: for example, by currents that move POC horizontally in the ocean, or by transport of POC of terrestrial origin to the oceans by river outflow. Though POC is typically treated as a single pool, there is growing awareness of the importance of different particles, such as defined by their size, because of their variety of biogeochemical functions, and their effect on ocean optical properties. For example, it has been shown that around 40% of POC concentration in the oligotrophic regions may be associated with bacteria alone (Cho and Azam, 1990) and submicron detrital particles can also make a significant contribution to the POC pool (Mel'nikov, 1976). Similarly, relatively large particles (generally larger than a few micrometers) can play an important role in POC export (e.g., Boyd and Newton, 1995; Dall'Olmo et al., 2009). The importance of particles characteristics in determining the optical signal of POC has also been recognized.

The theoretical work of Stramski and Kiefer (1991), assuming spherical and homogenous particles, indicated that small particles can make an important contribution to the backscattering signal in the oceans. Further work has shown the impact of non-sphericity and intracellular structures on optical properties, particularly backscattering (Meyer, 1979; Kitchen and Zaneveld, 1992; Quirantes and Bernard, 2004, 2006; Clavano et al., 2007; Matthews and Bernard, 2013; Robertson Lain et al., 2017). Work by Cetinić et al. (2012) linked variation in the beam attenuation coefficient with plankton community composition, and variability in particle backscattering with changes in particle composition due to remineralization. They also highlighted how measurement artifacts might influence the observed relationships between POC and optical properties. Further work has explored separation of the phytoplankton component in POC—through both indirect (Behrenfeld et al., 2005; Kostadinov et al., 2016) and direct methods (Graff et al., 2012, 2015). The contribution of phytoplankton to the POC pool leads to the covariance between chlorophyll a concentration ([Chl]) and POC concentrations, although some scatter exists in these relationships, as a result of variability in the phytoplankton community composition, physiological factors that can affect the carbon-chlorophyll ratio in phytoplankton, and the variable contribution of substances other than phytoplankton (including detritus and bacteria) to POC (see discussion and references within Stramska and Stramski, 2005; Sathyendranath et al., 2009).

POC is readily quantifiable by filtering seawater samples, and forms a key component of many biological ocean models. However, *in situ* samples are expensive to collect, leading to a scarcity of data that hinders efforts to both validate ocean models and develop a complete understanding of POC dynamics. Satellite ocean color data offers the opportunity to quantify POC at the global scale on an almost daily basis. Ocean color or more specifically the water-leaving radiance and corresponding remote sensing reflectance spectra, and derived [Chl] are recognized as Essential Climate Variables (ECVs) by the Global Climate Observing System (GCOS, 2011). This is in recognition of their importance for studying various biological variables and processes in the ocean. In fact, of all the oceanic ECVs that are amenable to remote sensing, ocean color is the only one that targets a biological property. In response to the GCOS requirements, the European Space Agency (ESA) Ocean Color Climate Change Initiative (OC-CCI) has generated a time series of merged satellite products for climate research, using data from the ESA satellite sensor MERIS (MEdium spectral Resolution Imaging Spectrometer) and NASA (National Aeronautics and Space Administration) satellite sensors SeaWiFS (Sea-viewing Wide Field-of-view Sensor) and MODIS-Aqua (Moderate-resolution Imaging Spectroradiometer-Aqua). The products include the normalized remote-sensing reflectances, R_{rs} at SeaWiFS wavelengths and [Chl], as well as some additional inherent optical properties (IOPs) such as absorption and backscattering coefficients of phytoplankton and other particulate matter, and diffuse attenuation coefficient for downward plane irradiance, K_d at 490 nm. There is a recognized need in the user community for additional products from ocean color that deal directly with POC, including separation of the contribution of phytoplankton, and the size distribution of particles. Further, these products need to be regionally optimized and their uncertainties well characterized. Both of these requirements may be addressed via optical classification (e.g., Moore et al., 2009), whereby waters are classified according to their spectral or bio-optical properties. Optical classification allows specific algorithms to be applied to the different optical water types (resulting in a global, merged product) and provides provides a method (Moore et al., 2009; Jackson et al., in press) that can be used to calculate per pixel errors. Other methods exist: e.g., formal error propagation (Lee et al., 2010), or estimation of uncertainties based on model-observation comparison (Maritorena et al., 2010). However, the users consulted within the OC-CCI project expressed a preference for uncertainties based on comparison with *in situ* data (Sathyendranath et al., in press).

Remote sensing of POC through ocean color radiometry requires the exploitation of some optical signal that is associated with the material. In fact, optically, the beam attenuation coefficient of particles (c_p), particle scattering coefficient (b_p), backscattering coefficient (b_{bp}) and the attenuation coefficient of downward irradiance K_d are all sensitive to particle abundance (to a first order), and to particle composition (through refractive index), size, shape and internal structure. It has been demonstrated that POC is correlated with *in situ* c_p measured using transmissometers (Gardner et al., 1993; Bishop, 1999; Claustre et al., 1999; Stramska and Stramski, 2005), which has provided a robust optical method for measuring POC using *in situ* devices. Although b_p and c_p are not among the data products that are routinely retrieved from remote sensing, satellite-based algorithms exist for retrieving POC from all the optical properties listed above, as well as from remote-sensing reflectance values.

This paper compares five different algorithms for estimating POC concentrations, selected as being representative of varied approaches that are prevalent for POC retrieval from ocean-color data. Each algorithm is applied to different optical properties derived from satellite ocean color data, and each uses different formulations for linking the parameters to the POC concentration. Matchups between *in situ* measurements of POC and satellite ocean color allow for the validation, intercomparison of global performance, and estimation of uncertainties associated with the POC calculated using these algorithms.

2. METHODS

2.1. Collation of an *In situ* Database

For this study, POC concentration data were collected from a number of existing databases and from individual contributors. Databases collated included PANGAEA (https://www.pangaea.de/) and SeaBASS (Werdell and Bailey, 2005), and those compiled by Martiny et al. (2014) and the Biological and Chemical Oceanography Data Management Office (BCO-DMO, USA). Further data were included from the Atlantic Meridional Transect (AMT) (including data derived from both CTD and the ship's clean water supply) and other cruises in the Southern Ocean (see a description of the Good Hope line and associated data collection in Thomalla et al., 2017). Operationally, POC is defined as all the organic carbon that is retained on GF/F filters (nominal pore size of 0.7 μm). To measure POC, samples are collected on pre-combusted (450°C) GF/F filters and dried overnight at 65°C before analysis. To remove particulate inorganic carbon, filters are acidified either by adding low-carbon HCl directly or by overnight exposure to the fumes of a concentrated HCl solution in a desiccator. Filters are then dried, packed in pre-combusted tin capsules, combusted at 960°C in an elemental analyser to convert the organic carbon in CO_2. The liberated CO_2 is finally detected by thermal conductivity (Sharp, 1974). Acetanilide is used as a standard. The procedure for applying a blank however is not always consistent across studies, and as such could be a source of bias within the data set collated here. Cetinić et al. (2012) (and references therein) have studied the consequences of different methodologies for treating POC blanks, summarizing that the effect of DOC adsorption on filters (if not accounted for with an adequate blank correction) can cause substantial bias at low POC concentrations. In the database used here, a large quantity of the samples from low POC regions (i.e., the oligotrophic gyres) are from the AMT cruise programme, where a multiple-volume intercept blank methodology is used to reduce potential bias from blanks. Where data was provided at depth, measurements were averaged over 10m to provide the "surface" value for the matchup. Optical weighting of these measurement were considered, however given the variability of the water types sampled and associated mixed layer depths, and the necessary assumptions to apply an optical model for this purpose, it was decided not to introduce additional sources of uncertainty for these data points.

2.2. Extraction of Satellite-*In situ* Matchups

Matchup extraction was based on the procedure developed for the OC-CCI. The daily, 4 km, sinusoidally projected OC-CCI version 2 data (Sathyendranath et al., 2016) were searched to find satellite data associated with each *in situ* data point. The OC-CCI data is a merged product from three sensors, each with a different overpass time. However these overpass times are generally around 12 p.m. $\pm$ 2.5 h, meaning a maximum time difference between the *in situ* and satellite data of <12 h. The OC-CCI data used contain all of the relevant optical and biogeochemical properties necessary for implementation of the different algorithms under consideration, as well as water class membership of each pixel which quantifies the similarities between the remote-sensing reflectance spectrum at that pixel and the characteristic mean and covariance spectra associated with each of the optical classes (Jackson et al., in press), see also the OC-CCI product user guide (http://www.esa-oceancolor-cci.org/?q=documents). The OC-CCI data were interrogated to assess the availability of data covering the latitude and longitude of the *in situ* data point, on the same date as the *in situ* data collection. If the central pixel contained valid data, the data surrounding eight pixels are also extracted from the selected data (a 3 × 3 pixel box, corresponding to a 12 km × 12 km region) . The central value and the mean, median, and standard deviation, number of valid pixels out of the nine pixels, the optical class with the dominant membership, and the calculated POC products using various algorithms applied to the central and mean pixel values, are returned as output, along with the *in situ* data values and metadata.

2.3. Candidate Algorithms

Five different algorithms for determination of POC concentration were considered. For consistency of comparison, all algorithms were implemented using the appropriate variables from the OC-CCI product suite. Another reason for using the OC-CCI products is that a rigorous algorithm selection procedure for atmospheric correction (Müller et al., 2015) and in-water properties (Brewin et al., 2015) (including derivation of [Chl], IOPs and the diffuse attenuation coefficient), had been put in place, to ensure quality of products. Furthermore, being a merged product, OC-CCI coverage is higher than that

available from single-sensor products, ensuring a higher number of match-up points with *in situ* data. However, we recognize that all the five candidate algorithms were initially developed, implemented and tested using other ocean-color products, and that any systematic differences between the OC-CCI products and the datasets used by the algorithm developers, could be a potential source of difference in performance. Therefore, in the description of the algorithms, we also provide details of how the algorithms were implemented in the original work.

We also note that there are differences in how the five algorithms compared here were developed and implemented in the original work. For example, two of the algorithms (algorithms A and B presented below) were derived solely from coincidentally collected *in situ* data, whereas Algorithm D (G06—described below) combined *in situ* measurements of POC and beam attenuation, with satellite-derived measurements of diffuse attenuation coefficient. Algorithms A and B are based on some 50 measurements, whereas the POC—beam attenuation coefficient relationship used in Algorithm D was based on over 3,000 measurements. Algorithm C (described below) relies on a large *in situ* database of [Chl] and backscattering ratio to estimate total particle scattering coefficient from particle backscattering coefficient, and then relies on an extensive literature review to find a conversion factor between particle scattering coefficient and POC. Here we use a common set of satellite data (OC-CCI) to compute POC using the different algorithms and compare the products against a common set of *in situ* data. Where there are differences between how the various steps in the algorithms were implemented in the original work, and how OC-CCI treated similar steps in its product generation, they are highlighted, since such differences could have potential impact on algorithm performance.

2.3.1. Algorithm A Based on Remote-Sensing Reflectance

This algorithm (designated $C_o(A)$) proposed by Stramski et al. (2008) uses remote-sensing reflectance at 443 and 555 nm (equation 1) as inputs, and takes the following form:

$$C_o(A)(\mathrm{mg\ m^{-3}}) = 203.2\left[\frac{\mathrm{R_{rs}}(443)}{\mathrm{R_{rs}}(555)}\right]^{-1.034}. \qquad (1)$$

The model parameters were determined using some 53 pairs of co-located *in situ* measurements of both POC and spectral values of R_{rs} from oligotrophic and upwelling waters of the East Atlantic and the South Pacific. The authors have provided various fits to the models for different pairs of wavebands for R_{rs} and for different selections of data. The one we have used here is based on all data, and for the 443–555 waveband pair, as recommended by the authors. This algorithm is currently used by the NASA Ocean Biology Processing Group to generate the global POC data product from ocean color data. The R_{rs} values in the OC-CCI product suite were used as input to this algorithm for the validation exercise presented here. We note that a similar POC algorithm based on the blue-to-green reflectance ratio has been developed and validated with the data from the Southern Ocean (Allison et al., 2010). For the Southern Ocean algorithm the best fit coefficients of the power function of the same form as Equation (1) were determined to be 189.29 and -0.87. A power function was found to provide better error statistics compared with those using modified fits of those typically used in maximum band ratio algorithms for [Chl].

2.3.2. Algorithm B Based on Backscattering by Particles

This algorithm ($C_o(B)$), also proposed by Stramski et al. (2008) uses $b_{bp}(555)$, the particulate backscattering coefficient at 555 nm (2), as input. The equation has the form:

$$C_o(B)(\mathrm{mg\ m^{-3}}) = 53606.7 \times \mathrm{b_{bp}}(555) + 2.468\,. \qquad (2)$$

Stramski et al. (2008) tested two approaches for calculating b_{bp} for the eventual determination of POC: firstly, the method of Maffione and Dana (1997) as refined by Boss and Pegau (2001) was used to derive spectral backscattering coefficient from *in situ* measurements of volume scattering function at 140°. Next, using a two-step empirical approach, Stramski et al. (2008) then calculated b_{bp} by removing the backscattering coefficient for pure water [they used pure-water backscattering coefficients proposed by Buiteveld et al. (1994) and by Morel (1974), and reported that the difference between the measurements of b_{bw} was likely within the range of errors associated with the measurements themselves, and therefore did not substantially impact the performance of the algorithm for deriving POC]. Then, b_{bp} was empirically related to POC concentrations (Equation 2). Stramski et al. (2008) provide various fits to this equation, but the parameters selected above correspond to the results (number of observations = 54) excluding upwelling data, which provided better uncertainty metrics, and may better reflect the data within the database used here. Secondly, to provide the remote-sensing context, Stramski et al. (2008) used either a direct empirical relationship between the backscattering coefficient (b_b) and R_{rs} or the Quasi Analytical Algorithm (QAA) approach of Lee et al. (2002) to derive b_{bp} from R_{rs}. They found that the comparison with measured b_{bp} was improved considerably when an empirical correction based on their measurements of b_{bp} was applied to the QAA algorithm. Thus, this algorithm uses a two-step approach: first, b_{bp} is derived from R_{rs}, and then that b_{bp} is used in Equation (2) to calculate POC.

In the computation of OC-CCI b_{bp} products, the QAA model was used, along with Zhang et al. (2009) for the pure-water backscattering coefficient (Sathyendranath et al., 2016). In QAA the spectral backscattering coefficient is calculated empirically using R_{rs} at 440 nm and at 555 nm. A power law is then used to calculate the particle backscattering coefficients at other wavelengths given the value of the spectral backscattering coefficient and $b_{bp}(555)$, estimated from an analytical relationship using $R_{rs}(555)$ and the absorption coefficient ($a(555)$). These values of b_{bp} were used to compute POC in the calculations presented here, without any empirical correction. For this algorithm a linear fit provided the best error statistics in the original study of Stramski et al. (2008).

2.3.3. Algorithm C Based on both Backscattering Coefficient and Chlorophyll-a Concentration

This algorithm ($C_o(C)$) is based on a combination of $b_{bp}(490)$ and [Chl], and was proposed by Loisel et al. (2002) (3). They derived b_{bp} and [Chl] from SeaWIFS data using the method of Loisel and Stramski (2000) and the OC4 algorithm (based on O'Reilly et al., 1998) respectively. They then used a relationship proposed by Twardowski et al. (2001) to compute the total particle scattering coefficient given the particle backscattering coefficient (b_{bp}) and [Chl] (designated B in Equation 3). They then adopted a conversion value of 400 mg C m^{-2} to go from total scattering coefficient to POC concentration. Combining these steps yields the following algorithm:

$$C_o(C)(\mathrm{mg\,m^{-3}}) = 41666.7 \times \mathrm{b_{bp}}(490) \times \mathrm{B}^{0.25}. \quad (3)$$

Algorithm C is implemented here using the Lee et al. (2002) QAA approach (for b_{bp}), and a more recent OC4 (v6) for [Chl].

2.3.4. Algorithm D Based on the Diffuse Attenuation Coefficient for Irradiance and Beam Attenuation Coefficient

This algorithm ($C_o(D)$), based on Gardner et al. (2006), uses a two-step relationship relating $K_d(490)$ to beam attenuation coefficient for particles (c_p), and then the beam attenuation coefficient to POC:

$$c_p = \exp(1.124 \times \log 10(K_d(490) + 1.1361)), \quad (4)$$

and

$$C_o(D)(\mathrm{mg\,m^{-3}}) = 12 \times (31.7 \times \mathrm{c_p} + 0.785). \quad (5)$$

In Gardner et al. (2006), $K_d(490)$ is obtained from SeaWIFS data where the algorithm of Mueller (2000) is used, based on water-leaving radiances at 490 and 555 nm. Gardner et al. (2006) used an extensive database (number of observations = 3,462) of concurrent measurements of POC and beam attenuation coefficient from Atlantic, Pacific and Indian Oceans, to derive the parameters of Equations (4) and (5) empirically.

The $K_d(490)$ values used in the comparison presented here are obtained from the OC-CCI version 2 data, which uses the method of Lee et al. (2005).

2.3.5. Algorithm E Based on Spectral Backscattering Coefficient:

This algorithm ($C_o(E)$) was developed by Kostadinov et al. (2009) and Kostadinov et al. (2016). It correlates the slope (η) of backscattering as a function of wavelength to the slope (ξ) of the particle size distribution (PSD) that is assumed to follow a power-law. The method has three steps: firstly, η is calculated from the spectral b_{bp} values at 490, 510, and 555 nm extracted from the OC-CCI matchup data (for which OC-CCI uses Lee et al. (2002)). Note that in Kostadinov et al. (2009) and Kostadinov et al. (2016), b_{bp} is instead retrieved using the formulation of Loisel and Stramski (2000). Secondly, look-up-tables (LUTs) are used to retrieve the parameters of the PSD, namely the slope (ξ) and the differential number concentration at a reference diameter of 2 μm, (N_o), given η and b_{bp} at 443 nm. The LUTs are constructed using theoretical forward simulations using Mie code (Bohren and Huffman, 1983). Finally, to compute C_o, the PSD is integrated to calculate particle volume in the 0.5 to 50 μm diameter range, and then volume is converted to carbon using existing allometric relationships derived from phytoplankton cultures (Menden-Deuer and Lessard, 2000), assuming biogenic origin for all the scattering particles. An empirical correction is applied to N_o (based on PSD validation statistics) to achieve more realistic absolute carbon concentration values (Kostadinov et al., 2016).

2.4. Separation of Matchups by Optical Water Class and Calculation of Uncertainties

The OC-CCI product suite includes memberships of each pixel in 14 optical classes, following the fuzzy logic classification methodology of Moore et al. (2009), with some modifications as described in Jackson et al. (in press). The memberships of the 14 optical water classes associated with the satellite matchups were extracted alongside the radiometric and biogeochemical properties required for the validation. Each matchup point was then assigned to the optical class that had the dominant membership in the central match-up pixel. The statistical analyses were carried out for the global dataset as well as for subsets of the data grouped according to dominant optical class. These uncertainty metrics per optical class were then used to assign uncertainties at each pixel, by calculating the weighted average of the metrics associated with each of the water classes, with the membership of the classes in that pixel providing the weighting function.

Statistical analysis used in the assessment of each algorithm was based on that used by OC-CCI (see Brewin et al., 2015). The Kolmogorov-Smirnov test for normality of the *in situ* matchup data showed a significant deviation of normality for $\log_{10}$ transformed and un-transformed data (p <0.001). Therefore, for completeness, the statistical analysis was conducted for both $\log_{10}$ transformed using parametric tests and for un-transformed data using non-parametric, rank-based, statistics. Statistical metrics computed were:

- Pearsons correlation coefficient for $\log_{10}$ transformed data, and Spearman's correlation for un-transformed data (r_p and r_s respectively),
- Root mean square differences for $\log_{10}$ transformed and un-transformed data (RMSD - Ψ, in $\log_{10}$ for the transformed data and mg m^{-3} of POC for the untransformed data),
- Bias for $\log_{10}$ transformed data and untransformed data ((δ), in $\log_{10}$ for the transformed data and mg m^{-3} of POC for the untransformed data),
- Median absolute percentage deviation between predictions and observations (MAPD in %), an estimate of bias and precision was estimated as the interquartile range (IQR) of the absolute percentage deviation for the untransformed data,

- Center pattern root mean square deviation (Δ), in mg m^{-3}), which is the error of the predicted values with respect to the observations, irrespective of the average bias between the two distributions. It is related to the RMSD and bias as follows: $\Delta^2 = \Psi - \delta^2$, and was calculated for both the $\log_{10}$ transformed and un-transformed data,
- Slope and intercept (S, I) from a Type-II linear regression (Reduced Major Axis) for $\log_{10}$ transformed and un-transformed data.

To provide an indication of the stability of the statistics and to compute confidence intervals on them, bootstrapping (Efron, 1979; Efron and Tibshirani, 1993) with random re-sampling and replacement was used to construct 1,000 different datasets from which confidence intervals were computed for some statistics. Statistics were computed for the whole dataset, and also after segregating the data according to dominant water class, at the central matchup pixel.

3. RESULTS

3.1. Distribution of Validation Data and Matchups

Geographic distribution of the data in the *in situ* data base is shown in **Figure 1**, with the color scale representing the average POC concentration (mg m^{-3}) over measurements made in the top 10 m (total $N = 63,704$, depth averaged $N = 19,282$). **Figure 2** shows the *in situ* data with valid satellite data matchups ($N = 3891$), colored according to the concentration of POC (mg m^{-3}) measured. A number of expected patterns can be observed. Firstly, the number of valid matchups is much reduced relative to the total number of data points in the *in situ* database. This is a result of several factors including: the averaging of the *in situ* data over the top 10m, elimination of data points outside of the OC-CCI period (1997–2012); and failure of matchup when there were no satellite observations corresponding to the date and location of the *in situ* sampling. Other factors are spatially heterogenous, such that there is some regional skew in the likely success of obtaining a matchup, for example cloud cover has a spatially and temporally variable influence on matchup availability, such that in some areas it is less likely that a matchup will be found, e.g., in the tropics, or during winter in the mid latitudes—where satellite coverage was relatively poor. There is a high concentration of data in the Atlantic—as a result of the substantial AMT cruise data. Similarly, there is a high concentration of data from some coastal regions, particularly in the northern hemisphere. Observed concentrations of POC follow an expected distribution, with higher values in coastal and shelf regions, and lower concentrations in the oligotrophic gyres (**Figure 2**). A bimodal distribution can be observed in the *in situ* data, as a result of the high numbers of data from the AMT cruises which go predominantly through the oligotrophic gyres, and from coastal regions, which tend to be predominantly high-POC areas (**Figure 3**).

3.2. Algorithm Performance

The histogram of the *in situ* data frequency distribution is replicated to a degree by all the algorithms (**Figure 3**). The histogram of Algorithm A is very similar to that of the *in situ* data, with both histogram shapes and peak locations reproduced. Algorithm B tends to overestimate the POC values relative to *in situ* data at the lower end of POC concentrations, resulting in the first peak in the histogram being offset toward the right of the figure (i.e., toward higher POC concentrations). By contrast, Algorithm C underestimates POC slightly, at the lower end. The range of estimates provided by Algorithm D is narrower than that of the *in situ* data. Algorithm E also has a narrower range, and in general underestimates POC concentrations. Both algorithms D and E show significant shifts in both peaks of the distribution relative to those of the *in situ* data.

As could be expected from the histograms (**Figure 3**), there is generally a high correlation between the algorithm estimates and *in situ* measurements of POC, with *r* values between 0.75 and 0.82 for all algorithms tested (**Figures 4A–E**). It is of note that all algorithms show a small cloud of underestimated concentrations associated with high *in situ* POC. These points are from different data sources, and from different regions, and as such, outliers do not appear to be related to any systematic errors in the *in situ* measurements. Furthermore, not all of these substantial underestimates are associated with common data points across all the algorithms. It is important to note, however, that although we compare the satellite-derived POC with *in situ* POC over a very broad range, extending to very high values >1,000 mg m^{-3}, the original formulations of the algorithms were not, in general, implemented with such high POC data.

The scatter plots and uncertainty statistics for Algorithm A shown in **Figure 4A** and **Table 1** suggest that this algorithm, on average, performs better than the other algorithms shown in **Figures 4B–E**. Algorithm A has the smallest bias, a linear fit that is closest to the 1:1 line, with a slope of 0.92 and a relatively small intercept value (the second smallest after Algorithm C), and also the highest (albeit slightly) correlation coefficient and the smallest values (albeit slightly) of the Root Mean Square Difference. The statistical parameters are consistently good for Algorithm A, though the results for Algorithm C also present some advantages. The performance of Algorithm C has a slope of 1 and intercept closest to zero, but has a slightly higher RMSD, CRMSD and bias compared with Algorithm A (**Figure 4C**). Algorithm B has relatively small bias, however other parameters are clearly inferior compared with Algorithm A (**Figure 4B**). Some statistical parameters associated with the performance of Algorithm D are significantly inferior compared with those associated with algorithms A and C, especially the slope and intercept of linear fit which indicate a large deviation of this fit from the 1:1 line (**Figure 4D**). As a result, low values of POC are overestimated, and high values underestimated, compared with the *in situ* data. The statistical performance for Algorithm E is poorest with significant negative bias and the best fit regression deviating greatly from the 1:1 line (**Figure 4E**). Performance statistics are summarized for all algorithms in the first section of **Table 1**. A statistical analysis was also conducted for the non-transformed data, and provided in the second section of **Table 1**.

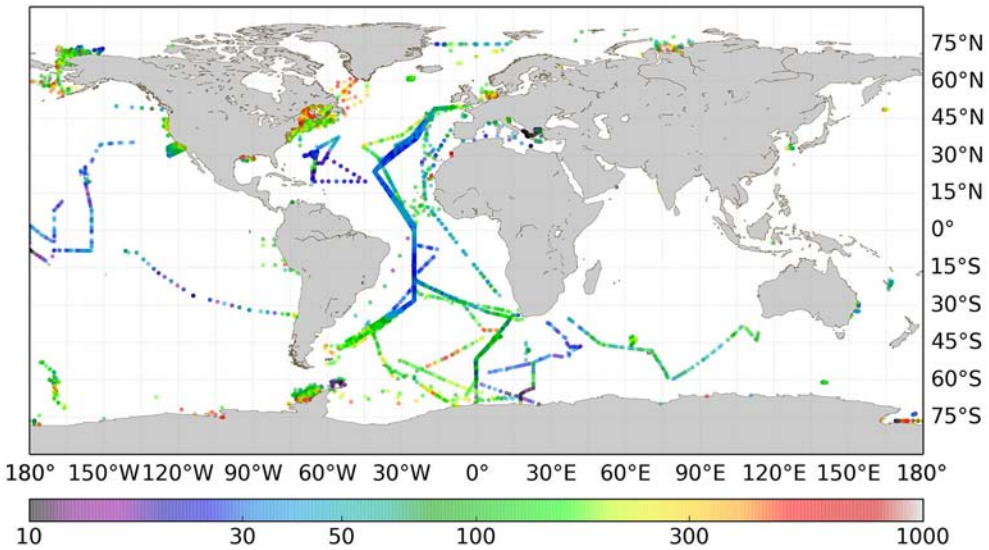


FIGURE 1 | Geographic distribution of Particulate Organic Carbon (POC) measurements within the *in situ* database. The color scale represents the averaged POC concentration (mg m^{-3}) over the top 10 m.

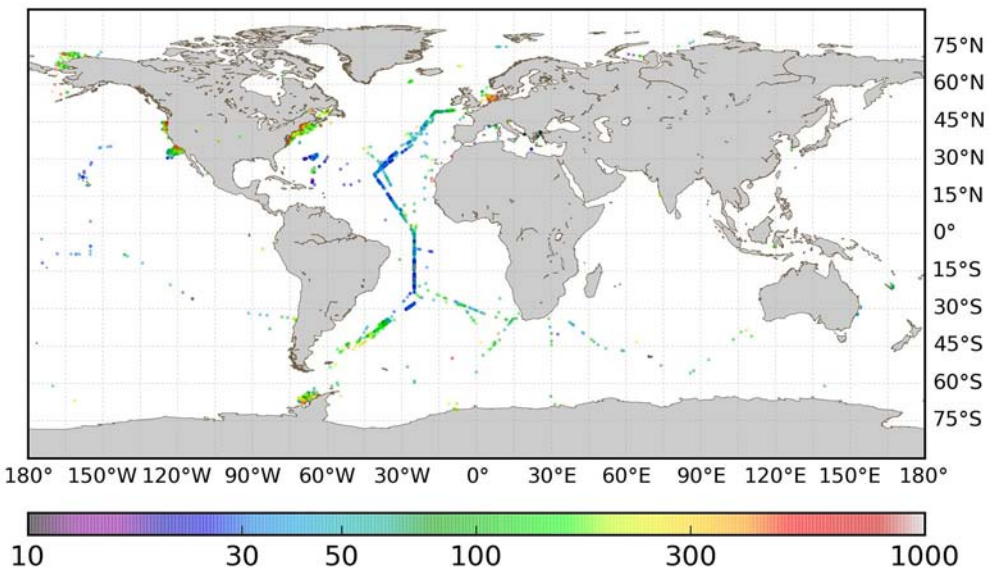


FIGURE 2 | Locations of OC-CCI satellite data matchups and associated *in situ* concentrations of Particulate Organic Carbon (mg m^{-3}).

3.3. Performance Per Water Class

To further understand algorithm performance, the matchups were separated by their optical water class. The total number of matchups per water class, and their distribution spatially, is shown in **Figures 5**, **6** respectively. The lower water classes are associated with oligotrophic regions, such as in the gyres, whilst the higher classes correspond to progressively more turbid shelf and coastal waters. Statistical performance of the algorithms across the different water classes is summarized in **Figure 7**.

Performing the statistical analysis across the different water classes reveals some similarities in performance across all algorithms, and some consistency with the overall performance (**Figure 7**). Algorithms A, C, and D show little differences between them in terms of RMSD across the water classes, with the exception of water class 14 (i.e., the most optically complex waters) (**Figure 7A**). The RMSD associated with Algorithm E follows the same broad pattern as algorithms A and D across the water classes, although its RMSD is substantially higher than those algorithms in all instances. Algorithm B shows slightly

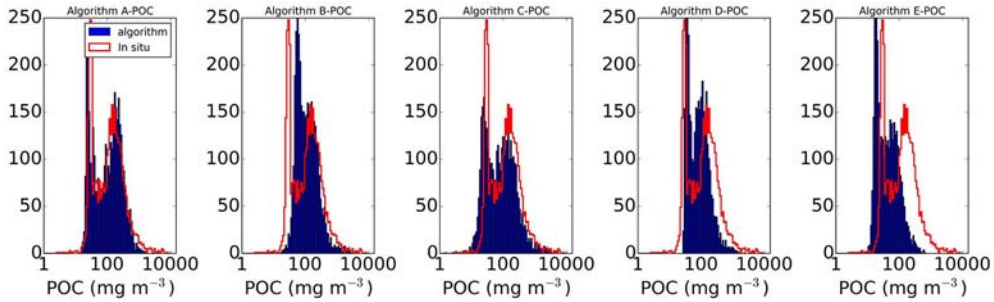


FIGURE 3 | Histograms summarizing distribution of *in situ* POC data and associated satellite based estimates for **(A)** Algorithm A (Stramski et al., 2008, R_{rs} based), **(B)** Algorithm B (Stramski et al., 2008, b_{bp} based), **(C)** Algorithm C (by Loisel et al., 2002) , **(D)** Algorithm D (by Gardner et al., 2006), **(E)** Algorithm E (by Kostadinov et al., 2016).

TABLE 1 | Summary of statistics of the algorithm performances for log_{10} and untransformed data.

Statistic	A	B	C	D	E
log_{10} TRANSFORMED DATA					
r_S	***0.80***	0.73	0.79	0.80	0.78
Ψ	***0.29 ± 0.01***	0.33±0.01	0.34± 0.01	0.32± 0.01	0.53±0.01
Δ	***0.29 ± 0.01***	0.32± 0.01	0.31± 0.01	0.29± 0.01	0.30± 0.01
δ	-0.04± 0.01	***0.03 ± 0.01***	-0.13± 0.01	-0.13± 0.01	-0.43± 0.01
S	0.92± 0.01	0.63± 0.01	***1.00 ± 0.01***	0.63± 0.01	0.66± 0.01
I	0.14± 0.02	0.76± 0.02	***−0.13 ± 0.03***	0.63± 0.02	0.27± 0.02
UNTRANSFORMED DATA					
r_p	0.84	0.80	0.84	***0.85***	0.83
Ψ	420± 12.7	443± 13.3	463± 13.9	***413 ± 12.46***	444± 13.1
Δ	417± 12.4	442±13.1	460± 13.8	***401 ± 11.8***	417± 12.1
δ	-51.9± 13.1	***-41.1 ± 14.0***	-48.5± 14.5	-98.8 ± 12.6	-154± 13.1
S	0.15	0.27	***0.49***	0.13	0.05
I	126	112	58.8	83.3	***45.3***
MAPD±IQR	***25.5*** ± 37.3	42.5± 58.0	31.4± 36.5	31.7± 34.6	60.9 ± ***30.0***

For log_{10}: r_s is Spearman's correlation; slope and intercept for were calculated with a Type-II linear regression model (Major Axis) and the statistics provided have uncertainty estimates (95% confidence interval), derived from 1,000 bootstrap realizations. For untransformed data: r_p is Pearson's correlation; Ψ, δ and Δ are provided with uncertainty estimates (95% confidence interval), derived from 1000 bootstrap realizations; slope and intercept for were calculated with a Reduced Major Axis regression model; MAPD is the median absolute percent deviation between predictions and observations and is a measure of bias, and IQR is the interquartile range of the absolute percent error, and is a measure of precision. Bold italic numbers are the best results for each statistic, for some is the highest value (e.g., r_s or r_p), for some is the lowest (e.g., Ψ, δ, Δ, Intercept and MAPD) and for some is the closest to one (e.g., Slope).

higher RMSD than the other algorithms in some of the more oligotrophic water classes (1–5), then largely follows the same patterns as algorithms A and D. The cloud of outlier points observed in the overall comparisons are associated with water classes 6, 7, and 8, where RMSD is relatively high for most of the algorithms. The estimates of bias for each of the water classes is consistent with the results from the global application e.g., Algorithm A shows very little bias, and this is consistent across water classes, whilst Algorithm E generally underestimates across all classes (**Figure 7B**). Algorithm C shows slight negative bias across most of the water classes, except 12 and 13, whilst algorithms D and B show slight (larger) positive bias in the more oligotrophic classes. Algorithm B shows large positive bias in water class 13, as with Algorithm C; however, both algorithms show negative bias in class 14. The center-pattern (or unbiased) RMSD in **Figure 7C** shows the largest uncertainties are associated with water classes 6, 7, and 8 whilst the largest differences in algorithm performance across the water classes are found in water classes 12, 13, and 14. The regional variability in algorithm performance, which can be associated with the optical water classes, is discussed further in Sections 3.4, 3.5, and 4.2 below.

3.4. Mapped Products

In addition to application to the matchups points, the POC algorithms can also be applied to global satellite data to compare algorithm performance at synoptic scales. POC concentrations

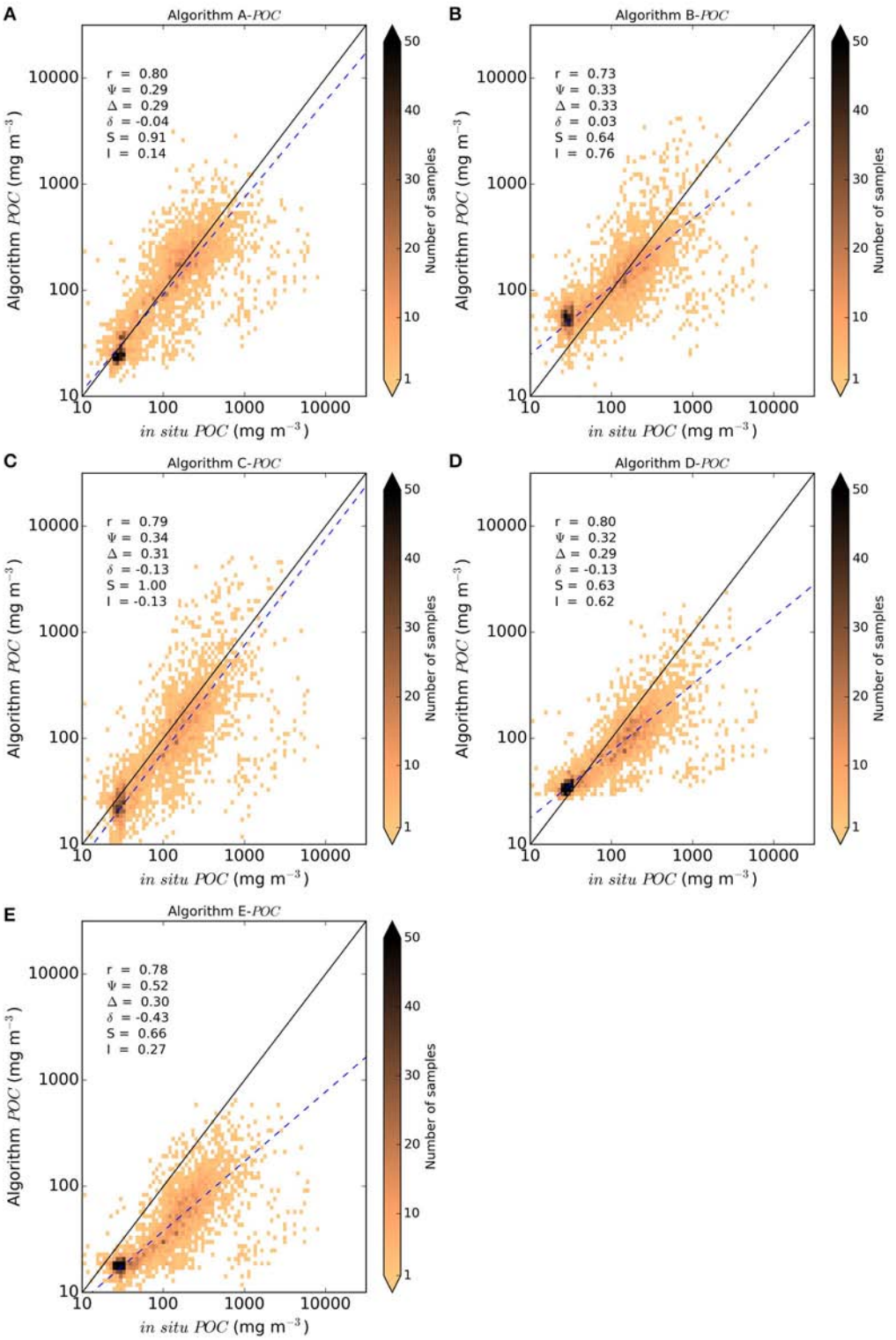


FIGURE 4 | Scatter plots and statistics detailing performance of **(A)** Algorithm A by Stramski et al. (2008) (using R_{rs}), **(B)** Algorithm B by Stramski et al. (2008) (using b_{bp}), **(C)** Algorithm C by Loisel et al. (2002), **(D)** Algorithm D by Gardner et al. (2006), and **(E)** Algorithm E by Kostadinov et al. (2016). Statistical parameters are as follows: correlation coefficient (r), bias (δ), Root Mean Square Difference (Ψ), Center Patterned Root Mean Square Difference (Δ), slope (s), intercept (I). The solid line is the 1:1 line, and the dashed line is the line of best fit for the linear regression.

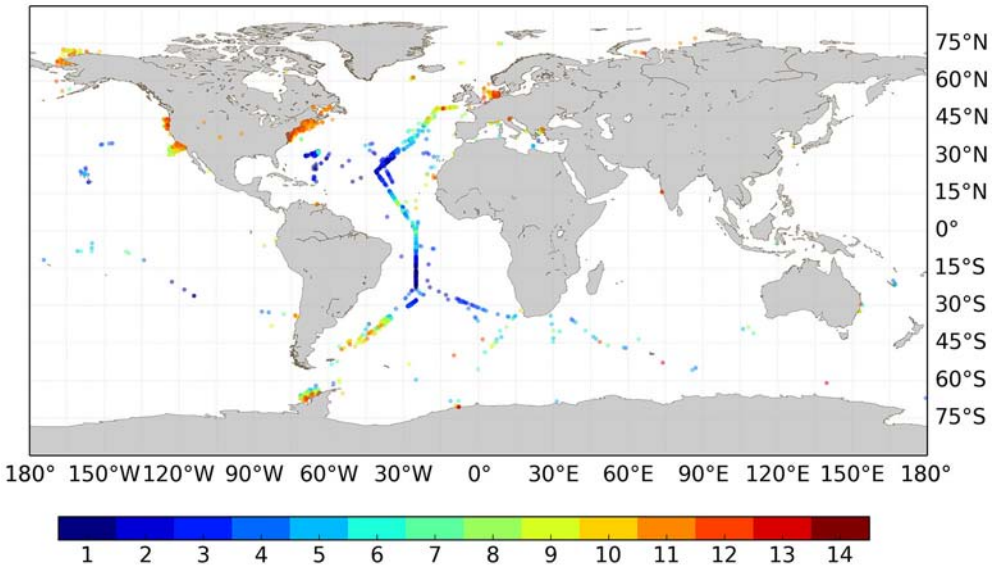


FIGURE 5 | Locations and associated water class for each satellite-*in situ* matchup.

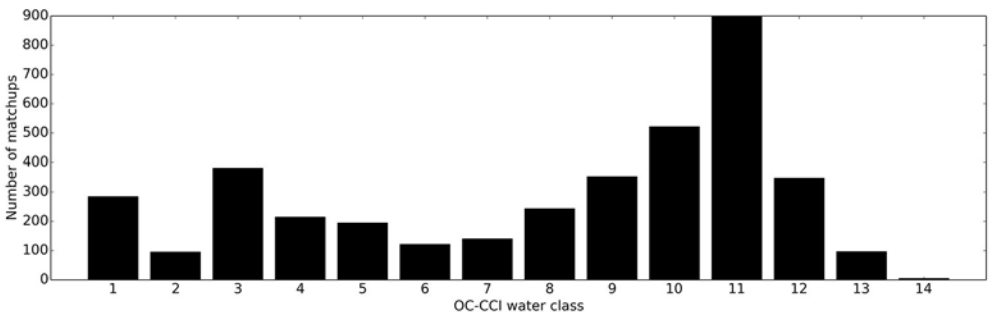


FIGURE 6 | Histogram showing number of matchups associated with each OC-CCI water class.

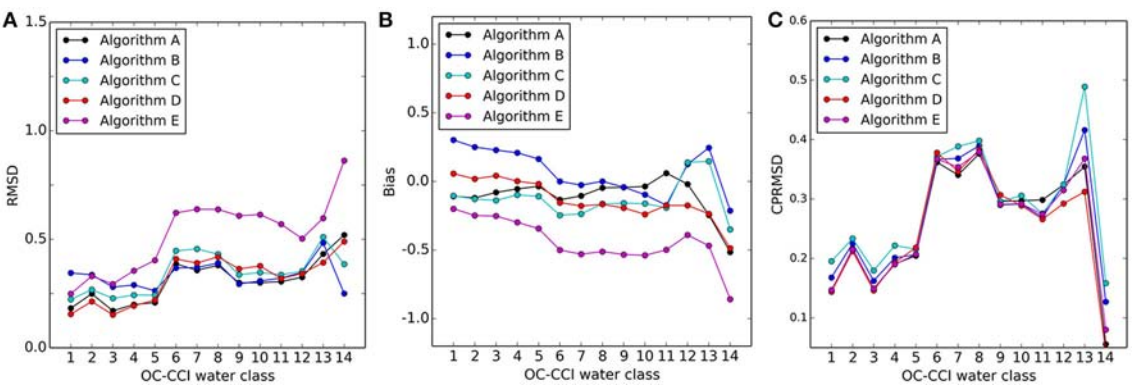


FIGURE 7 | Calculated RMSD **(A)**, Bias **(B)**, and Center Patterned RMSD **(C)** for each algorithm per water class based on the matchup statistics.

were estimated by applying algorithms A-E to a sample set of OC-CCI monthly products from May 2005 (**Figure 8**). All algorithms produce the broad patterns that were observed in the *in situ* measurements and would be expected to be associated with POC, i.e., increased POC associated with regions of high [Chl] in upwelling zones, lower concentrations in the oligotrophic

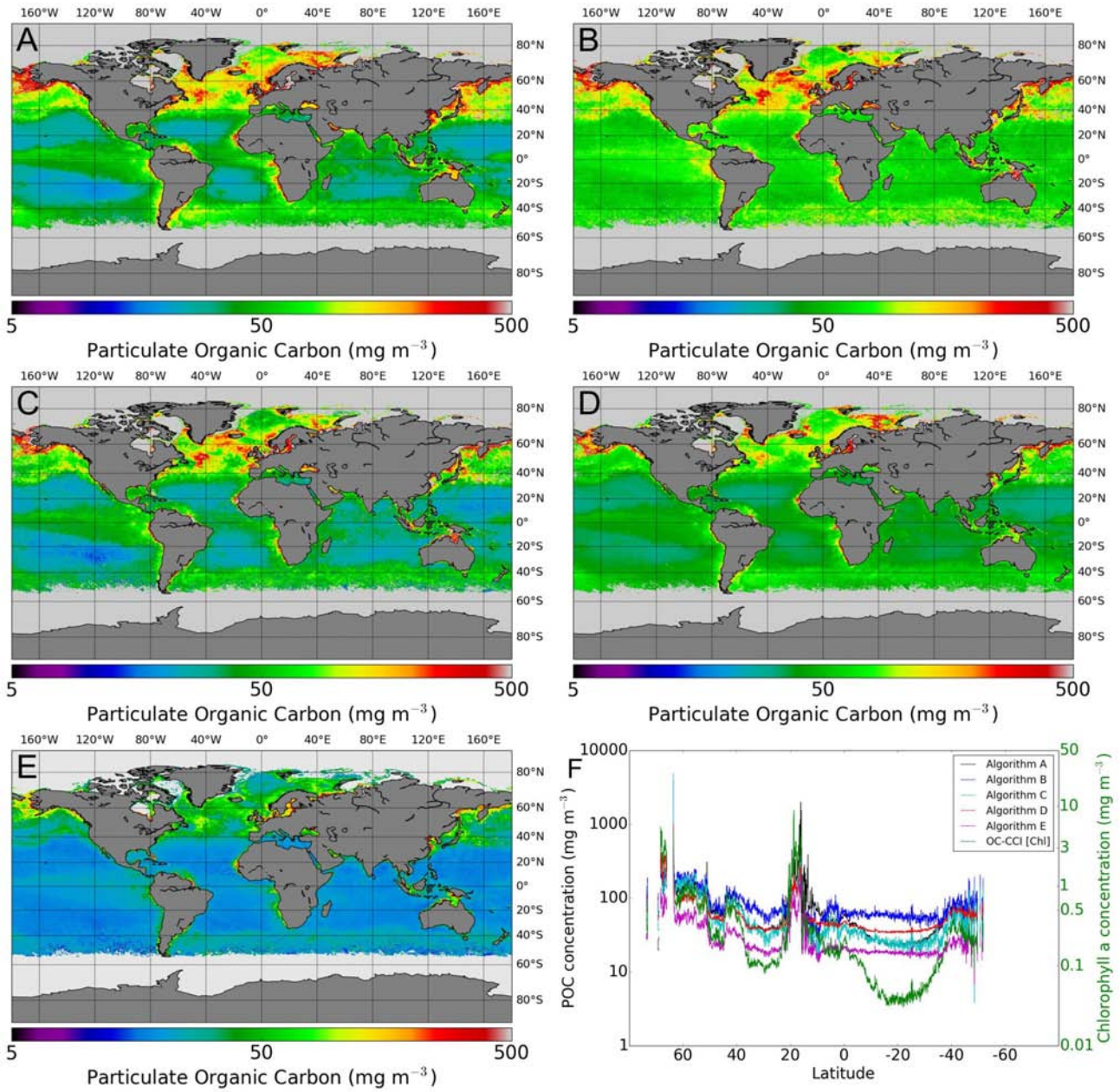


FIGURE 8 | POC estimated using the five candidate algorithms applied to a monthly composite of OC-CCI data from May 2005 **(A)** Stramski et al. (2008) (R_{rs}), **(B)** Stramski et al. (2008) (b_{bp}), **(C)** Loisel et al. (2002), **(D)** Gardner et al. (2006), **(E)** Kostadinov et al. (2016), and **(F)** POC associated with an extracted transect through the Atlantic at 20^{o} W for each algorithm, and the associated [Chl] from the OC-CCI data.

gyres, and higher concentrations in turbid shelf and coastal regions. However, there are some notable differences between the POC concentrations estimated by the different algorithms. Algorithm A and C perform similarly (**Figures 8A,C**). Algorithm B (**Figure 8B**) produces estimates that are generally higher relative to Algorithm A and C, particularly at low POC concentrations. Algorithm D (**Figure 8D**) underestimates at higher POC concentrations relative to all other algorithms, whilst at low concentrations its estimates are generally higher than algorithms A and C, and lower than algorithm B. In contrast, Algorithm E (**Figure 8E**) estimates lower POC concentrations relative to the other methods. A transect, extracted along from 20^{o} west, shows the regional differences in algorithm estimates for POC, and the associated OC-CCI [Chl] for reference (**Figure 8F**). Histograms of these products (not shown), show a similar range to the *in situ* data used for validation ($\approx$ 10–1,000), though values greater than 1,000 mg m^{-3} are scarce in the satellite products (though higher values

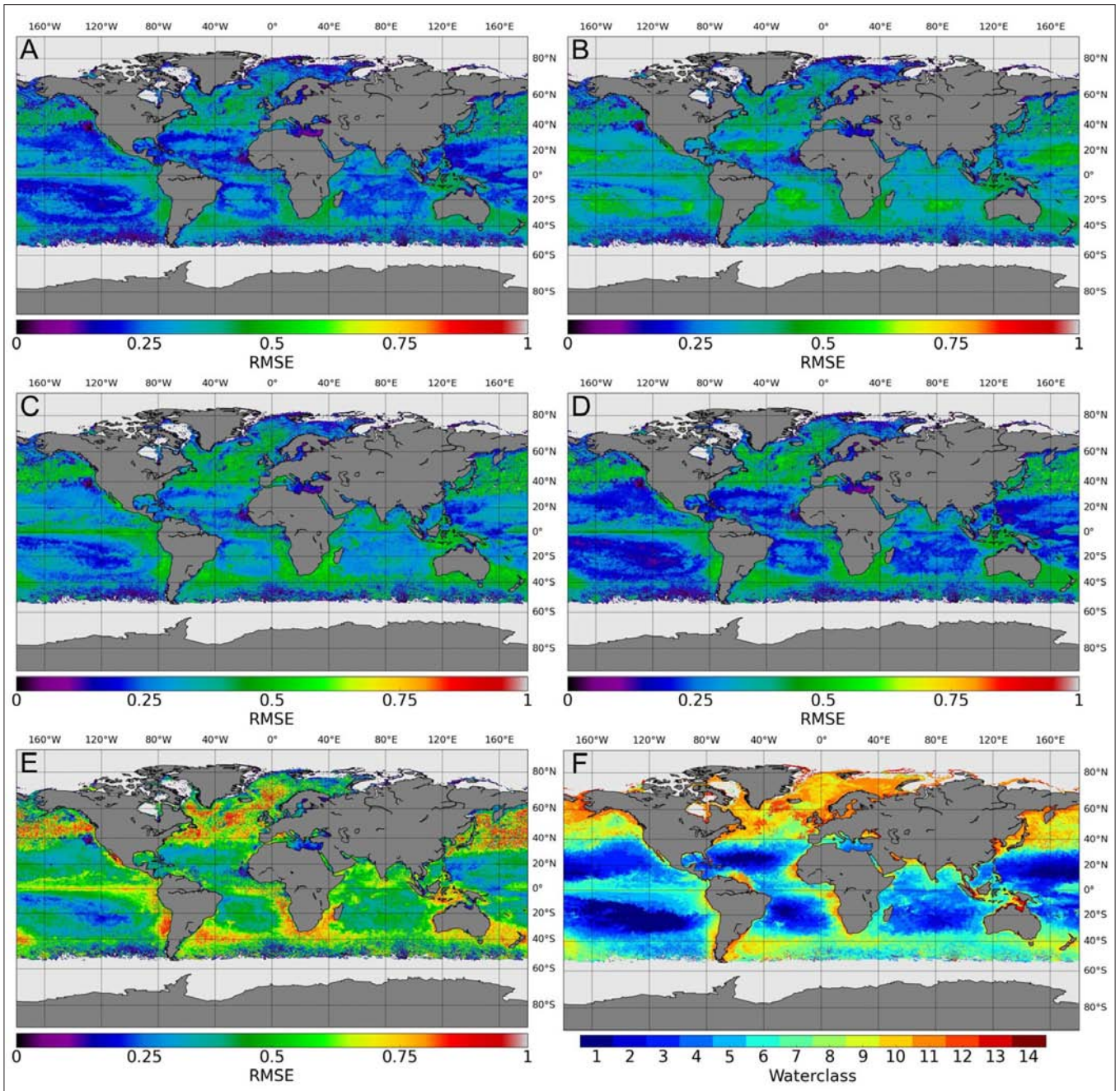


FIGURE 9 | Root Mean Squared Difference calculated for POC as estimated using water-class specific performance of the the five candidate algorithms applied to monthly composite OC-CCI data from May 2005 **(A)** Stramski et al. (2008) (R_{rs}), **(B)** Stramski et al. (2008) (b_{bp}), **(C)** Loisel et al. (2002), **(D)** Gardner et al. (2006), **(E)** Kostadinov et al. (2016), and **(F)** the associated OC-CCI water classes.

are more frequent with Algorithm C than with Algorithm A). Also, the pronounced bimodal frequency distribution of the *in situ* data is absent in the satellite products, which show a unimodal distribution. This difference between the frequency distribution of *in situ* data and satellite products could have had an impact on the comparative statistics presented here.

3.5. Mapped Uncertainties

Each algorithm has uncertainties associated with its performance for each water class, calculated from the validation exercise. These values can be used to estimate uncertainties for pixels outside of direct matchup locations, using a weighted average based on the percent membership to each of the classes. This procedure was applied to the data in **Figure 8** to calculate

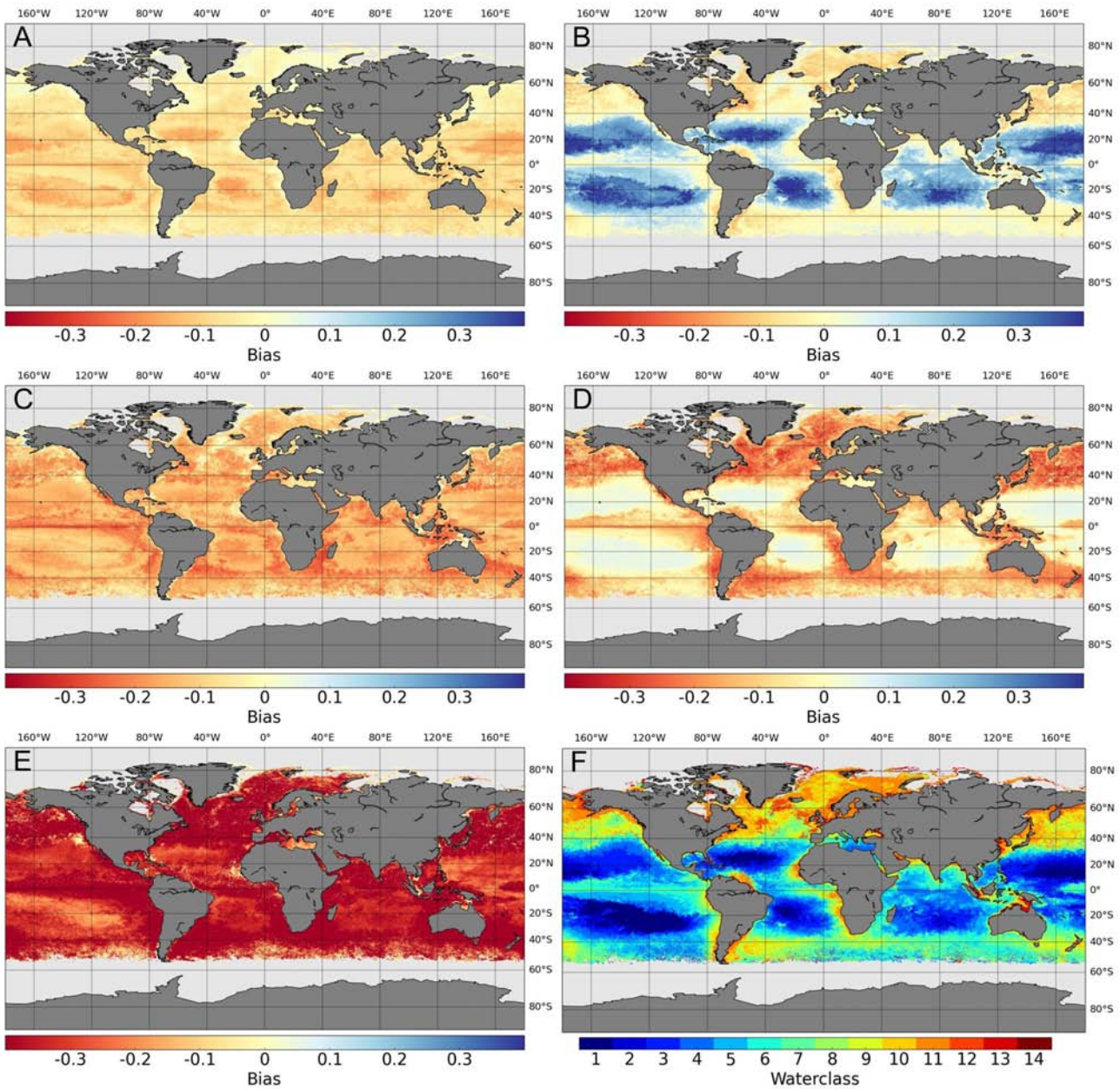


FIGURE 10 | Bias as estimated using per water class performance of the five candidate algorithms applied to monthly composite OC-CCI data from May 2005 **(A)** Stramski et al. (2008) (R_{rs}), **(B)** Stramski et al. (2008) (b_{bp}), **(C)** Loisel et al. (2002), **(D)** Gardner et al. (2006), **(E)** Kostadinov et al. (2016), and **(F)** the associated OC-CCI water classes.

per pixel RMSD and bias. As could be expected from the performance of the algorithms across the substantial *in situ* data set, Algorithm A (R_{rs} based algorithm of Stramski et al., 2008) shows low RMSD and bias when uncertainties are calculated (**Figures 9**, **10**). Algorithms B and particularly Algorithm C show higher RMSE, particularly in the gyre regions, consistent with the distribution of the matchup-based estimates in **Figure 3**. Algorithm D has low RMSD and bias in the oligotrophic gyres, but relatively higher values appear in the more productive (upwelling and coastal) areas. Algorithm E shows high RMSE throughout the image. Highest positive bias estimates are associated with algorithm B, relating to overestimation in the gyre regions (**Figure 10B**). Bias for Algorithm E (**Figure 10E**) shows negative bias globally, consistent with the general trend toward underestimation shown in **Figures 4E**, **8E**.

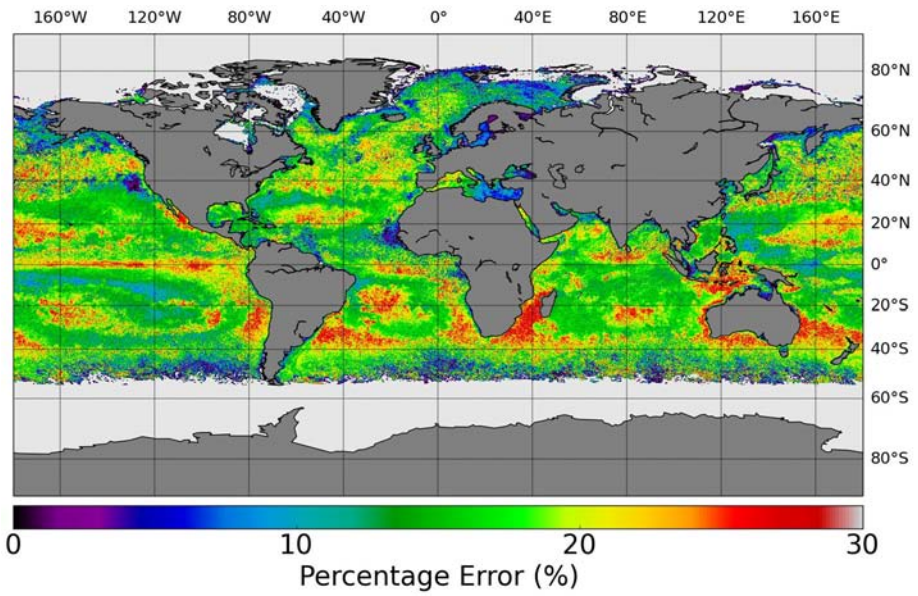


FIGURE 11 | Percentage error for POC as estimated using per water class performance of the Stramski et al. (2008) (R_{rs}) applied to monthly composite OC-CCI data from May 2005.

4. DISCUSSION

4.1. Variability in Algorithm Performance As a Result of Input Satellite Data, Choice of Optical Models, and Regional Optical Properties

The strength of the relationships between bio-optical properties and POC concentration has been quantified in the original studies where the algorithms examined above were formulated, and in some cases, validated against satellite data. To the extent that some of the satellite data used in the original studies are included in the match-up dataset used here, the impact on the results is likely to be small because the size of the match-up data used here is much larger than that used in any of the previous studies. Furthermore, the comparisons presented here are based on a common satellite product suite (OC-CCI). However, further insights are gained from the bigger *in situ* data base assembled for this study, and from the climate-quality data set provided by OC-CCI.

Results from the OC-CCI based validation are generally consistent with the observations made by Stramski et al. (2008). The *in situ* data used by Stramski (Stramski et al., 2008) were limited to the south eastern Pacific and eastern Atlantic Oceans and covered a POC range of 12—270 mg m^{-3}, whilst the range for the data used here cover a broader range (2.7–8,097 mg m^{-3}). The waters sampled by Stramski et al. (2008) ranged from upwelling to oligotrophic, with significant contributions of mineral particle matter to the particle assemblage at some stations. Consistent with the results here, Stramski et al. (2008) found that empirical relationships between R_{rs} and POC performed better than two-step approaches where an inherent optical property (IOP) is derived from the R_{rs} and then related to the POC. They also indicated relatively better performance of the R_{rs} relationship over that derived from b_{bp}, highlighting the uncertainties in the derivation of b_{bp} as one source of error in estimation of POC from IOPs. Additionally, the relationships between POC and IOPs would be expected to vary as a result of the particle size distribution (PSD), the refractive index of particles, and the fractional POC concentrations within different particle types in the assemblage. For example, significant variations in the POC-specific backscattering coefficient has been reported for different water bodies of the Southern Ocean (see Figure 1 in Stramski et al., 1999). Whilst variability in the POC-specific backscattering introduces uncertainty in total POC estimates, a better understanding of the relationship between particle characteristics and IOPs has the potential to provide further insight into the composition of the POC pool, and therefore to improve POC algorithms. Hence, it is important to pursue this line of algorithm development, even if the current performance of these methods might not be as good as that of some more empirical approaches. The line of investigation that accounts for the contributions of different types of POC to their optical properties is already yielding fruit (Stramski et al., 2008).

The algorithm of Loisel et al. (2002) (Algorithm C) is also a two-step approach, drawing on the relationship between b_p and POC, via a relationship between b_{bp} and [Chl]. Though Loisel et al. (2002) did not directly validate their POC estimation from SeaWiFS data, they found a good match between retrieved b_{bp} and that measured *in situ* in a previous study (Loisel et al., 2001). Loisel et al. (2002) did indicate variability in the b_{bp}:[Chl] relationship, linked to changes in the particulate pool; they highlighted the variable influence of small particles consisting of dead cells, grazers, and minerals. Gardner et al. (2006)

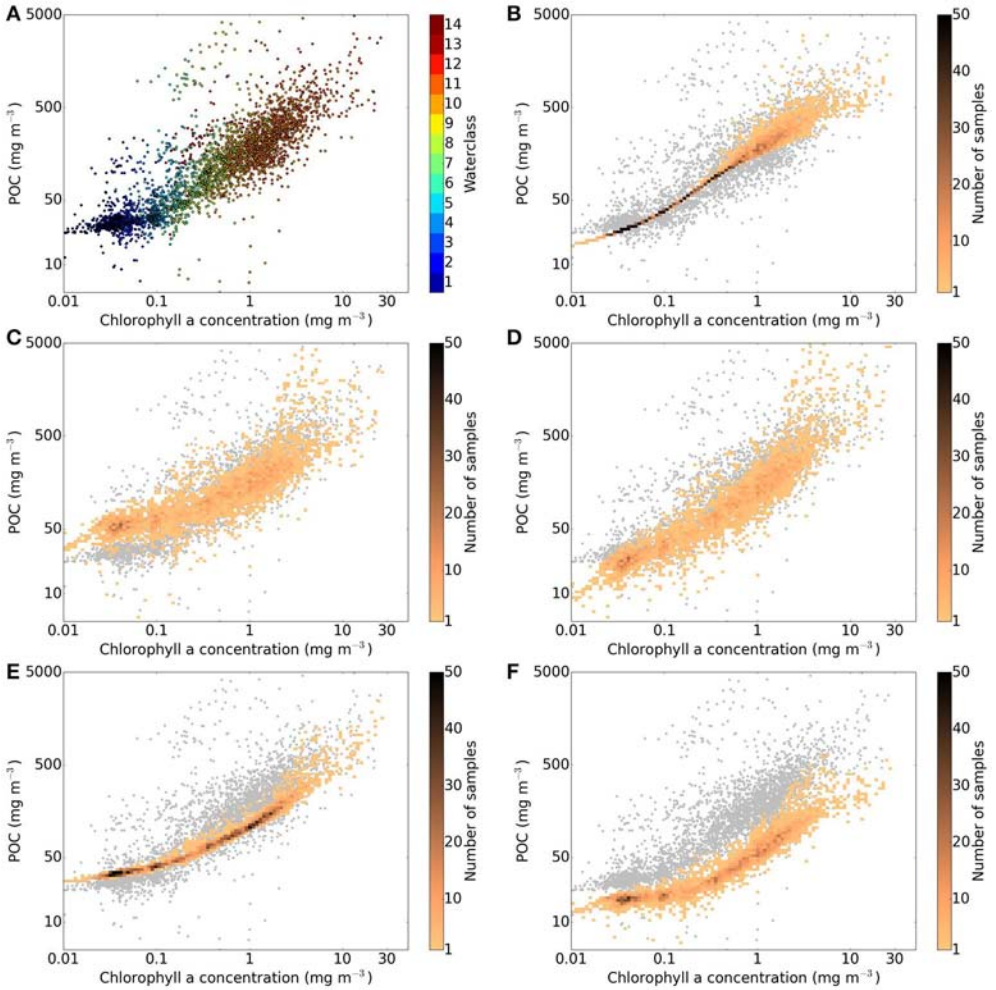


FIGURE 12 | Covariance between POC and [Chl] extracted from the OC-CCI matchups with the *in situ* database for **(A)** *in situ* POC data, **(B)** POC estimated using Stramski et al. (2008) (R_{rs}), **(C)** POC estimated using Stramski et al. (2008) (b_{bp}), **(D)** POC estimated using Loisel et al. (2002), **(E)** POC estimated using Gardner et al. (2006), **(F)** POC estimated using Kostadinov et al. (2016).

(Algorithm D) also uses a two-step approach, exploiting the relationship between the beam attenuation coefficient (c_p) and POC. This relationship was shown to be strong, when *in situ* POC was compared with transmissometer profiles. Although no fully-validated algorithm exists for routine derivation of c_p from satellite ocean color measurements, Gardner et al. (2006) showed *in situ* c_p was strongly correlated with [Chl] (r = 0.845–0.897) and $K_d(490)$) (r = 0.846-0.878) derived from SeaWIFS data over different oceanic regions.

Algorithm E, by Kostadinov et al. (2016), addresses some sources of variability between optical properties and POC, such as the influence of the particle size distribution, which was also identified as being important by Stramski et al. (2008). The method of Kostadinov et al. (2016) uses spectral values of b_{bp} to derive a PSD, which is then converted to POC (and phytoplankton carbon) using allometric relationships. The focus of the Kostadinov et al. (2016) paper was on phytoplankton carbon, computed as 1/3 of POC. Relationships between the phytoplankton carbon estimated from *in situ* PSD measurements and direct analytical determinations, showed r values between 0.5 and 0.714, depending on the limits of integration of the PSD, with wider limits resulting in the lower r. As discussed by Stramski et al. (2008), Kostadinov et al. (2016) also notes the impact of uncertainties in retrieved backscattering arising from both measurement and theory. In particular, assumptions of sphericity and homogeneity used in Mie theory are likely to be violated in real seawater particle assemblages, particularly for backscattering and in coastal and more productive areas (which are included in the database used here). For a more detailed discussion of the sphericity and homogeneity assumption, see Kostadinov et al. (2009) and refs. therein. Future work needs to focus on developing and more widely adopting bio-optical models that relax the Mie assumptions (e.g., Quirantes and Bernard, 2004, 2006; Clavano et al., 2007; Matthews and Bernard, 2013; Robertson Lain et al., 2017). Understanding of PSD variability, how it relates to backscattering, and how

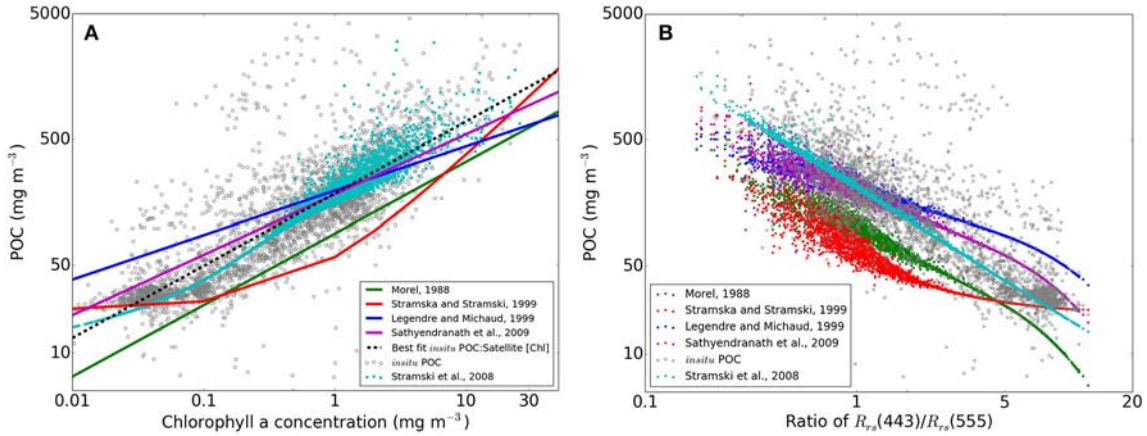


FIGURE 13 | Summary of **(A)** POC to [Chl] relationships including the *in situ* POC data collated in this study and the extracted OC-CCI [Chl (gray dots)] and the relationships proposed by Morel (1988) (green line), Stramska and Stramski (2005) (red line), Legendre and Michaud (1999) (blue line), and Sathyendranath et al. (2009) (pink line), a best fit line for the *in situ* data vs. the OC-CCI chlorophyll (dashed line), and the estimated POC using Stramski et al. (2008) (cyan dots) are provided for context, **(B)** estimated POC from the previously listed approaches, plotted against the R_{rs} (443) to R_{rs}(555) to show relationship to this ratio commonly used in algorithms to derive [Chl].

particle composition affects scattering over broad marine regions are required to develop further such detailed mechanistic approaches.

General sources of error associated with any ocean-color product include differences introduced by choice of sensor, sensor calibration, and the atmospheric correction procedure used to retrieve R_{rs}. In addition to these, a further consideration, particularly in the cases where algorithms use IOPs, is the methods used to derive the IOP product from the R_{rs} data. The OC-CCI processing uses the Quasi-Analytical Algorithm (QAA) of Lee et al. (2002) to calculate IOPs, including the b_{bp} values used in this study. The original study by Kostadinov et al. (2016) used the method of Loisel and Stramski (2000) to estimate b_{bp}. Stramski et al. (2008) also used different formulations to calculate b_{bp} from R_{rs}, finding a corrected version of QAA produced a better estimate of b_{bp}, and a strong relationship with POC (r = 0.933). The effect of the choice of method to derive b_{bp} on the POC estimates requires further consideration, which goes beyond the scope of this study, as this IOP is particularly poorly understood and validated (Lee et al., 2002). The differences in algorithm performance across the different water classes indicate that regional variability in performance of the different algorithms can be expected. This is confirmed in the mapped regional distribution of uncertainties (**Figures 9**, **10**). These results suggest that algorithms either need to be selected carefully for applications in different regions, or that a selection of optimal algorithms may have to be blended for a global product (as done in Jackson et al., in press). This point is also raised in Stramski et al. (2008), where different formulations are provided for global application, and excluding upwelling data. Uncertainties in the underlying satellite data may also be responsible for a portion of this variability: for example, an IOP model may be more or less suitable to derive backscattering. It should also be noted that there can also be uncertainties in the *in situ* data and the validation process that can affect the assessment of uncertainties in algorithm performance. Ideally, multiple replicates will be taken to quantify uncertainties in the *in situ* measurement, and instruments will have a well-characterized calibration history, and be processed with community endorsed methodologies. For POC, the issue of blank correction was already highlighted in Section 2.1. Uncertainties resulting from variable methods used for the *in situ* data collated for this study may influence the results presented here, particularly at low POC concentrations. In terms of comparison to matchups, further uncertainties can be introduced by comparing values at different scales, i.e., point measurements may not represent the average over a pixel (in this case of 4 km in size). These uncertainties will limit the ultimate accuracy to which any satellite based product can be derived and validated. However, issues of spatial mismatch are beginning to be addressed with the use of underway systems (for example, Brewin et al., 2016).

Despite the difficulties highlighted above, the overall performance of the algorithms studied here is encouraging. Percentage error estimates based on the OC-CCI methodology show how well these algorithms can generate products suitable for the needs of the scientific community. For example, the percentage errors associated with the Stramski et al. (2008) R_{rs} algorithm applied to OC-CCI data in May 2005 (**Figure 11**), show that a majority of pixels fall within an error range that is widely accepted by the ocean color community for [Chl] (30%; GCOS, 2011).

4.2. Variability in the Ratio of Particulate Organic Carbon to Chlorophyll-a

Further perspective on the performance of the different algorithms can be gained by considering the covariance between POC and [Chl]. The relationship between the *in situ* POC data and satellite [Chl] is shown in **Figure 12A**, where the color indicates the associated dominant optical water class. These data

then forms the background for each of the subsequent panels of **Figure 12**, which show the relationship between the POC estimated by each algorithm, and the satellite [Chl]. Algorithm A shares a commonality in method with the algorithms used to derive satellite [Chl], in that the same reflectance ratios are used to derive POC, and [Chl] (at lower concentrations); hence it shows a very constrained relationship in this domain of the parameter space (**Figure 12B**). Other algorithms capture the scatter in the POC:[Chl] relationship to a greater or lesser degree, though offsets can be seen, associated with the behavior identified in the validation exercise, i.e., overestimation of POC relative to lower [Chl] in the case of Algorithm B (**Figure 12C**), and underestimation of the ratio at low [Chl] using POC from Algorithm C (**Figure 12D**—though it should also be noted that this algorithm is also dependent on [Chl] to derive POC). As with Algorithm A, Algorithm D shows a relatively constrained relationship between POC and [Chl]. Algorithm E produces similar variability between POC and [Chl] as seen in the *in situ* data, in terms of shape and scatter of the curve, but the bias of this algorithm toward lower estimates of satellite derived POC is clear (**Figure 12D**).

The ratio of POC to [Chl] is important in the context of the discussion here for two reasons. Firstly, this ratio is important in the context of biogeochemical modeling, and the ecological and physiological processes that influence this ratio. Secondly, empirical relationships between POC and chlorophyll have been developed, which can be applied to satellite derived estimates of [Chl]. As mentioned above, these algorithms are typically similar to those employing blue:green reflectance ratios (e.g., Algorithm A from Stramski et al., 2008), and as such were not initially considered in the algorithm intercomparison here. **Figure 13A** shows a number of these empirical relationships, against a background of the same *in situ* POC and [Chl] data as shown in **Figure 12**. **Figure 13B** shows POC estimated using these [Chl]-based algorithms on OC-CCI [Chl] as a function of the blue-green R_{rs} reflectance ratio. The same reflectance ratio is employed by Stramski et al. (2008) to derive POC, and is also used in a number of empirical [Chl] algorithms. A linear regression of the *in situ* POC concentrations, against the satellite derived [Chl], results in an r^2 value of 0.70. Using the various relationships shown in **Figure 13A** to estimate POC based on the satellite [Chl] returns r^2 values between 0.63 and 0.69, lower than those returned for all the other algorithms assessed. The [Chl] based approaches show in **Figure 13** produce RMSD values (between 0.27 and 0.47) and bias (between −0.03 and 0.117) in the same range as the other algorithms.

Even though to first order Chl and POC are positively correlated in the global ocean, a residual scatter in the relationship remains (e.g., in satellite observations—**Figure 12A**, and *in situ* observations as well—e.g., Kostadinov et al., 2012). Ideally, a POC algorithm should be able to retrieve POC independently of [Chl] and capture the variable POC/[Chl] ratio correctly. Note that this ratio can vary due to both variability in the fraction of living phytoplankton carbon in the total POC pool, due to the physiology and photoacclimation of the phytoplankton component of POC (Geider, 1987; Geider et al., 1998; Behrenfeld et al., 2005), and species specific differences among phytoplankton themselves (Stramski, 1999). Therefore, independent knowledge of total POC, living phytoplankton carbon, and [Chl a] should be the goal of future bio-optical algorithm development.

4.3. Estimates of Total Pools of Carbon

The OC-CCI archive can be used to estimate total pools of POC in the mixed layer, taking into account interannual and regional variability, which is well captured by this merged dataset. Algorithms A-D were applied to the monthly OC-CCI version 2 data, and the values integrated over the mixed-layer depth (derived from MIMOC, Schmidtko et al., 2013), assuming homogeneity over the mixed layer. These were then averaged over all the months and for all the years of the OC-CCI version 2 (1998-2012) to provide estimates of the average standing pool of POC as follows: Algorithm A: 0.86 Pg C, Algorithm B: 1.3 Pg C, Algorithm C: 0.87 Pg C, Algorithm D: 0.77 Pg C. These are larger than the estimate of Gardner et al. (2006) and smaller than the estimate of Stramska (2009). Comparison of these estimates with those of phytoplankton carbon pools estimated in a parallel study (Martinez-Vicente et al., in review), indicates that phytoplankton carbon represents between 17 and 48% of the total POC pool. Whilst this ratio shows considerable variability, the often assumed value of 1/3 for phytoplankton carbon:POC falls within this range. High levels of variability in the phytoplankton carbon to POC ratio were also observed *in situ* by Graff et al. (2015). Satellite based estimates calculated by Kostadinov et al. (2016) (using a different set of mixed layer depth values) suggest a phytoplankton carbon standing stock of around 0.24 Pg C, implying a corresponding POC stock of around 0.72 Pg C when using the 1/3 assumption. Kostadinov et al. (2016) showed the estimated phytoplankton standing stock to be similar to estimates derived from the application of both the Stramski et al. (2008) b_{bp} based algorithm, combined again with a 1/3 assumption and the method of Behrenfeld et al. (2005) to SeaWIFS data, and to model estimates from the Coupled Model Intercomparison Project 5 (CMIP5). The estimate of phytoplankton carbon standing stock from Kostadinov et al. (2016) is similar to that estimated by other size class based approaches, such as that of Roy et al. (2017) which used size classes derived from absorption to estimate a total phytoplankton carbon stock of 0.26 Pg C. Though the global estimates of POC from the different approaches assessed here are quite similar to each other, the differences are more pronounced at smaller scales, as can be seen in **Figure 8F**.

5. CONCLUSIONS

A variety of POC algorithms were applied to matchup pixels extracted from the satellite OC-CCI ocean color data, and validated against *in situ* data. The database used here represents the largest collection of *in situ* POC data available, to the author's knowledge. The five algorithms showed strong predictive capacity for estimating POC, with Algorithm A (based on R_{rs}—Stramski et al., 2008) and C (based on Loisel et al., 2002) performing well across the broad range of the *in situ* dataset. Algorithms A and C performed consistently across different water types as defined in the OC-CCI data. From the water class

based validation, errors can be estimated per pixel. For Algorithm A and C, the errors were mostly within the range requested by the user community. These results suggest a maturity in POC algorithms and their suitability for production of long term time series for climate related studies. However, several key points of development are highlighted from the inter comparison of the different algorithms and the various studies reviewed here. Greater knowledge of the composition of the particulate pool, and how it affects the IOPs of the oceans, may allow increased accuracy of POC algorithms (within the constraints of the sensitivities of current satellite ocean color radiometry), as well as providing further information on different types of particles, many of which play important roles in water quality and ocean biogeochemistry. To support this aim, further *in situ* data should be collected, including additional measurements to provide detail on phytoplankton community size structure, physiology, and photoacclimation. Further, it is recommended that future work seeks to use consistent methodology for blank correction of POC measurements, and clarify any trends in the low POC region which may be influenced by these uncertainties. Further understanding of the sources of variability between POC and optical parameters can then be incorporated in to future, semi-analytical algorithms. New understanding of these relationships may also inform future sensor development (e.g., hyperspectral sensors) and optical modeling techniques.

AUTHOR CONTRIBUTIONS

HE: All the calculations, preparation of figures, lead on writing the manuscript. VM: Project manager for Pools of Carbon project; development of matchup processing and statistical analysis. RB: Provision of code for statistical analysis based on OC-CCI methodology. GD: Provision of *in situ* data, content on b_{bp} uncertainties and impact of particle sizing. AH: Perspectives on community requirements, particularly for ecosystem modeling. TJ: Provision of code for calculation of per optical water class uncertainties, based on the OC-CCI methodology. HK: Collation of the *in situ* database. TK: Algorithm provider, content on algorithm performance relating to particle size distributions. HL: Algorithm provider, content on regional variability in algorithm performance. SR: Input on statistical analysis and relative contributions of phytoplankton to the POC pool. RR: Collation of the *in situ* database. ST: Data contribution and guidance on POC data for the Southern Ocean. DS: Algorithm provider, content on history of POC algorithm development, and interpretation of comparative analysis of different algorithms. TP: Project leader, scientific advice. SS: Development of the concept and work plan, guidance of HEK, review and rewriting of various sections of manuscript. All authors reviewed and provided comments on the draft manuscript.

FUNDING

The POCO project is funded by the European Space Agency (ESA) under the program of Science Exploitation of Operational Missions (SEOM) following Contract: 4000113692/15/I-LG. This study is a contribution to the Ocean Color Climate Change Initiative of the European Space Agency, and to the activities of the National Center for Earth Observations (NCEO) of the Natural Environmental Research Council (NERC) of UK. This study is also a contribution to the international IMBER project and was supported by the UK Natural Environment Research Council National Capability funding to Plymouth Marine Laboratory and the National Oceanography Center, Southampton. AMT data were supported by the UK Natural Environment Research Council National Capability funding to Plymouth Marine Laboratory and the National Oceanography Center, Southampton. This is contribution number 309 of the AMT programme. TK was supported on NASA grant #NNX13AC92G and by the Division of Hydrologic Sciences, Desert Research Institute. The contribution of HL was funded by the CNES/TOSCA program in the frame of the COYOTE project.

ACKNOWLEDGMENTS

The authors would like to thank Peter Regner for his support and management of the POCO project. We would like to thank Oliver Fisher for his contributions to the project during his student internship. The authors would like to thank the participants of the Color and Light in the ocean from Earth Observation (CLEO) workshop, for their valuable discussions on POC, which contributed substantially to refining the approaches presented in this work. The authors would also like to thank the two reviewers who provided detailed and constructive comments which substantially improved this manuscript.

REFERENCES

Allison, D. B., Stramski, D., and Mitchell, B. G. (2010). Empirical ocean color algorithms for estimating particulate organic carbon in the Southern Ocean. *J. Geophys. Res.* 115, C10044. doi: 10.1029/2009JC006040

Behrenfeld, M. J., Boss, E., Siegel, D. A., and Shea, D. M. (2005). Carbon-based ocean productivity and phytoplankton physiology from space. *Glob. Biogeochem. Cycles* 19:GB1006. doi: 10.1029/2004GB002299

Bishop, J. (1999). Transmissometer measurement of poc. *Deep Sea Res. I* 46, 353–369. doi: 10.1016/S0967-0637(98)00069-7

Bohren, C. F., and Huffman, D. (1983). *Absorption and Scattering of Light by Small Particles*. New York, NY: Wiley.

Boss, E., and Pegau, W. S. (2001). Relationship of light scattering at an angle in the backward direction to the backscattering coefficient. *Appl. Opt.* 40, 5503–5507. doi: 10.1364/AO.40.005503

Boyd, P., and Newton, P. (1995). Evidence of the potential influence of planktonic community structure on the interannual variability of particulate organic carbon flux. *Deep-Sea Res. I* 42, 619–639. doi: 10.1016/0967-0637(95)00017-Z

Brewin, R. J. W., Dall'Olmo, G., Pardo, S., van Dongen-Vogels, V., and Boss, E. S. (2016). Underway spectrophotometry along the Atlantic Meridional Transect reveals high performance in satellite chlorophyll retrievals. *Remote Sens. Environ.* 183, 82–97. doi: 10.1016/j.rse.2016.05.005

Brewin, R. J. W., Sathyendranath, S., Müller, D., Brockmann, C., Deschamps, P.-Y., Devred, E., et al. (2015). The ocean colour climate change initiative: Iii. a round-robin comparison on in-water bio-optical algorithms. *Remote Sens. Environ.* 162, 271–294. doi: 10.1016/j.rse.2013.09.016

Buiteveld, H., Hakvoort, J. H. M., and Donze, M. (1994). "Optical properties of pure water," in *Proc. SPIE, 2258, Ocean Optics XII*, ed J. S. Jaffe (Bergen: International Society for Optics and Photonics), 174–185.

Cetinić, I., Perry, M. J., Briggs, N. T., Kallin, E., D'Asaro, E. A., and Lee, C. M. (2012). Particulate organic carbon and inherent optical properties during 2008 North Atlantic Bloom Experiment. *J. Geophys. Res. Oceans* 117:C06028. doi: 10.1029/2011JC007771

Cho, B. C., and Azam, F. (1990). Biogeochemical significance of bacterial biomass in the ocean's euphotic zone. *Mar. Ecol. Prog. Series* 63, 253–259. doi: 10.3354/meps063253

Claustre, H., Morel, A., Babin, M., Cailliau, C., Marie, D., Marty, J. C., et al. (1999). Variability in particle attenuation and chlorophyll fluorescence in the tropical pacific: scales, patterns, and biogeochemical implications. *J. Geophys. Res. Oceans* 104, 3401–3422. doi: 10.1029/98JC01334

Clavano, W. R., Boss, E., and Karp-Boss, L. (2007). Inherent optical properties of non-spherical marine-like particles - From theory to observation. *Oceanogr. Mar. Biol.* 45, 1–38. doi: 10.1201/9781420050943.ch1

Dall'Olmo, G., Westberry, T., Behrenfeld, M. J., Boss, E., and Slade, W. (2009). Significant contribution of large particles to optical backscattering in the open ocean. *Biogeosciences* 6, 947–967. doi: 10.5194/bg-6-947-2009

Ducklow, H., Steinberg, D. K., and Buesseler, K. O. (2001). Upper ocean carbon export and the biological pump. *Oceanography* 14, 50–58. doi: 10.5670/oceanog.2001.06

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Ann. Stat.* 7, 1–26. doi: 10.1214/aos/1176344552

Efron, B., and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. Boca Raton, FL: Chapman and Hall/CRC.

Gardner, W., Mishonov, A., and Richardson, M. J. (2006). Global poc concentrations from *in situ* and satellite data. *Deep Sea Res. II* 53, 718–740. doi: 10.1016/j.dsr2.2006.01.029

Gardner, W., Walsh, I., and Richardson, M. J. (1993). Biophysical forcing of particle production and distribution during a spring bloom in the north atlantic. *Deep Sea Res. II* 40, 171–195. doi: 10.1016/0967-0645(93)90012-C

GCOS (2011). *Systematic Observation Requirements from Satellite-Based Data Products for Climate 2011 Update. Supplemental Details to the Satellite-Based Component of the "Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC*. Technical report, No. 154, World Meteorological Organisation (WMO), Geneva 2, Switzerland.

Geider, R. J. (1987). Light and temperature dependence of the carbon to chlorophyll a ratio in microalgae and cyanobacteria: implications for physiology and growth of phytoplankton. *New Phytol.* 106, 1–34. doi: 10.1111/j.1469-8137.1987.tb04788.x

Geider, R. J., Maclntyre, H. L., and Kana, T. M. (1998). A dynamic regulatory model of phytoplanktonic acclimation to light, nutrients, and temperature. *Limnol. Oceanogr.* 43, 679–694. doi: 10.4319/lo.1998.43.4.0679

Graff, J. R., Milligan, A. J., and Behrenfeld, M. J. (2012). The measurement of phytoplankton biomass using flow-cytometric sorting and elemental analysis of carbon. *Limnol. Oceanogr. Methods* 10, 910–920. doi: 10.4319/lom.2012.10.910

Graff, J. R., Westberry, T. K., Milligan, A. J., Brown, M. B., Dall'Olmo, G., Dongen-Vogels, V. V., et al. (2015). Analytical phytoplankton carbon measurements spanning diverse ecosystems. *Deep-Sea Res. Part I Oceanogr. Res. Pap.* 102, 16–25. doi: 10.1016/j.dsr.2015.04.006

Jackson, T., Sathyendranath, S., and Mélin, F. (in press). An improved optical classification scheme for the Ocean Colour Essential Climate Variable and its applications. *Rem. Sens. Environ*. doi: 10.1016/j.rse.2017.03.036

Kitchen, J. C., and Zaneveld, J. R. (1992). A three-layered sphere model of the optical properties of phytoplankton. *Limnol. Oceanogr.* 37:1680. doi: 10.4319/lo.1992.37.8.1680

Kostadinov, T. S., Milutinović, S., Marinov, I., and Cabré, A. (2016). Carbon-based phytoplankton size classes retrieved via ocean color estimates of the particle size distribution. *Ocean Sci.* 12, 561–575. doi: 10.5194/os-12-561-2016

Kostadinov, T. S., Siegel, D. A., and Maritorena, S. (2009). Retrieval of the particle size distribution from satellite ocean color observations. *J. Geophys. Res. Oceans* 114:C09015. doi: 10.1029/2009JC005303

Kostadinov, T. S., Siegel, D. A., Maritorena, S., and Guillocheau, N. (2012). Optical assessment of particle size and composition in the santa barbara channel, california. *Appl. Opt.* 51, 3171–3189. doi: 10.1364/AO.51.003171

Lee, Z., Arnone, R., Hu, C., Werdell, P. J., and Lubac, B. (2010). Uncertainties of optical parameters and their propagations in an analytical ocean color inversion algorithm. *Appl. Opt.* 49, 369–381. doi: 10.1364/AO.49.000369

Lee, Z., Carder, K. L., and Arnone, R. (2002). Deriving inherent optical properties from water color: a multiband quasi-analytical algorithm for optically deep waters. *Appl. Opt.* 41, 5755–5772. doi: 10.1364/AO.41.005755

Lee, Z., Du, K., and Arnone, R. (2005). A model for the diffuse attenuation coefficient of downwelling irradiance. *J. Geophys. Res.* 110:C02016. doi: 10.1029/2004JC002275

Legendre, L., and Michaud, J. (1999). Chlorophyll a to estimate the particulate organic carbon available as food to large zooplankton in the euphotic zone of oceans. *J. Plankton Res.* 21, 2067–2083. doi: 10.1093/plankt/21.11.2067

Loisel, H., Nicolas, J. M., Deschamps, P. Y., and Frouin, R. (2002). Seasonal and inter-annual variability of particulate organic matter in the global ocean. *Geophys. Res. Lett.* 29, 491–494. doi: 10.1029/2002GL015948

Loisel, H., and Stramski, D. (2000). Estimation of the inherent optical properties of natural waters from the irradiance attenuation coefficient and reflectance in the presence of Raman scattering. *Appl. Opt.* 39, 3001–3011. doi: 10.1364/AO.39.003001

Loisel, H., Stramski, D., Mitchell, B. G., Fell, F., Fournier-Sicre, V., Lemasle, B., et al. (2001). Comparison of the ocean inherent optical properties obtained from measurements and inverse modeling. *Appl. Opt.* 40, 2384–2397. doi: 10.1364/AO.40.002384

Maffione, R. A., and Dana, D. R. (1997). Instruments and methods for measuring the backward-scattering coefficient of ocean waters. *Appl. Opt.* 36, 6057–6067. doi: 10.1364/AO.36.006057

Maritorena, S., d'Andon, O. H. F., Mangin, A., and Siegel, D. A. (2010). Merged satellite ocean color data products using a bio-optical model: characteristics, benefits and issues. *Remote Sens. Environ.* 114, 1791–1804. doi: 10.1016/j.rse.2010.04.002

Martiny, A. C., Vrugt, J. A., and Lomas, M. W. (2014). Concentrations and ratios of particulate organic carbon, nitrogen, and phosphorus in the global ocean. *Sci. Data* 1:140048. doi: 10.1038/sdata.2014.48

Matthews, M., and Bernard, S. (2013). Using a two-layered sphere model to investigate the impact of gas vacuoles on the inherent optical properties of M. aeruginosa. *Biogeosciences* 10, 10531–10579. doi: 10.5194/bgd-10-10531-2013

Mel'nikov, I. A. (1976). Finely dispersed fraction of suspended organic matter in eastern Pacific ocean waters. *Oceanology* 15, 182–186.

Menden-Deuer, S., and Lessard, E. J. (2000). Carbon to volume relationships for dinoflagellates, diatoms, and other protist plankton. *Limnol. Oceanogr.* 45, 569–579. doi: 10.4319/lo.2000.45.3.0569

Meyer, R. A. (1979). Light scattering from biological cells: dependence of backscatter radiation on membrane thickness and refractive index. *Appl. Opt.* 18, 585–588. doi: 10.1364/AO.18.000585

Moore, T. S., Campbell, J. W., and Dowell, M. D. (2009). A class-based approach to characterizing and mapping the uncertainty of the MODIS ocean chlorophyll product. *Remote Sens. Environ.* 113, 2424–2430. doi: 10.1016/j.rse.2009.07.016

Morel, A. (1974). "Optical properties of pure water and pure seawater," in *Optical Aspects of Oceanography,* eds N. G. Jerlov and N. E. Steemann (New York, NY: Academic Press), 1–24.

Morel, A. (1988). Optical modeling of the upper ocean in relation to its biogenous matter content (Case I waters). *J. Geophysi. Res.* 93, 10749–10768. doi: 10.1029/JC093iC09p10749

Mueller, J. L. (2000). "SeaWiFS algorithm for the diffuse attenuation coefficient, K (490), using water-leaving radiances at 490 and 555 nm," in *SeaWiFS Postlaunch Calibration and Validation Analyses, Part,* eds S.B. Hooker and E. R. Firestone (Greenbelt, MD: NASA), 24–27.

Müller, D., Krasemann, H., Brewin, R. J. W., Brockmann, C., Deschamps, P.-Y., Doerffer, R., et al. (2015). The Ocean Colour Climate Change Initiative: I. A methodology for assessing atmospheric correction processors based on *in situ* measurements. *Remote Sens. Environ.* 162, 242–256. doi: 10.1016/j.rse.2013.11.026

O'Reilly, J. E., Maritorena, S., Mitchell, B. G., Siegel, D. A., Carder, K. L., Garver, S. A., et al. (1998). Ocean color chlorophyll algorithms for SeaWiFS. *J. Geophys. Res. Oceans* 103, 24937–24953. doi: 10.1029/98JC02160

Quirantes, A., and Bernard, S. (2004). Light scattering by marine algae: two-layer spherical and nonspherical models. *J. Quant. Spectros. Radiat. Trans.* 89, 311–321. doi: 10.1016/j.jqsrt.2004.05.031

Quirantes, A., and Bernard, S. (2006). Light-scattering methods for modelling algal particles as a collection of coated and/or nonspherical scatterers. *J. Quant. Spectrosc. Radiat. Trans.* 100, 315–324. doi: 10.1016/j.jqsrt.2005.11.048

Robertson Lain, L., Bernard, S., and Matthews, M. W. (2017). Understanding the contribution of phytoplankton phase functions to uncertainties in the water colour signal. *Opt. Express* 25, A151–A165. doi: 10.1364/OE.25.00A151

Roy, S., Sathyendranath, S., and Platt, T. (2017). Size-partitioned phytoplankton carbon and carbon-to-chlorophyll ratio from ocean colour by an absorption-based bio-optical algorithm. *Remote Sens. Environ.* 194, 177–189. doi: 10.1016/j.rse.2017.02.015

Sathyendranath, S., Brewin, R. J. W., Jackson, T., Mélin, F., and Platt, T. (in press). Ocean-colour products for climate-change studies: what are their ideal characteristics? *Remote Sens. Environ.* doi: 10.1016/j.rse.2017.04.017

Sathyendranath, S., Groom, S., Grant, M., Brewin, R. J. W., Thompson, A., Chuprin, A., et al. (2016). *ESA Ocean Colour Climate Change Initiative (Ocean-Colour-cci): Version 2.0 Data.* Centre for Environmental Data Analysis. doi: 10.5285/b0d6b9c5-14ba-499f-87c9-66416cd9a1dc

Sathyendranath, S., Stuart, V., Nair, A., Oka, K., Nakane, T., Bouman, H., et al. (2009). Carbon-to-chlorophyll ratio and growth rate of phytoplankton in the sea. *Mar. Ecol. Prog. Series* 383, 73–84. doi: 10.3354/meps07998

Schmidtko, S., Johnson, G. C., and Lyman, J. M. (2013). MIMOC: a global monthly isopycnal upper-ocean climatology with mixed layers. *J. Geophys. Res. Oceans* 118, 1658–1672. doi: 10.1002/jgrc.20122

Sharp, J. (1974). Improved analysis for "particulate" organic carbon and nitrogen from seawater. *Limnol. Oceanogr.* 19, 984–989. doi: 10.4319/lo.1974.19.6.0984

Stramska, M. (2009). Particulate organic carbon in the global ocean derived from SeaWiFS ocean color. *Deep-Sea Res. Part I Oceanogr. Res. Pap.* 56, 1459–1470. doi: 10.1016/j.dsr.2009.04.009

Stramska, M., and Stramski, D. (2005). Variability of particulate organic carbon concentration in the north polar Atlantic based on ocean color observations with Sea-viewing Wide Field-of-view Sensor (SeaWiFS). *J. Geophys. Res.* 110:C10018. doi: 10.1029/2004JC002762

Stramski, D. (1999). Refractive index of planktonic cells as a measure of cellular carbon and chlorophyll a content. *Deep-Sea Res Part I-Oceanogra. Res. Pap.* 46, 335–351. doi: 10.1016/S0967-0637(98)00065-X

Stramski, D., and Kiefer, D. A. (1991). Light-scattering by microorganisms in the open ocean. *Prog. Oceanogr.* 28, 343–383. doi: 10.1016/0079-6611(91)90032-H

Stramski, D., Reynolds, R. A., Babin, M., Kaczmarek, S., Lewis, M. R., Rottgers, R., et al. (2008). Relationships between the surface concentration of particulate organic carbon and optical properties in the eastern south pacific and eastern atlantic oceans. *Biogeosciences* 5, 171–201. doi: 10.5194/bg-5-171-2008

Stramski, D., Reynolds, R. A., Kahru, M., and Mitchell, B. G. (1999). Estimation of particulate organic carbon in the ocean from satellite remote sensing. *Science* 285, 239–242. doi: 10.1126/science.285.5425.239

Thomalla, S. J., Ogunkoya, A. G., Vichi, M., and Swart, S. (2017). Using optical sensors on gliders to estimate phytoplankton carbon concentrations and chlorophyll-to-carbon ratios in the Southern Ocean. *Front. Mar. Sci.* 4, 133–119. doi: 10.3389/fmars.2017.00034

Twardowski, M. S., Boss, E., and Macdonald, J. B. (2001). A model for estimating bulk refractive index from the optical backscattering ratio and the implications for understanding particle composition in case I and case II waters. *J. Geophys. Res.* 106, 14129–14142. doi: 10.1029/2000JC000404

Volk, T., and Hoffert, M. I. (1985). Ocean carbon pumps: analysis of relative strengths and efficiencies in ocean-driven atmospheric CO_2 changes. *Carbon Cycle Atmosph. CO_2 Nat. Variat. Archean Present* 32, 99–110.

Werdell, P. J., and Bailey, S. W. (2005). An improved *in situ* bio-optical data set for ocean color algorithm development and satellite data product validation. *Remote Sens. Environ.* 98, 122–140. doi: 10.1016/j.rse.2005.07.001

Zhang, X., Hu, L., and He, M.-X. (2009). Scattering by pure seawater: effect of salinity. *Opt. Exp.* 17, 5698–5710. doi: 10.1364/OE.17.005698

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Evers-King, Martinez-Vicente, Brewin, Dall'Olmo, Hickman, Jackson, Kostadinov, Krasemann, Loisel, Röttgers, Roy, Stramski, Thomalla, Platt and Sathyendranath. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Marine Science

ORIGINAL RESEARCH
published: 22 August 2017
doi: 10.3389/fmars.2017.00272

Phytoplankton Group Identification Using Simulated and *In situ* Hyperspectral Remote Sensing Reflectance

*Hongyan Xi *, Martin Hieronymi, Hajo Krasemann and Rüdiger Röttgers*

Department of Remote Sensing, Institute of Coastal Research, Helmholtz-Zentrum Geesthacht, Center for Materials and Coastal Research, Geesthacht, Germany

OPEN ACCESS

Edited by:
Astrid Bracher, Alfred-Wegener-Institute Helmholtz Center for Polar and Marine Research, Germany

Reviewed by:
Maycira Costa, University of Victoria, Canada
Aleksandra Wolanin, Helmholtz-Zentrum Potsdam Deutsches Geoforschungszentrum (GFZ), Germany

***Correspondence:**
Hongyan Xi
hongyan.xi@hzg.de

Specialty section:
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

Received: *15 March 2017*
Accepted: *07 August 2017*
Published: *22 August 2017*

Citation:
Xi H, Hieronymi M, Krasemann H and Röttgers R (2017) Phytoplankton Group Identification Using Simulated and In situ Hyperspectral Remote Sensing Reflectance. Front. Mar. Sci. 4:272. doi: 10.3389/fmars.2017.00272

In the present study we investigate the bio-geo-optical boundaries for the possibility to identify dominant phytoplankton groups from hyperspectral ocean color data. A large dataset of simulated remote sensing reflectance spectra, $R_{rs}(\lambda)$, was used. The simulation was based on measured inherent optical properties of natural water and measurements of five phytoplankton light absorption spectra representing five major phytoplankton spectral groups. These simulated data, named as C2X data, contain more than 10^5 different water cases, including cases typical for clearest natural waters as well as for extreme absorbing and extreme scattering waters. For the simulation the used concentrations of chlorophyll *a* (representing phytoplankton abundance), Chl, are ranging from 0 to 200 mg m^{-3}, concentrations of non-algal particles, NAP, from 0 to 1,500 g m^{-3}, and absorption coefficients of chromophoric dissolved organic matter (CDOM) at 440 nm from 0 to 20 m^{-1}. A second, independent, smaller dataset of simulated $R_{rs}(\lambda)$ used light absorption spectra of 128 cultures from six phytoplankton taxonomic groups to represent natural variability. Spectra of this test dataset are compared with spectra from the C2X data in order to evaluate to which extent the five spectral groups can be correctly identified as dominant under different optical conditions. The results showed that the identification accuracy is highly subject to the water optical conditions, i.e., contribution of and covariance in Chl, NAP, and CDOM. The identification in the simulated data is generally effective, except for waters with very low contribution by phytoplankton and for waters dominated by NAP, whereas contribution by CDOM plays only a minor role. To verify the applicability of the presented approach for natural waters, a test using *in situ* $R_{rs}(\lambda)$ dataset collected during a cyanobacterial bloom in Lake Taihu (China) is carried out and the approach predicts blue cyanobacteria to be dominant. This fits well with observation of the blue cyanobacteria *Microcystis* sp. in the lake. This study provides an efficient approach, which can be promisingly applied to hyperspectral sensors, for identifying dominant phytoplankton spectral groups purely based on $R_{rs}(\lambda)$ spectra.

Keywords: ocean color, remote sensing, phytoplankton spectral groups, light absorption, extreme case-2 waters

INTRODUCTION

Phytoplankton play a fundamentally key role in oceans, seas, and freshwater basin ecosystems, as well as in related biogeochemical cycles. Phytoplankton communities are characterized by large taxonomic diversity that strongly determines their role in the ecosystem and their biogeochemical functioning (Uitz et al., 2015). The aquatic environment, whether inland, coastal, or open-ocean waters, is rarely comprised of a single algal class (IOCCG, 2014). Different phytoplankton groups adapt to environmental conditions such as high or low light, temperature, nutrient availability, and turbulence level (Aiken et al., 2008). Specific phytoplankton groups are characterized by some specific pigments—biomarkers—and can, thus, be identified from pigment inventories derived from *in situ* samples (Alvain et al., 2005). Recently, different bio-optical and ecological models have been developed for identifying phytoplankton functional types (PFTs), phytoplankton taxonomic composition, and specific phytoplankton species (e.g., Craig et al., 2006; Astoreca et al., 2009) by means of light absorption spectra, spectral response based on reflectance anomalies, backscatter-based derivation of the particle size distribution, phytoplankton abundance, or through look-up table of $R_{rs}(\lambda)$ that incorporates the range of absorption and scattering variability (e.g., Ciotti and Bricaud, 2006; Alvain et al., 2008, 2012; Hirata et al., 2008; Bracher et al., 2009; Kostadinov et al., 2009; Mouw and Yoder, 2010; Brewin et al., 2015; Lorenzoni et al., 2015). Two recent review articles provide an overview of the different methodological approaches, remote sensing algorithms, and a gap analysis for obtaining phytoplankton diversity from ocean color (Bracher et al., 2017; Mouw et al., 2017). One technical requirement for better phytoplankton identification comprises the utilization of hyperspectral ocean color data over the full visible range between 400 and 700 nm. A limited traceability of uncertainties in connection with phytoplankton group information for all water types has been identified as a current gap of knowledge (Bracher et al., 2017).

With recent advances in optical measurements and future improvements in satellite sensors, approaches of phytoplankton group discrimination have been proposed based on various types of data from *in situ* measurements, model simulations and satellite sensors (Hunter et al., 2008; Lubac et al., 2008; Nair et al., 2008; Taylor et al., 2011; Isada et al., 2015). The rapid development of hyperspectral sensors allows providing more comprehensive remote sensing data of water reflectance spectral properties, attributable to the full range of visible light, i.e., to more wavebands, and higher spectral resolution. The increasing quantity of hyperspectral satellite missions, from existing Hyperion (Folkman et al., 2001), CHRIS (Barnsley et al., 2004), and HICO (Corson et al., 2008) (terminated in 2014) to the expected missions such as EnMAP (Foerster et al., 2015), PRISMA (Meini et al., 2015), HyspIRI (Lee et al., 2015), HYPXIM (Michel et al., 2011), and PACE (Gregg and Rousseaux, 2017), has and will provide much potential for applications of hyperspectral satellite data in aquatic ecosystems (Guanter et al., 2015; Xi et al., 2015). Band placement for improving PFTs retrieval from remote sensing data was investigated by analyzing dominant spectral features in the absorption spectra of the PFTs determined with different methods, with recommendations of using continuous hyperspectral data as they will provide better results (Wolanin et al., 2016). Attempts on hyperspectral identification and differentiation of phytoplankton taxonomic groups have been carried out with various approaches (e.g., Bracher et al., 2009; Torrecilla et al., 2011; Sadeghi et al., 2012; Uitz et al., 2015; Xi et al., 2015; Kim et al., 2016). Progresses achieved so far have not only provided recommendations on the directions into which more effort need to be put, but also suggested the constraints and difficulties lying in these approaches. Our previous study has shown that identification of phytoplankton taxonomic groups is successful when using light absorption spectra, but the identification performance varies in different water types when using remote sensing reflectance, $R_{rs}(\lambda)$, as variability in water optical components changes $R_{rs}(\lambda)$ spectra significantly, both in magnitude and spectral shape (Xi et al., 2015). Light absorption spectra of phytoplanktonic algae are determined by pigment composition and pigment cell concentrations, both can alter e.g., with light condition during growth (photoacclimation). Modeling and identification approaches that are based on phytoplankton absorption features also need to take these intra-taxa and intra-species variability into account, but are due to computer performance issues usually based on just a few single spectra representing a taxonomic or spectral group.

Given that a commonly used parameter obtained directly from hyperspectral Earth observation sensors is the remote sensing reflectance of the water surface, we focused on phytoplankton identification using $R_{rs}(\lambda)$ only. In a former study (Xi et al., 2015) we have also shown that absorption features of pure water in $R_{rs}(\lambda)$ affect the identification performance when phytoplankton concentration is low. In the present study, based on five standard absorption spectra representing five phytoplankton spectral groups, an extensive dataset including 10^5 $R_{rs}(\lambda)$ spectra was simulated using HydroLight with various water optical conditions. This simulated dataset is part of a database compiled within the ESA SEOM C2X project (C2X, 2015). An identification approach is proposed to determine phytoplankton groups with the use of the C2X database. The objectives of this study are (i) to test the skill of the identification, (ii) to investigate how and to what extend other water optical constituents impact the accuracy of this identification, and (iii) to show the applicability of this approach in natural waters using *in situ* data.

DATA AND METHODS

Absorption Data

In order to obtain spectral absorption coefficient of different phytoplankton groups, 128 cultures of various algal species from six major phytoplankton taxonomic groups were prepared. Cultures had been prepared from 68 different species, these included 19 diatom species [*Heterokontophyta* (*Bacillariophyceae*)], 13 species of dinophytes [*Dinophyta* (*Dinophyceae*)], four species of prymnesiophytes [*Haptophyta*

(*Prymnesiophyceae*)], three species of cryptophytes [*Cryptophyta* (*Cryptophyceae*)], 23 species of chlorophytes [*Chlorophyta* (*Chlorophyceae, Picocystophyceae,* and *Trebouxiophyceae*)], and six species of cyanobacteria (*Cyanophyceae*). Culture preparation, growth and light conditions are detailed in Xi et al. (2015) and included different light conditions for each species to introduce some spectral variability for each species. The absorption coefficient spectrum of each culture, $a_{ph}(\lambda)$ (m^{-1}), was measured with a Point-Source Integration-Cavity Absorption Meter (PSICAM) following the procedures outlined in Röttgers et al. (2007). All measurements were done at least in triplicate against pure water as the reference. The PSICAM offers accurate determinations of the absorption coefficient without errors induced by light scattered on the algal cells. $a_{ph}(\lambda)$ spectra were measured and area-normalized in the full spectral range of photosynthetically active radiation, i.e., 400–700 nm (Xi et al., 2015).

Datasets of Simulated Remote Sensing Reflectance

In-water radiative transfer simulations have been carried out using HydroLight (version 5.2; Sequoia Scientific, Inc., USA; Mobley, 1994). The numerical model computes radiance distributions and other related quantities such as remote sensing reflectance, $R_{rs}(\lambda)$, for any given water body. Optical properties of the homogeneous water body are varied in a controlled light environment, i.e., clear maritime atmosphere, moderate wind, and the sun is at its zenith. Two datasets of $R_{rs}(\lambda)$ were modeled using HydroLight's "Case-2" model, assuming the same external conditions but differ in the number of representative spectra for the phytoplankton spectral groups. Thus, they have some similarities but are quasi-independent. The so-called C2X dataset (from ESA's Case-2 Extreme Water Project) that is based on five phytoplankton absorption spectra representing five spectral (taxonomic) groups is used as the standard database and the second one, which comprises optical situations based on 128 phytoplankton absorption spectra from cultures, is for testing; detailed descriptions of the datasets are provided in Hieronymi et al. (2017) and Xi et al. (2015), respectively (the test dataset used here contains more different CDOM absorption and non-algal particles, NAP, concentrations than in the previous study of Xi et al., 2015). Basic information about the HydroLight input for the two datasets is provided in **Table 1**.

The main feature of the C2X database is that it covers most water types, from clearest oceanic Case-1 waters to CDOM-dominated (extreme absorbing) and sediment-dominated (extreme scattering) Case-2 waters. For example, the total (organic and inorganic) particulate backscattering coefficient at 510 nm, $b_{bp}(510)$, varies between 0.0007 and 15.4 m^{-1} and the combined absorption coefficient of detritus and *gelbstoff* at 412 nm, $a_{dg}(412)$, is between 0.004 and 120.2 m^{-1}. On the

TABLE 1 | Specifications of the two used $R_{rs}(\lambda)$ datasets, C2X database and test data, both simulated with the "Case-2" model of HydroLight (with references in Mobley and Sundman, 2013).

	Description	Notation	Unit	C2X database	Test data
	Number of R_{rs} spectra	N	[–]	100,000	15,360
	Wavelengths	λ	[nm]	380: 2.5: 1,100	400: 2.5: 700
Water	Pure water absorption and scattering coefficients	a_w, b_w	[m^{-1}]	data from WOPP v2 by Röttgers et al. (2016)	
	Water temperature	T	[°C]	[0 30]	10
	Water salinity	S	[PSU]	[0 35]	30
Chlorophyll-bearing particles	Chlorophyll-a (Chl) concentration	[Chl]	[mg m^{-3}]	[0.02 200]	0.1, 0.3, 0.5, 1, 5, 10, 50, 100
	Number of chlorophyll-specific absorption coefficient spectra, a_{ph}^* [m^2 mg^{-1}]			5 (as mixtures of two of these 5 spectra with 80 and 20%)	128
	Scattering coefficient by Chl	b_{ph}	[m^{-1}]	Standard power law	
	Scattering phase function	β_{ph}	[sr^{-1}]	Fourier-Forand with specified backscatter fraction	Petzold phase function for "average particle"
	Particle backscatter fraction of phase functions	B_{ph}	[–]	[0.002 0.022]	0.018
CDOM	Absorption coefficient of colored dissolved organic matter at 440 nm	$a_{CDOM}(440)$	[m^{-1}]	[0 20]	0, 0.05, 0.1, 0.5, 1, 2
	Slope of CDOM absorption between 300 and 440 nm	S_{CDOM}	[nm^{-1}]	[0.0004 0.032]	0.014
Non-algal particles	Concentration of non-algal particles (minerals)	[NAP]	[g m^{-3}]	[0 1,500]	0, 0.5, 1, 5, 10, 50
	Mass-specific NAP absorption coefficient at 440 nm	$a_{NAP}^*(440)$	[m^2 g^{-1}]	0.0615 (from average of *in situ* spectra)	0.051 (from one *in situ* spectrum)
	Mass-specific NAP scattering coefficient at 440 nm	$b_{NAP}^*(440)$	[m^2 g^{-1}]	[0.273 1.093]	0.8902 (standard average)
	NAP scattering phase function	β_{NAP}	[sr^{-1}]	Fourier-Forand with specified backscatter fraction	Petzold phase function for "average particle"
	NAP backscatter fraction of phase functions	B_{NAP}	[–]	[0.001 0.035]	0.018

Two values in square brackets refer to a range.

basis of the Xi et al. (2015) study, five fundamental spectral shapes of chlorophyll-specific absorption were selected. The five absorption spectra (of different species of algae) are supposed to have the highest potential for identification of these five different spectral groups from a remote sensing reflectance spectrum. These normalized spectra are shown in **Figure 1A** and stand for: (1) a "brown spectral group" representing *Heterokontophyta*, *Dinophyta*, and *Haptophyta*, (2) a "green spectral group" representing *Chlorophyta*, (3) a group for *Cryptophyta*, (4) a blue-green cyanobacteria, and (5) a red cyanobacteria. The first four spectra are absorption spectra from single cultures that are close to the mathematical mean for all spectra of cultures from this group (from the 128 measured culture absorption spectra). As an example, the absorption spectrum for the "brown spectral group" was chosen from all the cultures in the brown group in **Figure 1B**. These culture-spectra are realistic as very similar spectra can be found in the HZG *in situ* database (unpublished data). The spectrum of the red cyanobacteria was obtained from field measurements in the Baltic Sea during a bloom of cyanobacteria (most likely *Nodularia* sp.). Culture spectra (e.g., a red *Synechococcus* sp.) of this type mostly exhibit much higher phycobilin-related absorption peaks around 570 nm. In order to account for natural variability in the simulation for the C2X database, the actually used $a_{ph}^*(\lambda)$ spectra are always mixtures from two of the five groups with individual contributions of 80 and 20%, respectively. The total phytoplankton absorption, a_{ph}, is related to the spectral chlorophyll-specific absorption and chlorophyll *a* concentration (denoted as [Chl] hereafter), $a_{ph}(\lambda) = a_{ph}^*(\lambda) \times$ [Chl]. The natural variability of phytoplankton absorption is very high (e.g., Bricaud et al., 2004); and the full range of observed natural variability is included in the simulations (**Figure 2**). Basis for estimating distributions, ranges, and covariances of optical properties and concentrations are several *in situ* datasets (e.g., Valente et al., 2016), but mainly our HZG *in situ* data from the North and Baltic Sea. The simulated data have been compared with *in situ* observations, e.g., $b_{bp}(510)$ and $a_{dg}(412)$ vs. different reflectance band ratios (Hieronymi et al., 2016), and we generally found a good agreement. But we have also found some discrepancies partly related to plausible measuring uncertainties and possibly due to model simplifications. In this context, it should be mentioned that the model assumptions for spectral scattering properties are identical for all five phytoplankton groups, i.e., the particle backscatter fraction depends on chlorophyll *a* concentration (Twardowski et al., 2001), but not on algae-specific (back-) scattering properties.

For the HydroLight simulations, the considered $R_{rs}(\lambda)$ is fully normalized, i.e., the sun is at zenith and the viewing angle is perpendicular; the water is infinitely deep; inelastic scattering, i.e., Raman scattering and Chl and CDOM fluorescence, are taken into account. Nonetheless, how inelastic scattering processes and their natural variability influence the results is

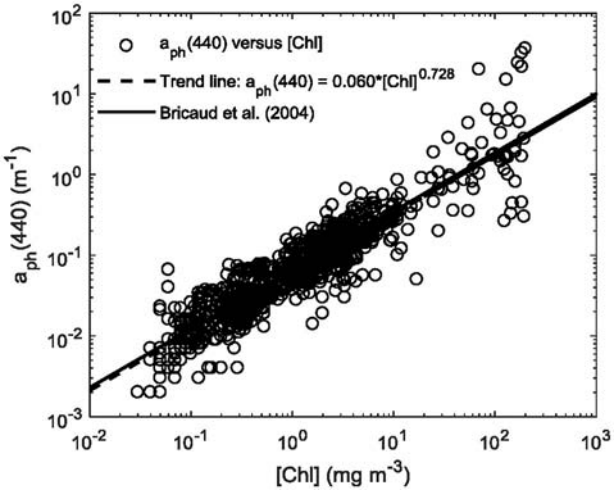


FIGURE 2 | Phytoplankton absorption coefficient at 440 nm, $a_{ph}(440)$, vs. chlorophyll *a* concentration [Chl] used in simulations for the C2X database. The trend line is also shown in comparison to that by Bricaud et al. (2004).

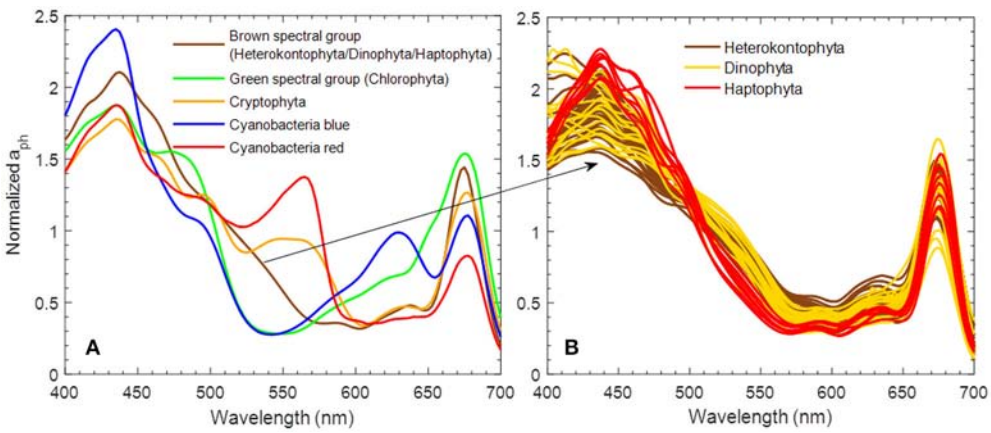


FIGURE 1 | (A) Five area-normalized absorption spectra of phytoplankton used in $R_{rs}(\lambda)$ simulation for the C2X database, and **(B)** corresponding spectra of cultures representing the "brown spectral group."

out of scope of this work. Ultimately, the C2X database built with HydroLight simulations includes in total 1×10^5 $R_{rs}(\lambda)$ spectra with five phytoplankton groups with various water optical conditions; while the test dataset includes 15,360 $R_{rs}(\lambda)$ spectra, for 120 different water conditions ([Chl] varys from 0.1 to 100 mg m^{-3}, [NAP] from 0 to 50 g m^{-3}, and CDOM from 0 to 2 m^{-1}) with 128 phytoplankton absorption spectra (**Table 1**). The corresponding concentration values were listed in **Table 1**, and rational of the water condition settings was described and discussed in Xi et al. (2015) where this dataset was firstly used.

Phytoplankton Group Identification

The general scheme of the identification approach is illustrated in **Figure 3**. At first, all $R_{rs}(\lambda)$ spectra are area-normalized and then second-order derivative is calculated. Details on the normalization and derivative transformation are described in Xi et al. (2015). To identify the corresponding phytoplankton groups in a test data set, each $R_{rs}(\lambda)$ spectrum in the test data set is compared to all spectra in the C2X database using the similarity index (SI) as an angular distance (Millie et al., 1997):

$$SI = 1 - \frac{2}{\pi} \times \arccos\left(\frac{x_1 \cdot x_2}{|x_1|\,|x_2|}\right) \quad (1)$$

where x_1 is a second-derivative spectrum of $R_{rs}(\lambda)$ in the C2X database, and x_2 is one in the test dataset. The SI is a number between 0 and 1, where 0 indicates no similarity and 1 indicates perfect similarity between the two spectra. It is noteworthy that only the second-derivative spectra of $R_{rs}(\lambda)$ in the range of 420–620 nm was used for SI calculation to minimize the influence of noises at shorter wavelengths, where reflectance is often low, and that of strong water absorption features at longer wavelengths (Xi et al., 2015).

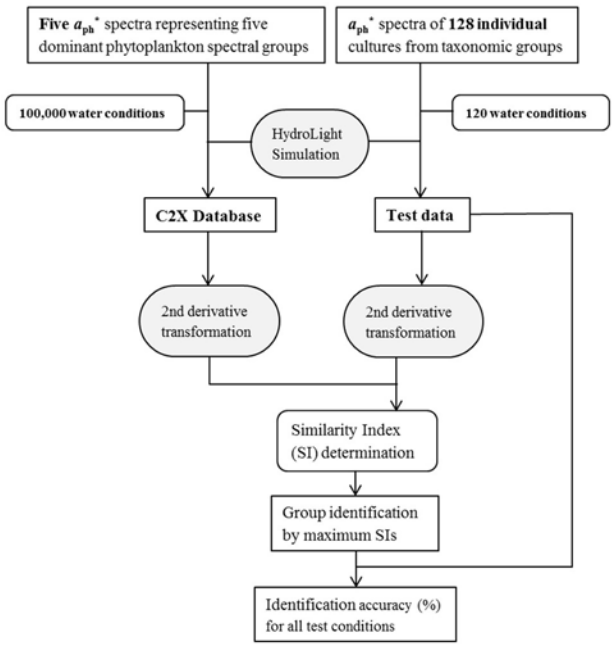


FIGURE 3 | Flowchart showing the approach for dominant phytoplankton group identification.

This approach produces 10^5 SI values for each spectrum of the test dataset. The first 20 spectra in the C2X database providing highest SI values for each test spectrum are selected and their corresponding known phytoplankton spectral group is recorded. The group that is dominant in these 20 spectra is taken as the identified spectral group for this test spectrum. For each test spectrum one of the five spectral groups is identified as being dominant. Each taxonomic group in the test data is represented by five to 48 $R_{rs}(\lambda)$ spectra (from 5 to 48 different cultures), and all taxonomic groups are categorized into five spectral groups, the spectral group identification accuracy is thus determined by calculating the percentage of the correctly identified $R_{rs}(\lambda)$ spectra in each spectral group. In the end, given that there are 120 water optical conditions in the test data, 120 values for the identification accuracy of each group are calculated.

In situ $R_{rs}(\lambda)$ Data of Lake Taihu

An investigation campaign was carried out from 5th to 17th October 2008 in Lake Taihu (China). A set of *in situ* $R_{rs}(\lambda)$ spectra was obtained by measuring the water-leaving radiances and sky radiances with a dual channel spectrometer, ASD FieldSpec Pro Dual VNIR (FieldSpec 931, ASD Inc., USA), following NASA ocean optics protocols (Mueller et al., 2003). When performing the measurements, the viewing angels of the two channels from the water surface at the zenith angle and the azimuth angle were 40° and 135°, respectively. Radiances of a 25 cm by 25 cm plaque with 25% reflectivity, water and sky radiances (each preceded by a dark offset reading) were measured and repeated five times. The measurements were performed at a location that minimized shading, reflections from superstructure, ship's wake, foam patches, and whitecaps. Moreover, the location was also pointed away from the sun to reduce the sunglint effect. Upon the upward radiance (Lu), sky radiance (Lsky), gray plaque radiance (Lplaq), and the water-air interface reflectivity determined based on the lake state at that time, $R_{rs}(\lambda)$ were calculated referring to the method proposed by Mobley (1999). Details of the approaches for radiance measurements and $R_{rs}(\lambda)$ calculation are illustrated in Ma et al. (2006). Water samples were taken simultaneously with the spectrometer for lab measurements of Chl, NAP, and CDOM concentrations. Absorption spectra by the total particles and the NAP were determined by quantitative filter technique (QFT) method (Mitchell, 1990) and $a_{ph}(\lambda)$ was obtained by subtracting $a_{NAP}(\lambda)$ from $a_p(\lambda)$. $a_{CDOM}(\lambda)$ was also measured spectrophotometrically in a 10 cm cuvette using 0.7 mm Whatman GF/F-filtered water sample pads by the same UV-2401 spectrophotometer. More details on the above determinations are described in Xi (2011).

As one of the biggest freshwater lakes in China, Lake Taihu covers an area of 2427.8 km^2 with highly varying water quality from area to area. Water types in Lake Taihu are mainly classified into two categories: optically deep waters (ODWs) and optically shallow waters (OSWs) (Xi, 2011). ODWs cover most area of the lake with highly eutrophicated and turbid waters and frequent occurrence of cyanobacteria blooms, while the southeastern area is mostly OSWs with clear waters and abundant aquatic plants.

Data used here are from ODWs only, as $R_{rs}(\lambda)$ from OSWs has much influence from the submerged aquatic plants and the lake bottom and are thus not suitable for use in this study. Due to the large area of the lake, water optical conditions in ODWs are also diverse. Variations of the water components are known: [Chl] varied from 4.0 to 180 mg m^{-3}, [NAP] from 9.5 to 95 g m^{-3} and CDOM from 0.4 to 1.7 m^{-1}. For the present study, $R_{rs}(\lambda)$ spectra together with other optical parameters for 66 stations in ODWs are obtained. This additional "Taihu dataset" is used to test the applicability of the presented approach in natural waters.

RESULTS

Spectral Analysis of C2X Reflectances

The C2X database contains simulated $R_{rs}(\lambda)$ spectra that included as model input a standard absorption spectrum for each of five different phytoplankton spectral groups. These $R_{rs}(\lambda)$ spectra show different spectral features reflecting various water optical conditions. Prior to utilizing the C2X data in the identification approach, $R_{rs}(\lambda)$ spectra in the database are firstly normalized and transformed to the second derivative spectra. To have an general understanding on the C2X database, representative $R_{rs}(\lambda)$ spectra of the five phytoplankton groups and their second derivatives are selected for a few water cases. For each water case, five $R_{rs}(\lambda)$ spectra with similar water optical conditions representing the five phytoplankton groups are chosen. **Figure 4** shows examples of $R_{rs}(\lambda)$ spectra and their second derivative spectra, for different phytoplankton groups in five water cases. The five water cases are however not exhaustive. According to Hieronymi et al. (2017), 13 different water optical classes in total are classified by a fuzzy logic classification approach, but they are not completely included here as this study is not focusing on water type interpretation. Only examples of five water cases are chosen to show spectral variations in different scenarios. These five water cases possess the following conditions:

(1) low [Chl], low [NAP], and low CDOM (**Figure 4A–C**);
(2) low [Chl], low [NAP], and moderate CDOM (**Figure 4D–F**);
(3) low [Chl], moderate [NAP], and moderate CDOM (**Figure 4G–I**);
(4) high [Chl], moderate [NAP], and moderate CDOM (**Figure 4J–L**); and
(5) moderate [Chl], extremely high [NAP], and moderate CDOM (**Figure 4M–O**).

The corresponding water optical conditions of the five water cases and the variation of water optical components are listed in **Table 2**. Note that the ranges of "low," "moderate," and "high" concentrations are a bit varying from case to case which may cause slight difference in the spectral magnitude as well as the chlorophyll fluorescence. Water case (1) represents clear Case-1 waters (**Figure 4A**) where phytoplankton, CDOM and pure water are the main contributors to the $R_{rs}(\lambda)$; phytoplankton absorption contributes to the suppression at 440 nm and the reflectance in the blue spectral region is high; low scattering and the high absorption by water at red and near infrared wavelengths results in low reflectance in this region. In absorbing waters such as water case (2) where CDOM is moderate but other concentrations are low (**Figure 4D**), reflectance is lower in the blue band suggests a strong CDOM absorption, and the peak at about 682 nm is due to the fact that Chl fluorescence was included in the simulations. In scattering dominated waters as water case (3) (**Figure 4G**), NAP are the dominating component; the reflectance is high in the whole visible region and the maximum is shifted to longer wavelength with the increase of NAP concentrations; peaks and troughs attributable to pigment absorption are suppressed. In high [Chl] waters as water case (4), the contribution by different phytoplankton pigments in the $R_{rs}(\lambda)$ spectra is clearly seen (**Figure 4J**), suggesting that it is relatively easy to identify phytoplankton groups in such waters. Scattering by NAP results in higher reflectance at longer wavelengths (>550 nm), therefore when sediment load is extremely high, as shown in **Figure 4M**, the reflectance shows an increasing pattern with wavelengths in the visible region. In NAP-dominated waters, little spectral difference can be observed among the different phytoplankton groups (**Figure 4N**); absorption and scattering by sediments mask the algae pigment features. This masking effect is a generally known limitation and uncertainty source for remote sensing of biomass in turbid Case-2 waters (e.g., IOCCG, 2000).

Though phytoplankton groups exhibit distinct spectral features in some water cases, the corresponding second derivative spectra of $R_{rs}(\lambda)$ in **Figure 4** (third column) show much variation in different water cases even for the same dominating phytoplankton group, due to the different contribution of other water optical constituents to the reflectance spectra. This indicates a possible difficulty in identifying phytoplankton groups for highly variable natural waters, by only inter-comparing the reflectance spectra without references. Given that, our theoretical basis is the C2X database that in the following is used as a look-up table (LUT) of standard reflectance spectra with information about the dominating phytoplankton groups, so that any test spectrum can be spectrally compared to the LUT and a certain phytoplankton group can be allocated to it. With the use of the simulated test data, the performance by the LUT identification approach can be evaluated for various water optical conditions.

Phytoplankton Spectral Group Identification

In order to investigate how accurate the phytoplankton groups can be identified under different water optical conditions, $R_{rs}(\lambda)$ spectra of the test data are compared with those in the C2X database by following the identification approach described above. Identification accuracy based on the 120 water conditions is generated for each spectral group via the proposed approach. The identification accuracy is presented for each spectral group as a function of the different water conditions using a ternary plot (**Figure 5**). Since a triangular diagram displays the proportion of three variables that sum to a constant, absorption coefficients by phytoplankton and NAP are used here to represent [Chl] and [NAP], respectively. CDOM itself is typically represented by the absorption of CDOM. The sum of these three absorption coefficients can be normalized to be 1, ignoring the absorption

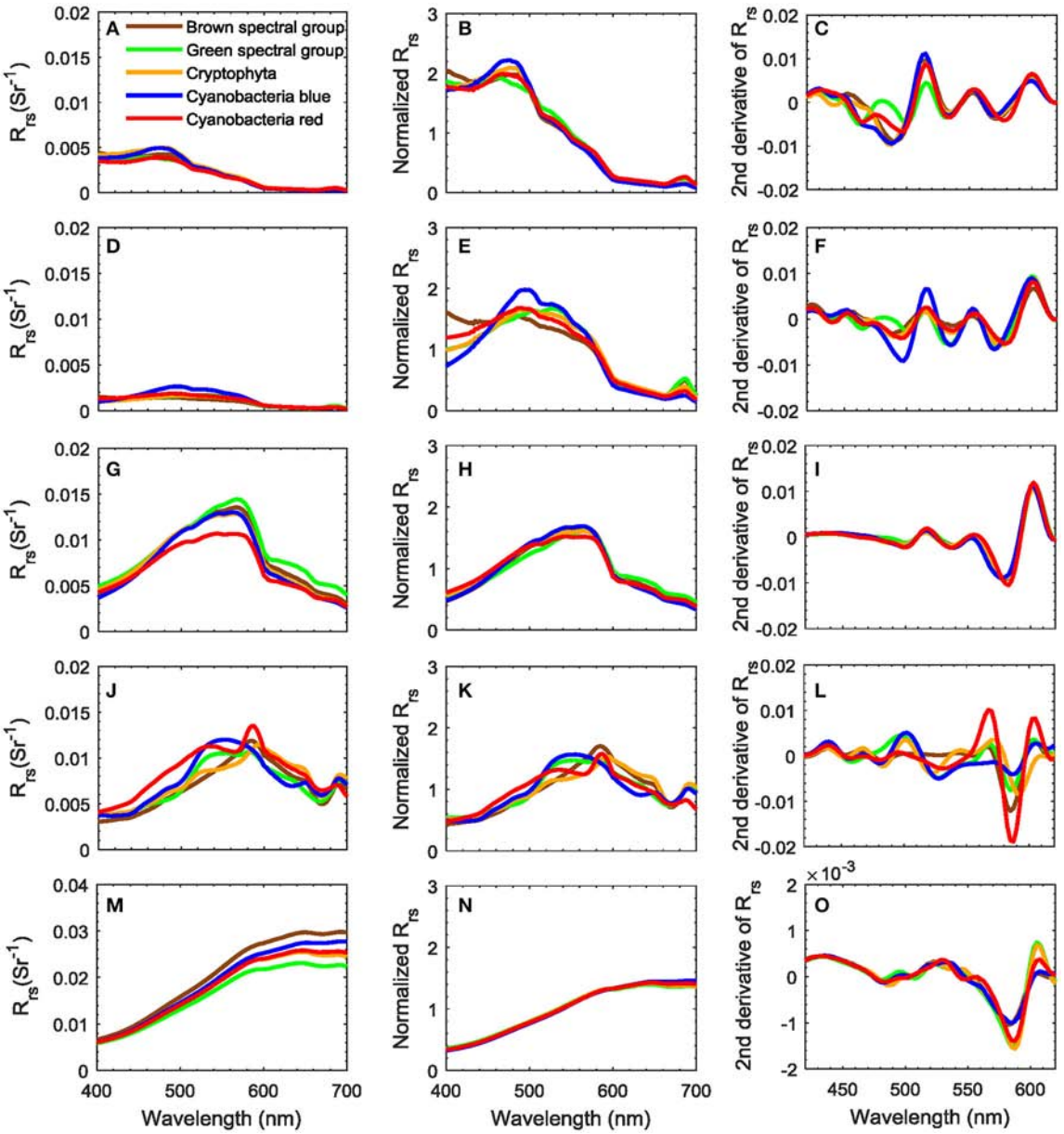


FIGURE 4 | Representatives of $R_{rs}(\lambda)$ spectra (first column), the corresponding area-normalized $R_{rs}(\lambda)$ spectra (second column) and second derivative spectra (third column) for five phytoplankton spectral groups and different water conditions. **(A–C)** Low [Chl], [NAP], and CDOM; **(D–F)** low [Chl] and [NAP], moderate CDOM; **(G–I)** low [Chl], moderate [NAP] and CDOM; **(J–L)** high [Chl], moderate [NAP] and CDOM; and **(M–O)** moderate [Chl], extremely high [NAP], and moderate CDOM. Note that the second derivative spectra were only for 420–620 nm, i.e., the spectral range used in the identification approach.

TABLE 2 | Corresponding variations of [Chl], [NAP] and CDOM for $R_{rs}(\lambda)$ spectra in **Figure 4**.

Figure 4	[Chl] (mg m^{-3})	[NAP] (g m^{-3})	CDOM (m^{-1})
A–C	Low: 0.25–0.27	Low: 0.1–0.2	Low: 0.01–0.015
D–F	Low: 0.25–0.3	Low: 0.1–0.2	Moderate: 0.1–0.15
G–I	Low: 0.25–0.27	Moderate: 5–8	Moderate: 0.1–0.15
J–L	High: 30–60	Moderate: 10–20	Moderate: 0.1–0.2
M–O	Moderate: 5–10	Extremely high: 200–400	Moderate: 0.1–0.3

of pure water. Their proportions for all the 120 water conditions can be well displayed in a ternary plot. [Chl] and [NAP] are thus transformed to the corresponding absorption coefficients at 440 nm, $a_{ph}(440)$ and $a_{NAP}(440)$, and used together with $a_{CDOM}(440)$. Four hundred and forty nanometers is chosen as all components do significantly absorb light at this wavelength, and contribution by pure water is low. The transformation from [Chl] to $a_{ph}(440)$ is based on the relationship shown in **Figure 2**: $a_{ph}(440) = 0.06 \times [\text{Chl}]^{0.728}$ and $a_{NAP}(440) = a_{NAP}^{*}(440) \times$ [NAP], where the mass-specific NAP absorption at 440 nm,

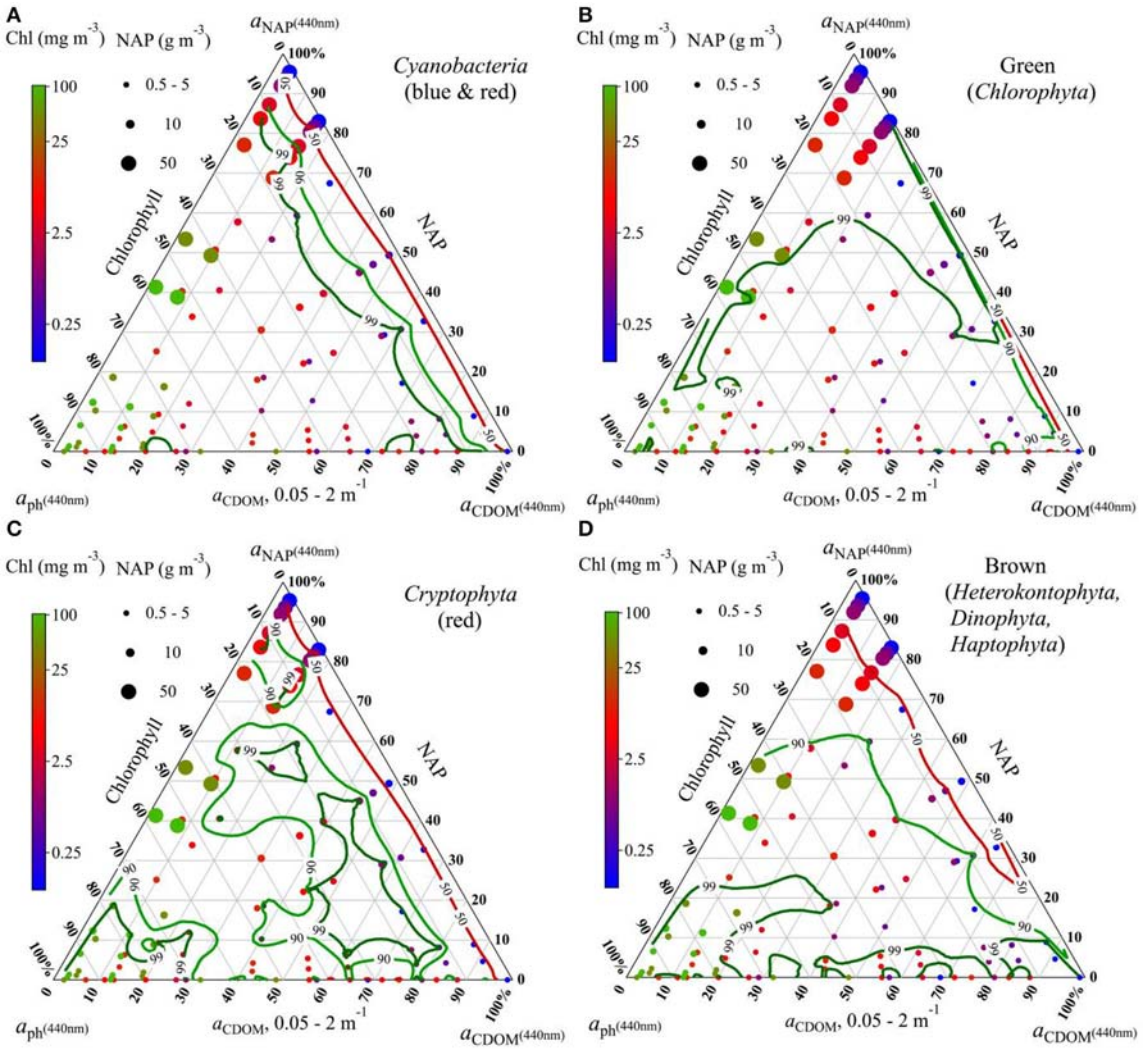


FIGURE 5 | Ternary plots showing the identification accuracy of phytoplankton groups **(A)** Cyanobacteria (red and blue), **(B)** *Chlorophyta* (green spectral group), **(C)** *Cryptophyta* (red), and **(D)** Brown spectral group in different water conditions with respect to fractions of phytoplankton, NAP, and CDOM absorption at 440 nm: Contour lines indicate the accuracy of identification. Colors of dots indicate Chl concentrations; sizes of dots indicate NAP concentrations.

$a_{NAP}{}^{*}(440) = 0.0615\ m^2\ g^{-1}$ as shown in **Table 1**. Different colors and sizes of the dots are used to represent the actual [Chl] and [NAP] respectively, as the position in the ternary plot only shows each relative contribution.

Contour lines of the identification accuracy are plotted for each spectral group (**Figure 5**). The contour lines indicate different distribution of the identification accuracy for the different groups. However, a main finding in common is that low identification (50% contour line) for all groups located in the plot area where [NAP] is high (bigger dots) and [Chl] (blue dots) is low. However, this is not always true. If taking a further look on the plots one can see that the identification accuracy is also dependent on the absorption contribution of each water components but not only on concentrations. In **Figure 5**, for simplification, blue and red cyanobacteria are combined, as they show distinct spectral features compared to other groups and their identification results are highly similar. Cyanobacteria blue and red (combined in **Figure 5A**) show a relatively distinct contour pattern: the contour lines are roughly parallel with low Chl contributions, e.g., the identification rate of 90% is approximately at $a_{ph}(440)$ taking up only 10% of the total non-water absorption ($a_{ph+CDOM+NAP}$), and the 99% contour line is between 10 and 20%, meaning that cyanobacteria can be successfully identified when $a_{ph}(440)$ takes up more than 20% of the total non-water absorption. The approach also performs well on the green spectral group (*Chlorophyta*), as **Figure 5B** shows that the identification rates falls in 90% only when [Chl] is extremely low [$a_{ph}(440)$ is <5%] and [NAP] is as high as 50 g m^{-3}, in all other conditions chlorophytes are correctly identified. The identification contour lines for cryptophytes in **Figure 5C** show more variations for different water conditions, due to the fact that there are only five cultures for *Cryptophyta* in the test dataset and one culture showed higher similarity with red cyanobacteria and is thus misidentified at some water conditions.

This has lowered 20% of the identification rate. The overall performance of identifying species of the brown spectral group show that 90% of the cultures from *Heterokontophyta*, *Dinophyta,* and *Haptophyta* are correctly identified as being brown when $a_{ph}(440)$ contributes more than 20% and $a_{NAP}(440)$ contributes <60% to the total non-water absorption (**Figure 5D**).

Applicability Test Using Taihu *In situ* Dataset

Though the performance of using the simulated test dataset showed satisfactory results at specific water conditions, the ultimate goal of our approach is to identify phytoplankton groups by using reflectance spectra of natural waters. A set of *in situ* $R_{rs}(\lambda)$ spectra of Lake Taihu is taken to perform an additional test on the applicability of the proposed approach. The second derivative spectra of *in situ* $R_{rs}(\lambda)$ in 420–620 nm are compared with that in C2X database to produce highest SIs leading us to find the corresponding dominating phytoplankton groups. Results show that only two phytoplankton spectral groups are found, blue cyanobacteria and *Chlorophyta*, In the examined 66 stations, for 52 (80%) the dominating phytoplankton are identified as cyanobacteria (blue) and 14 (20%) as chlorophytes. **Figure 6** shows the classified spectra of *in situ* $R_{rs}(\lambda)$, the corresponding normalized $R_{rs}(\lambda)$ in 400–700 nm, and the second derivative spectra. It is clearly seen that most spectra identified as *Chlorophyta* exhibit higher reflectance in 500–600 nm and lower chlorophyll fluorescence peaks around 690 nm, indicating higher sediment concentrations and lower Chl concentrations. On the contrary, spectra that are identified as cyanobacteria show distinct absorption peaks at 675 nm and more pronounced fluorescence (**Figure 6B**). **Table 3** lists the identification results for Lake Taihu when stations are selected by different [Chl], showing that the identification rate of cyanobacteria increases with the increasing minimum [Chl]: 90% when [Chl] > 10 mg m^{-3}, 98% when [Chl] > 20 mg m^{-3}, and 100% when [Chl] > 30 mg m^{-3}.

It has been known in the context that [Chl] has an order of two in magnitude, and water optical conditions highly varied from station to station in Lake Taihu. The information we had from the campaign was that in waters where [Chl] was roughly 30 mg m^{-3} or higher, cyanobacteria aggregated obviously and were the dominating group. Whereas, lower [Chl] waters could either be green algae or cyanobacteria dominated according to the identification. To explore whether the two identified groups are relating to the absorption contributions of each water component, proportions of $a_{ph}(440)$, $a_{NAP}(440)$, and $a_{CDOM}(440)$ to the total non-water absorption at 440 nm were statistically summarized for cyanobacteria identified stations and green algae identified stations, respectively (**Table 4**). The overall [Chl] and contribution of aph(440) at cyanobacteria dominating stations are higher than that at green algae dominating stations. Mean $a_{ph}(440)$ contribution is 16.1% for cyanobacteria while only 9.8% for green algae. $a_{NAP}(440)$ contribution shows the opposite with $a_{ph}(440)$, with lower mean value (58.1%) for cyanobacteria but slightly higher for green algae; no difference is found in CDOM contribution between the two groups. These statistical results in **Table 4** also reveal that CDOM has little influence in the identification, in agreement with results from the simulated test dataset. Lower phytoplankton contribution in green algae dominating stations might suggest that the identification of green algae is less accurate.

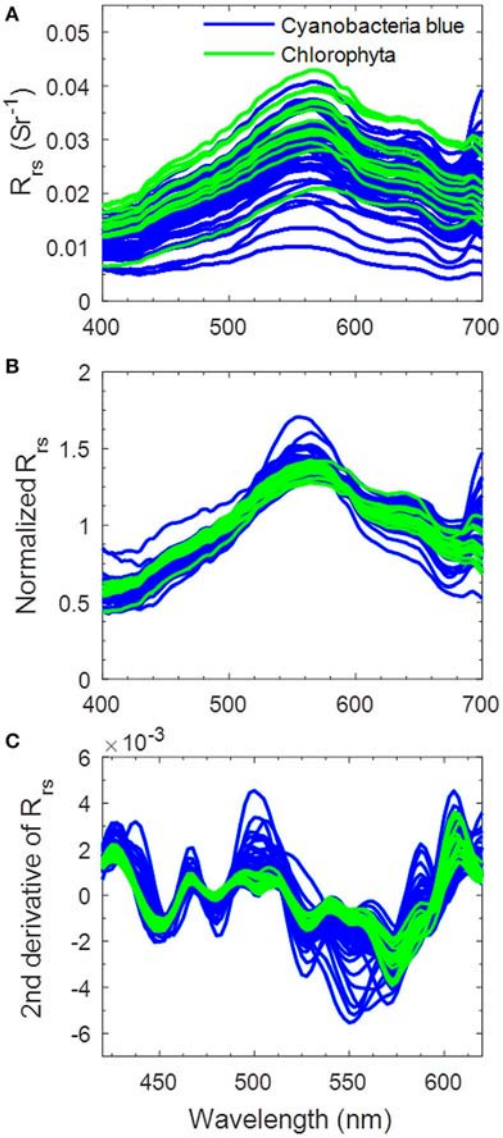


FIGURE 6 | Phytoplankton-group-classified spectra of **(A)** *in situ* $R_{rs}(\lambda)$ in Lake Taihu in 400–700 nm, **(B)** the corresponding area-normalized $R_{rs}(\lambda)$, and **(C)** second derivative spectra in 420–620 nm.

TABLE 3 | Phytoplankton groups identification in Lake Taihu with different [Chl] ranges.

Number of stations	[Chl] (mg m^{-3})	Identification rate of cyanobacteria	Identification rate of chlorophyta
66 (all)	4–180	80% (52 stations)	20% (14 stations)
56	10–180	90%	10%
39	20–180	98%	2%
25	30–180	100%	0

TABLE 4 | Identified phytoplankton groups with relation to contributions of $a_{ph}(440)$, $a_{NAP}(440)$, and $a_{CDOM}(440)$ to the total non-water absorption.

	Number of stations	[Chl] (mg m^{-3})	% $a_{ph}(440)$			% $a_{NAP}(440)$			% $a_{CDOM}(440)$		
			Min	Max	Mean (SD)	Min	Max	Mean (SD)	Min	Max	Mean (SD)
Cyanobacteria	52	8–180	5.5	46.1	16.1 (8.6)	27.9	84.2	58.1 (11.3)	8.6	54.3	27.4 (9.7)
Chlorophyta	14	4–26	5.2	16.5	9.8 (3.1)	53.6	77.0	62.4 (6.1)	14.9	39.5	27.8 (6.2)

DISCUSSION

Simulated C2X Data

Generally, the simulated reflectance spectra in the C2X database are plausible and fit to *in situ* measurements. For example, many lakes have signal-dominating CDOM fractions and exhibit reflectance spectra as shown in **Figure 4D** (e.g., Eleveld et al., 2017) and **Figure 4M** shows similar spectra as measured in a turbid estuary (e.g., Knaeps et al., 2012). Nonetheless, the simulations are based on some spectral assumptions that may lead to inaccuracy and therefore uncertainties in the group identification. One simplification regards the scattering properties of phytoplankton; some natural variability is included in the simulations, but due to the lack of reliable specific information, all phytoplankton groups are modeled with the same scattering assumptions. But due to different particle shapes and size distribution, it is evident that the spectral scattering properties vary (e.g., Morel, 1987; Evers-King et al., 2014; Harmel et al., 2016). Robertson Lain et al. (2017) showed the potentially considerable influence of different phytoplankton phase functions on modeled remote sensing reflectance over the entire visible range. A second point of model uncertainties, particularly in the range 650–700 nm, regards inelastic scattering effects such as CDOM and phytoplankton fluorescence; in nature, the quantum yield efficiency of phytoplankton varies significantly depending on nutrient- and light-availability and algae species (e.g., Greene et al., 1994). However, in the HydroLight simulations, the standard quantum yield efficiency was used.

Identification Approach and Its Skill

The proposed approach was chosen for phytoplankton group identification based on the idea of whether we can identify phytoplankton groups by only knowing $R_{rs}(\lambda)$, which is a directly obtained parameter from satellite sensors. All other inversion models require information on water inherent optical properties and the retrieval accuracy can be various (e.g., Werdell et al., 2014; Wang et al., 2016). In our previous study, we have made a performance comparison between using $R_{rs}(\lambda)$ directly and using QAA-inverted absorption spectra from $R_{rs}(\lambda)$ for phytoplankton group differentiation (Xi et al., 2015). Results show that the inverted absorption spectra performed less precise compared to the $R_{rs}(\lambda)$. Due to the retrieval algorithm constraints, pigment information in the derived absorption spectra might be lost or distorted as theoretical or empirical relationships between the IOPs and AOPs are normally used in the retrieval algorithm. Upon the simulated extensive $R_{rs}(\lambda)$ database, a direct comparison in spectral shapes between the second derivatives of a test $R_{rs}(\lambda)$ spectrum and that of the spectra in the database is carried out by the current approach, omitting the knowledge of optical properties as well as the retrieval errors introduced by inversions. The benefit of this approach would be that we provide a straightforward way allowing us to know the dominant phytoplankton group (if it is one of the five) once $R_{rs}(\lambda)$ is obtained from either *in situ* measurements or hyperspectral satellite data.

Skill of the proposed approach varies in different water conditions. Results derived by using the test dataset for various water optical properties indicated that the identification accuracy was highly subject to the water optical conditions; the identification was effective for waters with high phytoplankton contribution but less effective in NAP dominated waters, whereas CDOM has little influence even when it is extremely high. It is not only in agreement with the results by Xi et al. (2015) that phytoplankton groups differentiation is unsuccessful in waters with [Chl] lower than 1 mg m^{-3}, but also suggested the low efficiency in high [NAP] waters. However, regarding the optical boundaries of the successful identifications, it should be clarified that the accuracy of the identification is not only dependent on the concentrations of water components but also on the contribution of absorption by each water component to the total absorption. That means the identification accuracy can possibly be high both in case 1 clear waters and in highly turbid productive (phytoplankton abundant) waters. There are no clear concentration boundaries. And the optical boundaries in terms of absorption contribution are nicely shown in the ternary plots (**Figure 5**). Regarding this matter, ternary plots exhibit clearly the contour line distribution of the identification accuracy for all water optical conditions generated by 120 points representing 120 water optical conditions. We can roughly wrap up some general findings from the ternary plots, that are—the identification accuracy is higher than 90% when the absorption by phytoplankton is taking up more than 20% and the NAP contribution is <60% to the total absorption; for the groups of cyanobacteria and green algae, the identification accuracy is even higher at the above boundaries.

Though there were only 120 water optical conditions considered in the test dataset, it included most of the natural aquatic environment from clear to moderate turbid and productive waters. The findings have provided us a basic knowledge that the proposed approach for phytoplankton group identification performs well except for waters where phytoplankton contribution to the overall absorption is low and for NAP dominated waters. Regarding the five spectral groups included in C2X database, they are not exhaustive but are chosen to represent natural common groups. In addition,

due to modeling and computing constraints, the phytoplankton groups that were taken into account in the simulation have to be representative and the number of the groups should be as low as possible to allow extensive simulations. This database is however adjustable, when absorption features of other phytoplankton groups (or species with typical features) are important.

Phytoplankton Group Identification in Lake Taihu

To test the applicability of the proposed approach in natural waters, *in situ* remote sensing reflectance data obtained from Lake Taihu were used. Though lacking the information of dominant phytoplankton species or pigment analysis for this campaign in October 2008, previous investigations on the phytoplankton community and composition in Lake Taihu can be taken as reference. Chen et al. (2003) revealed that phytoplankton groups commonly observed in Lake Taihu are cyanobacteria, *Chlorophyta*, *Bacillariophyta,* and flagellates. A study on phytoplankton community structure succession in Lake Taihu from 1992 to 2012 by Deng et al. (2014) showed that *Cryptomonas* (*Cryptophyta*) was the dominant species in spring during the early 1990s. Dominance then shifted to *Ulothrix* (*Chlorophyta*) in 1996 and 1997. However, *Cryptomonas* again dominated in 1999, 2000, and 2002, with *Ulothrix* regaining dominance from 2003 to 2006. The bloom-forming cyanobacterial species *Microcystis* sp., a typical blue cyanobacteria, dominated in 1995, 2001, and 2007–2012. Another study revealed that *Microcystis* sp. is the most commonly seen cyanobacteria species, approximately taking up 85% of algae biomass and forming algal blooms each summer (Zhu et al., 2007). More importantly, a year-long investigation in dominant phytoplankton species from October 2008 to October 2009 conducted in the lake showed that *Microcystis* sp. dominated in October 2008, when our *in situ* data was collected, contributing more than 90% of total biovolume in most area of the lake, coexisted with minor portion of the cyanobacteria *Dolichospermum flos-aquae* and the diatom *Cyclotella meneghiniana* (Ai et al., 2015). Our identification results showed good agreement with these investigations, except that chlorophytes were identified as the dominant group at some stations when [Chl] was moderate. Comparison in absorption contributions between cyanobacteria and green algae identified stations shows that phytoplankton contribute <10% on average to the total non-water absorption at green algae identified stations (**Table 4**), leading to lower identification accuracy as indicated in **Figure 5**. It is highly likely that this is a misinterpretation, as still cyanobacteria were codominant.

The *in situ* data of Lake Taihu are used as a first example. Coinciding data of spectral reflectance and information of phytoplankton taxonomic composition are still quite rare. However, this first example has given optimistic outcome and the fact that the approach is applicable in this optically complex lake. More datasets in different water types are under collection and processing, with expectations to testify further the identification approach in more natural waters.

CONCLUSIONS AND OUTLOOK

A database of $R_{rs}(\lambda)$ spectra, C2X database, based on five phytoplankton groups was built using HydroLight simulations for various water optical conditions. A similarity-index approach was proposed to identify phytoplankton groups, using remote sensing reflectance spectra only, by spectrally comparing an input test spectra with the $R_{rs}(\lambda)$ in C2X database. The performance of the approach was tested using another simulated $R_{rs}(\lambda)$ dataset with 128 spectra of phytoplankton algae from six taxonomic groups arranged into five spectral groups. For 120 water optical conditions, the identification was high at most occasions except for waters with a low phytoplankton contribution and for waters dominated by NAP. Whereas, the influence of CDOM is less pronounced and only significant at extremely high level. Though the proposed approach was based on simulated datasets, its applicability in natural waters was also tested by using *in situ* $R_{rs}(\lambda)$ spectra from Lake Taihu, China. Despite of possibly wrong identification of chlorophytes that could not be validated, cyanobacteria were successfully identified in Lake Taihu as a dominating group in high [Chl] waters, proving the applicability of the approach in natural waters when a single group is dominating.

The current approach is only capable of identifying spectral groups that are already presented in the C2X database. However, the database can be expanded by running the same HydroLight simulations with different absorption characteristics of other phytoplankton groups. The following aspects might be worth to investigate: (1) more validation with *in situ* data and measurements; (2) applicability in extreme events such as floating algae in highly turbid waters; (3) determination on the required or lowest spectral resolution for a wider use of hyperspectral $R_{rs}(\lambda)$; and (4) examination of using hyperspectral satellite data in consideration of influences of radiometric, spectral, and atmospheric effects on $R_{rs}(\lambda)$ from the space.

AUTHOR CONTRIBUTIONS

HX carried out the data analysis and the approach development with guidance from RR. MH simulated the hyperspectral remote sensing reflectance for the C2X database and the test data using HydroLight. HX drafted the main manuscript and MH refined the simulation part in Data and method. MH, RR, and HK all provided comments and suggestions and revised the manuscript.

FUNDING

The work of HX is supported by the project "Preparations for the scientific use of the EnMAP data: Coastal and inland waters (50EE1257)." EnMAP is funded under the DLR Space Administration with resources from the German Federal Ministry of Economic Affairs and Energy. MH is partly supported by ESA through a Living Planet Fellowship. The C2X database was established by MH in the ESA SEOM C2X project (ESA ESRIN/C-No. 4000113691/15/I-LG). We thank ESA for sponsoring the publication.

ACKNOWLEDGMENTS

We are thankful to Stephen Gehnke for measuring absorption spectra of various phytoplankton cultures. The *in situ* data of Lake Taihu was collected and measured by the Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences. We thank Dr. Hongtao Duan and Prof. Yuanzhi Zhang for providing *in situ* dataset under the support of the Provincial National Science Foundation of Jiangsu of China (BK20160049) and the National Key Research and Development Program of China (2016YFC1402000). The Open Fund at Jiangsu Key Laboratory of Remote Sensing of Ocean Dynamics and Acoustics in NUIST (KHYS1403) is also acknowledged. We finally thank the two reviewers for providing constructive comments to improve the manuscript.

REFERENCES

Ai, Y., Bi, Y., and Hu, Z. (2015). Response of predominant phytoplankton species to anthropogenic impacts in Lake Taihu. *J. Freshw. Ecol.* 30, 99–112. doi: 10.1080/02705060.2014.992052

Aiken, J., Hardman-Mountford, N. J., Barlow, R., Fishwick, J., Hirata, T., and Smyth, T. (2008). Functional links between bioenergetics and bio-optical traits of phytoplankton taxonomic groups: an verarching hypothesis with applications for ocean colour remote sensing. *J. Plankton Res.* 30, 165–181. doi: 10.1093/plankt/fbm098

Alvain, S., Loisel, H., and Dessailly, D. (2012). Theoretical analysis of ocean color radiances anomalies and implications for phytoplankton groups detection in case 1 waters. *Opt. Express* 20, 1070–1083. doi: 10.1364/OE.20.001070

Alvain, S., Moulin, C., and Dandonneau, Y. (2008). Seasonal distribution and succession of dominant phytoplankton groups in the global ocean: a satellite view. *Glob. Biogeochem. Cycles* 22, GB3001. doi: 10.1029/2007GB003154

Alvain, S., Moulin, C., Dandonneau, Y., and Breon, F. M. (2005). Remote sensing of phytoplankton groups in case 1 waters from global SeaWiFS imagery. *Deep Sea Res. Part I* 1, 1989–2004. doi: 10.1016/j.dsr.2005.06.015

Astoreca, R., Rousseau, V., Ruddick, K., Knechciak, C., von Mol, B., Parent, J., et al. (2009). Development and application of an algorithm for detecting *Phaeocystis globosa* blooms in the Case 2 Southern North Sea waters. *J. Plankton Res.* 31, 287–300. doi: 10.1093/plankt/fbn116

Barnsley, M. J., Settle, J. J. , Cutter, M. A., Lobb, D. R., and Teston, F. (2004). The PROBA/CHRIS MISSION: a low-cost smallsat for hyperspectral multiangle observations of the earth surface and atmosphere. *IEEE Trans. Geosci. Remote Sens.* 42, 1512–1520. doi: 10.1109/TGRS.2004.827260

Bracher, A., Bouman, H., Brewin, R. J., Bricaud, A., Brotas, V., Ciotti, A. M., et al. (2017). Obtaining phytoplankton diversity from ocean color: a scientific roadmap for future development. *Front. Mar. Sci.* 4:55. doi: 10.3389/fmars.2017.0005

Bracher, A., Vountas, M., Dinter, T., Burrows, J. P., Röttgers, R., and Peeken, I. (2009). Quantitative observation of *cyanobacteria* and diatoms from space using PhytoDOAS on SCIAMACHY data. *Biogeosciences* 6, 751–764. doi: 10.5194/bg-6-751-2009

Brewin, R. J. W., Sathyendranath, S., Jackson, T., Barlow, R., Brotas, V., Airs, R., et al. (2015). Influence of light in the mixed-layer on the parameters of a three-component model of phytoplankton size class. *Remote Sens. Environ.* 168, 437–450. doi: 10.1016/j.rse.2015.07.004

Bricaud, A., Claustre, H., Ras, J., and Oubelkheir, K. (2004). Natural variability of phytoplanktonic absorption in oceanic waters: influence of the size structure of algal populations. *J. Geophys. Res.* 109, C11. doi: 10.1029/2004JC002419

C2X (2015). *ESA SEOM C2X Project 'Case-2 Extreme Waters'*. Available online at: http://seom.esa.int/page_project014.php

Chen, Y., Qin, B., Teubner, K., and Dokulil, M. T. (2003). Long-term dynamics of phytoplankton assemblages:*Microcystis*-domination in Lake Taihu, a large shallow lake in China. *J. Plankton Res.* 25, 445–453. doi: 10.1093/plankt/25.4.445

Ciotti, A. M., and Bricaud, A. (2006). Retrievals of a size parameter for phytoplankton and spectral light absorption by colored detrital matter from water-leaving radiances at SeaWiFS channels in a continental shelf region of Brazil. *Limnol. Oceanogr. Methods* 4, 237–253. doi: 10.4319/lom.2006.4.237

Corson, M. R., Korwan, D. R., Lucke, R. L., Snyder, W. A., and Davis, C. O. (2008). "The Hyperspectral Imager for the Coastal Ocean (HICO) on the International Space Station," in *IGARSS 2008 - 2008 IEEE International Geoscience and Remote Sensing Symposium*, Vol. 4 (Boston, MA), IV-101–IV-104. doi: 10.1109/IGARSS.2008.4779666

Craig, S. E., Lohrenz, S. E., Lee, Z., Mahoney, K. L., Kirkpatrick, G. J., Schofield, O. M., et al. (2006). Use of hyperspectral remote sensing reflectance for detection and assessment of the harmful alga, *Karenia brevis*. *Appl. Opt.* 45, 5414–5425. doi: 10.1364/AO.45.005414

Deng, J., Qin, B., Paerl, H. W., Zhang, Y., Wu, P., Ma, J., et al. (2014). Effects of nutrients, temperature and their interactions on spring phytoplankton community succession in Lake Taihu, China. *PLoS ONE* 9:e113960. doi: 10.1371/journal.pone.0113960

Eleveld, M. A., Ruescas, A. B., Hommersom, A., Moore, T. S., Peters, S. W., and Brockmann, C. (2017). An optical classification tool for global lake waters. *Remote Sens.* 9:420. doi: 10.3390/rs9050420

Evers-King, H., Bernard, S., Lain, L. R., and Probyn, T. A. (2014). Sensitivity in reflectance attributed to phytoplankton cell size: forward and inverse modelling approaches. *Opt. Express* 22, 11536–11551. doi: 10.1364/OE.22.011536

Foerster, S., Carrère, V., Rast, M., and Staenz, K. (2015). Preface: the Environmental Mapping and Analysis Program (EnMAP) Mission: preparing for its scientific exploitation. *Remote Sens.* 8:957. doi: 10.3390/rs8110957

Folkman, M. A., Pearlman, J., Liao, L. B., and Jarecke, P. J. (2001). "EO-1/Hyperion hyperspectral imager design development characterization and calibration," in *Proceedings SPIE 4151, Hyperspectral Remote Sensing of the Land and Atmosphere* (Sendai), 40.

Greene, R. M., Kolber, Z. S., Swift, D. G., Tindale, N. W., and Falkowski, P. G. (1994). Physiological limitation of phytoplankton photosynthesis in the eastern equatorial Pacific determined from variability in the quantum yield of fluorescence. *Limnol. Oceanogr.* 39, 1061–1074. doi: 10.4319/lo.1994.39.5.1061

Gregg, W., and Rousseaux, C. (2017). Simulating PACE Global Ocean Radiances. *Front. Mar. Sci.* 4:60. doi: 10.3389/fmars.2017.00060

Guanter, L., Kaufmann, H., Segl, K., Foerster, S., Rogass, C., Chabrillat, S., et al. (2015). The EnMAP spaceborne imaging spectroscopy mission for Earth Observation. *Remote Sens.* 7, 8830–8857. doi: 10.3390/rs70708830

Harmel, T., Hieronymi, M., Slade, W., Röttgers, R., Roullier, F., and Chami, M. (2016). Laboratory experiments for inter-comparison of three volume scattering meters to measure angular scattering properties of hydrosols. *Opt. Express* 24, A234–A256. doi: 10.1364/OE.24.00A234

Hieronymi, M., Krasemann, H., Müller, D., Brockmann, C., Ruescas, A., Stelzer, K., et al. (2016). "Ocean colour remote sensing of extreme case-2 waters," in *Proceedings of Living Planet Symposium* (Prague: ESA SP-740).

Hieronymi, M., Müller, D., and Doerffer, R. (2017). The OLCI Neural Network Swarm (ONNS): a bio-geo-optical algorithm for open ocean and coastal waters. *Front. Mar. Sci.* 4:140. doi: 10.3389/fmars.2017.00140

Hirata, T., Aiken, J., Hardman-Mountford, N., Smyth, T. J., and Barlow, R. G. (2008). An absorption model to determine phytoplankton size classes from satellite ocean color. *Remote Sens. Environ.* 112, 3153–3159. doi: 10.1016/j.rse.2008.03.011

Hunter, P. D., Tyler, A. N., Présing, M., Kovács, A. W., and Preston, T. (2008). Spectral discrimination of phytoplankton colour groups: the effect of suspended particulate matter and sensor spectral resolution. *Remote Sens. Environ.* 112, 1527–1544. doi: 10.1016/j.rse.2007.08.003

Isada, T., Hirawake, T., Kobayashi, T., Nosake, Y., Natsuike, M., Imai, I., et al. (2015). Hyperspectral optical discrimination of phytoplankton community structure in Funka Bay and its implications for ocean color remote sensing of diatoms. *Remote Sens. Environ.* 159, 134–151. doi: 10.1016/j.rse.2014.12.006

IOCCG (2000). "Remote sensing of ocean colour in coastal, and other optically-complex, waters," in *Reports of the International Ocean-Colour Coordinating Group*, ed S. Sathyendranath (Dartmouth: IOCCG).

IOCCG (2014). "Phytoplankton functional types from space," in *Reports of the International Ocean-Color Coordinating Group*, No. 15, ed S. Sathyendranath (Dartmouth: IOCCG).

Kim, Y., Yoo, S., and Son, Y. B. (2016). Optical discrimination of harmfulCochlodinium polykrikoidesblooms in Korean coastal waters. *Opt. Express* 24, A1471–A1488. doi: 10.1364/OE.24.0A1471

Knaeps, E., Dogliotti, A. I., Raymaekers, D., Ruddick, K., and Sterckx, S. (2012). *In situ* evidence of non-zero reflectance in the OLCI 1020nm band for a turbid estuary. *Remote Sens. Environ.* 120, 133–144. doi: 10.1016/j.rse.2011.07.025

Kostadinov, T. S., Siegel, D. A., and Maritorena, S. (2009). Retrieval of the particle size distribution from satellite ocean color observations. *J. Geophys. Res.* 114, C09015. doi: 10.1029/2009JC005303

Lee, C. M., Cable, M. L., Hook, S. J., Green, R. O., Ustin, S. L., Mandl, D., et al. (2015). An introduction to the NASA Hyperspectral InfraRed Imager (HyspIRI) mission and preparatory activities. *Remote Sens. Environ.* 167, 6–19. doi: 10.1016/j.rse.2015.06.012

Lorenzoni, L., Toro-Farmer, G., Varela, R., Guzman, L., Rojas, J., Montes, E., et al. (2015). Characterization of phytoplankton variability in the Cariaco Basin using spectral absorption, taxonomic and pigment data. *Remote Sens. Environ.* 167, 259–268. doi: 10.1016/j.rse.2015.05.002

Lubac, B., Loisel, H., Guiselin, N., Astoreca, R., Artigas, L. F., and Meriaux, X. (2008). Hyperspectral and multispectral ocean color inversions to detect *Phaeocystis globosa* blooms in coastal waters. *J. Geophys. Res.* 113, C06026. doi: 10.1029/2007JC004451

Ma, R., Tang, J., and Dai, J. (2006). Bio-optical model with optical parameter suitable for Taihu Lake in water colour remote sensing. *Int. J. Remote Sens.* 27, 4305–4328. doi: 10.1080/01431160600857428

Meini, M., Fossati, E., Giunti, L., Molina, M., Formaro, R., Longo, F., et al. (2015). "The PRISMA mission hyperspectral payload," in *IAC-15-B1.3.7, 66th International Astronautical Congress* (Jerusalem).

Michel, S., Lefèvre-Fonollosa, M., and Hsford, S. (2011). "HYPXIM–A hyperspectral satellite defined for science, security and defence users," in *3rd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)* (Lisbon). doi: 10.1109/WHISPERS.2011.6080864

Millie, D. F., Schofield, O. M., Kirpatrick, G. J., Johnsen, G., Tester, P. A., and Vinyard, B. T. (1997). Detection of harmful algal blooms using photopigments and absorption signatures: a case study of the Florida red tide dinoflagellate, Gymnodinium breve. *Limnol. Oceanogr.* 42, 1240–1251. doi: 10.4319/lo.1997.42.5_part_2.1240

Mitchell, B. G. (1990). "Algorithms for determining the absorption coefficient of aquatic particulates using the quantitative filter technique (QFT)," in *Ocean Optics X* (Orlando, FL: USA SPIE), 137–148. doi: 10.1117/12.21440

Mobley, C. D. (1994). *Light and Water: Radiative Transfer in Natural Waters*. San Diego, CA: Academic Press.

Mobley, C. D. (1999). Estimation of the remote-sensing reflectance from above-surface measurements. *Appl. Opt.* 38, 7442–7455. doi: 10.1364/AO.38.007442

Mobley, C. D., and Sundman, L. K. (2013). *HydroLight 5.2 - EcoLight 5.2 Technical Documentation*. Sequoia Scientific, Inc.

Morel, A. (1987). Chlorophyll-specific scattering coefficient of phytoplankton. A simplified theoretical approach. *Deep Sea Res. Part A Oceanogr. Res. Pap.* 34, 1093–1105. doi: 10.1016/0198-0149(87)90066-5

Mouw, C. B., Hardman-Mountford, N. J., Alvain, S., Bracher, A., Brewin, R. W., Bricaud, A., et al. (2017). A consumer's guide to satellite remote sensing of multiple phytoplankton groups in the global ocean. *Front. Mar. Sci.* 4:41. doi: 10.3389/fmars.2017.00041

Mouw, C. B., and Yoder, J. A. (2010). Optical determination of phytoplankton size composition from global SeaWiFSimagery, *J. Geophys. Res.* 115, C12018. doi: 10.1029/2010JC006337

Mueller, J. L., Fargion, G. S., and McClain, C. R. (2003). O*cean Optics Protocols for Satellite Ocean Color Sensor Validation,* Vol. 1–4. Greenbelt, MD: NASA Goddard Space Flight Center.

Nair, A., Sathyendranath, S., Platt, T., Morales, J., Stuart, V., Forget, M.-H. E., et al. (2008). Remote sensing of phytoplankton functional types. *Remote Sens. Environ.* 112, 3366–3375. doi: 10.1016/j.rse.2008.01.021

Robertson Lain, L., Bernard, S., and Matthews, M. W. (2017). Understanding the contribution of phytoplankton phase functions to uncertainties in the water colour signal. *Opt. Express* 25, A151–A165. doi: 10.1364/OE.25.00A151

Röttgers, R., Doerffer, R., McKee, D., and Schönfeld, W. (2016). *The Water Optical Properties Processor (WOPP): Pure Water Spectral Absorption, Scattering, and Real Part of Refractive Index Model.* ATBD, Issue 1.8, ESA Water Radiance project.

Röttgers, R., Häse, C., and Doerffer, R. (2007). Determination of the particulate absorption of microalgae using a point-source integrating-cavity absorption meter: verification with a photometric technique, improvements for pigment bleaching and correction for chlorophyll fluorescence. *Limnol. Oceanogr. Methods* 5, 1–12. doi: 10.4319/lom.2007.5.1

Sadeghi, A., Dinter, T., Vountas, M., Taylor, B. B., Soppa, M. A., Peeken, I., et al. (2012). Improvement to the PhytoDOAS method for identification of coccolithophores using hyper-spectral satellite data. *Ocean Sci.* 8, 1055–1070. doi: 10.5194/os-8-1055-2012

Taylor, B. B., Torrecilla, E., Bernhardt, A., Taylor, M. H., Peeken, I., Röttgers, R., et al. (2011). Bio-optical provinces in the eastern Atlantic Ocean and their biogeographical relevance. *Biogeosciences* 8, 3609–3629. doi: 10.5194/bg-8-3609-2011

Torrecilla, E., Stramski, D., Reynolds, R. A., Millán-Núñez, E., and Piera, J. (2011). Cluster analysis of hyperspectral optical data for discriminating phytoplankton pigment assemblages in the open ocean. *Remote Sens. Environ.* 115, 2578–2593. doi: 10.1016/j.rse.2011.05.014

Twardowski, M. S., Boss, E., Macdonald, J. B., Pegau, W. S., Barnard, A. H., and Zaneveld, J. R. V. (2001). A model for estimating bulk refractive index from the optical backscattering ratio and the implications for understanding particle composition in case I and case II waters. *J. Geophys. Res.* 106, 14129–14142. doi: 10.1029/2000JC000404

Uitz, J., Stramski, D., Reynolds, R. A., and Dubranna, J. (2015). Assessing phytoplankton community composition from hyperspectral measurements of phytoplankton absorption coefficient and remote-sensing reflectance in open-ocean environments. *Remote Sens. Environ.* 171, 58–74. doi: 10.1016/j.rse.2015.09.027

Valente, A., Sathyendranath, S., Brotas, V., Groom, S., Grant, M., Taberner, M., et al. (2016). A compilation of global bio-optical *in situ* data for ocean-colour satellite applications. *Earth Syst. Sci. Data* 8, 235–252. doi: 10.5194/essd-8-235-2016

Wang, G., Lee, Z., Mishra, D. R., and Ma, R. (2016). Retrieving absorption coefficients of multiple phytoplankton pigments from hyperspectral remote sensing reflectance measured over cyanobacteria bloom waters. *Limnol. Oceanogr. Methods* 14, 432–447. doi: 10.1002/lom3.10102

Werdell, P. J., Roesler, C. S., and Goes, J. I. (2014). Discrimination of phytoplankton functional groups using an ocean reflectance inversion model. *Appl. Opt.* 53, 4833–4849. doi: 10.1364/AO.53.004833

Wolanin, A., Soppa, M. A., and Bracher, A. (2016). Investigation of spectral band requirements for improving retrievals of phytoplankton functional types. *Remote Sens.* 8:871. doi: 10.3390/rs8100871

Xi, H. (2011). *Water Optical Properties and Water Color Remote Sensing in Optically Deep and Shallow Waters of Lake Tahu, China*. Dissertation, The Chinese University of Hong Kong, Hong Kong.

Xi, H., Hieronymi, M., Röttgers, R., Krasemann, H., and Qiu, Z. (2015). Hyperspectral differentiation of phytoplankton taxonomic groups: a comparison between using remote sensing reflectance and absorption spectra. *Remote Sens.* 7, 14781–14805. doi: 10.3390/rs71114781

Zhu, G., Qin, B., Gao, G., Zhang, L., Luo, L., and Zhang, Y. (2007). "Effects of hydrodynamics on phosphorus concentrations in water of Lake Taihu, a large, shallow, eutrophic lake of China," in *Eutrophication of Shallow Lakes with Special Reference to Lake Taihu, China,* eds B. Qin, Z. Liu, and K. Havens (Dordrecht: Netherlands Springer), 53–61. doi: 10.1007/978-1-4020-6158-5_6

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Xi, Hieronymi, Krasemann and Röttgers. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Marine Science

ORIGINAL RESEARCH
published: 08 September 2017
doi: 10.3389/fmars.2017.00283

An Exact Solution For Modeling Photoacclimation of the Carbon-to-Chlorophyll Ratio in Phytoplankton

Thomas Jackson[1]*, *Shubha Sathyendranath*[1,2] *and Trevor Platt*[1]

[1] *Plymouth Marine Laboratory, Plymouth, United Kingdom,* [2] *National Centre for Earth Observation, Plymouth Marine Laboratory, Plymouth, United Kingdom*

A widely-used theory of the photoacclimatory response in phytoplankton has, until now, been solved using a mathematical approximation that puts strong limitations on its applicability in natural conditions. We report an exact, analytic solution for the chlorophyll-to-carbon ratio as a function of the dimensionless irradiance (mixed layer irradiance normalized to the photoadaptation parameter for phytoplankton) that is applicable over the full range of irradiance occurring in natural conditions. Application of the exact solution for remote-sensing of phytoplankton carbon at large scales is illustrated using satellite-derived chlorophyll, surface irradiance data and mean photosynthesis-irradiance parameters for the season assigned to every pixel on the basis of ecological provinces. When the exact solution was compared with the approximate one at the global scale, for a particular month (May 2010), the results differed by at least 15% for about 70% of Northern Hemisphere pixels (analysis was performed during the northern hemisphere Spring bloom period) and by more than 50% for 24% of Northern Hemisphere pixels (approximate solution overestimates the carbon-to-chlorophyll ratio compared with the exact solution). Generally, the divergence between the two solutions increases with increasing available light, raising the question of the appropriate timescale for specifying the forcing irradiance in ecosystem models.

Keywords: photoacclimation, phytoplankton, carbon-to-Chlorophyll, photo-physiology, primary production

OPEN ACCESS

Edited by:
Laura Lorenzoni, University of South Florida, United States

Reviewed by:
Christoph Voelker, Alfred-Wegener-Institut für Polar- und Meeresforschung, Germany
Greg M. Silsbe, University of Maryland Center For Environmental Sciences, United States

***Correspondence:**
Thomas Jackson
thja@pml.ac.uk

Specialty section:
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

Received: *01 March 2017*
Accepted: *21 August 2017*
Published: *08 September 2017*

Citation:
Jackson T, Sathyendranath S and Platt T (2017) An Exact Solution For Modeling Photoacclimation of the Carbon-to-Chlorophyll Ratio in Phytoplankton. Front. Mar. Sci. 4:283. doi: 10.3389/fmars.2017.00283

1. INTRODUCTION

When quantifying the standing stock of marine phytoplankton or its rate of change, various metrics can be used, depending on the application envisaged. The possibilities include cell count, cell volume, carbon content, nitrogen content and chlorophyll concentration. Primary production (rate of production of organic material by phytoplankton through photosynthesis) is typically measured in carbon units, a convenient measure in studies of the global carbon cycle. It is also a practical unit in calculations of fluxes of material through the food chain or through the water column. On the other hand, chlorophyll-a concentration is by far the most commonly-used measure of phytoplankton abundance. There are many reasons for this choice also, including its principal role in the photosynthetic apparatus and in primary production; its presence in all types of phytoplankton, either in its common form or as derivates such as divinyl chlorophyll-a; and the

ease with which it can be measured at a variety of scales, from single cells in the laboratory to ocean-basin scales using remote sensing by satellites.

The carbon-to-chlorophyll ratio, necessary to convert between these two common measures of phytoplankton biomass, is a dynamic, and highly-variable property of phytoplankton. Phytoplankton growing in high-light environments need to absorb only a small fraction of the available light, and they adapt to the ambient light field by reducing their pigment quota, resulting in a high carbon-to-chlorophyll ratio. The opposite is true in low-light conditions, for example in deep chlorophyll maxima in the ocean gyres, where chlorophyll concentration increases relative to the carbon concentration (Cullen, 1982, 2015; Morel and Berthon, 1989). Estimating such changes in carbon-to-chlorophyll ratio in response to variations in available light, i.e., due to photo-acclimation, is not a trivial task, but it is an essential step in many biogeochemical models. As reviewed by Halsey and Jones (2015), nutrients can also play a role in carbon-to-chlorophyll variations, although the sign of the change depends on the nutrient in question, with some nutrients being utilized for the production of pigments and others for photosystem reaction centers.

The links between carbon-to-chlorophyll-a ratios, photosynthesis and photo-acclimation are discussed in the works of Platt and Jassby (1976) and Geider (1987). Subsequently, Geider et al. (1996, 1997) developed a mechanistic model of photo-acclimation that has become commonly used to assign the chlorophyll:carbon ratio of phytoplankton populations in ecosystem models (Hickman et al., 2010; Dutkiewicz et al., 2015; Laufkötter et al., 2015). In a further development, Geider et al. (1998) dealt with the possible variations in photosynthetic parameters with nutrients and temperature. But the approximation used to derive the solution to the photoacclimation model (Geider et al., 1997) still limits the range of irradiance levels for which the solution holds. Some authors have addressed this problem by a numerical solution to the Geider et al. (1997) model rather than the approximation (e.g., Li et al., 2010), while others have imposed a numerical upper limit on the C:Chl ratio (Butenschön et al., 2016) to constrain model output.

Here, we present an exact solution that dispenses with the need for an approximation, removes the existing limitation and is therefore universally applicable. We examine conditions under which the differences between the approximate solution and the exact solution become significant, and discuss some of the implications for implementation of the model to compute carbon-to-chlorophyll ratios under natural environmental conditions. We show that, in some instances, the differences between the exact and approximate solutions depend on the assumptions in the model regarding the time scales on which photo-acclimation occurs in phytoplankton.

2. DATA

To demonstrate some applications of the new solution, a variety of datasets were used, which are described here briefly.

Monthly, climatological Photosynthetically Available Radiation (PAR) data from SeaWiFS (Frouin et al., 2002) are used for demonstrating an application of the new solution at large scales (http://oceancolor.gsfc.nasa.gov/cms/atbd/par). We used monthly composites to minimize data gaps. Climatological mixed-layer depth (MLD) was obtained from de Boyer Montégut et al. (2004) and also re-gridded onto a 9 km grid to match the input PAR data.

We used mean values of photosynthesis-irradiance parameters (the assimilation number P_m^B and the initial slope α^B, where the superscript B indicates normalisation to biomass B, in chlorophyll units; see **Table 1**) organized by season and by ecological provinces (as defined by Longhurst et al. 1995), from Mélin and Hoepffner (2004), which were then re-gridded, with a 30 × 30 pixel smoothing filter, to 9 km resolution to match the PAR data. These parameters can be used to calculate the chlorophyll-normalized production (P^B) at any value I of photosynthetic irradiance (PAR), in the absence of photoinhibition, as described by Platt et al. (1980):

$$P^B = P_m^B \left(1 - \exp(\frac{-\alpha^B I}{P_m^B})\right) . \quad (1)$$

The P_m^B and α^B values allow the calculation of the photoadaptation parameter I_k, defined as P_m^B/α^B. Surface Chl-a concentration from the Ocean Colour Climate Change Initiative (OC-CCI) dataset, Version 2.0 (European Space Agency, available online at http://www.esa-oceancolour-cci.org/) and the spectral light-transmission model of Sathyendranath and Platt (1988) were used to compute K_d, the diffuse attenuation coefficient for photosynthetically-active radiation for the mixed layer. The daily average irradiance in the mixed layer (I_m) was computed as

$$I_m = \frac{\bar{I}_0}{K_d Z_m}(1 - \exp(-K_d Z_m)), \quad (2)$$

where $\bar{I}_0$ is the daily (24 h) average PAR at the sea-surface and Z_m is the mixed-layer depth (Platt et al., 1991; Cloern et al., 1995).

An *in-situ* bio-optical dataset of particulate organic carbon (POC), chlorophyll, and photosynthesis-irradiance parameters (Sathyendranath et al., 2009) was also used in this work. This dataset lacked information on PAR and MLD, which were filled in using the climatological data mentioned above.

3. EXACT SOLUTION FOR THE CHLOROPHYLL-TO-CARBON RATIO (θ) IN THE GEIDER ET AL. (1997) MODEL

According to Geider et al. (1997), the chlorophyll-to-carbon ratio, θ, is a function of irradiance I:

$$\theta^2 = \theta_m a \left(1 - \exp\left(-\frac{\theta}{a}\right)\right) , \quad (3)$$

where (θ_m) is a prescribed model parameter, corresponding to the maximum attainable value of θ. The above equation is

equivalent to equation A12 in Geider et al. (1997), noting that there is a typographical error in the equation, such that the denominator of the argument to the exponential term should be a, and not $\alpha^B I$. For conditions of balanced growth, Geider et al. (1997) point out that their parameter k_{chl}, which represents the maximum proportion of photosynthesis that can be directed to chlorophyll-a synthesis, would be equivalent to the parameter θ_m. We have applied the equivalence here, such that the solution would be valid only for balanced growth. The model development also assumes that the specific respiration rates of carbon and chlorophyll are either negligible or equal to each other.

We note that $a = P_m^C/(\alpha^B I)$, where P_m^C is the carbon-specific, light saturated photosynthesis. By definition, $P_m^C = P_m^B\theta$, such that $a = P_m^B\theta/(\alpha^B I)$. Substitution into Equation (3) gives:

$$\theta^2 = \theta_m a\left(1 - \exp\left(-\frac{\theta\alpha^B I}{P_m^B\theta}\right)\right). \tag{4}$$

Applying the equivalence $I_k = P_m^B/\alpha^B$, we get

$$\theta^2 = \theta_m a\left(1 - \exp\left(-\frac{I}{I_k}\right)\right), \tag{5}$$

and setting $I/I_k = I_*$, a dimensionless irradiance, the equation becomes

$$\theta^2 = \theta_m\frac{\theta}{I_*}\left(1 - \exp\left(-I_*\right)\right). \tag{6}$$

Solution for θ is obtained by simplifying the equation above:

$$\theta = \frac{\theta_m}{I_*}\left(1 - \exp\left(-I_*\right)\right). \tag{7}$$

The solution expresses θ as a function I_*, such that the chlorophyll-to-carbon ratio can be calculated explicitly as a function of the dimensionless scaled irradiance (I_*). Note that the carbon-to-chlorophyll ratio $\chi = 1/\theta$. As I_* tends to zero, the exact solution (Equation 7) tends to θ_m. As I_* tends to infinity, the solution tends to zero. However, this limit for high values of I_* is approached very slowly, well beyond reasonable values of I_* that might be expected in the natural environment. The solution remains well-constrained for plausible values of I_*.

We note that the same solution is obtained when, instead of substituting $P_m^C = P_m^B\theta$, we make the equivalent change of $\alpha_B = \alpha^C/\theta$. The key to solution is consistency: both parameters have to be normalized to the same quantity, carbon or chlorophyll, it does not matter which. The solution is indifferent to the choice as (apart from θ) it contains only the dimensionless quantity I_*. However, ecosystem models are often formulated to use carbon-normalized P_m^C as input, along with α^B, in which case, Equation 7 becomes (see also Li et al., 2010):

$$\theta = ((\theta_m P_m^C)/(I\alpha^B\theta))(1 - \exp((-I\alpha^B\theta)/P_m^C)). \tag{8}$$

In this context, θ can be retrieved from the above equation iteratively.

It is is also possible to calculate the sensitivity (relative) of θ to changes (relative) in I_*; and we find

$$\left|\left(\frac{d\theta}{\theta}\Big/\frac{dI_*}{I_*}\right)\right| = \frac{\left(\exp(-I_*)(1+I_*) - 1\right)}{\left(1 - \exp(-I_*)\right)} \leq 1, \tag{9}$$

such that the relative error in θ will not be greater than that in I_*.

3.1. The Approximate Solution

Geider et al. (1997) provided an approximate solution for θ using the first three terms of the Taylor expansion of $\exp(-\theta/a)$:

$$\theta^2 = \theta_m a\left(1 - 1 + \frac{\theta}{a} - \frac{\theta^2}{2a^2}\right). \tag{10}$$

For comparison with the exact solution (Equation 7), we can rearrange terms in the approximate solution, such that it is also expressed as a function of I_*. Following an initial simplification:

$$\theta^2 = \theta_m a\left(\frac{\theta}{a} - \frac{\theta^2}{2a^2}\right); \theta = \theta_m\left(1 - \frac{\theta}{2a}\right). \tag{11}$$

We can then substitute for $a = P_m^B\theta/(\alpha^B I)$ to find

$$\theta = \theta_m\left(1 - \frac{I_*}{2}\right). \tag{12}$$

Geider et al. (1997) noted that the approximation holds for only for $I_* < 1$. This limitation is overcome by the analytic solution for θ (Equation 7), which is valid for all values of I_*.

The approximate solution (Equation 12) and the exact solution (Equation 7) are identical and equal to θ_m as I_* tends to zero. But the approximate solution θ becomes zero when $I_* = 2$, and becomes negative for higher values. Hence the limitation with using the approximate solution for high values of I_*.

3.2. Effects of Nutrients and Temperature

We see from the exact solution (Equation 7) that θ depends on P_m^B through I_k. In the Geider et al. (1998) model, effects of nutrient limitation and ambient temperature on P_m^B are accounted for, as follows:

$$P_m^C = P_{ref}^C\frac{N}{N + K_N}f(T), \tag{13}$$

where P_{ref}^C is the maximum C-specific rate of photosynthesis at a reference temperature, T is the ambient temperature, $f(T)$ is the Arrhenius function, N is the nitrate concentration and K_N is the half saturation constant for nitrate uptake.

P_m^B, defined as $P_m^C \times \theta$, therefore contains implicitly the effects of temperature and nutrients on photosynthetic rates. Consequently, Equation 7 accounts for their effects on θ through P_m^B. Since P_m^B is more readily measured in the field than P_m^C, the new solution facilitates the study of C:Chl ratio in the natural environment.

TABLE 1 | Definitions of symbols.

Symbol	Definition	Units
θ	Chlorophyll-a:Carbon ratio	dimensionless
θ_m	Maximum Chlorophyll-a:Carbon ratio	dimensionless
χ	Carbon:Chlorophyll-a ratio	dimensionless
B	Chlorophyll-a concentration	mg Chl-a m^{-3}
K_d	Downwelling attenuation coefficient	m^{-1}
Z_m	Mixed layer depth	m
C_p	Phytoplankton Carbon	mg C m^{-3}
POC	Particulate Organic Carbon	mg C m^{-3}
I_m	Mean daily irradiance in the mixed layer	Wm^{-2}
$\bar{I}_0$	Mean daily surface irradiance	Wm^{-2}
I_k	Photoadaptation parameter	Wm^{-2}
$I_\star$	Dimensionless scaled irradiance	dimensionless
I	Irradiance	Wm^{-2}
P_m^B	Assimilation number	mgC $mgChl^{-1}h^{-1}$
α^B	PI curve initial slope	mgC $mgChl^{-1}$ $(Wm^{-2})^{-1}$ h^{-1}
P_m^C	Carbon specific assimilation number	mgC $mgC^{-1}h^{-1}$
R^B	Respiration loss of Chlorophyll-a	d^{-1}
R^C	Respiration loss of Carbon	d^{-1}
μ	Growth Rate	d^{-1}
ξ	Cost of Biosynthesis	gC gN^{-1}

4. RESULTS

4.1. Comparison between Exact and Approximate Solutions

4.1.1. Theoretical Comparison

The approximate solution (Equation 12) and the exact solution (Equation 7) for $1/\theta = \chi$ are shown in **Figure 1** for three values of θ_m: 0.005, 0.01 and 0.02 (corresponding to carbon-to-chl ratios of 200, 100, and 50). For low values of I_* the exact and approximate solutions are practically indistinguishable from each other. But as I_* approaches and exceeds 0.8, the deviation between them becomes significant. For I_* close to 2.0 the approximate solution for θ tends to zero and the inverse of θ (the carbon-to-chlorophyll ratio, χ) tends to infinity, whereas the exact solution remains stable. **Figure 1A** shows that the absolute error is dependent on both θ_m and I_*. However, the relative error (**Figure 1B**) is independent of θ_m. The approximation overestimates the carbon-to-chlorophyll ratio by around 15% when $I_* = 0.8$, by 50% at $I_* = 1.235$ and by 100% at $I_* = 1.478$.

4.1.2. A practical example

To see whether the differences between the exact and approximate solutions are likely to be significant under conditions encountered in the natural environment, we made some calculations at the global scale, using a combination of satellite and *in situ* data. The sequence of images in **Figure 2** shows the input data fields (daily mean irradiance at the surface, mixed-layer depth, photoadaptation parameter I_k and chlorophyll-a concentration) and resultant daily mean irradiance in the mixed layer (I_m) and I_* for May 2010, where in this instance $I_* = I_m/I_k$. Of the valid ocean pixels in **Figure 2F**), 70.3% in the Northern hemisphere (which at the time would be the hemisphere of greater phytoplankton growth due to the spring bloom) have I_* values greater than 0.8, such that for these pixels the difference between the approximate and exact solutions would be greater than 15%. The error in the approximate solution is greater than 50% in some 24% of the Northern hemisphere pixels. During November a similar situation occurs in the Southern hemisphere, with I_* values greater than 0.8 in 61.5% of pixels (results not shown).

This demonstrates that phytoplankton in the surface oceans are frequently exposed to conditions in which the difference between the approximate and exact solution for θ is significant, and worth accounting for.

4.2. Computation of Phytoplankton Carbon in the Ocean

In this section, we first impliment the analytic solution using the *in situ* bio-optical data to compute phytoplankton carbon at the observation points. Since it is known that θ_m varies with phytoplankton type (Geider et al., 1997), we assigned values of θ_m according to phytoplankton size classes. First, based on the work of Brewin et al. (2010), the chlorophyll-a concentration at each data point was used to estimate the proportions of the three phytoplankton size classes (micro-, nano- and pico-plankton) present in the sample. Next, based on the C:Chl ratios given in Sathyendranath et al. (2009) for different phytoplankton types sampled in the natural environment, θ_m was set to 0.05, 0.02, 0.008 for micro-, nano- and pico-phytoplankton, corresponding to a minimum C:Chl ratio of 20, 50 and 125 for each size class. These values are consistent with θ_m values reported by Geider et al. (1997) for various phytoplankton species in culture and also by Li et al. (2010) in the natural marine environment. The θ_m for the populations was then computed as a weighted sum of the three components of the population. As θ_m dictates a maximum Chl:C ratio, it also sets a minimum C:Chl ratio. The photosynthesis-irradiance parameters (P_m^B and α^B) in the database were then used to compute I_k (*in situ*) and the daily average I_* for the mixed layer, given the daily average I_m for the layer.

For each sample in the *in situ* dataset taken at a depth within the climatological mixed-layer depth (410 samples), we calculated the C:Chl ratio χ using I_* and θ_m, and then multiplied χ by the chlorophyll concentration measured *in situ* to estimate total phytoplankton carbon (C_p). **Figure 3** shows measured POC plotted against computed phytoplankton carbon (C_p). The model imposes no upper limit on the C:Chl ratio. Therefore, if the model parameters were incorrectly assigned, it could lead to many C_p values being greater than the measured POC, which would clearly indicate an overestimation of phytoplankton carbon, since it should not exceed POC concentration. The C_p estimated using the analytical solution and estimated θ_m exceeds total POC in only 4 of the 410 points. Most of the C_p:POC ratios lie in the range of ≈10–70% with a mean of 31%, which is consistent with existing *in situ* measurements from the Atlantic and Pacific oceans (Martinez-Vicente et al.,

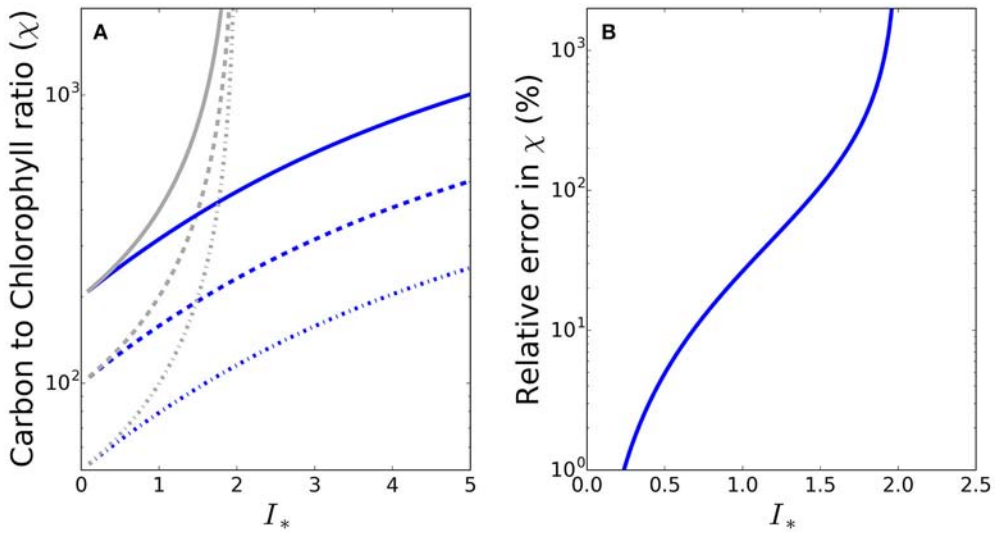


FIGURE 1 | (A) The divergence of the approximate (gray) and exact (blue) solution estimates of the C:Chl ratio as a function of I_*. Solid, dotted and dot-dash lines are for θ_m values of 0.005, 0.01, and 0.02 respectively. The exact solution is clearly stable across the full range of I_* values, while the approximate solution is not. **(B)** The relative difference between the approximate and the exact solutions as a function of I_*.

2013; Graff et al., 2015), suggesting that θ values are not grossly underestimated either. The results using the approximate solution are significantly higher ($I_* > 0.8$ and difference > 15%) in 130 of the 410 *in situ* measurements. The differences when using the *in situ* I_k values were greater than when the calculations were performed using the province-based average I_k values, demonstrating that sometimes, the errors from the approximate solution are reduced when using broadly-averaged fields of I_k, since averaging eliminates extreme values.

As the calculations yielded plausible values of phytoplankton carbon when compared with measured POC values, we applied the method to the I_* map and the satellite-derived chlorophyll field shown in **Figure 2** to produce global maps of C:Chl ratio and C_p. The results are compared with the approximate solution to the Geider et al. (1997) model and with the method of Sathyendranath et al. (2009) (see **Figure 4**), which implemented the equation $C_p = 64B^{0.63}$, where B is Chlorophyll-a concentration (see their **Figure 1B**). As expected, the C:Chl ratios from the exact solution are lower than those from the approximate solution, with the largest differences occurring in regions of high I_*. The corresponding C_p values are also lower for the exact solution. The distribution of C_p values using the analytical solution appears more natural than those using the approximate solution, with fewer artificial boundaries present in the output fields.

The exact solution for C_p is also closer (smaller mean absolute-difference) than the approximate one to the results from the empirical approach of Sathyendranath et al. (2009), but some of the similarities have to be attributed to the use of θ_m values from Sathyendranath et al. (2009) in this work. Both the exact solution and the method of Sathyendranath et al. (2009) show the anticipated increase in C:Chl ratio toward the subtropical gyres (associated with the dominance of pico-plankton in these areas), although the magnitudes differ. Similarly, in both these examples, the C:Chl ratio decrease toward the Southern Ocean. The similarities in patterns are encouraging. However, the exact solution provides a lower range for the C:Chl ratio globally, when compared with the outputs from the method of Sathyendranath et al. (2009). This is to be expected as the averaging of I_k by province and by season removes extreme values, as well as any small-scale variability that might otherwise be present in a dynamic assignment of I_k. On the other hand, we recognize that the method of Sathyendranath et al. (2009) is purely empirical and was designed to provide something of an upper limit to the carbon-to-chlorophyll ratio, whereas the Geider et al. (1997) model has a strong mechanistic basis and is able to account for the effects of photo-acclimation on θ. Clearly, more work is required to reconcile the differences between the empirical and theoretical approaches.

4.3. Application in Marine Ecosystem Models

In addition to the remote-sensing applications demonstrated above, the Geider et al. (1997) model is also used extensively in marine ecosystem models (Laufkötter et al., 2015). But to estimate the impact that the exact solution might have on the calculated fields of carbon-to-chlorophyll ratio, we have to consider the time scales over which light is averaged, before carbon-to-chlorophyll ratio is computed in the models. For example, in the European Regional Seas Ecosystem Model (ERSEM), the instantaneous light field is used to compute θ at each time step of the model (Butenschön et al., 2016). The common time step for ERSEM is 15 min. But other models, such as the "Darwin" model developed at MIT, perform these calculations at longer time steps (Dutkiewicz et al., 2015). A model with a 24 h time step might use daily-averaged light fields. Calculations that use short time-steps would have a greater range in I_* values, relative to those that use daily averages.

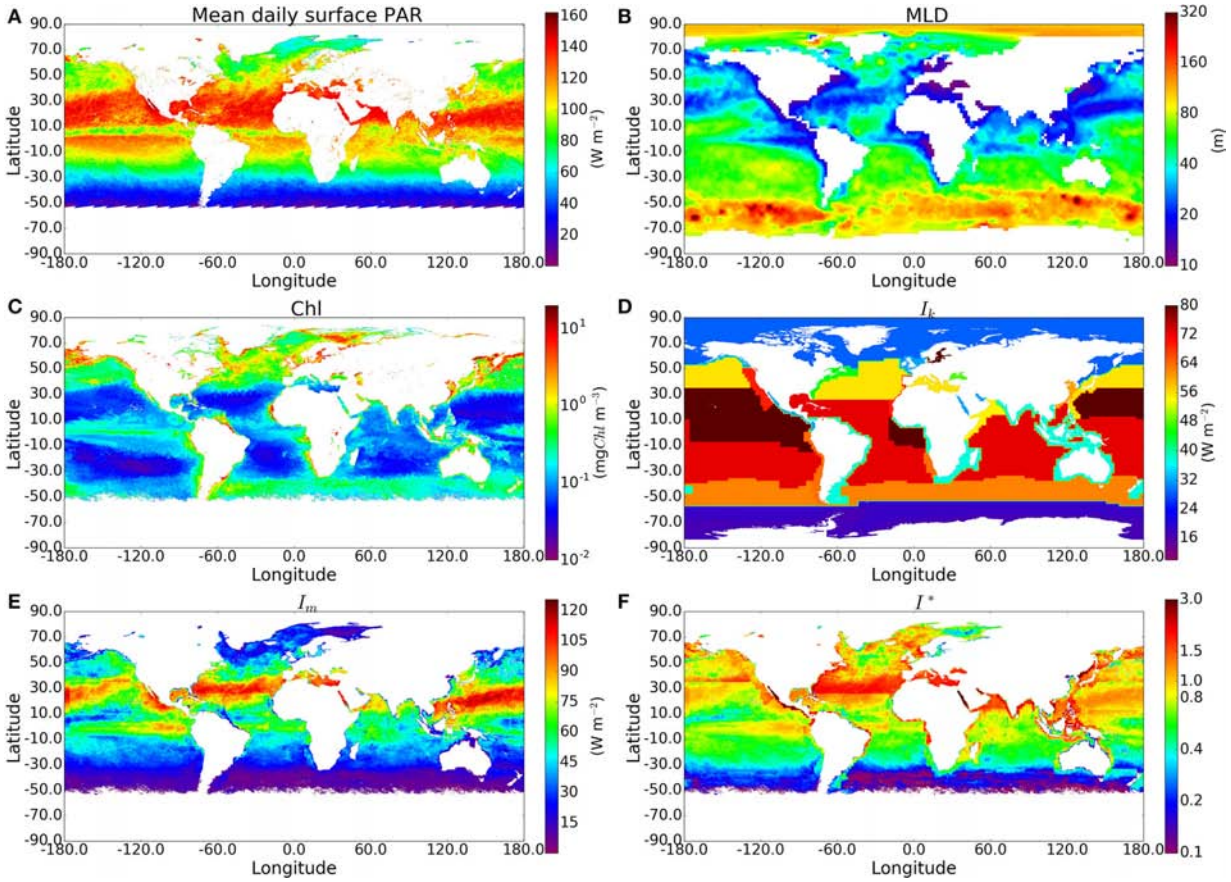


FIGURE 2 | Map showing the input data and resulting I_* estimates at the global scale during May 2010. **(A)** SeaWiFS PAR product converted into W m^{-2} (Morel and Smith, 1974) and averaged over the day (24 h), **(B)** MLD climatology (de Boyer Montégut et al., 2004), **(C)** OC-CCI v2.0 monthly composite of chlorophyll-a, **(D)** Biogeochemical-province based I_k (Mélin and Hoepffner, 2004), **(E)** daily-mean mixed-layer irradiance, and **(F)** daily-mean mixed-layer dimensionless irradiance. Values of I_* around 0.8 or greater (yellow and warmer colors) will give a significant difference between the approximate and exact solutions for C:Chl.

An example of a calculation of θ done at a 2-h time-step is shown in **Figure 5**, where results are plotted for optical depths of zero (surface) to 4. Note that one optical depth is the depth at which light is reduced to $1/e$ of the initial value, and that only 1% of the surface value remains at an optical depth of 4.6. In this example, we used a fixed I_k value of 50 Watts m^{-2}, and a noon-time maximum value of I at the surface of 400 W m^{-2}, and set $\theta_m = 0.01$. The total daily irradiance was allowed to vary, over a 12-h day, as described by a sine function. At noon, I_* values of 1.0 or greater occur even down to the first optical depth and the errors in the approximate solution are high in the surface waters for a large portion of the day. The value of irradiance averaged over 24 h at the optical depth of 1 (dashed lines shown for comparison) is well below the peak values seen at noon; and as expected, the difference between the exact and approximate solutions is reduced, though still significant (over 20%), for this case. Even in this instance, the errors would increase toward the surface, as average light increased. This is consistent with the findings of Moore et al. (2006) that for surface populations, the peak irradiance can be significantly higher than the measured I_k.

5. DISCUSSION AND CONCLUSION

In this paper we have presented a new, exact solution for the Geider et al. (1997) model for estimating the C:Chl ratio in phytoplankton as a function of a dimensionless irradiance scaled to the photoadaptation parameter, I_k. The result is directly applicable to remote-sensing and modeling of marine ecosystems, as demonstrated here, but finds further applications in modeling phytoplankton physiological properties, growth rates and stoichiometry (Sathyendranath et al., 2009; Dutkiewicz et al., 2015; Laufkötter et al., 2015). Using an *in situ* bio-optical database and the model, we have computed phytoplankton carbon, and shown that the derived ratios of phytoplankton carbon to POC were plausible.

The Geider et al. (1997) model was initially conceived to be implemented with P_m^C and α^B as inputs. The work presented here provides a new exact solution to the model. The advantage of the solution is that it allows the Geider et al. (1997) model to be implemented in any instance where there are direct measurements or indirect estimates of I_k. So the starting point for implementation of the new solution would be estimates of I_k or

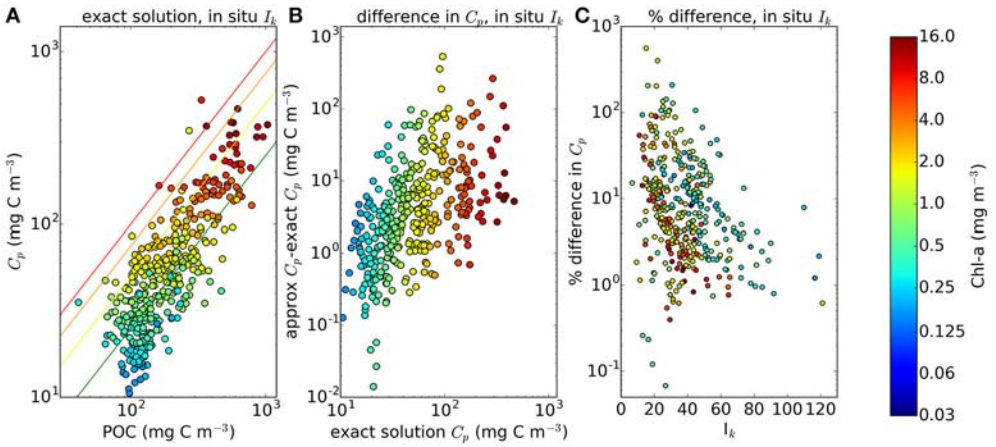


FIGURE 3 | Comparison of phytoplankton Carbon estimates using the approximate and exact solution with *in situ* I_k data from around 400 samples, mostly from the N.W Atlantic region. **(A)** Calculated Phytoplankton Carbon ($\theta * B$) in relation to POC measured for the BIO samples using the exact solution. Red, orange, yellow and green lines correspond to phytoplankton carbon equalling 100, 75, 50, and 25% of POC respectively. The θ_m values are calculated using an estimate of the community size structure calculated using the method of Brewin et al. (2010). **(B,C)** show the absolute and % difference between results from the exact and approximate solutions.

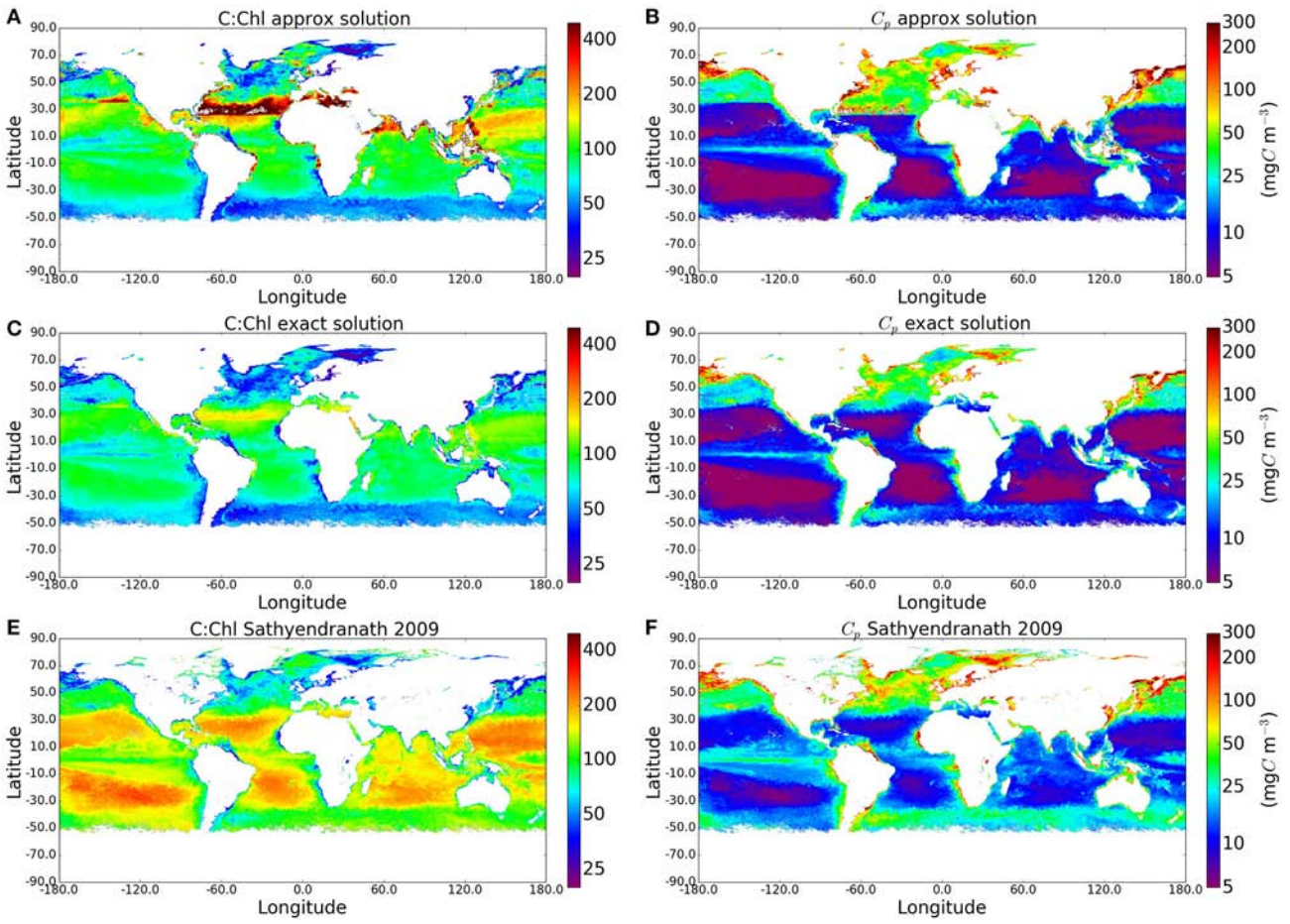


FIGURE 4 | Maps comparing the C:Chl and C_p estimates using the original approxmiate solution, the new exact solution, and the method of Sathyendranath et al. (2009) globally for May 2010. The $I_\star$ and Chl input fields can be seen in **Figure 2**.

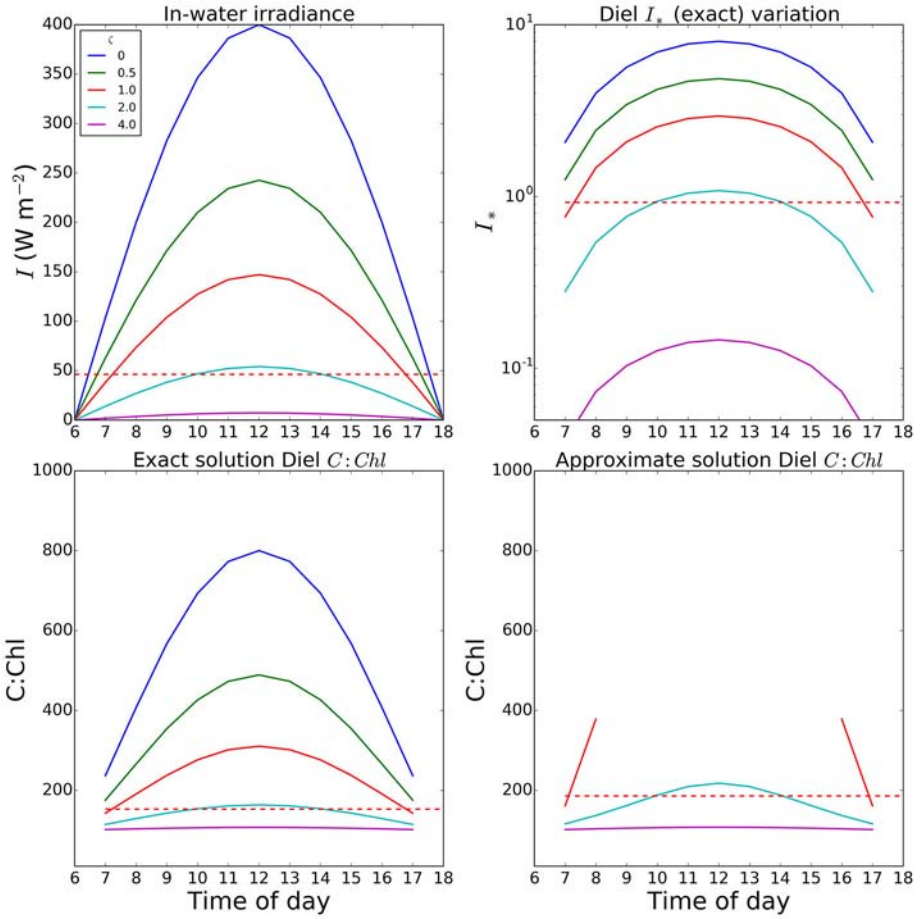


FIGURE 5 | Variation in I_* and resultant C:Chl ratios through the a diurnal cycle at various optical depths in a simple optical model. The value set for I_k (50.0 W m^2) was taken as a reasonable value from the fields seen in **Figure 2**. Both I_k and θ_m (0.01) were assumed uniform within a mixed layer extending to the euphotic depth. Dashed red lines show the value for the first optical depth when calculations are performed using a daily mean (24-h time step). Missing values in the final panel are due to values of I_* exceeding the limit of the Taylor expansion.

P_m^B and α^B. In this regard, the new solution takes the Geider et al. (1997) model in a new direction. However, in ecosystem models that are implemented with with P_m^C and α^B as inputs, the value of θ can be found from the exact solution iteratively (note that Li et al., 2010 have also proposed a numerical solution). The extra computation required for an iterative solution would certainly be worth the effort, especially for $I_* > 0.8$, when errors in the approximate solution begin to be greater than 15% (**Figure 1**).

Irradiance is a fundamental driver of phytoplankton growth, and phytoplankton employ a suite of strategies in response to the range of irradiance conditions in the global oceans. Some groups of cyanobacteria have genetically diversified into "high-light" and "low-light" variants (Moore et al., 1998) taking advantage of the stable irradiance conditions in the central gyres. In more dynamic regions it is essential for algae to be able to respond to changes in the light environment. Here we have presented a refinement of the Geider et al. (1997) mechanistic model of carbon-to-chlorophyll ratio allowing a smooth response in phytoplankton C:Chl ratios across a greater range of irradiance conditions. This allows a more accurate calculation of model results across a complete range of spatial and temporal scales.

Geider et al. (1997) give two solutions for the Chl:C ratio, both for balanced growth. One of them assumes that the chlorophyll-a losses due to respiration are zero ($R^B = 0$) or that the chlorophyll-a specific degradation has the same dependence on specific growth rate as cellular carbon specific respiration ($R^B = R^C = \mu\xi$, where μ is growth rate and ξ is the cost of biosynthesis). This is the option that has been pursued here, since it would be appropriate for use in models of gross primary production using photosynthesis-irradiance parameters that have already been corrected for respiration. If, instead, we were to use the model for the case where carbon respiration was not zero, an equivalent solution would exist, provided that a correction term were applied to θ_m as suggested by Geider et al. (1997). But, given the uncertainties in θ_m, and given that the correction term is typically found to be small, we can assume that the model discussed here is sufficient to cover such conditions as well, under our current state of knowledge. A more pertinent question is at what time scales the condition of balanced growth might be met. In fact, acclimation from one light level to another will take place over a finite period, with Geider et al. (1986) and Raven and Geider (2003) suggesting that the

appropriate time scale for acclimation is of the order of hours to days, implying that balanced growth would hold on daily time scales. Moore et al. (2003, 2006) have provided examples where photoacclimation timescales were longer than those for surface mixing, and Talmy et al. (2013) highlighted the importance of surface irradiance, depth of mixing, and light attenuation using a resource allocation based model of photoacclimation. It is also apparent that when numerical models are run at short time steps (less than an hour), it will be increasingly important to account in some manner for non-balanced growth during the transition phase.

The solution for C:Chl can produce both high C:Chl values, in line with those exceeding 300 observed in cultures (Cloern et al., 1995), and the low values (25–70) observed in ocean samples (Riemann et al., 1989). That said, a suitable θ_m is essential to obtain the correct result. In the example presented here (**Figure 3**), a three-component model of phytoplankton size classes is used in the assignment of θ_m. Although this allows a dynamic estimation of θ_m it is still derived from fixed values for each group. Refinements in the estimation of θ_m would also result in improved estimates of the realized C:Chl values.

Our application of the model at large scales using remote-sensing data (**Figure 2**) utilized average estimates of I_k (by season and province), whereas in reality the values would be more variable. Dynamic assignment of parameters would lead to a greater range of I_* values, increasing the potential for errors when using the approximate solution for θ. The concept of dynamic estimates of photosynthesis parameters using environmental variables, has been discussed by Platt and Sathyendranath (1993, 1995), Saux-Picart et al. (2013), and Silsbe et al. (2016).

The computed carbon-to-chlorophyll ratio depends strongly on available light. It raises the question of what would be the appropriate value of I to use in the calculations, given that phytoplankton experience changes in available light over a variety of time scales. These include changes at time scales of seconds, as the sun rises and sets and as clouds pass, to seasonal scale changes dictated by the Earth's declination. In addition, phytoplankton are at the mercy of vertical movement of the water column due to, for example, turbulence, internal waves or upwelling. But what would be the appropriate time scales for acclimation of carbon-to-chlorophyll ratio? As noted above, previous studies have indicated that it is of the order of 1 day. But further information on this point would be valuable. A related matter, from a modeling perspective is that the photosynthetic response of phytoplankton to available light is instantaneous. So it is clear that computation of photosynthesis within numerical ecosystem models has to be driven by instantaneous light. If, along with such calculations, we need light fields averaged over some yet-to-be-defined time scale for computation of θ, simulation models would have to be designed to keep track of at least two values of available light, to be used as required. This time scale would be related to that appropriate for balanced growth, as discussed above.

The Geider et al. (1997) model presented here is re-formulated as a function of I_*, which requires only the photosynthesis parameter I_k for implementation, in addition to data on available light. Bearing in mind the body of data on photosynthesis-irradiance parameters that exists, and the relative ease with which these parameters can be measured, compared with direct measurements of phytoplankton carbon in the field (see Casey et al., 2013; Graff et al., 2015), these results open up the possibility of significant augmentation of the information base on carbon-to-chlorophyll ratio in the marine environment. But when photosynthesis-irradiance parameters, available light and phytoplankton carbon are measured concurrently, we also have the possibility to estimate the parameter θ_m, about which we have so little information from the field.

AUTHOR CONTRIBUTIONS

SS conceived the problem. TJ, SS, and TP worked jointly to find an analytical solution. TJ made all calculations and figures. The preparation of the manuscript was led by TJ with all authors contributing significantly to the final text.

FUNDING

This work was funded through the European Space Agency's MAPPS (MArine primary Production: model Parameters from Space) project as part of the Support to Science Element (STSE) Pathfinders Program and the POCO (Pools of Carbon in the Ocean) Project, which is a part of the SEOM (Scientific Exploitation of Operational Missions) Programme. This work is also a contribution to activities of the National Centre for Earth Observation (NCEO), UK. Additional support from Simons Foundation through the CBIOMES project is acknowledged. Finally the contribution of The Jawaharlal Nehru Science Fellowship to TP in the course of this work is gratefully acknowledged.

ACKNOWLEDGMENTS

We thank ESA Ocean Colour Climate Change Initiative and NASA for the open-access remote-sensing products used in this work. We would like to thank all those who contributed to the in situ database used in this study, to Elizabeth Goult for performing sensitivity analysis checks. We thank Frédéric Mélin and Nicolas Hoepffner for making available their photosynthesis-irradiance parameter database for this work. We thank reviewers Christoph Voelker and Greg M. Silsbe for their helpful comments on the initial manuscript. We also thank Peter Regner and Diego Fernandez at ESA for their continuous support.

REFERENCES

Brewin, R. J. W., Sathyendranath, S., Hirata, T., Lavender, S. J., Barciela, R., and Hardman-Mountford, N. J. (2010). A three-component model of phytoplankton size class for the Atlantic Ocean. *Ecol. Model.* 221, 1472–1483. doi: 10.1016/j.ecolmodel.2010.02.014

Butenschön, M., Clark, J., Aldridge, J. N., Allen, J. I., Artioli, Y., Blackford, J., et al. (2016). Ersem 15.06: a generic model for marine biogeochemistry and

the ecosystem dynamics of the lower trophic levels. *Geosci. Model Develop.* 9, 1293–1339. doi: 10.5194/gmd-9-1293-2016

Casey, J. R., Aucan, J. P., Goldberg, S. R., and Lomas, M. W. (2013). Changes in partitioning of carbon amongst photosynthetic pico- and nano-plankton groups in the sargasso sea in response to changes in the north atlantic oscillation. *Deep Sea Res. II Topic. Stud. Oceanogr.* 93, 58–70. doi: 10.1016/j.dsr2.2013.02.002

Cloern, J. E., Grenz, C., and Vidergar-Lucas, L. (1995). An empirical model of the phytoplankton chlorophyll:carbon ratio-the conservation factor between productivity and growth rate. *Limnol. Oceanogr.* 40, 1313–1321. doi: 10.4319/lo.1995.40.7.1313

Cullen, J. J. (1982). The deep chlorophyll maximum: Comparing vertical profiles of chlorophyll a. *Canad. J. Fish. Aquat. Sci.* 39, 791–803. doi: 10.1139/f82-108

Cullen, J. J. (2015). Subsurface chlorophyll maximum layers: enduring enigma or mystery solved? *Annu. Rev. Marine Sci.* 7, 207–239. doi: 10.1146/annurev-marine-010213-135111

de Boyer Montégut, C., Madec, G., Fisher, A. S., Lazar, A., and Iudicone, D. (2004). Mixed layer depth over the global ocean: An examination of profile data and a profile-based climatology. *J. Geophys. Res.* 109:C12003. doi: 10.1029/2004JC002378

Dutkiewicz, S., Hickman, A. E., Jahn, O., Gregg, W. W., Mouw, C. B., and Follows, M. J. (2015). Capturing optically important constituents and properties in a marine biogeochemical and ecosystem model. *Biogeosciences* 12, 4447–4481. doi: 10.5194/bg-12-4447-2015

Frouin, R., Franz, B., and Werdell, P. (2002). The SeaWiFS PAR product. *NASA Techn. Memorand. 2003 206892* 22, 46–50.

Geider, R. J. (1987). Light and temperature dependence of the carbon to chlorophyll a ratio in microalgae and cyanobacteria: implications for physiology and the growth of phytoplankton. *New Phytol.* 106, 1–34. doi: 10.1111/j.1469-8137.1987.tb04788.x

Geider, R. J., Macintyre, H. L., and Kana, T. M. (1997). Dynamic model of phytoplankton growth and acclimation: responses of the balancedgrowth rate and chlorophyll a:carbon ratio to light, nutrient-limitation and termperature. *Marine Ecol. Progr. Ser.* 148, 187–200. doi: 10.3354/meps148187

Geider, R. J., Macintyre, H. L., Kana, T. M., and Jan, N. (1996). A Dynamic Model of Photoadaptation in Phytoplankton. *Limnol. Oceanogr.* 41, 1–15. doi: 10.4319/lo.1996.41.1.0001

Geider, R. J., Maclntyre, H. L., and Kana, T. M. (1998). A dynamic regulatory model of phytoplanktonic acclimation to light, nutrients, and temperature. *Limnol. Oceanogr.* 43, 679–694. doi: 10.4319/lo.1998.43.4.0679

Geider, R. J., Platt, T., and Raven, J. A. (1986). Size dependence of growth and photosynthesis in diatoms: a synthesis. *Marine Ecol. Progr. Ser.* 30, 93–104. doi: 10.3354/meps030093

Graff, J. R., Westberry, T. K., Milligan, A. J., Brown, M. B., Dall'Olmo, G., van Dongen-Vogels, V., et al. (2015). Analytical phytoplankton carbon measurements spanning diverse ecosystems. *Deep Sea Res. I Oceanogr. Res. Papers* 102, 16–25. doi: 10.1016/j.dsr.2015.04.006

Halsey, K., and Jones, B. (2015). Phytoplankton strategies for photosynthetic energy allocation. *Ann. Rev. Marine Sci.* 7, 265–297. doi: 10.1146/annurev-marine-010814-015813

Hickman, A., Dutkiewicz, S., Williams, R., and Follows, M. (2010). Modelling the effects of chromatic adaptation on phytoplankton community structure in the oligotrophic ocean. *Marine Ecol. Progr. Ser.* 406, 1–17. doi: 10.3354/meps08588

Laufkötter, C., Vogt, M., Gruber, N., Aita-Noguchi, M., Aumont, O., Bopp, L., et al. (2015). Drivers and uncertainties of future global marine primary production in marine ecosystem models. *Biogeosciences* 12, 6955–6984. doi: 10.5194/bg-12-6955-2015

Li, Q. P., Franks, P. J. S., Landry, M. R., Goericke, R., and Taylor, A. G. (2010). Modeling phytoplankton growth rates and chlorophyll to carbon ratios in california coastal and pelagic ecosystems. *J. Geophys. Res. Biogeosci.* 115:G04003. doi: 10.1029/2009JG001111

Longhurst, A. R., Sathyendranath, S., Platt, T., and Caverhill, C. M. (1995). An estimate of global primary production in the ocean from satellite radiometer data. *J. Plankton Res.* 17, 1245–1271. doi: 10.1093/plankt/17.6.1245

Martinez-Vicente, V., Dall'Olmo, G., Tarran, G., Boss, E., and Sathyendranath, S. (2013). Optical backscattering is correlated with phytoplankton carbon across the atlantic ocean. *Geophys. Res. Lett.* 40, 1154–1158. doi: 10.1002/grl.50252

Mélin, F., and Hoepffner, N. (2004). *Global Marine Primary Production: A Satellite View.* Ispra, VA: Institute for Environment and Sustainability.

Moore, C., Suggett, D., Holligan, P., Sharples, J., Abraham, E., Lucas, M., et al. (2003). Physical controls on phytoplankton physiology and production at a shelf sea front: a fast repetition-rate fluorometer based field study. *Marine Ecol. Progr. Ser.* 259, 29–45. doi: 10.3354/meps259029

Moore, C. M., Suggett, D. J., Hickman, A. E., Kim, Y.-N., Tweddle, J. F., Sharples, J., et al. (2006). Phytoplankton photoacclimation and photoadaptation in response to environmental gradients in a shelf sea. *Limnol. Oceanogr.* 51, 936–949. doi: 10.4319/lo.2006.51.2.0936

Moore, L. R., Rocap, G., and Chisholm, S. W. (1998). Physiology and molecular phylogeny of coexisting prochlorococcus ecotypes. *Nature* 393, 464–467. doi: 10.1038/30965

Morel, A., and Berthon, J.-F. (1989). Surface pigments, algal biomass profiles, and potential production of the euphotic layer: relationships reinvestigated in view of remote-sensing applications. *Limnol. Oceanogr.* 34, 1545–1562. doi: 10.4319/lo.1989.34.8.1545

Morel, A., and Smith, R. C. (1974). Relation between total quanta and total energy for aquatic photosynthesis1. *Limnol. Oceanogr.* 19, 591–600. doi: 10.4319/lo.1974.19.4.0591

Platt, T., Caverhill, C., and Sathyendranath, S. (1991). Basin-scale estimates of oceanic primary production by remote sensing: The north atlantic. *J. Geophys. Res. Oceans* 96, 15147–15159. doi: 10.1029/91JC01118

Platt, T., Gallegos, C. L., and Harrison, W. G. (1980). Photoinhibition of photosynthesis in natural assemblages of marine phytoplankton. *J. Marine Res.* 38, 687–701.

Platt, T., and Jassby, A. D. (1976). The relationship between photosynthesis and light for natural assemblages of coastal marine phytoplankton. *J. Phycol.* 12, 421–430. doi: 10.1111/j.1529-8817.1976.tb02866.x

Platt, T., and Sathyendranath, S. (1993). Estimators of primary production for interpretation of remotely sensed data on ocean color. *J. Geophys. Res. Oceans* 98, 14561–14576. doi: 10.1029/93JC01001

Platt, T., and Sathyendranath, S. (1995). "Latitude as a factor in the calculation of primary production," in *Ecology of Fjords and Coastal Waters*, eds H. Skjoldal, C. Hopkins, K. Erikstad, and H. Leinaas (Amsterdam: Elsevier Science), 3–13.

Raven, J. A., and Geider, R. J. (2003). *Adaptation, Acclimation and Regulation in Algal Photosynthesis* (Dordrecht: Springer).

Riemann, B., Simonsen, P., and Stensgaard, L. (1989). The carbon and chlorophyll content of phytoplankton from various nutrient regimes. *J. Plankton Res.* 11, 1037–1045. doi: 10.1093/plankt/11.5.1037

Sathyendranath, S., and Platt, T. (1988). The spectral irradiance field at the surface and in the interior of the ocean: a model for applications in oceanography and remote sensing. *J. Geophys. Res.* 93, 9270–9280. doi: 10.1029/JC093iC08p09270

Sathyendranath, S., Stuart, V., Nair, A., Oka, K., Nakane, T., Bouman, H., et al. (2009). Carbon-to-chlorophyll ratio and growth rate of phytoplankton in the sea. *Marine Ecol. Progr. Ser.* 383, 73–84. doi: 10.3354/meps07998

Saux-Picart, S., Sathyendranath, S., Dowell, M., Moore, T., and Platt, T. (2013). Remote sensing of assimilation number for marine phytoplankton. *Remote Sens. Environ.* 146, 87–96. doi: 10.1016/j.rse.2013.10.032

Silsbe, G. M., Behrenfeld, M. J., Halsey, K. H., Milligan, A. J., and Westberry, T. K. (2016). The cafe model: a net production model for global ocean phytoplankton. *Global Biogeochem. Cycles* 30, 1756–1777. doi: 10.1002/2016GB005521

Talmy, D., Blackford, J., Hardman-Mountford, N. J., Dumbrell, A. J., and Geider, R. J. (2013). An optimality model of photoadaptation in contrasting aquatic light regimes. *Limnol. Oceanogr.* 58, 1802–1818. doi: 10.4319/lo.2013.58.5.1802

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Jackson, Sathyendranath and Platt. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Marine Science

ORIGINAL RESEARCH
published: 08 November 2017
doi: 10.3389/fmars.2017.00349

Characterizing Spatial Variability of Ice Algal Chlorophyll *a* and Net Primary Production between Sea Ice Habitats Using Horizontal Profiling Platforms

Benjamin A. Lange[1,2*†], *Christian Katlein*[1], *Giulia Castellani*[1], *Mar Fernández-Méndez*[1,3], *Marcel Nicolaus*[1], *Ilka Peeken*[1,4] *and Hauke Flores*[1,2]

[1] *Alfred-Wegener-Institute Helmholtz Center for Polar and Marine Research, Bremerhaven, Germany,* [2] *Centre for Natural History (CeNak), Zoological Museum, University of Hamburg, Hamburg, Germany,* [3] *Norwegian Polar Institute, Fram Centre, Tromsø, Norway,* [4] *MARUM, Center for Marine Environmental Sciences, University of Bremen, Bremen, Germany*

Assessing the role of sea ice algal biomass and primary production for polar ecosystems remains challenging due to the strong spatio-temporal variability of sea ice algae. Therefore, the spatial representativeness of sea ice algal biomass and primary production sampling remains a key issue in large-scale models and climate change predictions of polar ecosystems. To address this issue, we presented two novel approaches to up-scale ice algal chl *a* biomass and net primary production (NPP) estimates based on profiles covering distances of 100 to 1,000 s of meters. This was accomplished by combining ice core-based methods with horizontal under-ice spectral radiation profiling conducted in the central Arctic Ocean during summer 2012. We conducted a multi-scale comparison of ice-core based ice algal chl *a* biomass with two profiling platforms: a remotely operated vehicle and surface and under ice trawl (SUIT). NPP estimates were compared between ice cores and remotely operated vehicle surveys. Our results showed that ice core-based estimates of ice algal chl *a* biomass and NPP do not representatively capture the spatial variability compared to the remotely operated vehicle-based estimates, implying considerable uncertainties for pan-Arctic estimates based on ice core observations alone. Grouping sea ice cores based on region or ice type improved the representativeness. With only a small sample size, however, a high risk of obtaining non-representative estimates remains. Sea ice algal chl *a* biomass estimates based on the dominant ice class alone showed a better agreement between ice core and remotely operated vehicle estimates. Grouping ice core measurements yielded no improvement in NPP estimates, highlighting the importance of accounting for the spatial variability of both the chl *a* biomass and bottom-ice light in order to get representative estimates. Profile-based measurements of ice algae chl *a* biomass identified sea ice ridges as an underappreciated component of the Arctic ecosystem

OPEN ACCESS

Edited by:
Victor Martinez-Vicente, Plymouth Marine Laboratory, United Kingdom

Reviewed by:
Jaume Piera, Institut de Ciències del Mar (CSIC), Spain
Thomas Jackson, Plymouth Marine Laboratory, United Kingdom
Karley Lynn Campbell, University of Manitoba, Canada

***Correspondence:**
Benjamin A. Lange
benjamin.lange@dfo-mpo.gc.ca

†Present Address:
Benjamin A. Lange, Fisheries and Oceans Canada, Freshwater Institute, Winnipeg, MB, Canada

Specialty section:
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

Received: *23 February 2017*
Accepted: *18 October 2017*
Published: *08 November 2017*

Citation:
Lange BA, Katlein C, Castellani G, Fernández-Méndez M, Nicolaus M, Peeken I and Flores H (2017) Characterizing Spatial Variability of Ice Algal Chlorophyll a and Net Primary Production between Sea Ice Habitats Using Horizontal Profiling Platforms. Front. Mar. Sci. 4:349. doi: 10.3389/fmars.2017.00349

because chl *a* biomass was significantly greater in this unique habitat. Sea ice ridges are not easily captured with ice coring methods and thus require more attention in future studies. Based on our results, we provide recommendations for designing an efficient and effective sea ice algal sampling program for the summer season.

Keywords: ice algae, ice core, chl *a*, remotely operated vehicle, surface and under-ice trawl, net primary production, spectral irradiance, bio-optics

INTRODUCTION

There is mounting evidence for an overall increase in Arctic-wide net primary production (NPP) as a result of the declining sea ice cover and increasing duration of the phytoplankton growth season (Arrigo and van Dijken, 2011, 2015; Fernández-Méndez et al., 2015). However, it remains uncertain how sea ice algae NPP will respond to continued changes of the sea ice environment. It has been suggested that a thinning Arctic sea ice cover, which will lead to increased light transmittance, will also result in increased sea ice algal NPP rates due to more available photosynthetically active radiation (PAR; Nicolaus et al., 2012; Fernández-Méndez et al., 2015). On the other hand, some forecasts predict increased snow precipitation in the Arctic (IPCC, 2013), which would result in less available light for bottom-ice algal growth during spring. Other than available light, other variables may have an equal or greater influence on Arctic primary production depending on region and season. Such variables include nutrient supply, temperature, and CO_2 intake (Tremblay et al., 2015). Declining sea ice may increase oceanic CO_2 intake, which would result in increased NPP, but could be counteracted by increased runoff and higher temperatures expected throughout the Arctic (Tremblay et al., 2015).

In the central Arctic Ocean sea-ice algae has been documented to contribute up to 60% of the NPP during summer (Gosselin et al., 1997; Fernández-Méndez et al., 2015). However, net sympagic (ice-associated) primary production is relatively low accounting for 1–10% of total NPP in the Arctic Ocean (Dupont, 2012; Arrigo and van Dijken, 2015). Regardless of the overall low contribution of sympagic NPP, both sympagic and pelagic organisms showed a high dependency on ice-algae produced carbon within the central Arctic Ocean (Budge et al., 2008; Wang et al., 2015; Kohlbach et al., 2016, 2017). The key role of sea ice algae in Arctic foodwebs, particularly in terms of reproduction and growth of key Arctic organisms, such as: *Calanus glacialis* (Michel et al., 1996; Søreide et al., 2010), highlights the importance of timing and duration of ice algal growth, and the availability of algal biomass throughout different times of the year.

Spatial variability of springtime ice algal chl *a* biomass has been related to the distribution of snow on first-year sea ice (FYI), due to the large influence of snow on light transmission by the reflection and scattering of light near the surface. This relationship explains the similar patch sizes observed for snow and sea ice algae biomass on the same study sites. Between study sites, however, patch sizes had a large range between 5 and 90 m, which was the result of differences in the snow distribution and drifting patterns over relatively level FYI (Gosselin et al., 1986; Rysgaard et al., 2001; Granskog et al., 2005; Søgaard et al., 2010). In contrast, the undulating surface topography of MYI plays an important role in the distribution of snow, which has been linked to the presence of high ice algal chl *a* biomass at the bottom of thick MYI hummocks with little or no snow cover (Lange et al., 2015, 2017). Gradinger et al. (2010) identified sea ice ridges as important accumulation regions of sea ice fauna during advanced melt. This further highlights the ecological importance of thick sea ice features. Using traditional coring methods, however, it is very difficult to sample ridges and hummocks resulting in sparse observations for ice algae at the bottom or within these features.

In summer when the snow is melted and melt ponds are present, light availability has a less important role in controlling the distribution of ice algal chl *a* biomass. This is due to increased melt induced algal losses during late-spring and early-summer, which becomes the limiting factor controlling the ability of algal communities to remain in the bottom-ice environment (Grossi et al., 1987; Lavoie et al., 2005). The spatial distribution of ice algal chl *a* biomass during mid- to late-summer, however, remains poorly understood and under-sampled, particularly in the central Arctic Ocean (Wassmann et al., 2011; Miller et al., 2015).

The high spatial and temporal variability of sea ice algae, in addition to sparse sampling, results in poorly constrained sea ice algal chl *a* biomass and PP estimates for the central Arctic Ocean (Miller et al., 2015). Large-scale estimates of sea ice algal chl *a* biomass and PP are limited to modeling studies as satellites are unable to observe the underside of sea ice. Lee et al. (2015) demonstrated that pelagic phytoplankton PP models for the Arctic Ocean were highly sensitive to uncertainties in chlorophyll *a* (chl a) and performed best with *in situ* chl *a* data. *In situ* ice algal chl *a* estimates used in models, however, are typically based on a small number of ice core observations (e.g., Fernández-Méndez et al., 2015). A recent study comparing ice core chl *a* biomass to sea ice algal chl *a* biomass derived from an 85 m ROV transect of under-ice spectral radiation measurements showed large differences, which could carry high uncertainties for large-scale estimates based on these ice core data alone (Lange et al., 2016).

Miller et al. (2015) reviewed the different methods for PP measurements with spatial sampling resolution on the order of 0.01 m for ice coring-based *in vitro* incubations (e.g., Gosselin et al., 1997; Gradinger, 2009; Fernández-Méndez et al., 2015) or *in situ* incubations (e.g., Mock and Gradinger, 1999; Gradinger, 2009). At larger scales the under-ice eddy covariance method integrates primary production over an area of 100 m^2 (Long et al., 2012). Thus there is a large gap in spatial coverage between the 0.01 to 100 m^2 scales, which is not resolved by these methods. It is within this spatial range that many sea ice and snow properties

(such as thickness, porosity, temperature) can vary, which can have a large influence on light availability, ice melt and growth, nutrient availability, and therefore, the spatial distribution of ice algae. Typical patch sizes of snow have been reported in the range 20–25 m (Gosselin et al., 1986; Steffens et al., 2006). Surface properties such as albedo have patch sizes of ∼10 m (Perovich et al., 1998; Katlein et al., 2015a) and sea ice draft can vary at scales of around 15 m (Katlein et al., 2015a).

Here we present a novel approach to fill this important gap in the spatial scales of ice algal chl *a* biomass and NPP estimates by combining *in vitro* photosynthetic parameters of ice algae with chl *a* biomass derived from under-ice spectral radiation measurements and under-ice available PAR measurements obtained from a moving under-ice profiling platform, the ROV. Furthermore, we investigate the spatial patterns of chl *a* biomass and NPP estimates, using two under-ice profiling platforms: the ROV and Surface and Under Ice Trawl (SUIT), with special emphasis on sea ice ridges, and evaluate potential discrepancies between the up-scaled and ice core-based estimates. Based on our results, we provide recommendations for designing an efficient and effective sea ice algal sampling program for the summer season.

MATERIALS AND METHODS

The Profiling Platforms

All surveys were conducted during the RV Polarstern expedition PS80 to the central Arctic Ocean in August and September 2012. Under-ice profiling platform surveys were conducted using an under-ice Remotely Operated Vehicle (ROV) V8Sii-ROV (Ocean Modules, Åtvidaberg, Sweden) and a SUIT (van Franeker et al., 2009), with mounted sensor arrays, described in Nicolaus and Katlein (2013), David et al. (2015), and Lange et al. (2016). Simplified diagrams and images showing the deployment of the under-ice profiling platforms were presented in Lange et al. (2016). The ROV is an under-water vehicle with mounted sensor array deployed through a small 2 × 2 m man made hole in the sea ice, and is attached by a 300 m long fiber optic cable. The ROV is controlled remotely from a sheltered base station (e.g., tent) located adjacent to the deployment hole. A detailed description of the ROV spectral measurements, calibration and calculations, and ROV operation was provided by Katlein et al. (2015b) and Nicolaus and Katlein (2013). The V8ii ROV was equipped with an altimeter (DST Micron Echosounder, Tritech, UK), a sonar (Micron DST MK2, Tritech, UK), one zoom-camera (Typhoon, Tritech, UK), and one fixed focal length camera (Ospray, Tritech, UK). The SUIT is a net developed for deployment in ice covered waters, typically behind an icebreaker, for sampling sea ice associated zooplankton and micronekton in the upper 2 m of the water within the ice-water interface. During this cruise the sensor array was specifically enhanced to measure the variability of sea ice algae chl *a* biomass within the sea ice and sea ice habitat properties along the SUIT hauls. The new sensor package included an Aquadopp Acoustic Doppler Current Profiler (ADCP; Nortek AS, Rud, Norway), a Conductivity Temperature Depth probe (CTD; Sea and Sun Technology, Trappenkamp, Germany) with a built-in Cyclops 7 fluorometer (Turner Designs, Sunnyvale, CA, USA), an PA500/6S altimeter (Tritech International Ltd., Aberdeen, UK), one RAMSES-ACC irradiance sensor (Trios, GmbH, Rastede, Germany), one RAMSES-ARC radiance sensor (Trios GmbH, Rastede, Germany) and a forward-looking video camera (GoPro Hero 2).

The ROV spectral surveys were conducted during seven ice stations (**Table 1**; **Figure 1**). The SUIT spectral surveys were conducted at 6 stations (**Table 1**; **Figure 1**). Stations conducted in relatively close proximity (<50 km) to each other were grouped into similar locations represented by the letters A to I (**Figure 1**). Two profiles separated by small distances were sampled using the SUIT (<10 km) at location B, and using the ROV (<500 m) at locations C and D. Incoming solar radiation observations were measured on-ice for ROV-based spectral measurements, and from a ship-mounted sensor for SUIT-based spectral measurements. To ensure high quality spectra, data were limited to observations at a distance to the ice-bottom of ≤1 m and with a pitch and roll between −10° and 10°, as suggested by Nicolaus and Katlein (2013) and Katlein et al. (2016). Reducing the pitch and roll, and distance to ice bottom also reduced the potential influence of spectral absorption by the water. Since the SUIT behaves less predictable near ridges (e.g., it hits the ridge and is redirected in an unpredictable direction), we manually inspected the spectra to identify reliable spectral measurements at sea ice ridges (e.g., noisy spectra). Less than 1% of the spectra were excluded from analyses.

Sea ice draft was calculated based on sensor measurements of depth and distance to ice bottom, and corrected for pitch and roll angles as described in Lange et al. (2016) and David et al. (2015). Sea ice ridges were identified from the SUIT ice draft profiles using the Rayleigh criteria, following procedures described by Rabenstein et al. (2010) and Castellani et al. (2014) for the sea ice surface topography, and Castellani et al. (2015) for the sea ice bottom profile. Ice draft local minima (thicker sea ice draft values are more negative) identified along the SUIT profiles with a threshold of 1.5 m deeper than the surrounding ice, following Castellani et al. (2015), were selected as potential ridges. Adjacent minima needed a separation distance between points which was less than half the depth of the first minima in order to be identified as two single elements not belonging to the same ridge. Ridge depth and width were measured in order to calculate ridge density (ridges km^{-1}) and percent coverage of ridges. Here, ridge depth was calculated as the width at half maximum. During one SUIT haul (station 358, location H) there were no altimeter measurements. Because the SUIT generally travels directly under the ice, the depth measurements can be used to reliably (R^2 = 0.78) derive level ice draft using a simple linear model (David et al., 2015). We could calculate ridge density, ridge coverage, and ridge width from these ice draft measurements without altimeter data. The absolute draft values at ridges, however, were less accurate and therefore excluded from analysis.

All profiling platform-derived observations (i.e., transmittance, sea ice algal chl *a*, NPP, draft) were divided into 5 ice classes based on the sea ice draft values in the following ranges: (1) 0–0.5 m; (2) 0.5–1.0 m; (3) 1.0–1.5 m; (4) 1.5–2.0 m; and (5) >2.0 m. Furthermore, we separated profiling

TABLE 1 | Summary of downwelling surface and bottom-ice light, chlorophyll *a* biomass (chl *a*), net primary production (NPP), and explained variance of NPP per location (shown in **Figure 1**) and sampling method (gear): ice cores (FM or LA), remotely operated vehicle (ROV) and surface and under-ice trawl (SUIT).

Group	Gear[a]	Sample size	Station	Downwelling surface PAR[b]	Scalar PAR (I)[b]	Chl *a*[c] mg m^{-2}	NPP[c] mg C m^{-2} d^{-1}	Explained variance (R^2) of NPP by	
				μmols photons m^{-2} s^{-1}				I	Chl *a*
A	SUIT	46	216	–	–	0.0 (0.0–0.2)	–	–	–
B	FFM	1	224	249 ± 90	40.8 ± 14.7	1.2	10.16	–	–
	LLA	8	224	–	–	0.3 (0.2–0.5)	–	–	–
	CORES	9	224	–	–	0.4 (0.2–0.7)*	–	–	–
	ROV	468	224	211 ± 72	51.2 ± 25.0	1.0 (1.0–1.2)	8.45 (5.59–12.29)	0.64 (0.54–0.70)	0.10 (0.07–0.19)
	SUIT	43	223	–	–	0.2 (0.0–0.7)	–	–	–
	SUIT-2	45	233	–	–	0.1 (0.0–0.4)	–	–	–
C	FM	1	237	174 ± 90	28.5 ± 14.7	1.7 (+)•	0.56 (–)•	–	–
	LA	12	237	–	–	0.6 (0.5–1.1)	–	–	–
	CORES	13	237	–	–	0.7 (0.5–1.2)*	–	–	–
	ROV	156	237a	137 ± 59	28.9 ± 23.2	1.0 (0.8–1.1)•	0.60 (0.30–0.98)	0.61 (0.37–0.82)	0.11 (0.02–0.24)
	ROV-2	1378	237b	137 ± 59	18.7 ± 8.2	1.3 (1.1–1.5)•	0.89 (0.62–1.03)•	0.61 (0.38–0.79)	0.09 (0.03–0.17)
D	FM	1	255	104 ± 71	26.7 ± 18.2	0.6 (–)•	0.62 (–)•	–	–
	LA	4	255	–	–	0.8 (0.7–1.2)*	–	–	–
	CORES	5	255	–	–	0.7 (0.6–1.2)	–	–	–
	ROV	186	255	93 ± 60	36.3 ± 20.3	1.4 (1.4–1.5)•	1.73 (1.48–1.91)•	0.12 (0.0–0.25)	0.70 (0.52–0.93)
E	FM	1	277	101 ± 57	25.9 ± 14.6	0.4 (–)•	0.45	–	–
	SUIT	91	285	–	–	0.1 (0.0–0.9)	–	–	–
F	FM	1	323	81 ± 63	24.2 ± 18.8	0.3 (–)•	0.02 (–)•	–	–
	LA	6	323	–	–	0.2 (0.1–0.2)*	–	–	–
	CORES	7	323	–	–	0.2 (0.0–0.3)	–	–	–
	ROV	1145	323	67 ± 49	7.7 ± 8.8	1.5 (1.3–1.7)•	0.14 (0.10–0.19)•	0.84 (0.72–0.90)	0.18 (0.14–0.23)
	SUIT	63	321	–	–	0.9 (0.0–1.7)	–	–	–
G	FM	1	335	49 ± 43	5.9 ± 5.2	0.4 (–)•	0.05 (–)•	–	–
	LA	6	335	–	–	0.9 (0.4–1.1)*	–	–	–
	CORES	7	335	–	–	0.8 (0.3–1.1)	–	–	–
	ROV	762	335m	46 ± 39	3.0 ± 7.6	2.3 (1.9–2.8)•	0.13 (0.07–0.22)•	0.93 (0.89–0.94)	0.01 (0.01–0.02)
	ROV-2	302	335f	46 ± 39	2.3 ± 2.7	2.7 (2.3–3.1)•	0.13 (0.08–0.23)•	070 (0.68–0.70)	0.09 (0.08–0.10)
	SUIT	18	345	–	–	1.9 (0.0–4.4)	–	–	–
H	FM	1	349	25 ± 15	1.4 ± 0.9	8.0 (+)•	1.00 (+)•	–	–
	LA	7	349	–	–	0.6 (0.3–2.3)	–	–	–
	CORES	8	349	–	–	0.8 (0.3–4.7)	–	–	–
	ROV	282	349	23 ± 13	2.4 ± 1.5	1.3 (1.2–1.5)•	0.14 (0.08–0.21)•	0.16 (0.11–0.21)	0.61 (0.56–0.68)
	SUIT	101	358	–	–	0.9 (0.4–1.7)	–	–	– –
I	FM	1	360	13 ± 7	0.9 ± 0.5	8.0 (+)•	0.39 (+)•	–	–
	LA	4	360	–	–	7.3 (4.9–9.0)	–	–	–
	CORES	5	360	–	–	8.0 (4.3–9.3)	–	–	–
	ROV	647	360	10 ± 5	0.4 ± 0.4	4.3 (2.8–6.6)•	0.07 (0.05–0.12)•	0.79 (0.78–0.80)	0.15 (0.15–0.16)

[a] *"FM" corresponds to FM-cores from Fernández-Méndez et al. (2015); "LA" correspond to LA-cores from Lange et al. (2016); "ROV" correspond to the up-scaled remotely operated vehicle estimates; and "SUIT" correspond to the up-scaled surface and under-ice trawl estimates.*

[b] *Downwelling surface PAR and bottom ice scalar PAR (I) are presented as mean ± sd to maintain consistency with Fernández-Méndez et al. (2015).*

[c] *Chl a and NPP are presented as median (interquartile range).*

• *Correspond to FM-cores not representative of the corresponding up-scaled ROV estimates for that location, i.e., FM-core estimate outside the interquartile range of ROV estimates. (+) indicates over-estimate; (–) under-estimate of the FM cores compared to up-scaled ROV estimates.*

* *Represents significant difference between the CORES (FM and LA cores combined) and the up-scaled ROV estimates.*

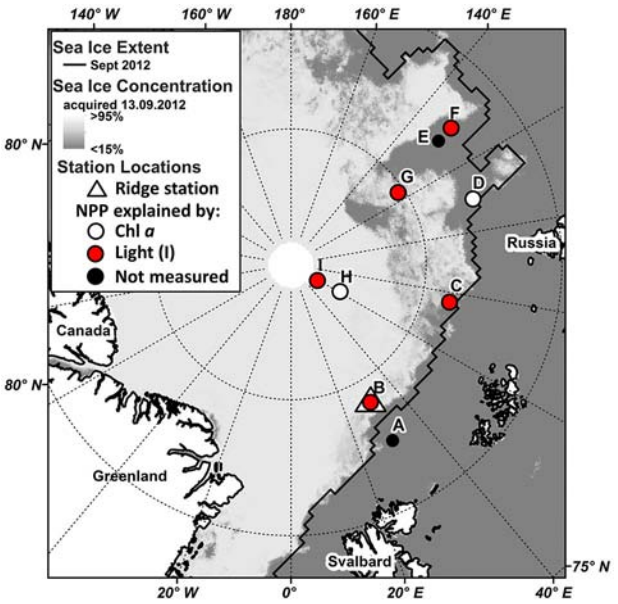


FIGURE 1 | Map of the Arctic Ocean with sea ice extent and concentration data, and the locations of the corresponding station groupings conducted during the 2012 PS80 cruise (station numbers for each grouping are listed in **Table 1**). Locations are color coded to identify which variable, light (red) or chl *a* biomass (white), explains the dominant portion of the net primary production variance. The ice station with identified ridges, station 224 (location B), is identified by a triangle. Sea ice concentration data acquired from www.meereisportal.de according to algorithms in Spreen et al. (2008). Sea ice extent correspond to the 2012 September monthly mean (extent data acquired from NSIDC, (Fetterer et al., 2002, updated 2011).

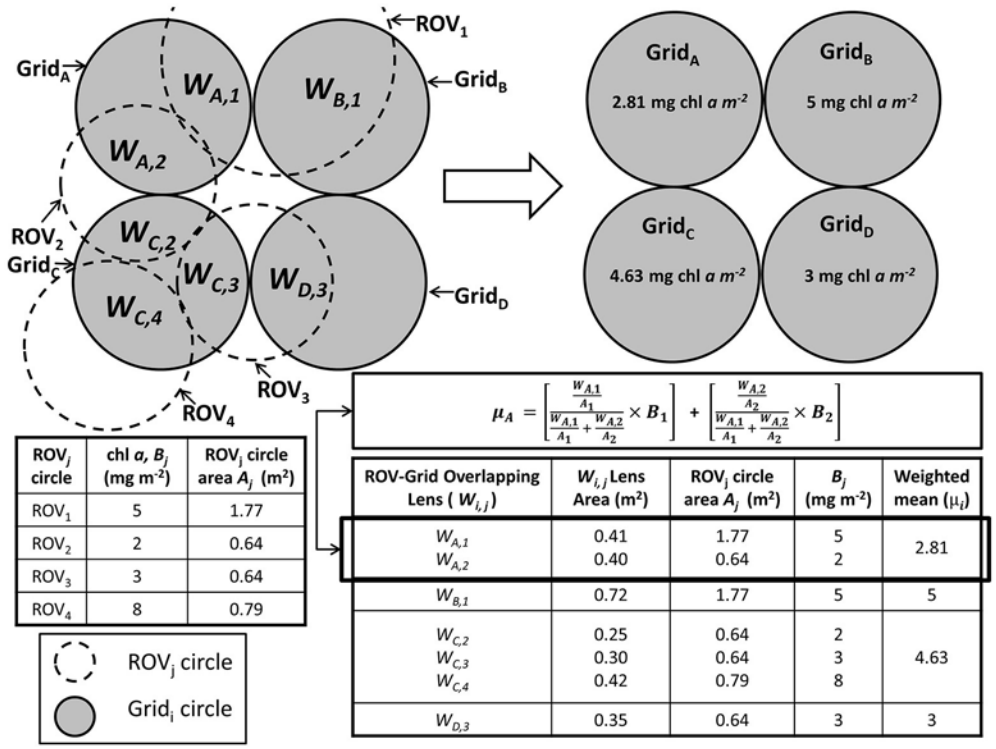
platform-derived observations into level ice and ridged ice. This was done by manually identifying all observations acquired under the identified ridges. We identified dominant ice classes for each location using the modal ice thickness (converted to draft by multiplying by 0.9) from electromagnetic induction sounding ice thickness surveys, using an EM31 instrument, of the entire floe (data presented in Boetius et al., 2013; Fernández-Méndez et al., 2015; Katlein et al., 2015b). We used these larger scale ice thickness surveys to assign the dominant ice class because these surveys were conducted specifically for the purpose of assessing the distribution of ice thickness at the ice floe. Ground-based EM surveys are a common method to representatively capture the spatial variability of ice thickness on floe scales (Haas et al., 1997; Haas, 2004).

Sea Ice Algal Chl *a* Biomass Estimates Derived from Under-Ice Spectral Radiation

Ice algal chl *a* biomass estimates were derived from under-ice profiling platform-based spectral transmittance observations using empirical orthogonal function (EOF) analysis combined with generalized linear models (GLM), as described in Lange et al. (2016). EOF analyses reduce the dimensionality of the data while maintaining the variability of key spectral absorption properties, which can then be used to relate to chl *a* concentrations or other environmental variables. GLMs were fitted using ice core chl *a* concentrations as a response variable and EOF modes as predictor variables. All ice cores were extracted along ROV spectral radiation profiles. The best set of EOF modes used as predictor variables was selected by searching all possible combinations of EOF modes and using the Bayesian Information Criterion (BIC) to assess the quality of the GLM. The EOFs used represented the spectral variability that can best be explained by the variability within the ice algal chl *a* biomass. Furthermore, EOF analyses captured variability within multiple regions of the PAR light spectrum (400–700 nm) where chl *a* light absorption occurs. In addition, we used mean robustness R^2 and true prediction error estimates as ranking criteria to find the best predictive model for our data set. Each model was applied to 5 data subsets not used to fit the model then we determined the predicted vs. observed R^2-value for each data subset then took the mean R^2-value as the mean robustness R^2. To determine the true prediction error estimate we used 10-Fold Cross-Validation (10 FCV). In 10 FCV, data are randomly separated into 10 data subsets then model fitting and error estimation are repeated 10 times. Each time the model is fitted to 9-folds then applied to the 10th-fold. This is repeated 100 times and the mean of all root mean square error (RMSE) values is used as the true prediction error estimate. Based on these criteria we determined that the combination of spectral transmittance, calculated according to Nicolaus et al. (2010), and the EOF approach resulted in the most reliable predictive model (EOF-Transmittance) with a predicted vs. observed chl *a* R^2 of 0.90, and a true prediction error estimate (10-fold cross validated root mean squared error, $RMSE_{CV}$), of 1.8 mg chl *a* m^{-2} (model M15 from Lange et al., 2016). In addition, the selected predictive model showed good agreement between chl *a* estimates derived from independent spectral data (spectra not used to fit the model) and ice core chl *a* concentrations, which were all extracted along the ROV profiles.

ROV Data Re-Sampling

We resampled the ROV chl *a*, ice draft and transmittance observations in order to account for potential spatial sampling biases (e.g., multiple or overlapping measurements at the same location; **Figure 2**), and variable footprint size of the under-ice ROV spectral measurements. Data were resampled to a grid (x, y) of equally spaced 1 m diameter circles (grid circles; **Figure 2**). A grid of circles was created for the ROV measurements (ROV circles) with each circle's center location determined by the measurement location (x, y) and the diameter determined by the footprint of the measurement (i.e. distance to ice bottom multiplied by 2, as described in Lange et al. (2016). For each grid circle with only one overlapping ROV circle, which had an overlapping area $\geq$0.2 m^2 (25% of the 1 m circle), the corresponding ROV-based transmittance and chl *a* were assigned to that grid circle. For each grid circle that had more than one overlapping ROV circle, of which at least one ROV circle had an overlapping area $\geq$0.2 m^2, weighted means of the corresponding ROV-based transmittance, draft and chl *a* were assigned to the grid circle. Weighting factors were calculated for all overlapping ROV circles in each grid as the overlapping area of each ROV circle with the corresponding



$$\mu_A = \left[\frac{\frac{W_{A,1}}{A_1}}{\frac{W_{A,1}}{A_1}+\frac{W_{A,2}}{A_2}} \times B_1\right] + \left[\frac{\frac{W_{A,2}}{A_2}}{\frac{W_{A,1}}{A_1}+\frac{W_{A,2}}{A_2}} \times B_2\right]$$

ROV$_j$ circle	chl *a*, B_j (mg m^{-2})	ROV$_j$ circle area A_j (m^2)
ROV$_1$	5	1.77
ROV$_2$	2	0.64
ROV$_3$	3	0.64
ROV$_4$	8	0.79

ROV-Grid Overlapping Lens ($W_{i,j}$)	$W_{i,j}$ Lens Area (m^2)	ROV$_j$ circle area A_j (m^2)	B_j (mg m^{-2})	Weighted mean (μ_i)
$W_{A,1}$	0.41	1.77	5	2.81
$W_{A,2}$	0.40	0.64	2	
$W_{B,1}$	0.72	1.77	5	5
$W_{C,2}$	0.25	0.64	2	4.63
$W_{C,3}$	0.30	0.64	3	
$W_{C,4}$	0.42	0.79	8	
$W_{D,3}$	0.35	0.64	3	3

FIGURE 2 | Detailed diagram and example calculation of the re-sampling process. A grid of circles was created for the ROV measurements (e.g., **ROV$_{1\text{-}4}$** circles) with each circle's center location determined by the measurement location (x, y) and the diameter determined by the footprint of the measurement. An additional grid of circles (e.g., **Grid$_{A\text{-}D}$** circles) was created where each adjacent circle was spaced 1 m apart and each had a diameter of 1 m. For each grid circle (e.g., **Grid$_{B\ and\ D}$**) with only one overlapping ROV circle (e.g., **ROV$_{1\ and\ 3}$**, respectively), which had an overlapping area ≥0.2 m^2 (e.g., $W_{B,1}$ and $W_{D,3}$, respectively), the corresponding ROV-based transmittance and chl *a* were assigned to that grid circle. For each grid circle (e.g., **Grid$_{A\ and\ C}$**) that had more than one overlapping ROV circle (e.g., ROV$_{2-4}$ and ROV$_{1-2}$, respectively), of which at least one ROV circle had an overlapping area ≥0.2 m^2 (e.g., $W_{A,1\text{-}2}$ and $W_{C,2\text{-}4}$, respectively), weighted means (e.g., μ_A for **Grid$_A$**chl *a*) of the corresponding ROV-based transmittance, draft and chl *a* were assigned to the grid circle. Weighting factors were calculated as the overlapping area of each ROV circle with the corresponding grid circle divided by the sum of all overlapping areas for that grid circle.

grid circle relative to the total ROV circle area. **Figure 2** shows a detailed diagram outlining the resampling process with an example calculation. SUIT data were not re-sampled because they represent a straight linear profile and therefore the measurements have no possibility to have overlapping footprints for the same regions.

ROV-Derived Net Primary Production Estimates

All NPP estimates were calculated based on the re-sampled ROV observations of chl *a* and transmittance. Up-scaled daily ice algal NPP estimates, P (mg C m^{-2} d^{-1}), were calculated using the photosynthesis equation from (Platt et al., 1980):

$$P = \int_t \left[\left(P_s^B\left[1 - e^{-\alpha^B I_t/P_s^B}\right]e^{-\beta^B I_t/P_s^B}\right) B\right]$$

where P_s^B is the chl *a*-normalized maximum fixation rate with no photoinhibition (mg C [mg chl *a*]$^{-1}$ h^{-1}); α^B is the initial slope of the saturation curve (mg C [mg chl *a*]$^{-1}$ h^{-1} [μmol photons m^2 s^{-1}]$^{-1}$); and β^B is strength of photoinhibition (same units as α). P_s^B, α^B, and β^B correspond to the photosynthetic parameters determined by Fernández-Méndez et al. (2015) using the ^{14}C method and incubating for 12 h, based on ice core samples collected from the same seven ice stations. Derivation of the photosynthetic parameters was conducted for upper-half and lower-half portions (mean: 0.58 m; range: 0.40–0.98 m) of the sea ice melted at 4°C in the dark for 24 h. NPP estimates were only calculated and compared for the bottom ice portions because previous *in situ* incubations studies demonstrated bottom-ice had the highest primary production rates, despite lower irradiance levels (Mock and Gradinger, 1999). Furthermore, because sea ice algal chl *a* biomass typically accumulates in the bottom-ice portion it is safe to assume a large majority of the primary production also occurred in the bottom-ice. Accordingly, we used only chl *a* biomass estimates for the lower portion, where 75% of the total chl *a* biomass was observed Fernández-Méndez et al. (2015). ROV-based chl *a* correspond to the total chl *a* biomass within the entire ice column, therefore we multiplied by 0.75 to get the appropriate fraction of the total chl *a* in the bottom ice portion. B represents the bottom-ice algae chl *a* concentrations derived from ROV-based spectral transmittance measurements. I_t is the hourly-averaged transmitted PAR (μmol photons m^2 s^{-1}) at the ice-water interface, converted to

bottom-ice scalar irradiance according to Katlein et al. (2014), and calculated for each hour (t) over a 24 h period ($t = 1, 2, \ldots 24$) by multiplying the ROV spectral (PAR) transmittance by hourly-averaged (t) incoming PAR (μmol photons m^2 s^{-1}) measured during each ice station.

Statistical Analyses

All statistical analyses were conducted using R software Version 2.15.2 with all relevant packages (R-Development-Core-Team, 2012) listed after the corresponding analysis description.

Ice core chl *a* data used for comparison were presented in Fernández-Méndez et al. (2015), hereafter referred to as "FM" cores (1 core per station), and were melted in filtered sea water. Since the FM-cores were used to characterize the NPP for each ice station (Fernández-Méndez et al., 2015) we assessed the representativeness of the single cores compared to the up-scaled ROV surveys of chl *a* biomass and NPP. NPP was not measured on the LM-cores, thus we only compared NPP estimates for FM-cores with the ROV estimates, which had both chl *a* biomass and under-ice light measurements. FM-cores were considered representative of the area if they were within the interquartile range (IQR; 25–75 percentiles) of the up-scaled ROV and SUIT estimates.

Cores from Lange et al. (2016), hereafter referred to as "LA" cores (4–12 cores per station) were directly melted. For comparisons of chl *a* biomass between ice core and ROV-derived estimates and between level ice and ridged ice, FM and LA cores were grouped together, referred to as CORES. The significance of differences between these groupings was assessed using the non-parametric Wilcoxon rank sum test (Wilcoxon, 1945). We used a non-parametric test because the assumption of normality required for parametric tests (e.g., *t*-test) could not be achieved for the entire datasets using common data transformation methods (e.g., log, square root, squared, cube-root).

The relative importance of each variable (B and I_t), in terms of explaining the variance of NPP for each ROV station, was assessed using the coefficient of determination (R^2) for all up-scaled NPP estimates (P_t) vs. chl *a* (B) estimates (i.e., explained variance due to chl *a*), and NPP estimates (P_t) vs. bottom-ice light (I_t) observations (i.e., explained variance due to light). The R^2 was calculated for each hour (t) of the 24 h period to capture the diurnal variability of light conditions. Values provided in **Table 1** correspond to the daily mean R^2.

Spatial Autocorrelation Analyses

Spatial autocorrelation was used to investigate the horizontal patchiness of sea ice draft, transmittance, chl *a* biomass and NPP measured at the seven ice stations (**Table 1**). Autocorrelation was estimated using Moran's I (Moran, 1950; Legendre and Fortin, 1989; Legendre and Legendre, 1998), which was calculated for each of the eight sites at equally spaced (3 m) distance classes. Individual autocorrelation coefficients or Moran's I estimates were plotted for each distance class in the form of a spatial correlogram (Legendre and Fortin, 1989; Legendre and Legendre, 1998). These analyses were conducted using the "R" software function *correlog* from the "pgirmess" package. Autocorrelation coefficients for each distance class were assigned a two-sided p-value following methods in Legendre and Fortin (1989) and Legendre and Legendre (1998). Global significance was determined on the correlogram using the Bonferroni-corrected significance level. The presence of spatial autocorrelation (i.e., spatial patterns or patchiness) was determined if the correlogram was globally significant at $p < 0.05$. We identified the first x-intercept of globally significant correlogram lines as the patch size (P) of the variables (Legendre and Fortin, 1989; Legendre and Legendre, 1998). Here, patches were identified for sea ice draft (P_d), transmittance (P_t), chl *a* biomass (P_c), and NPP (P_p). This methodology is consistent with spatial autocorrelation analyses used in other snow and sea ice studies to identify patch sizes of both biological and physical variables (e.g., Gosselin et al., 1986; Rysgaard et al., 2001; Granskog et al., 2005; Søgaard et al., 2010).

We classified the correlograms according to correlogram curve patterns described in Legendre and Legendre (1998): (i) multiple-bumps; (ii) wave-like structure; (iii) single bump; (iv) gradient; (v) step; or (vi) random. Because we do not have fully gridded data, it is difficult to differentiate between i) vs. ii), or iv) vs. (v), as the correlograms are very similar. Therefore, we combined these pattern types together resulting in four categories (1) multi-bump/wave; (2) one-bump; (3) gradient/step; and (4) random/noisy. Interpretations of the correlograms together with the *xy* gridded maps allowed for more detailed interpretation of the patterns (Legendre and Legendre, 1998). Patches or regions of high chl *a* biomass, high transmittance, thick draft, and high NPP were identified manually by visually inspecting the gridded maps. The identified patches were compared between variables to identify coincident patches for different variables.

RESULTS

Sea Ice Algal Chl *a* Biomass Estimates

The median chl *a* concentrations were generally low (<3.0 mg m^{-2}) at sampling locations A-H, irrespective of the method used (**Table 1**). Only at location I, median chl *a* concentrations were above 4 mg m^{-2} for ice core and ROV estimates (**Table 1**). The range of chl *a* concentrations observed, however, appeared to be greater at locations G to I compared to locations A to F (**Figure 3A**).

At 5 of the 7 locations sampled for ice cores and ROV measurements, (B-D, F-G), sea ice cores had significantly lower chl *a* biomass than ROV estimates (Wilcoxon test, $p < 0.05$). No significant differences were observed at locations H and I (Wilcoxon test, $p > 0.05$; **Table 1**; **Figure 3A**). On average, ice core-based estimates of chl *a* concentration were 63% of the ROV-based estimates from the same sampling sites. The range was 13–62% for locations B to H, however, location I was substantially larger at 182%. Excluding location H results in a mean underestimation of core based estimates of 43% compared to ROV based estimates. There was no significant difference between integrated estimates of sea ice chl *a* concentrations of ROV and nearby SUIT profiles (Wilcoxon test, $p < 0.05$).

FM-cores were not representative (i.e., within the IQR) of the ROV-derived chl *a* biomass estimates at all locations, except

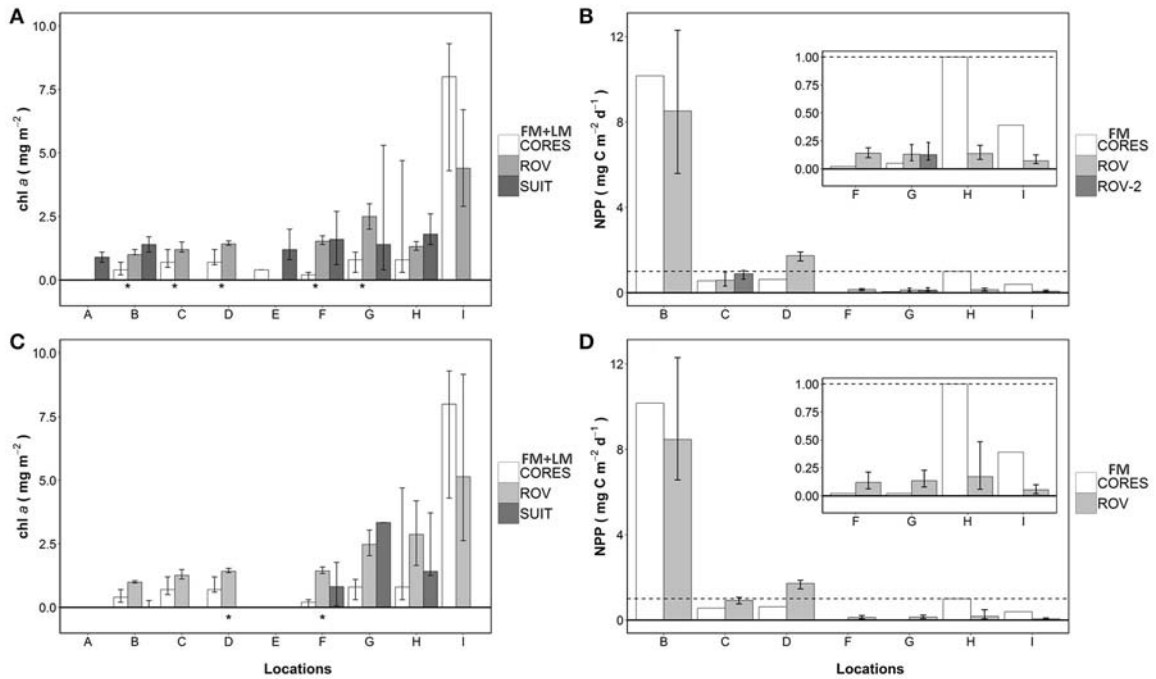


FIGURE 3 | Chl *a* and NPP summarized per sampling gear and location, and subset into dominant ice classes. **(A)** chl *a* biomass for the entire datasets of each sampling gear; **(B)** NPP for the entire datasets of each sampling gear; **(C)** chl *a* estimates from only the dominant ice class; and **(D)** NPP estimates from only the dominant ice class. Dominant ice classes for each location are listed in **Table 3**. Bars represent median and error bars the interquartile range. *Indicates significant Wilcoxon rank sum test at $p < 0.05$ between the CORES and ROV data for the corresponding location. **(A,C)** CORES are the combined datasets of FM-CORES, data from Fernández-Méndez et al. (2015); LA-CORES, data from Lange et al. (2016). **(B,D)** CORES are only the FM-CORES.

location B (**Table 1**). FM-cores at location C, H, and I, over-estimated chl *a* biomass compared to the ROV-derived estimates (**Table 1**). At locations D, E, F, and G the FM-cores under-estimated chl *a* biomass compared to ROV-derived estimates (**Table 1**). When chl *a* estimates were combined by FYI and MYI stations for each sampling method, mean FM-core chl *a* estimates were considerably lower than spatially integrated ROV- and SUIT-based estimates, but these differences were not significant due to the large variability of the datasets (Wilcoxon test, $p > 0.05$; **Table 2**). Regardless of the sampling method, MYI stations had consistently higher chl *a* concentrations and lower PP rates than FYI stations (**Table 2**).

All gridded ROV surveys of chl *a*, sea ice draft, transmittance and NPP are shown in Figures S1–S8. SUIT profiles of chl *a*, sea ice draft, and identified ridges are shown in Figures S9–S16.

ROV-Derived Sea Ice Algal NPP

We accounted for the spatial variability of NPP by combining the variability of both chl *a* and bottom-ice light in the calculations of the larger-scale NPP estimates. All gridded ROV surveys of NPP are shown in Figures S1–S8. We then determined the explained variance of NPP by each variable individually. At locations B, C, F, G, and I, the spatial variability of bottom-ice light explained most of the spatial variability of the up-scaled NPP estimates, whereas at locations D and H, chl *a* explained most of the spatial variability of NPP (**Table 1**; **Figure 4**).

The largest diurnal variabilities of light levels and explained variances were observed at locations with the highest mean bottom-ice light levels (**Table 1**; **Figure 4**). At all stations, the explained variance of chl *a* was inversely related to light, which is expected since NPP is a function of both variables and chl *a* estimates were constant over the diurnal cycle while only light varied. The inter-location differences regarding which variable (chl *a* or light) explained most of the variance in NPP cannot be stated for certain as we observed no significant correlations between the explained variance for each station and any other station variable (e.g., nutrient concentration, median and IQR chl *a* or bottom-ice light).

FM-core NPP estimates were representative (i.e., within the IQR) of the up-scaled estimates at station group B and one ROV survey at station group C (**Table 1**; **Figure 3B**). FM-cores under-estimated NPP at station groups C, D, F, and G, and over-estimated NPP at station groups H and I compared to the up-scaled ROV-based NPP estimates (**Table 1**; **Figure 3B**). The differences between methods were likely the result of differences in chl *a* and/or light. Location B had similar chl *a* biomass and NPP for both the FM-core and up-scaled estimates (**Table 1**; **Figures 3A,B**). Station groups D, F, and G had higher up-scaled chl *a* biomass and NPP estimates compared to FM-core estimates (**Table 1**; **Figures 3A,B**). Conversely, station groups H and I had lower up-scaled chl *a* biomass and NPP estimates compared to FM-core estimates (**Table 1**; **Figures 3A,B**). Only station group C had higher chl *a* biomass but lower NPP estimates for

TABLE 2 | Ice algal chlorophyll *a* biomass and NPP summarized for sampling gears into MYI and FYI. Means, range (min–max), and sample size [N] are provided for comparison to values presented in Fernández-Méndez et al. (2015).

Sampling Method	Summary statistics	Chl *a* (mg m^{-2})		Net Primary Production (mg C m^{-2} d^{-1})	
		MYI	FYI	MYI	FYI
FM-CORES	Mean (range) [N]	5.5 (0.4–8.0) [3]	0.84 (0.3 –1.7) [5]	0.48 (0.05 –1.0) [3]	2.36 (0.02 –10.16) [5]
ROV	Mean (range) [N]	3.4 (0.0 –19.8) [1,993]	1.46 (0.0 –18.5) [3,333]	0.18 (0.0 –4.45) [1,993]	2.05 (0.0 –141) [3,333]
SUIT	Mean (range) [N]	2.5 (0.3 –16.7) [132]	1.7 (0.0 –18.5) [242]	–	–
FM-CORES	Median (IQR)	8.0 (4.2 –8.0)	0.6 (0.4 –1.2)	0.39 (0.22 –0.70)	0.56 (0.45 –0.62)
ROV	Median (IQR)	2.6 (1.8 –3.9)	1.3 [1.1 –1.6]	0.11 (0.06 –0.20)	0.71 (0.17 –1.17)
SUIT	Median (IQR)	1.8 (1.4 –2.7)	1.3 (0.8 –2.1)	–	–

the FM-cores compared to the up-scaled estimates (**Table 1**; **Figures 3A,B**). Furthermore, light levels were comparable (237a) or slightly higher (237b) for the FM-core derived NPP estimates compared to the ROV surveys (**Table 1**; **Figures 3A,B**). When FM-cores and the up-scaled NPP estimates were pooled into FYI and MYI stations, we observed no significant differences between the methods (Wilcoxon test, $p > 0.05$; **Table 2**). The median and IQR-values had large differences between sampling methods for the MYI stations but the mean values were similar (**Table 2**).

Sea Ice Algal Chl *a* Biomass and NPP in Relation to Sea Ice Properties

Sea Ice Classes

Chl *a* biomass and NPP estimates were divided into the five different ice classes. The values showed large variability between ice classes and locations, and within ice classes and locations (**Figure 5**). ROV-derived chl *a* biomass estimates at locations B and I were highest in the thickest sea ice class (2.0 m +; **Figure 5A**). Locations B and C had high ROV-derived chl *a* biomass in the thinnest ice class (0.0–0.5 m; **Figure 5A**). The three middle ice classes generally had uniform ROV-derived biomass estimates, with the exception of location H which had the highest ROV-derived chl *a* biomass in the 1.5–2.0 m ice class (**Figure 5A**). The SUIT-derived estimates were very low at location B for all ice classes and highly variable within the ice classes for all other stations with no obvious patterns (**Figure 5C**). In general, at each location ROV-derived NPP estimates showed a decreasing trend with increasing range of ice class thickness values (**Figure 5B**).

The dominant ice class surveyed by the ROV was identified by the modal sea ice draft of ice floes based on EM31 measurements (**Table 3**). Ice core and ROV chl *a* biomass estimates for the dominant ice classes differed significantly (Wilcoxon test, $p < 0.05$) at 2 locations (F,G; **Table 3**; **Figure 3C**). NPP estimates derived from FM-cores and ROV observations showed no obvious changes and maintained the same patterns (i.e., non-representativeness) for all locations. Most obvious differences were observed between the entire chl *a* biomass surveys and dominant ice class subsets for the SUIT at locations B, F and G, and for the ROV at locations H and I (**Tables 1**, **3**; **Figures 3A,C**). Furthermore, the separation between low chl *a* biomass locations B to F and high chl *a* biomass locations G to I is more obvious from the large scale dominant ice class estimates (**Figure 3C**).

Two sea ice regimes were identified at station 349 of group H: one thicker sea ice region and one thinner region (Figure S7). The thicker region (median: 1.9, IQR: 1.2–3.5 mg chl *a* m^{-2}) had significantly higher (Wilcoxon test, $p < 0.05$) chl *a* biomass than the thinner region (median: 1.3, IQR: 1.2–1.4 mg chl *a* m^{-2}). NPP, however, was significantly lower at the thicker region (median: 0.07, IQR: 0.04–0.19 mg C m^{-2} d^{-1}) compared to the thinner region (median: 0.14, IQR: 0.12–0.21 mg C m^{-2} d^{-1}). Ice cores from the thicker region had higher chl *a* biomass (median: 0.3, IQR: 0.2–0.5 mg chl *a* m^{-2}) compared to ice cores from the thinner region (median: 4.6, IQR: 2.8–6.2 mg chl *a* m^{-2}) although the *p*-value of the Wilcoxon test was 0.06 due to the low sample size.

Sea Ice Ridges

At ice location B (station 224; **Figure 1**) we identified two sea ice ridges based on the ROV draft measurements (**Figure 6A**). Ridge 1 had a median sea ice draft of 4.5 m and ridge 2 had a median draft of 2.8 m based on ROV measurements (**Table 4**). Bottom-ice light was significantly higher in level ice compared to both ridges ($p < 0.05$; **Table 4**). Nonetheless, both ridges had significantly higher ice algal chl *a* biomass than the level ice ($p < 0.05$; **Table 4**; **Figure 6C**). Ridge 2, however, had significantly lower NPP compared to level ice, whereas ridge 1 had similar NPP compared to the level ice (**Table 4**; **Figure 6D**). Conversely, ridge 1 had both higher draft values and higher bottom-ice scalar irradiance values *I* at the bottom compared to ridge 2 (**Table 4**; **Figures 6A,B,D**). In the level ice, chl *a* biomass and bottom-ice light explained comparable amounts of the NPP variance. At ridges 1 and 2, however, chl *a* biomass explained relatively more variance compared to bottom-ice light (**Table 4**).

Based on the ridge identification analysis for all SUIT stations we calculated a mean (min–max) ridge density of 7.5 ridges km^{-1} (2.5–18.0), mean ridge width of 68.7 m (47.6–100.3), and a mean percent total ice coverage by ridges of 9.2% (2.5–15.4%). Ridge analysis summaries for each SUIT station are shown in **Table 5**. SUIT profiles with identified ridges are shown in **Figure 7** (station 223) and for all other stations in Figures S9–S16.

High chl *a* biomass sea ice ridges were also identified within three SUIT stations (station 223: **Figure 7**; stations 233, 285, and 358 Figures S11, S13, S16). These identified high chl *a* biomass ridges had chl *a* biomass estimates in the range 2–9 mg chl *a* m^{-2} (**Table 5**), which was larger than the overall SUIT profile median values in the range 1.2–1.9 mg chl *a* m^{-2}

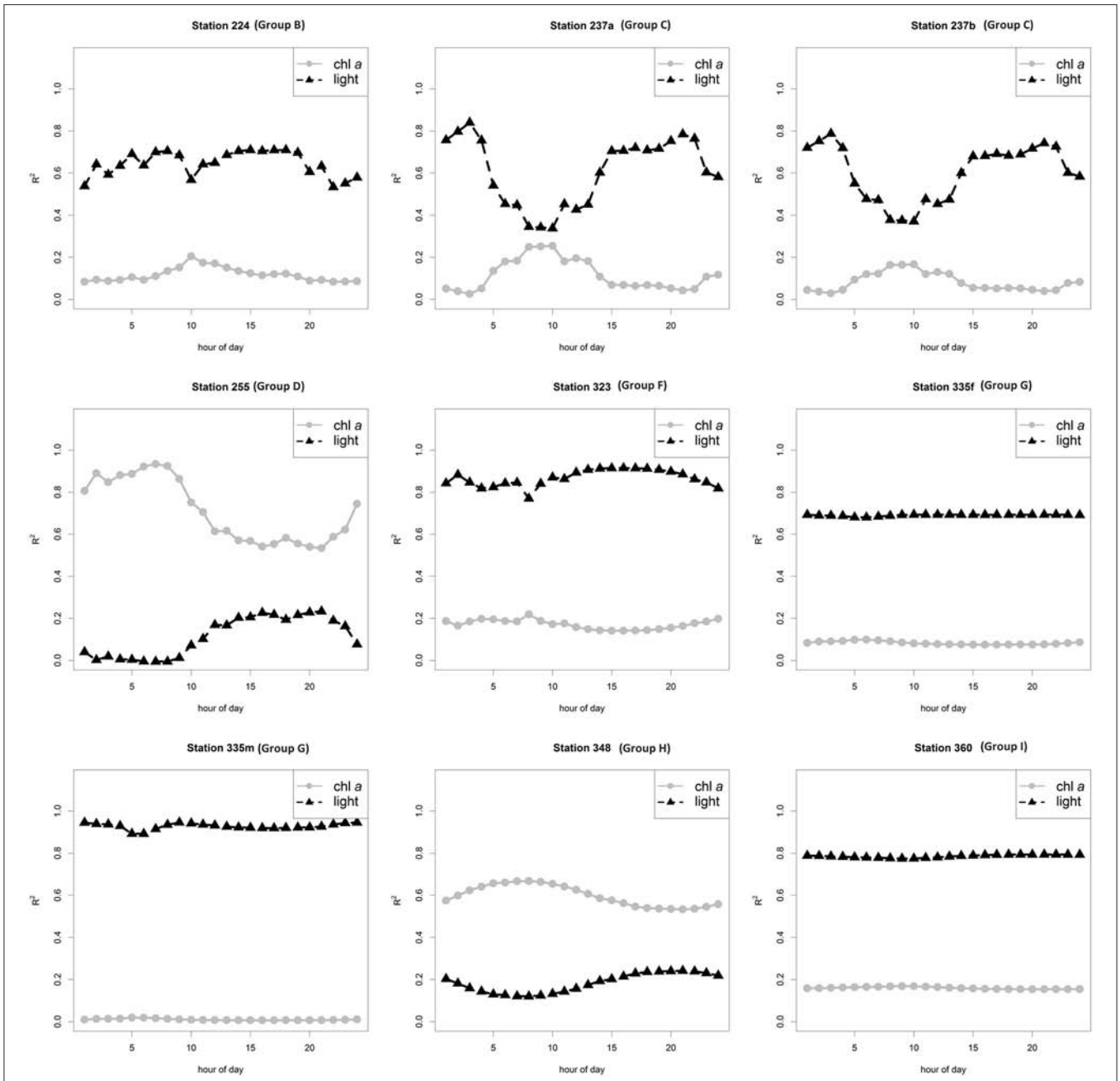


FIGURE 4 | Explained variance (R^2) of NPP by up-scaled chlorophyll *a* and bottom-ice PAR (light) per hour for each ROV station and survey listed in **Table 1.**

(**Table 1**). When comparing chl *a* biomass values at coincident identified sea ice ridges with chl *a* biomass at level ice for each SUIT haul separately, we observed significantly higher (Wilcoxon test, $p < 0.05$) sea ice ridge chl *a* biomass than level ice chl *a* biomass at 2 SUIT hauls (stations 223 and 233; **Table 5**). When comparing all SUIT observations combined, sea ice ridge chl *a* biomass (median: 0.7 and IQR: 0.2–1.4 mg chl *a* m^{-2}) was significantly higher (Wilcoxon test, $p < 0.05$) than level ice chl *a* biomass (median: 0.3 and IQR: 0.0–1.0 mg chl *a* m^{-2}).

Spatial Variability of Sea Ice Properties, Algae Chl *a* Biomass, and NPP

Autocorrelation analyses for each station were conducted using correlograms (i.e., Moran's I vs. distance classes), and were all globally significant at the Bonferonni corrected level ($p < 0.05/n$; $n =$ the number of distance classes). Patch sizes, identified as the distance class at which the first zero value of Moran's I occurred in the correlograms, were highly variable between stations and between measured variables (**Table 6**). Patch sizes for chl *a* (P_c) had a lower range of values between 7 and 30 m, whereas

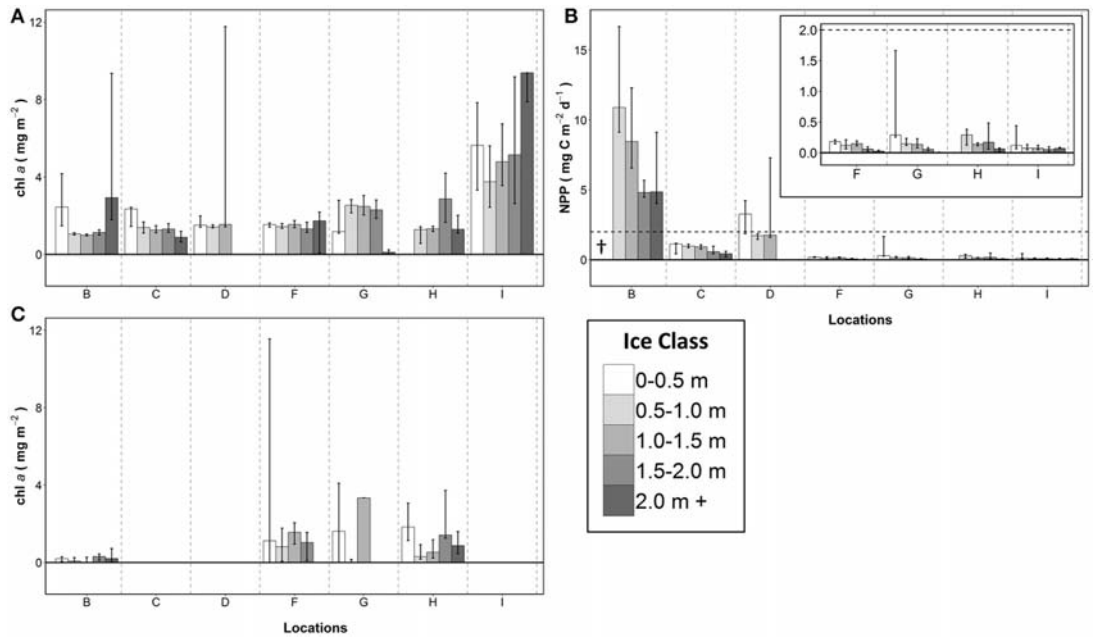


FIGURE 5 | Summary of chlorophyll *a* and NPP estimates per ice class and location for: **(A)** ROV derived chlorophyll *a* biomass; **(B)** ROV-derived NPP estimates; and **(C)** SUIT derived chlorophyll *a* biomass estimates. Bars represent median and error bars the interquartile range. †Indicates missing values.

patch sizes for transmittance (P_t), draft (P_d) and NPP (P_p) were slightly higher in the range 10–50 m (**Table 6**). P_t and P_p were comparable (within 5 m) at all ROV stations except 224, which had the two identified ridges. The shapes of correlogram curves were similar for transmittance and NPP for all station surveys (**Figure 8** and Figures S17–S24). Correlogram shape comparisons for all stations were highly variable with no obvious patterns for all other measured variables (**Figure 8** and Figures S17–S24).

Based on the manually identified patches within the gridded maps, coincident patches of high transmittance and high NPP were observed at all stations. Coincident patches of only high chl *a* and thick draft values were observed at stations 224 and 237b, although the patches at 237b were more subtle (**Figure 6** and Figure S18). The two draft patches observed at 224 correspond to ridge 1 and ridge 2 (**Figure 6**) described in the previous section Sea Ice Ridges. Coincident patches of only high chl *a*, transmittance and NPP were observed at stations 224, 335f,m, and 360 (**Figure 6**, Figures S5, S6, S8).Coincident patches of only high chl *a* and NPP were observed at stations 255 and 349 (Figures S3, S7).

DISCUSSION

Overall Representativeness of the Ice Algal Chl *a* Biomass and NPP Estimates Using Different Sampling Methods

Chl a Biomass

During land-based campaigns in coastal regions it is possible to achieve ice core sample sizes well over 50 ice cores (e.g., Gosselin et al., 1986; Rysgaard et al., 2001; Granskog et al., 2005; Mundy et al., 2007; Campbell et al., 2015). However, such studies are conducted over a period of weeks to months and are typically confined to a local study region. Furthermore, land-based studies are generally conducted on landfast sea ice, in regions dominated by seasonal sea ice. Thus, during the advanced melt stages in seasonally ice covered regions sampling sea ice is typically not done because it also coincides with the termination of the algal bloom, and/or due to logistical and safety constraints. Where sea ice survives into late-summer (e.g., the central Arctic Ocean), ship-based sampling is the most effective sampling approach. Although ship-based sampling has some advantages (e.g., bringing the equipment and lab to the study region), the main disadvantage is that sampling is generally time-limited. Thus, ice core sampling during ship-based campaigns is generally limited to <10 algal chl *a* biomass or NPP cores per ice station making it difficult to conduct spatial studies of sea ice algae (e.g., this study; Gosselin et al., 1997; Gradinger, 1999; Schünemann and Werner, 2005; Fernández-Méndez et al., 2015). Even during long term ship-based studies (e.g., Melnikov et al., 2002) ice core sampling was limited to a small number of cores for each sampling interval every 1–2 weeks.

Our results demonstrate large uncertainties in coring-based methods for capturing the larger-scale variability of ice algal chl *a* biomass observed by the ROV-based methods. However, assessing the magnitude of this uncertainty for other studies is not possible. In general, our ice coring results under-estimated ice algal chl *a* biomass at the relatively lower chl *a* biomass locations (B-F), which implies an overall under-estimation of total chl *a* biomass. Only at the higher chl *a* biomass locations (H and I) the ice cores accurately captured the variability of ice algal chl *a* biomass. The higher chl *a* biomass observed at

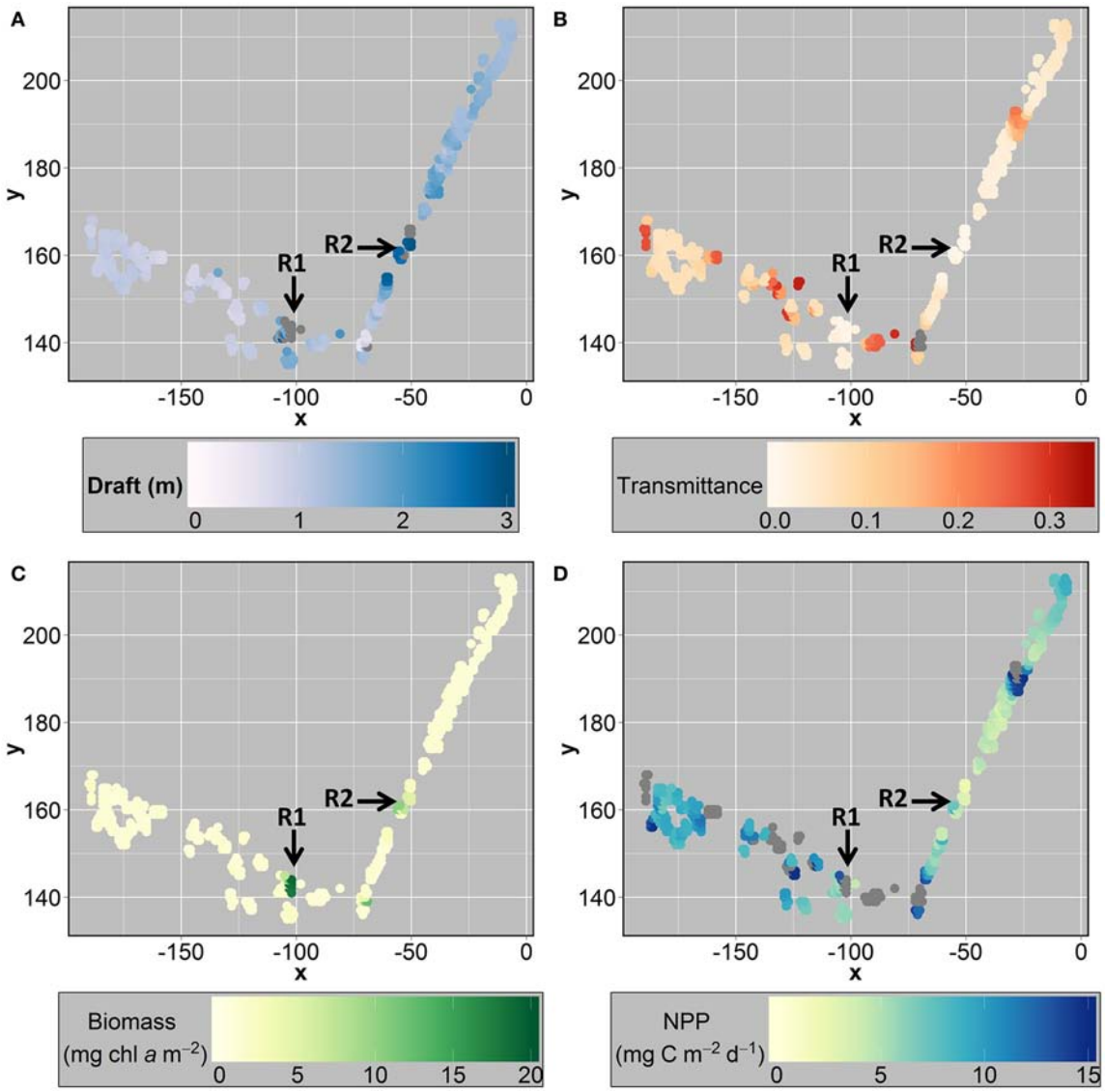


FIGURE 6 | Gridded x-y (meters) map of the remotely operated vehicle (ROV) station 224, showing: **(A)** draft (m); **(B)** transmittance; **(C)** chlorophyll *a* biomass (mg m^{-2}) derived from ROV spectral radiation measurements; and **(D)** net primary production-NPP (mg C m^{-2} s^{-1}) derived from ROV measurements. R1 and R2 depict ridge 1 and ridge 2, respectively. Gray circles represent values greater than the scale maximum value.

locations G-I was likely the result of less melt-induced algal losses due to thicker ice and lower melt rates at these high-latitude locations (Lange et al., 2016). This difference can be explained because at low chl *a* biomass stations relatively higher chl *a* biomass patches had a lower probability to be sampled by coring compared to higher chl *a* biomass stations, and hence were not accurately represented, whereas at high chl *a* biomass locations the probability of sampling higher chl *a* biomass locations was higher. We must also note that the possibility that the up-scaled spectrally derived estimates over-estimated the true chl *a* biomass is unlikely, because the model for spectrally deriving chl *a* biomass had no directional bias related to chl *a* concentration in sea ice (Lange et al., 2016).

The higher chl *a* biomass location I showed no significant difference between the cores and ROV-based chl *a* biomass estimates. In the individual core values, however (0.05, 6.46, 8.03, 8.00, and 11.83 mg chl *a* m^{-2}), only one core was within the IQR (2.96–6.70 mg chl *a* m^{-2}). In this sample size, one core with near-zero chl *a* biomass was highly influential and may have impeded the detection of significant differences. A similar pattern was also apparent at location H, which also showed no significant difference, but also had only one core within the IQR of the up-scaled chl *a* biomass estimates. The discrepancy between the ice core-based and ROV-derived chl *a* biomass estimates indicates the ice algal chl *a* biomass was highly variable at small scales (<2 m), which was difficult to

TABLE 3 | Modal sea ice draft from literature (Boetius et al., 2013; Katlein et al., 2015b) and ROV measurements, dominant ice class based on literature modal draft, interquartile range of chlorophyll *a* biomass observations for the dominant ice class using different gears (ice coring, ROV, and SUIT) summarized for each location.

Location	Dominant ice class	Floe draft[a] (m)	Modal ROV draft (m)	Modal SUIT draft	Gear	chl *a* (mg m^{-2})	NPP (mg C m^{-2} d^{-1})
B	1.0–1.5 m	1.1	1.0		CORES	0.4 (0.2–0.7)	10.16
					ROV	0.99 (0.95–1.05)	8.43 (6.73–12.29)
					SUIT	0.0 (0.0–0.27)	nd
C	1.0–1.5 m	1.1	1.2		CORES	0.7 (0.5–1.2)	0.56•(–)
					ROV	1.26 (1.11–1.49)	0.92 (0.76–1.06)
D	0.5–1.0 m	0.8	0.8		CORES	0.7 (0.6–1.2)	0.62•(–)
					ROV	1.42 (1.37–1.53)	1.72 (1.45–1.86)
F	0.5–1.0 m	0.7	1.3		CORES	0.2 (0.0–0.3)*	0.02•(–)
					ROV	1.43 (1.34–1.60)	0.12 (0.06–0.22)
					SUIT	0.8 (0.1–1.8)	nd
G	1.0–1.5 m	1.3	1.1/1.4		CORES	0.8 (0.3–1.1)*	0.02•(–)
					ROV	2.48 (2.04–3.03)	0.13 (0.08–0.23)
					SUIT	3.34 (3.34–3.34)	nd
H	1.5–2.0 m	1.7	1.2		CORES	0.8 (0.3–4.7)	1.00•(+)
					ROV	2.87 (1.65–4.19)	0.17 (0.06–0.48)
					SUIT	1.4 (1.3–3.7)	–
I	1.5–2.0 m	1.6	1.1		CORES	8 (4.3–9.3)	0.39•(+)
					ROV	5.14 (2.62–9.46)	0.05 (0.02–0.10)

[a]*Literature modal ice thickness converted to draft by multiplying by 0.9. "–" indicates no data. Noteworthy wilcoxon test results are indicated by * for a significant difference at $p < 0.05$ for comparisons between cores and ROV chl a biomass for observations on ice within the dominant ice class. •Indicates a CORES NPP estimate outside the IQR of the ROV NPP estimates for observations within the dominant ice class. (+) indicates CORES greater than ROV 75th percentile; and (–) indicates CORES smaller than ROV 25th percentile. nd refers to no data.*

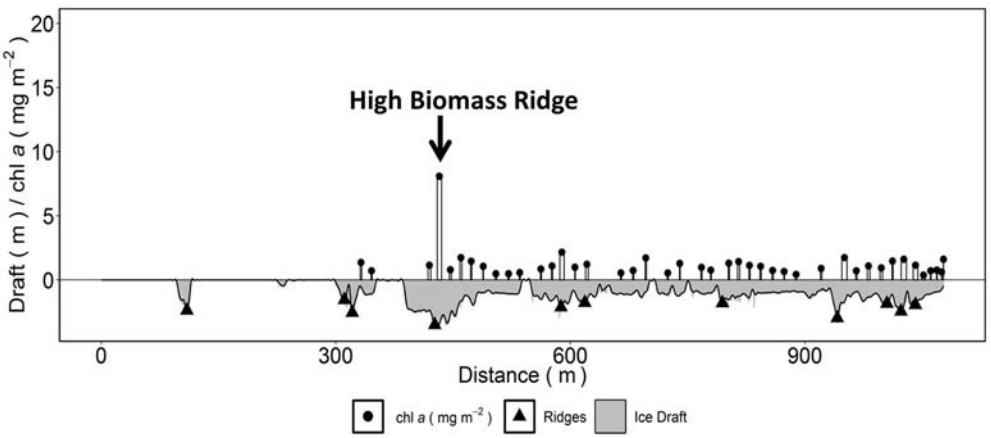


FIGURE 7 | Horizontal profile of Surface and Under-Ice Trawl (SUIT) station 223 showing sea ice draft, identified ridges and chlorophyll *a* biomass derived from spectral radiation measurements. Highlighted is an identified high chlorophyll *a* biomass sea ice ridge. Width of the white bars corresponds to the relative along-track footprint of spectral radiation measurements. The black line corresponds to the smoothed sea ice draft curve used for the ridge identification procedure and was determined from the ice draft measurements (gray shaded area).

capture with average measurement footprints between 1 and 2 m for ROV surveys. Individual data points of up-scaled estimates averaged chl *a* concentration over a larger area, and were thus less likely to capture small patches of extremely high chl *a* biomass or extremely low chl *a* biomass (i.e., values in the range 8–12 mg chl *a* m^{-2} or with near-zero chl *a* biomass). These considerations highlight two important sampling constraints. First, the cores did not capture the large-scale variability; and second, we were unable to assess the small-scale variability below 2 m. The second limitation is less drastic since the signal received from the sensor under the ice does capture the small-scale variability within its measurement by averaging it over a larger distance. Since little is known or has been reported on summertime spatial variability of ice algal chl *a* biomass we propose that observations from both core-based and under-ice spectral profiling systems should be combined when making

TABLE 4 | Comparison of chlorophyll *a* biomass and net primary production between sea ice ridges and level ice at station 224.

Variable	Ridge 1 (*N* = 20)	Ridge 2 (*N* = 21)	Level Ice (*N* = 427)
Draft (m)	4.5 (2.4–5.0)*	2.8 (2.7–2.9)*	1.1 (0.95–1.4)
Chl a (mg m^{-2})	1.8 (1.7–17.9)* [0.91][a]	3.4 (2.5–5.3)* [0.58][a]	1.0 (0.97–1.1) [0.78][a]
NPP (mg C m^{-2})	6.9 (5.7–17.9)	4.0 (2.9–4.2)*	8.7 (5.9–12.3)
I (μmol photons m^{-2} s^{-1})	11.6 (4.8–12.9)* [0.79][a]	2.5 (2.2–5.6)* [0.0][a]	39.0 (22.8–61.2) [0.71][a]

*Ridges are identified in **Figure 7**. *Indicates a statistically significant ($p < 0.05$) Wilcoxon test between the corresponding Ridge and Level Ice.*
[a]*Value within square brackets represents the explained variance of NPP by the corresponding variable and data subset of ridge or level ice. "I" is the bottom ice light levels (PAR).*

TABLE 5 | Summary of ridge identification analysis from the SUIT hauls conducted during PS80.

SUIT station (location)	Ridge count	Chl *a* in ridges (mg m^{-2})	Chl *a* in level ice (mg m^{-2})	Total Distance (km)	Density (ridges km^{-1})	Mean ridge width (m)	Ridge coverage (% of total ice)
216 (A)	4	0.2(0.2–0.8) [3]	0.0(0.0–0.2) [43]	1.6	2.5	47.6	3.0
223 (B)	12	0.6(0.5–1.2) [9]*	0.0 (0.0–0.4) [34]*	0.8	15.8	79.5	10.5
233 (B)	12	0.5(0.1–0.7) [10]*	0.1(0.0–0.3) [35]*	1.5	8.1	60.8	4.1
248 (C)	4	0.6(0.4–1.1) [3]	0.3(0.1–0.7) [58]	1.5	2.7	91.1	6.2
285 (E)	4	0.5(0–1.1) [7]	0.2(0.0–0.9) [84]	1.3	3.1	91.5	7.1
321 (F)	3	0.0 (0.0–0.0)[3]	1.0 (0.0–1.9) [60]	0.7	4.6	100.3	15.4
345 (G)	8	4.7 [1]	0.0 (0.0–3.2) [17]	1.2	6.6	49.3	4.1
358 (H)	11	2.7(1.2–2.8) [9]	0.8(0.4–1.4) [92]	2.0	5.6	48.7	2.5
376 (-)	3	4.6(2.3–7.7) [3]	0.9(0.6–1.9) [10]	0.2	18.0	49.5	29.6

**Indicates a statistically significant ($p < 0.05$) Wilcoxon test comparing chl a biomass in ridges and level ice.*

assumptions about multi-scale spatial variability of ice algal chl *a* biomass.

The fact that no statistical differences (Wilcoxon test, $p > 0.5$; **Table 2**) were observed between ROV-based and ice core-based estimates (both chl *a* and NPP) when they were grouped into MYI and FYI stations, an approach taken by Fernández-Méndez et al. (2015), may suggest an improvement because it increased the probability of the ice cores to be representative of the larger area. In this case the sample sizes and range of chl *a* biomass values were sufficient to obscure any differences at the station level. This method should only be considered when other options are not possible, because large uncertainties are still present even though significant differences were not found. For example, even though in this case the mean MYI FM-core chl *a* biomass values were not significantly different (Wilcoxon test, $p > 0.05$), each MYI FM-core value was higher or lower than the IQR of the larger-scale estimates. A similar pattern was observed with the FYI grouping comparison although not as drastic because overall the values were smaller, particularly the range of values. Potential uncertainties of grouping ice cores should also be considered depending on the objectives of your study. Grouping the ice cores into MYI and FYI would result in mean ice core MYI algal chl *a* biomass estimates 160% larger than the ROV MYI estimates. In contrast, FYI ice core estimates would be around 60% of the ROV FYI estimates. For grouped ice core-derived NPP the difference for MYI is even more pronounced at 270% larger compared to ROV estimates. FYI NPP values, however, were comparable between both methods.

Photoacclimation may be another potential factor influencing the chl *a* to carbon ratios, which could in turn explain the increased chl *a* biomass at higher latitude stations due to increased chl *a* production under lower light conditions. Fernández-Méndez et al. (2015) measured lower photoacclimation indices for the higher latitude stations (I_k; mean: 30 μmol photons m^{-2} s^{-1}, range 17–45) compared to the lower latitude stations (mean: 60 μmol photons m^{-2} s^{-1}, range: 34–77). However, we did not observe any variability (<1 g C: g chl *a*) between high and low latitude stations in the chl *a* to POC ratios (data not presented here), therefore it is unlikely that photoacclimation explains the regional chl *a* biomass differences.

NPP

In general, NPP sampling involves measuring available PAR levels through a hole in the ice (Gosselin et al., 1997), which may produce higher than expected values due to the hole. PAR available for bottom-ice algae may also be modeled by using simple light extinction models (Fernández-Méndez et al., 2015). Both methods are established and regularly employed, however, both are limited in the fact that they do not account for the spatial variability of the bottom-ice PAR levels. During spring, ice algae are typically light-limited and therefore have higher chl *a* biomass where light levels are higher (e.g., Gosselin et al., 1986), assuming there is no or limited photo-inhibition. During our summer sampling period, however, we found no strong correlation at any station between the ROV-derived chl *a* estimates and available under-ice light (maximum spearman correlation coefficient, $r = 0.22$). This means the under-ice light

TABLE 6 | Summary of the autocorrelation analyses per location and ROV survey.

Location	Station	P_c^*	P_t *	P_d^*	P_p^*	Pattern chl *a*	Pattern TM	Pattern draft	Pattern NPP	Similar Correlograms[a]	Coincident patches[b]
B	224	10	12	25	30	Bumps-waves	Bumps-waves	Bumps-waves	Bumps-waves	Chl *a*-TM-NPP	2 × chl *a*-draft; 4 × TM-NPP; 1 × chl *a*-TM-NPP;
C	237a	10	18	30	20	Random-noisy	Bumps-waves	Step-gradient	Bumps-waves	TM-NPP	3 × TM-NPP
	237b	23	15	19	15	Bumps-waves or 1-bump	Bumps-waves	Bumps-waves	Bumps-waves	Chl *a*-draft-NPP; TM-NPP	1 × chl *a*-NPP-draft; 3 × TM-NPP
D	255	7	10	12	10	1-bump or random	Bumps-waves	Bumps-waves	Bumps-waves	TM-NPP	1 × chl *a*-NPP; 1 × TM-NPP
F	323	14	31	47	35	Bumps-waves	1-bump	1-bump	1-bump	TM-NPP-Draft	1 × large/multi-patch TM-NPP
G	335m	13	14	24	14	Bumps-waves	Bumps-waves	1-bump	Bumps-waves	TM-NPP-chla	1 × chl *a*-TM-NPP; 2 × TM-NPP
	335f	15	50	47	51	Bumps-waves	Step-gradient	Step-gradient	Step-gradient	TM-NPP	1 × large/multi-patch chl *a*-NPP-TM
H	349	25	39	41	40	Bumps-waves	Step-gradient	Step-gradient	Step-gradient	TM-NPP	1 × large/mulit-patch chl *a*-NPP; 2 × TM-NPP
I	360	30	29	33	29	Bumps-waves	1-bump	1-bump	1-bump	TM-NPP-Draft	1 × chl *a* –TM; 1 × TM- NPP

Patch sizes for chl a, P_c; transmittance, P_t; draft, P_d; and NPP, P_p. TM corresponds to transmittance, and NPP to net primary production.

[*] All correlograms globally significant at the Bonferonni corrected level ($p < 0.05/n$; where n is number of distance classes; Legendre and Legendre, 1998).

[a] Identifies correlogram curves which are similar in shape to each other (e.g., chl a-TM-NPP means the correlogram curves are similar for the chl a, transmittance and net primary production).

[b] Manually identified patches that are coincident in location to each other. The number of patches per ROV survey is followed by which patches are coincident (e.g., TM-NPP refers to a transmittance patch coincident to an NPP patch). Large/multi refers to a larger area with multiple small patches in close proximity.

and chl *a* varied independently of each other. This behavior is expected, because in late-summer biomass losses due to high melt rates have a dominant influence on bottom-ice biomass (e.g., Grossi et al., 1987; Lavoie et al., 2005; Lange et al., 2016). These conditions would not have sustained a bottom-ice algal community, and therefore even if light conditions were suitable for high primary production rates the NPP would have been almost zero if no algae were present. With an additional variable (i.e., melt), which can influence NPP, the spatial distribution of NPP may be more complex in late-summer than during the spring to summer transition making it even more important to understand and account for the spatial variability of both chl *a* biomass and the bottom-ice light field.

Location B had similar NPP estimates for the FM-core and up-scaled observations (**Table 1**; **Figure 2**), which we attributed to the similar chl *a* biomass estimates (**Table 1**; **Figure 1**). Even though light levels and chl *a* biomass were only slightly larger at location B compared to groups C and D, group B had NPP estimates almost an order of magnitude larger than groups C and D. This was attributed to the substantially higher value of the photosynthetic parameter P_s^B determined for this station (Fernández-Méndez et al., 2015), compared to all other stations. This demonstrates that the combination of data from several stations, an approach described by Fernández-Méndez et al. (2015) and used by others (Mundy et al., 2011; Campbell et al., 2016), was not only able to improve the spatial representativeness of light and chl *a* but also accounted for the potential variability of the derived photosynthetic parameters. Our results suggest pooling ice core samples increases the chance for the samples to be representative of both chl *a* biomass and NPP estimates. Because of the large range of chl *a* biomass and NPP estimates, and the small number of samples, however, this approach can still carry a high risk of obtaining non-representative estimates (e.g., overestimates of up to 270% for MYI).

The same directional difference of chl *a* biomass and NPP observed between up-scaled and FM-core estimates for all station groups, except group C, suggests the differences between the FM-cores and up-scaled NPP estimates were driven by the differences in chl *a* biomass. This was further confirmed by the fact that the bottom-ice light levels used for each method were comparable for each station (**Table 1**). The opposing pattern of chl *a* biomass and NPP between up-scaled and FM-core estimates at location C, even though light levels were comparable, suggests that the spatial variability of both the chl *a* biomass and bottom-ice light had a combined influence on the observed differences that is not apparent from the overall survey estimates. The explained variance of NPP by chl *a* and light showed large diurnal variability and large inter-location variability, which indicates a complex and highly variable relationship between ice algal chl *a* biomass and light levels during our sampling period. These results emphasize the importance of accounting for both the spatial variability of ice algal chl *a* biomass and the bottom-ice light field in order to make representative NPP estimates. We must also note the possible influence of nutrients since we found a significant ($p < 0.05$) positive correlation ($r = 0.46$) between explained variance of NPP by chl *a* with sea

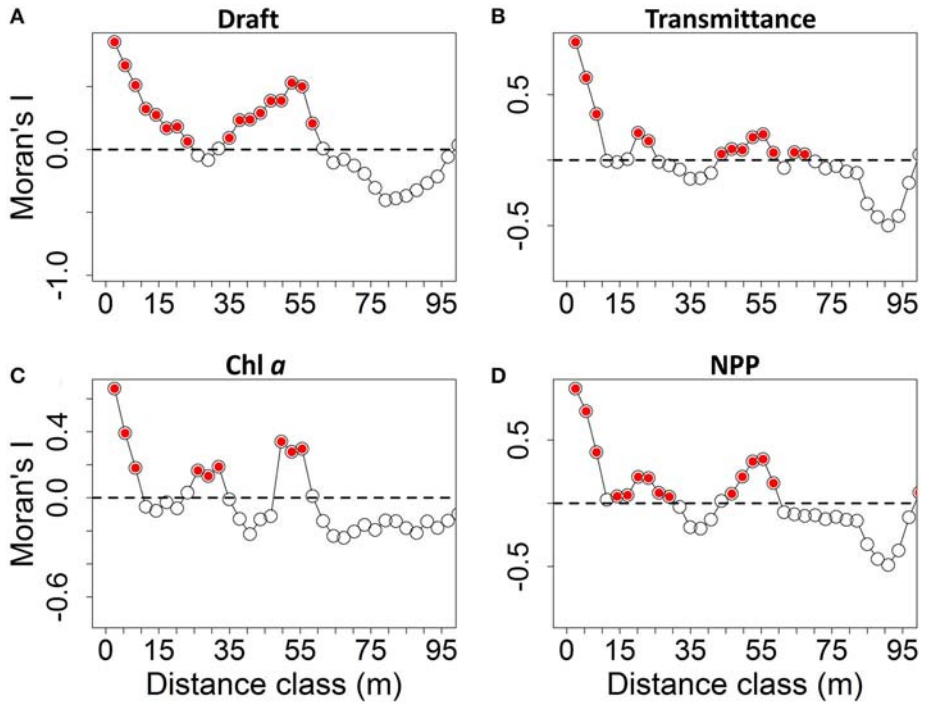


FIGURE 8 | Correlograms showing Moran's I vs. distance classes at ROV station 224 for: **(A)** draft (m); **(B)** transmittance; **(C)** chlorophyll *a* biomass (mg m^{-2}) derived from ROV spectral radiation measurements; and **(D)** net primary production-NPP (mg C m^{-2} s^{-1}) derived from ROV measurements. Red filled circles represent significant values at $p < 0.05$.

ice NO_3 concentrations (data from Fernández-Méndez et al., 2015), and a significant ($p < 0.05$) negative correlation between explained variance of NPP by bottom-ice light with sea ice NO_3 concentrations ($r = -0.55$). These correlations provide some indication that the sea ice nutrient regime could have also had some influence on the relative (inter-station) importance of chl *a* biomass vs. light on NPP.

Gosselin et al. (1997) measured ice algal NPP of up to 300 mg C m^{-2} d^{-1} in the high Arctic Ocean (>87°N) during August. Our results for September in the high Arctic Ocean (station 360) were over 3 orders of magnitude lower than those found by Gosselin et al. (1997) in August for the same area. The large difference in NPP estimates between the studies could partially be explained by the higher incoming solar irradiance in August compared to September. However, upscaling results from Fernández-Méndez et al. (2015) for August were also substantially lower with a mean (range) of 5.8 (0.06–42) mg C m^{-2} d^{-1} and with a similar range of daily mean incoming solar irradiance (101–249 μmols photons m^{-2} s^{-1}). Therefore, we applied the range of incoming irradiance values (∼125–214 μmols photons m^{-2} s^{-1}) observed at the high latitude stations (>87°N) during the Gosselin et al. (1997) study to this studies ROV survey at station 360. We used our observed chl *a* biomass and transmittance in order to calculate potential NPP under higher incoming irradiance conditions typical for August at these high latitudes. Overall NPP increased by nearly the same relative amount as the available light, however, with median values between 0.82 and 1.32 mg C m^{-2} d^{-2} (**Table 7**) this remains two orders of magnitude lower than ice algal NPP observed by Gosselin et al. (1997). This suggests that something other than available light is influencing these observed differences. This is likely explained by the fact that the Gosselin et al. (1997) estimates were dominated by the sub-ice algal species *Melosira arctica*, whereas Fernández-Méndez et al. (2015) measured primary production on ice samples with a lower contribution of *M. arctica*. Thus, our samples represent a good estimate for in-ice algal NPP, however, a conservative estimate for overall ice-associated NPP (Fernández-Méndez et al., 2015).

The explained variance of NPP by bottom-ice light compared to chl *a* using the increased incoming irradiance levels, which were observed in August at high latitudes by Gosselin et al. (1997), showed interesting differences (**Table 7**). As the incoming irradiance increased, the explained variance of chl *a* biomass also increased, while the explained variance of the bottom-ice light decreased to nearly equal values of 0.39 and 0.48, respectively, at an incoming irradiance of 214 μmols photons m^{-2} s^{-1} (**Table 7**). This indicates that under increased irradiance levels the spatial variability of chl *a* biomass becomes more important in terms of contribution to overall NPP estimates. Furthermore, these results suggest a complex spatio-temporal relationship of the relative importance of chl *a* biomass and available bottom-ice irradiance for NPP estimates, which can only be accounted for by characterizing biomass and under-ice light at spatial scales from

TABLE 7 | Net primary production estimates for the ROV survey at station 360 with observed downwelling surface irradiance (PAR) and using different downwelling surface irradiance conditions as observed for the same region (>87° N) earlier in the season (~ mid-August) by Gosselin et al. (1997).

Station	Chl *a*	Downwelling surface PAR	Scalar PAR (I)[b]	NPP	Explained Variance by:	
	(mg m-2)	μmols photons m^{-2} s^{-1}		mg C m^{-2} d^{-1}	I	Chl *a*
360	4.3 (2.8 –6.6)	10 ± 5	0.4 ± 0.4	0.07 (0.05–0.12)	0.78	0.15
		125[a]	4.8 ± 5.7	0.82 (0.55–1.39)	0.61	0.29
		214[a]	8.3 ± 9.7	1.32 (0.86–2.20)	0.48	0.39

[a] *Downwelling surface irradiance data presented in Gosselin et al. (1997) from the same region as station 360.*
[b] *The bottom-ice scalar irradiance used to calculate NPP.*

meters to 100s of kilometers, and temporal scales accounting for diurnal and seasonal variations.

Sea Ice Algal Chl *a* Biomass and NPP in Relation to Sea Ice Properties

Sea Ice Classes

Electromagnetic (EM) sea ice thickness surveys are commonly used to representatively characterize the overall ice thickness distribution (Eicken, 2001; Haas and Eicken, 2001; Haas, 2004) and thus represent a reliable characterization of the dominant ice class for the surveyed floe and overall region. The differences between the ROV-derived and EM-derived modal draft values at several ice stations warranted the use of the EM data to determine dominant ice types. The range of modal ice thicknesses for locations B-G dominated by FYI (0.8–1.3 m) were consistent with previous studies that conducted large-scale airborne and floe-scale ground-based electromagnetic ice thickness surveys for the same region and season (Haas et al., 1997; Haas and Eicken, 2001; Rabenstein et al., 2010). The two locations H and I dominated by MYI had modal thicknesses between 1.6 and 1.8 m, which were also consistent with modal ice thickness values for second-year sea ice from the same region and season (Haas and Eicken, 2001).

Since the dominant ice type thickness value (i.e., modal ice thickness) is a commonly used metric to characterize the sea ice environment it stands to reason that sea ice algal chl *a* biomass from the dominant ice class would also provide a representative metric to describe the overall sea ice algal chl *a* biomass. Comparing the ice algal chl *a* biomass estimates solely from the dominant ice classes showed better agreement between ROV and ice core-derived values (**Figure 3C**). Therefore, we suggest that using chl *a* biomass estimates from the dominant ice class only may be an improvement on providing a single value, which is representative of the large scale sea ice algal chl *a* biomass for that region. There remain some limitations to this approach, since these estimates do not account for the chl *a* biomass of the other ice types/classes. Sampling other ice types/classes may be of particular importance in regions of low chl *a* biomass (e.g., station 224) where high chl *a* biomass features such as ridges may have a substantial contribution to the overall large-scale ice algae chl *a* biomass. A further step to improve these overall chl *a* biomass values could be to use the larger-scale ice thickness density distributions (data not available for this study) to provide weighting factors for chl *a* biomass values of each ice type/class.

The observed trend of higher chl *a* biomass at higher latitude stations was more obvious within the dominant ice class estimates (**Figure 3C**). This was previously attributed to enhanced melt-induced algal losses at lower latitude stations, although based on a smaller number of stations (Lange et al., 2016). Here we have a larger sample size covering a larger geographic region and confirmed the pattern related to latitude and the presence of thicker ice. Enhanced melt is a common mechanism for substantial losses of bottom-ice algae in summer (Grossi et al., 1987; Lavoie et al., 2005). Gosselin et al. (1997) also observed a shift from low to high bottom-ice chl *a* biomass with a shift from low to high latitude, which is consistent with our observed trend. Furthermore, the higher dominant ice class chl *a* biomass estimates between 2.5 and 5.1 mg chl *a* m^{-2} observed at the three high latitude, thicker ice (1.4–1.9 m) locations G–I is consistent with previous studies from high latitude regions of the central Arctic Ocean with bottom-ice algae concentrations in the range of 3–14 mg chl *a* m^{-2} (Gosselin et al., 1997), and up to 22 mg chl *a* m^{-2} (Melnikov, 1997).

Castellani et al. (2017) introduced a pan-Arctic Sea Ice Model for Bottom-Algae (SIMBA) coupled with a 3D sea-ice-ocean model and also showed that within the eastern Eurasian basin during late-summer (this studies sampling region/period) there was an increasing trend in bottom-ice algal chl *a* biomass from lower to higher latitudes. The SIMBA model, however, showed the opposite trend with increasing chl *a* biomass from higher to lower latitudes in the region from the North Pole toward the northern coast of Canada and Greenland (the Lincoln Sea) where the thickest ice in the Arctic Ocean is located. During late-summer Castellani et al. (2017) identified sea ice thickness as a main factor controlling bottom-ice chl *a* biomass by limiting basal melt-induced algal losses during a period of advanced melt. Based on observations alone, a purely latitudinal effect may have been identified as driving the large-scale spatial patterns of sea ice algae. Therefore, the modeling results of Castellani et al. (2017) emphasize the need for more observations in the region between Canada, Greenland and the North Pole (a region coined: the "Last Ice Area"), and that modeling studies are essential in the context of interpreting large-scale patterns of sea ice algae observations. Furthermore, the SIMBA model included different ice classes, such as sea ice ridges. The SIMBA results showed that ridges, though exhibiting lower peak chl *a* biomass compared to level ice, the maximum chl *a* biomass was reached later in the season compared to level ice. More information is required in

order to accurately parameterize sea ice features such as ridges in pan-Arctic models, however, SIMBA is a big step forward in terms of including ice classes within models, which we have identified as an important component of sea ice algal spatial variability.

In contrast to chl *a* biomass, NPP estimates showed no improvement when comparing only the dominant ice class (**Figure 3D**). This suggests that NPP estimates require a different approach for up-scaling and parameterizing models. The complex and highly variable relationship between ice algal chl *a* biomass and light levels during our sampling period suggests that more representative sea ice algal NPP estimates may be achieved by accounting for the relative contribution of NPP within each ice type. This would involve using larger scale ice thickness estimates to assign weighting factors to each ice classes' NPP estimate. In the absence of larger scale observations it is not possible to discover the spatial patterns of sea ice algal chl *a* biomass and NPP, or assess if the ice cores are actually representative of the area. To further improve upon the large scale pan-Arctic NPP and chl *a* biomass estimates we suggest to integrate our five ice classes, together with weighting factors for each ice class (based on large-scale ice thickness surveys), into pan-Arctic studies (e.g., Fernández-Méndez et al., 2015). Accurately assessing sea ice associated NPP in models is of particular importance since it can represent a dominant portion of total (water plus sea ice) NPP in regions covered by sea ice for most of the year (e.g., Gosselin et al., 1997; Fernández-Méndez et al., 2015). In general, pan-Arctic models of NPP, which include sea ice contributions to NPP are very limited (e.g., Lee et al., 2015) highlighting the need for improved sea ice algae model parameterizations.

Sea Ice Ridges

One source of variability in sea ice chl *a* concentrations, light transmittance and derived NPP may be topographical features of sea ice, such as ridges. Sea ice ridges are often under-sampled due to the logistical challenges to sampling this type of ice. Despite this fact, sea ice ridges have been reported to host high abundances of sea ice fauna during advanced melt (Gradinger et al., 2010). Furthermore, in the northern Baltic Sea high chl *a* biomass were observed within the ice along the upper sides of sea ice ridges and within the interstitial spaces, typically present within the unconsolidated aggregation of ice blocks that form ridge keels (Kuparinen et al., 2007). Therefore, we specifically investigated sea ice ridges with the hypothesis that they could host high abundances of ice algae during advanced melt due to lower melt rates in these locations.

We showed that the identified sea ice ridges at ROV station 224 and all SUIT stations (measurements grouped together) had significantly higher chl *a* biomass than measurements under relatively more level ice (e.g., areas that are not ridges). It can be assumed that ridges were under-represented in the ROV sampling due to a preference for relatively uniform sampling sites. In SUIT profiles, the natural distribution of ridges was likely well-represented, because the sampled profile cannot be chosen after the deployment of the net. The overall difference between median level ice chl *a* biomass and median ridge chl *a* biomass from the SUIT surveys, however, was relatively small (0.4 mg chl *a* m^{-2}). The small difference is likely the result of not all ridges having high chl *a* biomass.

Our results of sea ice ridge densities between 2.5 and 18.0 ridges km^{-1} are within the range of larger scale airborne surveys with mean ridge sail densities between 4.3 and 7.2 ridges km^{-1} (Rabenstein et al., 2010). With the high resolution (0.5 m) under-ice topography measurements, we were able to accurately estimate the widths of the ridge and determined that these features represented up to 10% of the total sea ice area. Together with the higher chl *a* biomass observed at sea ice ridges, this indicates that these features require more in-depth investigations and may have a significant impact on overall chl *a* biomass estimates and availability of food for under-ice organisms.

Gradinger et al. (2010) showed sea ice ridges had elevated concentrations of ice meiofauna and under-ice amphipods, which was attributed to the flushing of the sea ice and low-salinity stress imposed at the thinner sea ice environment. Sea ice ridges may also extend into higher salinity water below the highly stratified, fresher surface melt water, which accumulates adjacent to the ridges under thinner ice (Gradinger et al., 2010). These results suggest that higher ice algal chl *a* biomass at ridges may be the result of reduced flushing and lower environmental stress. Furthermore, the presence of high algal chl *a* biomass as a food source may provide an additional explanation for the observed accumulation of organisms at ridges by Gradinger et al. (2010).

In addition to the possibility of reduced flushing and lower environmental stress at ridges, we suggest that the thicker ice experienced lower melt rates than the surrounding level ice resulting in lower algal losses. Perovich et al. (2003) indicated that sea ice ridges experienced an overall greater amount of melt than the surrounding undeformed sea ice, which may appear to contradict our premise. The higher overall melt observed at ridges by Perovich et al. (2003), however, was partially attributed to a few very thick ridges extending deep into the water, which were experiencing melt the entire year even during winter. Except for one weekly measurement in August, the melt rates for ridges were lower than the mean and were among the lowest of all ice types for that entire month during advanced melt (Perovich et al., 2003).

NPP estimates for sea ice ridges showed some interesting patterns at ROV station 224. Although both ridges had significantly higher chl *a* biomass than the level ice, ridge 2 had significantly lower NPP rates than ridge 1 and the level ice, whereas ridge 1 and level ice were not significantly different. These differences were due to the available light measured under the different types of sea ice. The higher chl *a* biomass at ridge 1 compensated for lower light levels compared to the level ice, resulting in similar NPP estimates compared to the level ice. However, the chl *a* biomass at ridge 2 was not sufficient to compensate for the lower bottom-ice light levels. Even though ridge 2 had a thinner median draft (2.8 m) value compared to ridge 1 it still had lower light levels. This shows that ridges can have a considerable impact on the complex relationship between chl *a* biomass and available PAR for NPP estimates at larger spatial scales. Furthermore, these results imply that sea ice features such as ridges have a different and perhaps more complex relationship between available light and chl *a* biomass

than the surrounding sea ice. As a consequence, ridges must be sampled representatively, and both the variability of bottom-ice light levels and the variability of chl *a* biomass are required to make representative large-scale ice algal chl *a* biomass and NPP estimates.

The identification of sea ice ridges as potential chl *a* biomass and NPP hotspots warrants further dedicated research of these features. Further work should include dedicated modeling of the (bio)optical properties of sea ice ridges, which would require ice core chl *a* biomass estimates from ridges and high spatial resolution spectral radiation measurements under ridges.

Spatial Variability and Patchiness of Sea Ice Properties, Algae Chl *a* Biomass, and NPP

Our results indicated high variability of patch sizes between locations, which suggests that there is large regional and temporal variability of ice algal chl *a* biomass. Patch sizes of algal chl *a* biomass were within the range of springtime chl *a* biomass patch sizes between 5 and 90 m (Gosselin et al., 1986; Rysgaard et al., 2001; Granskog et al., 2005; Søgaard et al., 2010). However, the upper limit of this range is nearly the scale of some ROV surveys. The above mentioned studies were limited to the spring and found that ice algal chl *a* biomass and NPP typically followed the light regime (Gosselin et al., 1986; Rysgaard et al., 2001; Granskog et al., 2005), which is primarily controlled by the overlying snow pack (Perovich, 1996). Furthermore, these studies were conducted on uniform, landfast sea ice from coastal regions and thus are not representative of sea ice from the central Arctic Ocean. During our study, we also found that patch sizes and spatial variability of NPP was controlled primarily by light availability, albeit in the absence of snow. This was evident by the high explained variance of NPP by bottom-ice light, the similarity of NPP and transmittance correlogram curves, and the coincidence of high NPP patches with high light transmittance. Similar to a recent study by Campbell et al. (2017) that demonstrated a disjoint in ice algal carbon biomass and NPP over the spring to summer transition period, our chl *a* biomass patches did not always follow the light and NPP regimes, which clearly illustrates one key difference between the spring and summer ice algal communities.

We also demonstrated that patches of high NPP were associated with patches of high chl *a* biomass in the absence of high light availability. The fact that both chl *a* and transmittance show spatial patterns consistent with NPP patterns is not surprising given the fact that NPP estimates were calculated from light and chl *a* biomass. However, this emphasizes the need to account for the spatial variability of both the bottom-ice light and chl *a* biomass to properly characterize the spatial variability of NPP in order to make accurate large-scale estimates. At a few stations (most notably 360), however, we did observe high chl *a* biomass patches directly adjacent to high transmittance locations (e.g., melt ponds). NPP was also high at the high transmittance locations and the adjacent high chl *a* biomass patches creating one high NPP patch. We propose that the presence of high chl *a* biomass adjacent to high transmittance regions could be explained by a combination of lower melt rates in the thicker ice adjacent to high transmittance regions and increased bottom-ice light levels due to horizontal light scattering from e.g., melt ponds. This would have allowed for higher NPP rates and increased accumulation of chl *a* biomass while having reduced melt-induced losses, however, we note that more work is needed to confirm this hypothesis.

Sea Ice Algae Sampling Recommendations

In this section we provide some recommendations for conducting the most representative sea ice algae sampling possible under the typical time limitation of an ice station on this cruise of ∼8 h. We assume that the dominant ice class (e.g., modal ice thickness) is known before sampling. Knowledge of the dominant ice class is important to ensure representative sampling; however, this depends on the objectives of the study. Knowledge of the spatial distribution for all ice types and classes will provide the best sampling protocol since a representative sample of each ice type/class will provide the most accurate and reliable estimates for the region.

Ice Core Chl *a* Biomass and NPP

A nested approach has been outlined in Miller et al. (2015), which identifies four hierarchical levels of sea ice sampling. We suggest, however, some modifications for sampling during summer. If the main objective of the study is to acquire one representative ice algal chl *a* biomass value for that ice station, then we suggest following the quaternary scale of the nested approach according to Miller et al. (2015) by extracting replicate cores ($N = 3$) within a small area (<2 m). This accounts for the small scale variability. The tertiary scale of the nested approach should be selected based on ice type/class. Here we have demonstrated that a representative chl *a* biomass value may be best estimated by the dominant ice class (NOTE: this should only be considered if additional sampling is not possible). Therefore, ice cores should be sampled in triplicate (quaternary scale) at three different dominant ice class locations (tertiary scale).

To capture the spatial variability of chl *a* biomass using ice coring alone, all ice classes should be considered. The nested approach should sample triplicate ice cores (quaternary scale) at 10 m intervals (tertiary scale) based on our observed patch sizes between 10 and 30 m. We further suggest that the replicates and direction of tertiary scale transects should be designed to capture all ice classes. A systematic approach would be to classify the sea ice using 0.5 m interval classes (as presented here). The sample design must also consider other ice types such as melt ponds, bare ice and thick ice features (e.g., ridges and hummocks). We must also note that the time requirements for conducting a spatial variability study using ice coring will be highly variable depending on season and ice conditions. For example, to quantify the spatial variability of thick MYI in early spring over a distance of 100 m (e.g., 3 cores at 11 sites = 30 cores) would take 30 h (based on previous experience coring spring MYI). This same task could be accomplished by an ROV with a typical deployment time of 8 h for two perpendicular survey transects of 100 m.

We demonstrated that NPP estimates have a complex relationship between light and chl *a* biomass. Therefore, in order to acquire a representative estimate the spatial variability of both the under ice light field and chl *a* biomass must be accounted for. We suggest a nested approach similar to that proposed for assessing the spatial variability of ice algal chl *a* biomass. Triplicate ice cores (quaternary scale) should be sampled at 10 m intervals (tertiary scale). In general, nested NPP sampling schemes should be conducted at the 5 different ice classes, as proposed earlier ($N = 15$).

ROV Chl *a* Biomass and NPP

ROV surveys should be conducted either over a grid or perpendicular transects with at least 60 m axis lengths in both directions for chl *a* biomass (two times maximum patch size of chl *a*) and at least 100 m for NPP estimates (two times maximum patch sizes for TM and NPP ∼50 m). This ensures you cross the boundary of the patch at least once. The survey should be chosen so that it covers these dimensions depending on the objectives of the study. However, the main criteria for setting transect/grid dimensions should be so that all 5 ice classes are surveyed (or all identified ice classes for the study site), of particular importance is the inclusion of unique and under-sampled sea ice features such as ridges or MYI hummocks (Lange et al., 2017). As we have shown, even over distances of 100 m the dominant ice classes and all ice features such as ridges were not representatively sampled. This is partly due to the location chosen and also due to filtering of the data resulting in inconsistent sample spacing-coverage. This in turn may over-sample some regions compared to others and result in non-representative surveys for certain ice classes. Therefore, we recommend that care is taken to choose ROV transects or grids that cover all ice classes, and to ensure ROV measurements are conducted while minimizing distance to the ice bottom, and pitch and roll angles.

During data analyses one should always consider the dominant ice class for the corresponding region based on larger scale ice thickness surveys. Because universal algorithms are not yet available for deriving chl *a* biomass from spectral radiation, ice cores should always be conducted at as many locations as possible along the ROV surveys for training bio-optical models, deriving photosynthetic parameters for up-scaling ROV NPP estimates and subsequently to parameterize algae models. Because the time requirements for ice coring (this does not include laboratory processing times) at the distances required for spatial variability studies (e.g., >100 m) are likely much greater than a typical ROV deployment of 8 h, we strongly recommend to conduct both ROV and ice core sampling particularly for spatial variability studies of both chl *a* biomass and NPP.

This method does have limitations in terms of assessing the temporality of ice algal chl *a* biomass and NPP due to the limited period of sampling and logistical constraints. This is a common drawback in observational sea ice biogeochemistry, which results from the limitations of sample processing and incubation times, and the shear difficulty of sampling within the Arctic Ocean in order to cover the necessary periods of weeks to months. However, our approach showed the successful application on spatially extensive datasets and thus is an ideal approach that should be applied to long-term studies (e.g., ice-tethered sensor arrays; Nicolaus et al., 2010) in order to assess the short- to long-term temporal variability of ice algal chl a biomass and NPP.

CONCLUSIONS

We provided, for the first time, a detailed multi-scale comparison of ice-core based ice algal chl *a* biomass and NPP estimates with estimates derived from under-ice spectral radiation measurements conducted over distances of tens to thousands of meters. These approaches demonstrated substantial improvements regarding representative sea ice algae observations. Our results showed that ice core-based estimates of summertime ice algal chl *a* biomass and NPP do not representatively capture the spatial variability compared to the spatially more extensive estimates of moving platforms. This may carry similar uncertainties, with an overall negative bias of ∼60%, for pan-Arctic estimates based on ice core observations alone.

Our autocorrelation analyses showed patch sizes of algal chl *a* biomass (10–30 m) and NPP (10–50 m) that were highly variable between locations and with scales of variability unlikely to be captured by ice coring alone. Based on our results we presented sampling recommendations depending on the objectives of the study. To estimate ice algal chl *a* biomass alone, taking a representative sample ($N = 3$) of each ice type/class using the ice core method should provide a reliable estimate of the overall area if there is also knowledge/observations of the ice thickness distribution on large scales (>1 km). Upscaling chl *a* biomass estimates would benefit from sampling all ice classes and factoring in weights for the spatial coverage of different ice classes in the region of interest. For NPP estimates, however, a combination of larger scale (>100 m) under ice light and ice algal chl *a* biomass is required because of the independent relationship between light and chl *a* biomass during the end of summer. In order to get the most representative estimates and to address the spatial variability of chl *a* biomass and NPP, we recommend that future sea ice sampling should combine ice-core based methods with the larger-scale under-ice spectral profiling approaches presented and described here and in Lange et al. (2016). This combined approach is also logistically justified since the time requirements for ice coring, which does not include processing times, at the distances required for spatial variability studies are typically much greater than a typical ROV deployment of 8 h.

We also identified high chl *a* biomass ridges within several up-scaled surveys, which have been generally neglected in sea ice biogeochemical studies. Sea ice ridges had significantly higher chl *a* biomass than the level ice and accounted for up to 10% of the total areal ice coverage. This suggests that these features may represent important regions for sea ice algal growth that are not easily captured by ice coring methods due to logistical difficulties of coring such thick sea ice. Further dedicated sea ice ridge studies are warranted particularly in terms of ice algal chl *a* biomass, nutrients, primary production and bio-optical properties.

AUTHOR CONTRIBUTIONS

This study was designed by BL, CK, MN, IP, and HF. Data acquisition were performed by BL, CK, MF-M, MN, IP, and HF. Data analyses were performed by BL, GC, CK, MF-M, MN, IP, and HF. Interpretation of the results were performed by BL, CK, GC, and HF. Drafting the first version of the manuscript was done by BL with critical revisions and important intellectual additions by all other authors (CK, MN, GC, IP, MF-M, and HF) during all stages of manuscript preparation. All authors give final approval for the publication of this manuscript in its current form.

ACKNOWLEDGMENTS

We thank Captain Uwe Pahl, the crew, and scientific cruise leader Antje Boetius of RV Polarstern expedition PS80.3 (ARK27-3; IceArc), for their excellent support and guidance with work at sea. We thank Martin Schiller for his technical expertise and operational support during ROV deployments. We thank Jan Andries van Franeker (IMARES) for kindly providing the SUIT and Michiel van Dorssen for technical support. SUIT was developed by IMARES with support from the Netherlands Ministry of EZ (project WOT-04-009-036) and the Netherlands Polar Program (project ALW 866.13.009). We acknowledge the collaboration and technical support by Ocean Modules, Sweden for development and deployment of the ROV. This study is part of the Helmholtz Association Young Investigators Group Iceflux: Ice-ecosystem carbon flux in polar oceans (VH-NG-800). We also acknowledge the Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar- und Meeresforschung for essential financial and logistical support. All data are available from the PANGAEA databases: doi: 10.1594/PANGAEA.833292; and doi: 10.1594/PANGAEA.834221.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2017.00349/full#supplementary-material

REFERENCES

Arrigo, K. R., and van Dijken, G. L. (2011). Secular trends in Arctic Ocean net primary production. *J. Geophys. Res.* 116, C09011. doi: 10.1029/2011JC007151

Arrigo, K. R., and van Dijken, G. L. (2015). Continued increases in Arctic Ocean primary production. *Prog. Oceanogr.* 136, 60–70. doi: 10.1016/j.pocean.2015.05.002

Boetius, A., Albrecht, S., Bakker, K., Bienhold, C., Felden, J., Fernández-Méndez, M., et al. (2013). Export of algal biomass from the melting Arctic Sea ice. *Science* 339, 1430–1432. doi: 10.1126/science.1231346

Budge, S. M., Wooller, M. J., Springer, A. M., Iverson, S. J., McRoy, C. P., and Divoky, G. J. (2008). Tracing carbon flow in an arctic marine food web using fatty acid-stable isotope analysis. *Oecologia* 157, 117–129. doi: 10.1007/s00442-008-1053-7

Campbell, K., Mundy, C. J., Barber, D. G., and Gosselin, M. (2015). Characterizing the sea ice algae chlorophyll a–snow depth relationship over Arctic spring melt using transmitted irradiance. *J. Mar. Syst.* 147, 76–84. doi: 10.1016/j.jmarsys.2014.01.008

Campbell, K., Mundy, C. J., Gosselin, M., Landy, J. C., Delaforge, A., and Rysgaard, S. (2017). Net community production in the bottom of first-year sea ice over the Arctic spring bloom. *Geophys. Res. Lett.* 44, 8971–8978. doi: 10.1002/2017GL074602

Campbell, K., Mundy, C. J., Landy, J. C., Delaforge, A., Michel, C., and Rysgaard, S. (2016). Community dynamics of bottom-ice algae in Dease Strait of the Canadian Arctic. *Prog. Oceanogr.* 149, 27–39. doi: 10.1016/j.pocean.2016.10.005

Castellani, G., Gerdes, R., Losch, M., and Lüpkes, C. (2015). "Impact of Sea-ice bottom topography on the ekman pumping," in *Towards an Interdisciplinary Approach in Earth System Science: Advances of a Helmholtz Graduate Research School*, eds G. Lohmann, H. Meggers, V. Unnithan, D. Wolf-Gladrow, J. Northolt, and A. Bracher (Cham: Springer International Publishing), 139–148. doi: 10.1007/978-3-319-13865-7_16

Castellani, G., Losch, M., Lange, B. A., and Flores, H. (2017). Modeling Arctic sea-ice algae: physical drivers of spatial distribution and algae phenology. *J. Geophys. Res.* 122, 7466–7487. doi: 10.1002/2017JC012828

Castellani, G., Lüpkes, C., Hendricks, S., and Gerdes, R. (2014). Variability of Arctic sea-ice topography and its impact on the atmospheric surface drag. *J. Geophys. Res.* 119, 6743–6762. doi: 10.1002/2013JC009712

David, C., Lange, B., Rabe, B., and Flores, H. (2015). Community structure of under-ice fauna in the Eurasian central Arctic Ocean in relation to environmental properties of sea-ice habitats. *Mar. Ecol. Prog. Ser.* 522, 15–32. doi: 10.3354/meps11156

Dupont, F. (2012). Impact of sea-ice biology on overall primary production in a biophysical model of the pan-Arctic Ocean. *J. Geophys. Res.* 117, C00D17. doi: 10.1029/2011JC006983

Eicken, H. (2001). Indirect measurements of the mass balance of summer Arctic sea ice with an electromagnetic induction technique. *Ann. Glaciol.* 33, 193–200. doi: 10.3189/172756401781818356

Fernández-Méndez, M., Katlein, C., Rabe, B., Nicolaus, M., Peeken, I., Bakker, K., et al. (2015). Photosynthetic production in the Central Arctic during the record sea-ice minimum in 2012. *Biogeosciences* 12, 2897–2945. doi: 10.5194/bgd-12-2897-2015

Fetterer, F., Knowles, K., Meier, W., and Savoie, M. (2002, updated 2011). *Sea Ice Index*. Boulder, CO: National Snow and Ice Data Center. Digital Media.

Gosselin, M., Legendre, L., Therriault, J. C., Demers, S., and Rochet, M. (1986). Physical control of the horizontal patchiness of sea-ice microalgae. *Mar. Ecol. Prog. Ser.* 29, 289–298. doi: 10.3354/meps029289

Gosselin, M., Levasseur, M., Wheeler, P. A., Horner, R. A., and Booth, B. C. (1997). New measurements of phytoplankton and ice algal production in the Arctic Ocean. *Deep Sea Res. Part II Top. Stud. Oceanogr.* 44, 1623–1644. doi: 10.1016/S0967-0645(97)00054-4

Gradinger, R. (1999). Vertical fine structure of the biomass and composition of algal communities in Arctic pack ice. *Mar. Biol.* 133, 745–754. doi: 10.1007/s002270050516

Gradinger, R. (2009). Sea-ice algae: major contributors to primary production and algal biomass in the Chukchi and Beaufort Seas during May/June 2002. . *Deep Sea Res. Part II Top. Stud. Oceanogr.* 56, 1201–1212. doi: 10.1016/j.dsr2.2008.10.016

Gradinger, R., Bluhm, B., and Iken, K. (2010). Arctic sea-ice ridges—Safe heavens for sea-ice fauna during periods of extreme ice melt? *Deep Sea Res. Part II Top. Stud. Oceanogr.* 57, 86–95. doi: 10.1016/j.dsr2.2009.08.008

Granskog, M. A., Kaartokallio, H., Kuosa, H., Thomas, D. N., Ehn, J., and Sonninen, E. (2005). Scales of horizontal patchiness in chlorophyll a, chemical and physical properties of landfast sea ice in the Gulf of Finland (Baltic Sea). *Polar Biol.* 28, 276–283. doi: 10.1007/s00300-004-0690-5

Grossi, S., Kottmeier, S., Moe, R., Taylor, G., and Sullivan, C. (1987). Sea ice microbial communities. 6. Growth and primary production in bottom ice under graded snow cover. *Mar. Ecol. Prog. Ser.* 35, 153–164. doi: 10.3354/meps035153

Haas, C. (2004). Late-summer sea ice thickness variability in the Arctic Transpolar Drift 1991–2001 derived from ground-based electromagnetic sounding. *Geophys. Res. Lett.* 31:L09402. doi: 10.1029/2003GL019394

Haas, C., and Eicken, H. (2001). Interannual variability of summer sea ice thickness in the Siberian and central Arctic under different atmospheric circulation regimes. *J. Geophys. Res.* 106, 4449–4462. doi: 10.1029/1999JC000088

Haas, C., Gerland, S., Eicken, H., and Miller, H. (1997). Comparison of sea-ice thickness measurements under summer and winter conditions in the Arctic using a small electromagnetic induction device. *Geophysics* 62, 749–757. doi: 10.1190/1.1444184

IPCC (2013). "Climate change 2013: the physical science basis," in *Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change,* eds T. F. Stocker, D. Qin, G.-K. Plattner, M. M. B. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley (Cambridge, UK; New York, NY: IPCC), 15–35.

Katlein, C., Arndt, S., Nicolaus, M., Perovich, D. K., Jakuba, M. V., Suman, S., et al. (2015a). Influence of ice thickness and surface properties on light transmission through Arctic sea ice. *J. Geophys. Res.* 120, 5932–5944. doi: 10.1002/2015JC010914

Katlein, C., Fernández-Méndez, M., Wenzhöfer, F., and Nicolaus, M. (2015b). Distribution of algal aggregates under summer sea ice in the Central Arctic. *Polar Biol.* 38, 719–731. doi: 10.1007/s00300-014-1634-3

Katlein, C., Nicolaus, M., and Petrich, C. (2014). The anisotropic scattering coefficient of sea ice. *J. Geophys. Res.* 119, 842–855. doi: 10.1002/2013JC009502

Katlein, C., Perovich, D. K., and Nicolaus, M. (2016). Geometric effects of an inhomogeneous Sea ice cover on the under ice light field. *Front. Earth Sci.* 4:6. doi: 10.3389/feart.2016.00006

Kohlbach, D., Graeve, M., Lange, B. A., David, C., Peeken, I., and Flores, H. (2016). The importance of ice algae-produced carbon in the central Arctic Ocean ecosystem: food web relationships revealed by lipid and stable isotope analyses. *Limnol. Oceanogr.* 61, 2027–2044. doi: 10.1002/lno.10351

Kohlbach, D., Schaafsma, F. L., Graeve, M., Lebreton, B., Lange, B. A., David, C., et al. (2017). Strong linkage of polar cod (*Boreogadus saida*) to sea ice algae-produced carbon: evidence from stomach content, fatty acid and stable isotope analyses. *Prog. Oceanogr.* 152, 62–74. doi: 10.1016/j.pocean.2017.02.003

Kuparinen, J., Kuosa, H., Andersson, A., Autio, R., Granskog, M. A., Ikävalko, J., et al. (2007). Role of sea-ice biota in nutrient and organic material cycles in the northern Baltic Sea. *AMBIO* 36, 149–154. doi: 10.1579/0044-7447(2007)36[149:ROSBIN]2.0.CO;2

Lange, B. A., Flores, H., Michel, C., Beckers, J. F., Bublitz, A., Casey, J. A., et al. (2017). Pan-Arctic sea ice-algal chl a biomass and suitable habitat are largely underestimated for multi-year ice. *Glob. Change Biol.* 23, 4581–4597. doi: 10.1111/gcb.13742

Lange, B. A., Katlein, C., Nicolaus, M., Peeken, I., and Flores, H. (2016). Sea ice algae chlorophyll a concentrations derived from under-ice spectral radiation profiling platforms. *J. Geophys. Res.* 121, 8511–8534. doi: 10.1002/2016JC011991

Lange, B. A., Michel, C., Beckers, J. F., Casey, J. A., Flores, H., Hatam, I., et al. (2015). Comparing springtime ice-algal chlorophyll a and physical properties of multi-year and first-year sea ice from the Lincoln Sea. *PLoS ONE* 10:e0122418. doi: 10.1371/journal.pone.0122418

Lavoie, D., Denman, K., and Michel, C. (2005). Modeling ice algal growth and decline in a seasonally ice-covered region of the Arctic (Resolute Passage, Canadian Archipelago). *J. Geophys. Res.* 110, C11009. doi: 10.1029/2005JC002922

Lee, Y. J., Matrai, P. A., Friedrichs, M. A. M., Saba, V. S., Antoine, D., Ardyna, M., et al. (2015). An assessment of phytoplankton primary productivity in the Arctic Ocean from satellite ocean color/*in situ* chlorophyll-a based models. *J. Geophys. Res.* 120, 6508–6541. doi: 10.1002/2015JC011018

Legendre, P., and Fortin, M.-J. (1989). Spatial Pattern and ecological analysis. *Vegetatio* 80, 107–138. doi: 10.1007/BF00048036

Legendre, P., and Legendre, L. (1998). "Spatial analysis," in *Numerical Ecology, 2nd Edn*., eds P. Legendre and L. Legendre (Amsterdam: Elsevier Science, BV), 853.

Long, M. H., Koopmans, D., Berg, P., Rysgaard, S., Glud, R. N., and Søgaard, D. H. (2012). Oxygen exchange and ice melt measured at the ice-water interface by eddy correlation. *Biogeosciences* 9, 1957–1967. doi: 10.5194/bg-9-1957-2012

Melnikov, I. (1997). *The Arctic Sea Ice Ecosystem*. Amsterdam: Gordon and Breach Science Publishers.

Melnikov, I. A., Kolosova, E. G., Welch, H. E., and Zhitina, L. S. (2002). Sea ice biological communities and nutrient dynamics in the Canada Basin of the Arctic Ocean. *Deep Sea Res. Part I Oceanogr. Res. Pap.* 49, 1623–1649. doi: 10.1016/S0967-0637(02)00042-0

Michel, C., Legendre, L., Ingram, R. G., Gosselin, M., and Levasseur, M. (1996). Carbon budget of sea-ice algae in spring: evidence of a significant transfer to zooplankton grazers. *J. Geophys. Res.* 101, 18345–18360.

Miller, L. A., Fripiat, F., Else, B. G., Bowman, J. S., Brown, K. A., Collins, E. R., et al. (2015). Methods for biogeochemical studies of sea ice: the state of the art, caveats, and recommendations. *Elementa* 3:000038. doi: 10.12952/journal.elementa.000038

Mock, T., and Gradinger, R. (1999). Determination of Arctic ice algal production with a new *in situ* incubation technique. *Mar. Ecol. Prog. Ser.* 177, 15–26. doi: 10.3354/meps177015

Moran, P. A. (1950). Notes on continuous stochastic phenomena. *Biometrika* 37, 17–23. doi: 10.1093/biomet/37.1-2.17

Mundy, C. J., Ehn, J. K., Barber, D. G., and Michel, C. (2007). Influence of snow cover and algae on the spectral dependence of transmitted irradiance through Arctic landfast first-year sea ice. *J. Geophys. Res.* 112, C03007. doi: 10.1029/2006JC003683

Mundy, C. J., Gosselin, M., Ehn, J. K., Belzile, C., Poulin, M., Alou, E., et al. (2011). Characteristics of two distinct high-light acclimated algal communities during advanced stages of sea ice melt. *Polar Biol.* 34, 1869–1886. doi: 10.1007/s00300-011-0998-x

Nicolaus, M., and Katlein, C. (2013). Mapping radiation transfer through sea ice using a remotely operated vehicle (ROV). *Cryosphere* 7, 763–777. doi: 10.5194/tc-7-763-2013

Nicolaus, M., Hudson, S. R., Gerland, S., and Munderloh, K. (2010). A modern concept for autonomous and continuous measurements of spectral albedo and transmittance of sea ice. *Cold Reg. Sci. Technol.* 62, 14–28. doi: 10.1016/j.coldregions.2010.03.001

Nicolaus, M., Katlein, C., Maslanik, J., and Hendricks, S. (2012). Changes in Arctic sea ice result in increasing light transmittance and absorption. *Geophys. Res. Lett.* 39, 2699–2700. doi: 10.1029/2012GL053738

Perovich, D. K. (1996). *The Optical Properties of Sea Ice.* Cold Regions Research and Engineering Laboratory.

Perovich, D. K., Grenfell, T. C., Richter-Menge, J. A., Light, B., Tucker Iii, W. B., and Eicken, H. (2003). Thin and thinner: sea ice mass balance measurements during SHEBA. *J. Geophys. Res.* 108:8050. doi: 10.1029/2001JC001079

Perovich, D. K., Roesler, C. S., and Pegau, W. S. (1998). Variability in Arctic sea ice optical properties. *J. Geophys. Res.* 103, 1193–1208. doi: 10.1029/97JC01614

Platt, T., Gallegos, C., and Harrison, W. (1980). Photoinhibition of photosynthesis in natural assemblages of marine phytoplankton. *J. Mar. Res.* 38, 103–111.

Rabenstein, L., Hendricks, S., Martin, T., Pfaffhuber, A., and Haas, C. (2010). Thickness and surface-properties of different sea-ice regimes within the Arctic Trans Polar Drift: data from summers 2001, 2004 and 2007. *J. Geophys. Res.* 115:C12059. doi: 10.1029/2009JC005846

R-Development-Core-Team (2012). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Rysgaard, S., Kühl, M., Glud, R. N., and Hansen, J. W. (2001). Biomass, production and horizontal patchiness of sea ice algae in a high-Arctic fjord (Young Sound, NE Greenland). *Mar. Ecol. Prog. Ser.* 223, 15–26. doi: 10.3354/meps223015

Schünemann, H., and Werner, I. (2005). Seasonal variations in distribution patterns of sympagic meiofauna in Arctic pack ice. *Mar. Biol.* 146, 1091–1102. doi: 10.1007/s00227-004-1511-7

Søgaard, D. H., Kristensen, M., Rysgaard, S., Glud, R. N., Hansen, P. J., and Hilligsøe, K. M. (2010). Autotrophic and heterotrophic activity in Arctic first-year sea ice: seasonal study from Malene Bight, SW Greenland. *Mar. Ecol. Prog. Ser* 419, 31–45. doi: 10.3354/meps08845

Søreide, J. E., Leu, E. V. A., Berge, J., Graeve, M., and Falk-Petersen, S. (2010). Timing of blooms, algal food quality and *Calanus glacialis* reproduction and growth in a changing Arctic. *Glob. Change Biol.* 16, 3154–3163. doi: 10.1111/j.1365-2486.2010.02175.x

Spreen, G., Kaleschke, L., and Heygster, G. (2008). Sea ice remote sensing using AMSR-E 89-GHz channels. *J. Geophys. Res.* 113, C02S03. doi: 10.1029/2005JC003384

Steffens, M., Granskog, M., Kaartokallio, H., Kuosa, H., Luodekari, K., Papadimitriou, S., et al. (2006). Spatial variation of biogeochemical properties of landfast sea ice in the Gulf of Bothnia, Baltic Sea. *Ann. Glaciol.* 44, 80–87. doi: 10.3189/172756406781811169

Tremblay, J.-É., Anderson, L. G., Matrai, P., Coupel, P., Bélanger, S., Michel, C., et al. (2015). Global and regional drivers of nutrient supply, primary production and CO2 drawdown in the changing Arctic Ocean. *Prog. Oceanogr.* 139, 171–196. doi: 10.1016/j.pocean.2015.08.009

van Franeker, J. A., Flores, H., and Van Dorssen, M. (2009). *The Surface and Under Ice Trawl (SUIT). Frozen Desert Alive—The Role of Sea Ice for Pelagic Macrofauna and Its Predators*. Ph.D. thesis, University of Groningen, Groningen.

Wang, S. W., Budge, S. M., Iken, K., Gradinger, R. R., Springer, A. M., and Wooller, M. J. (2015). Importance of sympagic production to Bering Sea zooplankton as revealed from fatty acid-carbon stable isotope analyses. *Mar. Ecol. Prog. Ser.* 518, 31–50. doi: 10.3354/meps11076

Wassmann, P., Duarte, C. M., Agusti, S., and Sejr, M. K. (2011). Footprints of climate change in the Arctic marine ecosystem. *Glob. Change Biol.* 17, 1235–1249. doi: 10.1111/j.1365-2486.2010.02311.x

Wilcoxon, F. (1945). Individual Comparisons by Ranking Methods. *Biometr. Bull.* 1, 80–83. doi: 10.2307/3001968

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer TJ and handling Editor declared their shared affiliation.

Copyright © 2017 Lange, Katlein, Castellani, Fernández-Méndez, Nicolaus, Peeken and Flores. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Marine Science

ORIGINAL RESEARCH
published: 05 December 2017
doi: 10.3389/fmars.2017.00387

Primary Production: Sensitivity to Surface Irradiance and Implications for Archiving Data

Trevor Platt[1], *Shubha Sathyendranath*[2]*, *George N. White III*[3], *Thomas Jackson*[1], *Stéphane Saux Picart*[4] *and Heather Bouman*[5]

[1] *Plymouth Marine Laboratory, Plymouth, United Kingdom,* [2] *Plymouth Marine Laboratory, National Centre for Earth Observation, Plymouth, United Kingdom,* [3] *Bedford Institute of Oceanography, Dartmouth, NS, Canada,* [4] *Centre de Météorologie Spatiale, Météo-France, Lannion, France,* [5] *Department of Earth Sciences, Oxford University, Oxford, United Kingdom*

An equation is derived to express the sensitivity of daily, watercolumn production by phytoplankton in the ocean to variations in irradiance at the sea surface. Assuming no spectral effects, and a vertically uniform chlorophyll profile, the sensitivity is a function only of the dimensionless irradiance. Spectral effects can be accounted for as a function of the chlorophyll concentration. At the global scale, the relative reduction in daily production consequent on halving the surface irradiance (representing the expected scope for variation in surface irradiance under natural conditions) is found to be from 30 to 40%. Choice of data source for irradiance may incur a further systematic error of up to 15%. Given that local irradiance (the principal forcing for primary production) may vary from day to day, the issue of how to archive production data for the most generality is discussed and recommendations made in this regard.

Keywords: primary production, data archiving, sensitivity analysis, irradiance, photophysiology, photosynthesis measurements

OPEN ACCESS

Edited by:
Chris Bowler,
École Normale Supérieure, France

Reviewed by:
Oliver Zielinski,
University of Oldenburg, Germany
Thomas Lacour,
UMI3376 TAKUVIK, Canada

***Correspondence:**
Shubha Sathyendranath
ssat@pml.ac.uk

Specialty section:
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

Received: *21 March 2017*
Accepted: *16 November 2017*
Published: *05 December 2017*

Citation:
Platt T, Sathyendranath S, White GN III, Jackson T, Saux Picart S and Bouman H (2017) Primary Production: Sensitivity to Surface Irradiance and Implications for Archiving Data. Front. Mar. Sci. 4:387. doi: 10.3389/fmars.2017.00387

1. INTRODUCTION

Conventionally, there are two approaches to the measurement of phytoplankton production at sea, both of which require incubation of phytoplankton samples for a finite time (**Table 1**). One is the so-called *in situ* method (or its variant, the simulated *in situ* method) (Lohrenz et al., 1992; Lohrenz, 1993). The object here is to produce data representing the vertical distribution of phytoplankton production through the photic zone. The irradiance that drives the photosynthesis is solar irradiance, attenuated by the sea itself (or by other filters in the case of the simulated *in situ* method). The estimated vertical profile of phytoplankton photosynthesis can be integrated over depth to calculate production in the water column during the period of the incubation. If the duration of incubation is less than that of the light day, the data may be extrapolated to obtain daily water-column production. The result expresses the daily, autotrophic carbon flux under unit area of sea surface at the place and time from which the sample was drawn, under all prevailing conditions, physical and biological.

The other method is through construction of photosynthesis–light curves. Here, the samples are incubated in artificial light at a sufficient number of irradiance levels that the curve of photosynthetic response can be established, fitted to a standard equation, and the parameters (minimum two for a normal range of irradiances) extracted (Platt and Jassby, 1976).

TABLE 1 | Comparison of two approaches (*in situ* method and photosynthesis-irradiance curve) to measuring primary production of aquatic samples.

	***In situ* method**	**Photosynthesis–irradiance curve**
Incubation time	Longer	Shorter
Demands on ship's programme	Severe	Minimal
Light source?	Natural sunlight	Artificial & constant
What is the output?	Watercolumn production	Photosynthesis parameters
Results valid for what irradiance?	Irradiance at time of sampling	All irradiances
Utility of results for Modeling?	Limited to model validation	Broader, key model parameter
Generality of results?	Less general (local)	More general
How to archive results?	Straightforward	Less straightforward

These parameters index the photosynthetic performance of the phytoplankton present at the time and place from which the sample was drawn. The results can be applied in mathematical models to calculate daily watercolumn production (Platt et al., 1977; Sathyendranath and Platt, 1989; Sathyendranath et al., 1989; Morel, 1991; Morel et al., 1996). This serves the same purpose as *in situ* measurements of primary production, but calculations using the photosynthesis parameters can be tailored for any reasonable time interval, to suit any application. Light-dependent models of primary production are used in remote sensing (Platt and Sathyendranath, 1988; Mélin and Hoepffner, 2004, 2011) and in large-scale simulation models of the marine ecosystem (Laufkötter et al., 2015). The photosynthesis parameters are fundamental bio-optical properties of phytoplankton and have many other applications beyond estimation of water-column production, for example in the calculation of chlorophyll-to-carbon ratio in phytoplankton (Jackson et al., 2017).

The photosynthesis parameters may therefore be considered to contain more information than estimates of *in situ* production (which they subsume through application of a mathematical model). In the planning of present-day oceanographic expeditions on which phytoplankton production estimates are required, a choice is usually made in favor of the photosynthesis–parameter method (which also imposes fewer restrictions on the movement of the ship than does the *in situ* method).

The question then arises: How should these data be archived? In the case of *in situ* production estimates, the picture is clear, except possibly for the time scale involved. The data represent the photosynthetic carbon flux per unit area of the sea surface for the duration of the incubation, perhaps extrapolated to a time scale of 1 day, and can be archived as such. For the (preferred) photosynthesis–response method, the parameters can be archived but they do not in themselves constitute an estimate of phytoplankton production. If they are to be applied to the calculation of daily water-column production, what should be taken as the forcing irradiance? Does it matter? Here, we analyse the sensitivity of daily, water-column production to variable surface irradiance, and discuss the archiving of data on primary production, for example as used in climate studies.

2. THEORETICAL BACKGROUND: NON-SPECTRAL-LIGHT, UNIFORM-BIOMASS CASE

The photosynthesis–irradiance curve relates phytoplankton photosynthesis P to irradiance I. It is convenient, and more general, to work with the normalized photosynthetic rate P^B, where the superscript B indicates normalization to the chlorophyll biomass. Thus,

$$P^B(I) = p^B(I; \text{parameters}), \tag{1}$$

where p^B is a function to be specified and where various choices are available for the parameter sets. We have shown that, regardless of the parameter set chosen and for any plausible choice of p^B, the various manifestations of Equation (1) can all be recast into a single common form (Platt and Sathyendranath, 1993). The discussion to follow is thus robust against alternative choices of p^B and the parameters. Therefore, with no loss of generality, we shall use as parameters the assimilation number P^B_m and the initial slope α^B of the photosynthesis–irradiance curve. Then,

$$P^B(I) = p^B(I; \alpha^B, P^B_m). \tag{2}$$

In the sea, irradiance is a function of depth z and of time of day t, such that we can write

$$P^B(I) = p^B\big(I(z,t); \alpha^B, P^B_m\big). \tag{3}$$

The desired result is the double integral of Equation (3) over time and depth, which is the daily watercolumn production $P_{Z,T}$.

$$P_{Z,T} = B \int_0^D \int_0^\infty p^B\big(I(z,t); \alpha^B, P^B_m\big)\, \mathrm{d}z\, \mathrm{d}t, \tag{4}$$

where D is the length of day in hours from sunrise and where it is assumed that there is no production outside the interval $0 \leq t \leq D$. The factor B is taken outside the integral signs because we assume that the chlorophyll biomass is independent of depth. The choice of infinity as the upper limit on the integral over z avoids any ambiguity over depth of the photic zone, without invoking

any simplifying assumption. Contributions to the integral from all depths below the photic zone are negligible for all practical purposes.

Under clear-sky conditions, the irradiance at the sea surface $I(0,t)$ can be calculated for any latitude and date according to standard astronomical functions (Bird, 1984; Bird and Riordan, 1986; Sathyendranath and Platt, 1988; Platt et al., 1990). Analytically, the time course of clear-sky irradiance $I(t)$ can be described by a function with two parameters (D and the irradiance at local noon I_0^m). Thus, $I(0,t) = I_0^m g(t;D)$, where $g(t)$ is a function to be specified. The effect of clouds may be represented through a reduction in the magnitude of I_0^m according to the proportion of the sky covered by clouds.

Given observations of α^B and P_m^B (and a choice of p^B) we can evaluate the double integral of Equation (4) for any station and time. The result will depend upon the magnitude of I_0^m. For archival purposes, is it appropriate to use the clear-sky value of I_0^m (which would return the maximum possible value of production under the prevailing conditions at that time and place)? Or is it more appropriate to use the value of I_0^m corresponding to the sky conditions at the station and time (which would return the best value of production for the conditions, but which would lack any generality for other sky conditions)? Would the differences be significant?

To address these issues, we need to look at the derivative of Equation (4) with respect to irradiance. Let us first select $p^B(I) = P_m^B\big(1 - \exp(-I/I_k)\big)$ as the formulation of the photosynthesis–light curve, where the photoadaptation parameter $I_k = P_m^B/\alpha^B$. The ratio I/I_k is a dimensionless irradiance that we designate as I_*. The dimensionless noon irradiance is I_*^m. We assume a vertically-homogeneous water column so that $B(z) = B$, and light attenuation can be characterized by a single coefficient K such that $I(z) = I(0)\exp(-Kz)$ where the sea surface is at $z = 0$ and z is positive downwards. Then Equation 4 may be written as

$$P_{Z,T} = BP_m^B \int_0^D \int_0^\infty \Big(1 - \exp\big(-I_*^m g(t)\exp(-Kz)\big)\Big)\,\mathrm{d}z\,\mathrm{d}t. \quad (5)$$

With a change of variable $x = I_*^m g(t)\exp(-Kz)$, we have

$$P_{Z,T} = \frac{BP_m^B}{K} \int_0^D \int_0^{I_*^m g(t)} \left(\frac{1-\exp(-x)}{x}\right) \mathrm{d}x\,\mathrm{d}t. \quad (6)$$

The inner integral (on x) is a standard form, the entire exponential integral Ein $\big(I_*^m g(t)\big)$, so that

$$P_{Z,T} = \frac{BP_m^B}{K} \int_0^D \mathrm{Ein}\left(I_*^m g(t)\right) \mathrm{d}t. \quad (7)$$

In the specification of the surface irradiance, the dimensionless irradiance at noon I_*^m can be considered as a scale factor, with the function $g(t)$ describing the time course of variation through the day. Our immediate interest is on the derivative of $P_{Z,T}$ with respect to this scale factor I_*^m,

$$\frac{\mathrm{d}P_{Z,T}}{\mathrm{d}I_*^m} = \frac{BP_m^B}{K} \int_0^D \frac{\mathrm{d}}{\mathrm{d}I_*^m} \mathrm{Ein}\left(I_*^m g(t)\right) \mathrm{d}t. \quad (8)$$

By virtue of the definition of the function Ein (.) as an integral, the differentiation returns the integrand of the definition:

$$\frac{\mathrm{d}P_{Z,T}}{\mathrm{d}I_*^m} = \frac{BP_m^B}{K} \int_0^D g(t) \left(\frac{1-\exp(-I_*^m g(t))}{I_*^m g(t)}\right) \mathrm{d}t. \quad (9)$$

The functions $g(t)$ in numerator and denominator cancel to give the simple result

$$\frac{\mathrm{d}P_{Z,T}}{\mathrm{d}I_*^m} = \frac{BP_m^B}{KI_*^m} \int_0^D \Big(1 - \exp\big(-I_*^m g(t)\big)\Big) \mathrm{d}t. \quad (10)$$

The choice of function $g(t) = 1 - (1 - 2t/D)^2$ gives a good representation of the time course of surface irradiance through the day. With this choice, and with the substitution $\theta = \pi t/D$, Equation 10 becomes

$$\frac{\mathrm{d}P_{Z,T}}{\mathrm{d}I_*^m} = \frac{BP_m^B}{KI_*^m} \frac{D}{\pi} \int_0^\pi \Big(1 - \exp\big(-I_*^m[1-(1-2\theta/\pi)^2]\big)\Big) \mathrm{d}\theta. \quad (11)$$

Given the symmetry of $g(\theta)$ about $\theta = \pi/2$, we can express Equation (11) as

$$\frac{\mathrm{d}P_{Z,T}}{\mathrm{d}I_*^m} = \frac{BP_m^B}{KI_*^m} \frac{D}{\pi} \Big\{\pi - 2\int_0^{\pi/2} \exp\big(-I_*^m[1-(1-2\theta/\pi)^2]\big)\,\mathrm{d}\theta\Big\}. \quad (12)$$

At this point it is convenient to make the substitution $s = \sqrt{I_*^m}(1 - 2\theta/\pi)$. Then,

$$\frac{\mathrm{d}P_{Z,T}}{\mathrm{d}I_*^m} = \frac{BP_m^B}{KI_*^m} \frac{D}{\pi} \Big\{\pi - \frac{\pi e^{-I_*^m}}{\sqrt{I_*^m}} \int_0^{\sqrt{I_*^m}} e^{s^2}\,\mathrm{d}s\Big\}. \quad (13)$$

Using the definition of Dawson's integral:

$$\mathrm{Daw}(x) = e^{-x^2} \int_0^x e^{s^2}\,\mathrm{d}s \quad (14)$$

with $x = \sqrt{I_*^m}$, we see that Equation (13) can be written as

$$\frac{\mathrm{d}P_{Z,T}}{\mathrm{d}I_*^m} = \frac{BP_m^B D}{KI_*^m} \Big\{1 - \frac{1}{\sqrt{I_*^m}} \mathrm{Daw}\big(\sqrt{I_*^m}\big)\Big\}. \quad (15)$$

This form is convenient because there exist robust numerical forms for computing Daw(x).

We want to assess the sensitivity of the daily production estimates to changes (errors) in the surface irradiance I_0^m. Noting that $\mathrm{d}I_0^m = I_k\,\mathrm{d}I_*^m$, we can rewrite Equation (15) as

$$\frac{\mathrm{d}P_{Z,T}}{\mathrm{d}I_0^m} = \frac{BP_m^B D}{KI_0^m} \Big\{1 - \frac{1}{\sqrt{I_*^m}} \mathrm{Daw}\big(\sqrt{I_*^m}\big)\Big\}. \quad (16)$$

It was shown in Platt and Sathyendranath (1993) that all solutions to Equation (4) can be written in the canonical form

$$P_{Z,T} = Af(I_*^m), \quad (17)$$

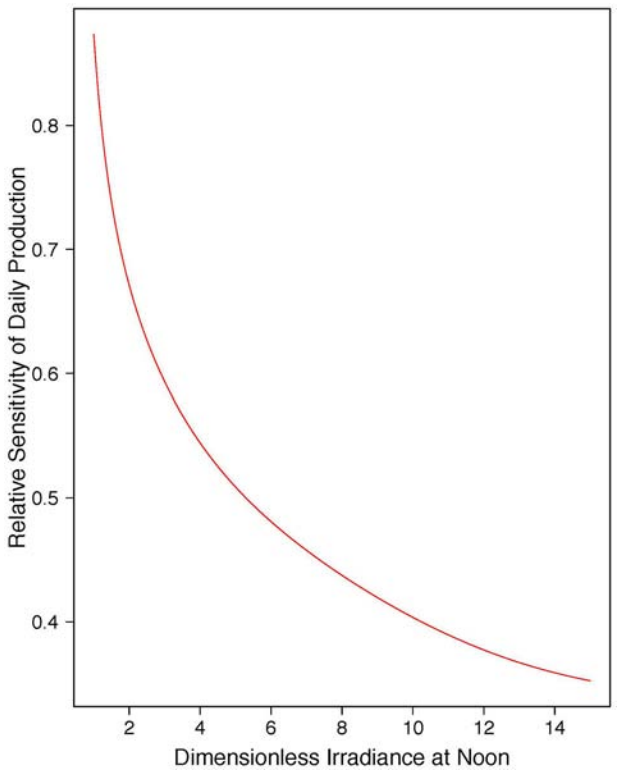


FIGURE 1 | Relative sensitivity of daily, watercolumn primary production ($P_{Z,T}$) to changes in dimensionless irradiance at noon (I_*^m), non-spectral solution, Equation (18).

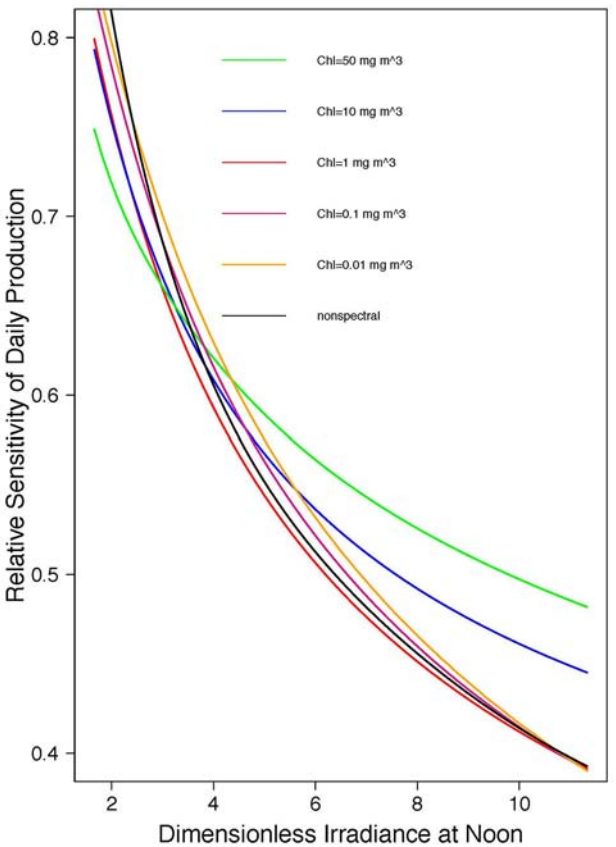


FIGURE 2 | Relative sensitivity of daily, watercolumn primary production ($P_{Z,T}$) to changes in dimensionless irradiance at noon (I_*^m), spectral results by numerical integration for different chlorophyll concentrations. Non-spectral result also shown, for comparison.

where $A = BP_m^B D/K$ is a scale factor with dimensions of daily production, and $f(I_*^m)$ is a (known) function of the dimensionless irradiance whose particular form depends on the choice of the photosynthesis-light curve. Dividing both sides of Equation (16) by $Af(I_*^m)$, we obtain

$$\left(\frac{\mathrm{d}P_{Z,T}}{P_{Z,T}}\right)\Big/\left(\frac{\mathrm{d}I_0^m}{I_0^m}\right) = \frac{1}{f(I_*^m)}\left\{1 - \frac{1}{\sqrt{I_*^m}}\,\mathrm{Daw}(\sqrt{I_*^m})\right\}. \quad (18)$$

Here, we can see that the relative change in daily primary production for a given relative change in surface irradiance depends only on a function of the scaled irradiance.

3. ESTIMATING DAILY, WATER-COLUMN PRODUCTION

In the previous section, we found an expression for the sensitivity of primary production to variations in surface irradiance, of an idealized water column chosen for mathematical tractability. We next consider the detailed calculation of primary production in operational mode and the sensitivity of the results to changes in the surface irradiance.

Various procedures exist for estimating daily, water-column production by phytoplankton, given information on the two photosynthesis parameters P_m^B and α^B (Platt et al., 1977, 1990, 2008; Platt and Sathyendranath, 1988, 1993). All require an estimate of surface irradiance. Briefly, the differences among them depend mainly on whether a depth-independent biomass profile is assumed and on whether the (known) wavelength dependence of light penetration and photosynthetic response is included or suppressed. A depth-independent, spectrally-neutral treatment provides a frame of reference (Sathyendranath et al., 1989; Platt and Sathyendranath, 1991); it has the canonical solution given in Equation (17).

At the other extreme, a wavelength-dependent and non-uniform biomass treatment has no analytic solution (although Equation 17 remains a reliable guide) and must be integrated numerically (Sathyendranath et al., 1989; Platt and Sathyendranath, 1991). This is the preferred calculation, and its reliability has been demonstrated (Platt and Sathyendranath, 1988). Applications of the various models have been compared (Platt et al., 1991; Kyewalyanga et al., 1992). In the most detailed models, the angular distribution of the irradiance is included (Platt and Sathyendranath, 1988; Sathyendranath and Platt, 1989; Sathyendranath et al., 1989), the light field being separated into its direct and diffuse components. The direct component is

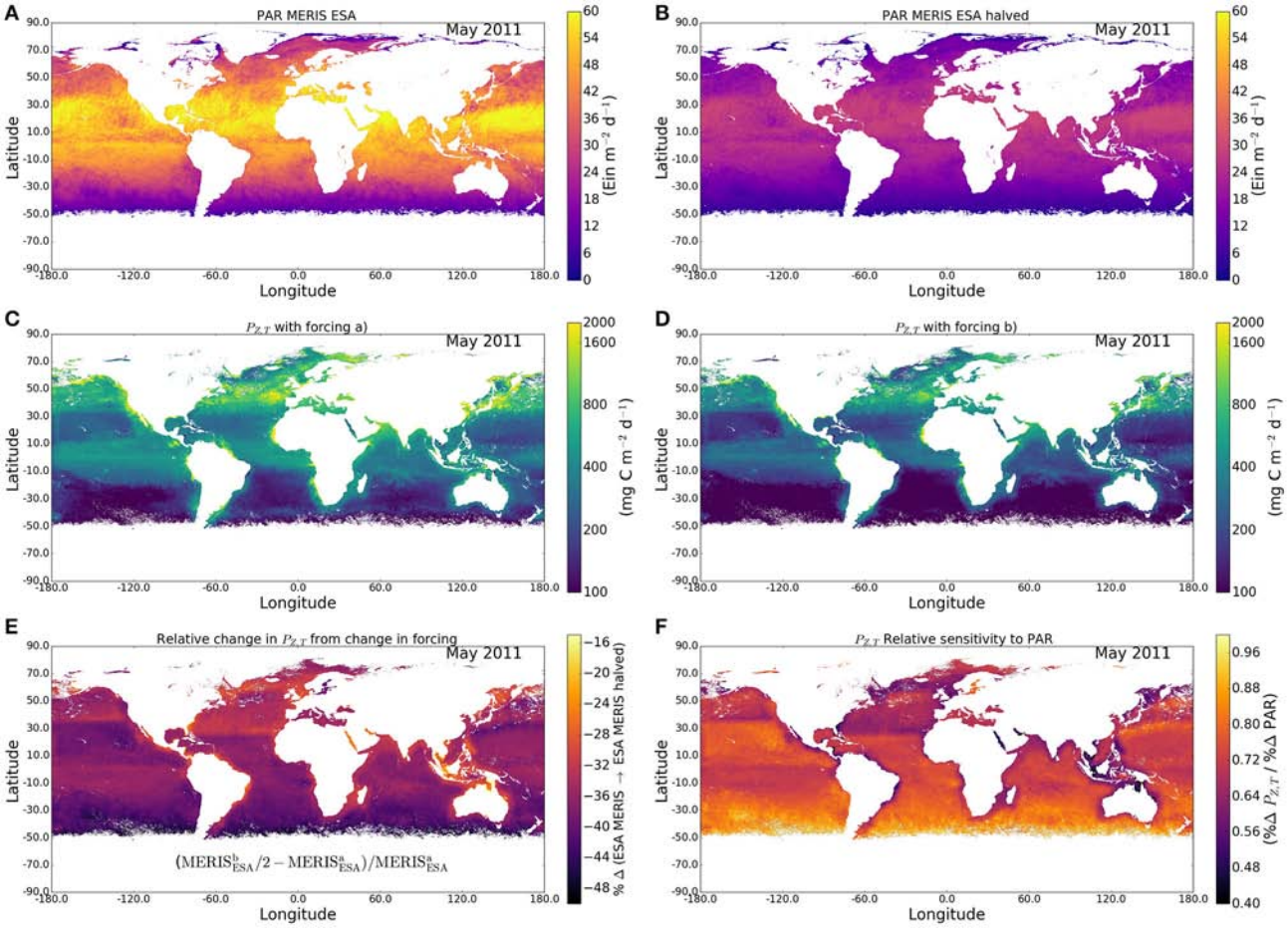


FIGURE 3 | Effect on daily, watercolumn production of a 50% reduction in surface irradiance for May 2011. **(A)** Surface PAR field before reduction; **(B)** Surface PAR field after reduction; **(C)** daily production before reduction of surface light; **(D)** daily production after reduction of surface light; **(E)** relative change in primary production following reduction of surface light; and **(F)** relative change in primary production following reduction of surface light, normalized to the relative change in surface light.

affected much more strongly by the presence of clouds, and the diffuse component (which may be one half of the total) will still be available for photosynthesis even under 100% cloud cover. For this reason, separation of the forcing irradiance into direct and diffuse components is a key step in addressing the sensitivity of water column production to variations in cloud cover.

For the calculations presented here, we use a spectral- and angular-distribution-resolving primary-production model (Platt and Sathyendranath, 1988; Longhurst et al., 1995; Platt et al., 1995; Sathyendranath et al., 1995), forced by remotely-sensed chlorophyll and light data, and information derived from ship-based *in situ* measurements on phytoplankton physiology and vertical biomass profile parameters. These model parameters were organized according to season and ecological province, as in Longhurst et al. (1995). Photophysiological parameter (P_B^m and α^B) values for defined biogeochemical provinces were taken from the work of Mélin and Hoepffner (2004). The vertical structure of the phytoplankton biomass profile was described by three parameters: the depth of maximum chlorophyll concentration (Z_m), the thickness of the subsurface peak in chlorophyll concentration (σ) and the ratio of the peak chlorophyll concentration to the background chlorophyll concentration (ρ), also provided on a provincial basis following Mélin and Hoepffner (2011). To provide a smooth transition of parameter estimates across province boundaries, a smoothing filter was applied to the province values of Mélin and Hoepffner (2004) where values at a given pixel were averaged from a 30 × 30 pixel box, with a pixel size of 9 km, centered on the pixel of interest.

The chlorophyll profile parameters were used in conjunction with the remotely-sensed chlorophyll data from the OC-CCI v1.0 products (Sathyendranath et al., 2016), to create chlorophyll profiles for each 9 km pixel. Average sea-surface irradiance (Photosynthetically-Active Radiation) for May 2011 was taken from three data sources. The first PAR data source was the NASA MODIS PAR product (OBPG, 2014b); the second data source was the NASA MERIS PAR product (OBPG, 2014a); the third source was a new ESA processing of MERIS as documented in the Algorithm Theoretical Baseline Document (ATBD) for the PAR and Primary Production (PPP) SEOM project.

In each of the three cases, the remotely-sensed PAR was used to scale the results of a spectral clear-sky model to provide spectrally-resolved irradiance for input to the primary production model. The propagation of spectral light to various depths in the water column accounted for attenuation by water, phytoplankton and other colored and scattering substances, assuming all optical properties were related to chlorophyll concentration (open-ocean conditions). The profile of light was then combined with the vertical profile of chlorophyll and photosynthetic parameters to obtain estimates of depth-resolved primary production. The calculations were repeated for 12 time steps during half the day length. The results were then integrated over time and depth to yield total primary production per day and per unit area.

4. SENSITIVITY OF $P_{Z,T}$ TO VARIATIONS IN SURFACE IRRADIANCE

For the non-spectral approximation, with vertically-uniform biomass profile, the sensitivity of daily, water-column production is given by Equation (18), shown in **Figure 1**. The spectral analog, calculated by numerical integration for the case of uniform biomass, is shown in **Figure 2**. The spectral results lie in the same range as that for the non-spectral case, varying slightly according to the chlorophyll concentration. The differences are greater for higher values of chlorophyll and higher values of normalized irradiance. We see that, in both spectral and non-spectral calculations, sensitivity is high and approaches one (corresponding to the case when relative change in primary production is the same as the relative change in surface irradiance) when the dimensionless noon-time irradiance approaches zero. As the scaled noon-time irradiance increases, the sensitivity decreases, reaching values less than 0.5 for I_*^m greater than 10. For the spectral model, there is an additional dependence of the answer on chlorophyll concentration: the sensitivity is less than that of the non-spectral case at low I_*^m, but the opposite holds for high I_*^m values. The additional dependence of the results on chlorophyll concentration arises from the effect of phytoplankton absorption on the spectral quality of the underwater light field, which is not taken into account in the non-spectral model. When phytoplankton absorption increases (high-chlorophyll conditions), the water turns progressively green, and less suitable for phytoplankton absorption, introducing a decrease in light available for photosynthesis. The effect on primary production is equivalent to decreasing I_*^m in a non-spectral model. For chlorophyll concentration less than 1 mg m^{-3}, which would be typical of most open-ocean waters, the spectral and non-spectral models yield results that are quite close to each other, such that the analytical solution may be taken to be a reasonable guide to realistic open-ocean conditions.

In global-scale computations, in addition to the spectral effects, we also have to explore the effect of vertical structure in chlorophyll concentration. We have estimated the global-scale effect of variations in surface irradiance as follows. We selected the month of May, 2011 to illustrate the results. First, we calculated primary production using a detailed, spectral model integrated numerically (**Figure 3C**), forced by the typical irradiance at each pixel for the month concerned (**Figure 3A**), using a spectral model, and with the photosynthesis parameters and chlorophyll profile parameters assigned according to provinces and season, as noted earlier. Then, we repeated the calculation (**Figure 3B**), but with irradiance only one half of that in the previous calculation (**Figure 3D**). The difference between values of $P_{Z,T}$ calculated for the two irradiances is a measure of the sensitivity of the daily, water-column production to changes in surface light (**Figure 3E**) and the relative sensitivity (change in production normalized to change in surface irradiance) is shown in **Figure 3F**.

The results, which vary with region (especially with latitude) are not symmetrical about the equator: they represent Spring in the northern hemisphere, but Winter in the southern hemisphere, with corresponding changes to the surface light field (**Figure 3A**). The results also reflect the regional assignment of photosynthesis parameters. The relative reduction in daily, water-column production lies generally in the range from 30 to 40%, consequent on halving the surface irradiance (**Figure 3E**).The sensitivity of daily production to changes in surface irradiance is typically in the range from 60 to 80% (**Figures 3F**, **4**), showing that the relative change in primary production is almost always less than the relative change in incoming solar radiation. The relative effect is stronger in areas with low surface irradiance, consistent with the diagnosis of the analytical solution.

Systematic errors may be introduced into the calculation of primary production through the choice of the surface field of photosynthetically-active radiation (PAR). We have examined this possibility by comparing the results arising from use of

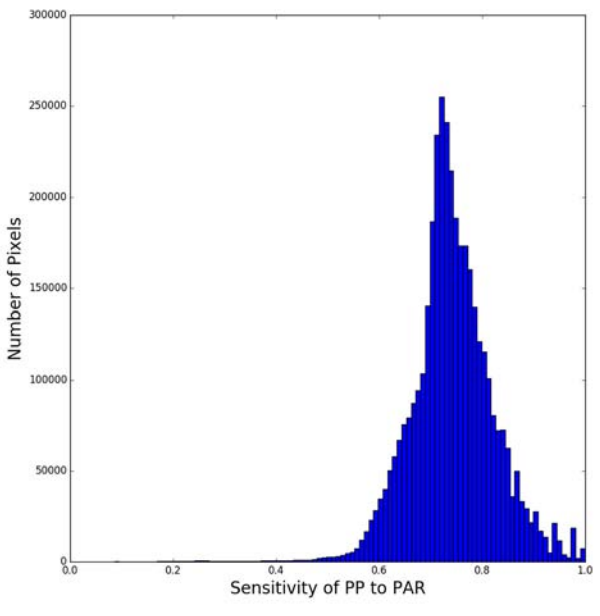


FIGURE 4 | Histogram of relative change in daily production for a 50% reduction in surface irradiance on all sea pixels in **Figure 3E**.

MERIS (**Figures 5A,C**) and MODIS (**Figure 5E**) fields of PAR (**Figures 5B,D,F**). We find that the differences in surface PAR between the two data streams are usually in the range from 5 to 20% (**Figure 5E**), implying a potential systematic error in primary production of some 15% or less arising from the choice of PAR field. Though these differences may be small, they are systematic; their effect could be significant in studies associated with climate change: when results from multiple sensors are merged to create a long time series, differences between PAR data from individual sensors may lead to an erroneous conclusion on trends in primary production, which is to be avoided.

5. DISCUSSION

In this paper, we have examined the sensitivity of modeled primary production to changes in the solar radiation available at the sea surface. Using an analytical solution derived for a non-spectral model assuming uniform chlorophyll concentration in the water column, using a spectral model with uniform chlorophyll concentration, and finally using a spectral model with non-uniform chlorophyll profile, we have demonstrated that the relative changes in primary production are likely to be always less than the relative change in surface irradiance. The effect of surface light on primary production decreases as the scaled surface irradiance increases. In the non-spectral model, the only parameter that affects the sensitivity is the photoadaptation parameter I_k. In the spectral model, the chlorophyll concentration has an additional modulating effect on the sensitivity.

In the light of these results, how might we enhance the management of primary production data? Given an archived set of photosynthesis-light parameters what is the most useful information that could be archived about the corresponding phytoplankton production? We suggest the following: First, calculate the phytoplankton production under the climatological value of the clear sky irradiance for the day number and latitude concerned. This represents the maximum phytoplankton production that could be achieved with the given photosynthesis

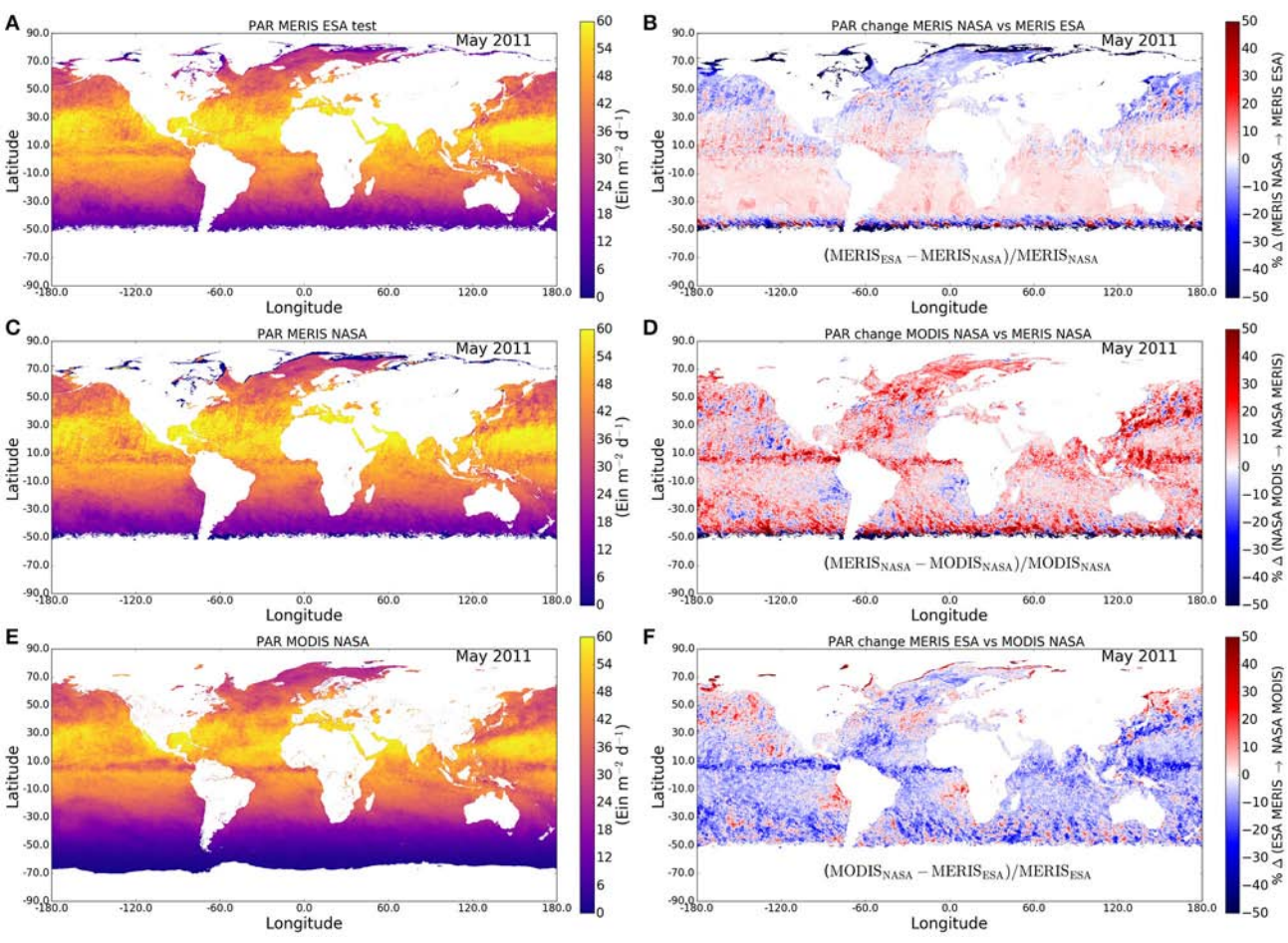


FIGURE 5 | Systematic differences in surface PAR arising from choice of data stream, exemplified by difference between MERIS and MODIS streams for May 2011. **(A)** PAR field from MERIS using ESA protocol (test product); **(C)** PAR field from MERIS using NASA product; **(E)** PAR field from MODIS Aqua using NASA product; the differences between the products are shown in **(B,D,F)**.

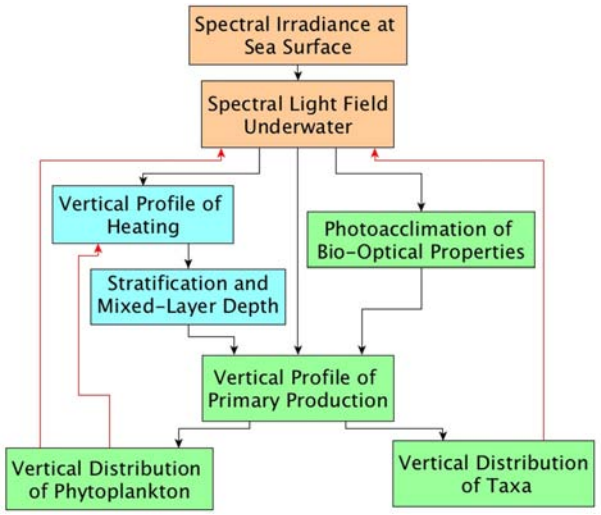


FIGURE 6 | Schematic showing principal factors affecting spectral irradiance field under water, and consequences for watercolumn production. The underwater light field is the resultant of a complex and reciprocal interplay between physics and biology.

parameters on that day of the year at that latitude. Next, calculate the production under 100% cloud cover, representing the minimum production that could be achieved under the prevailing conditions. These two calculations would establish the range of possible values of phytoplankton production expected at the relevant time and place under most of the possible values of surface irradiance. In fact, because the surface irradiance has both diffuse and direct components, 100% cloud cover will correspond to a reduction in I_0^m of nominally 50% (loss of direct component); all light reaching the sea surface will be in the diffuse component. The effect on daily production of halving the surface irradiance varies with region and season. Taking 50% as a rough estimate of the scope for variation in surface irradiance at a given time and place, we have shown that the implied relative variation in daily, water column is roughly 30–40%. The relative sensitivity, as given by Equation (18), will be independent of the magnitude chosen for the reduction in surface irradiance. A further systematic difference of about 15% may arise from the choice of data source for the surface irradiance field, indicating the value of improving the satellite-derived estimates of PAR.

The most useful particular value of phytoplankton production would be that at the climatologically-averaged value of surface irradiance at the time and place concerned. This average would take into account the effect of latitude on day length and also the local effect of cloud cover. As an example, for the Northwest Atlantic Ocean, we have developed from the SeaWiFS irradiance an empirical function that yields the climatological total daily irradiance as a function of latitude and day number (Platt et al., 2009). If longitudinal differences in cloud cover are to be taken into account, one could use the climatological data on cloud cover to estimate the relevant irradiance.

Paradoxically, the estimate of phytoplankton production made with the irradiance measured at the same time as the parameters themselves were measured (if available) is often the least useful (lowest generality) of all the estimates. Of course, the matter depends on the question being addressed. If the goal is to close a local carbon budget or a local energy budget over a short time period including the day of measurement, then the estimate of phytoplankton production made with the actual irradiance for that particular day would be the one to choose.

However, we expect that most applications of the archived data would be more general than this. In such cases, we would seek estimates of phytoplankton production that were suitably general, and the options we have indicated above would probably be preferred.

The light field in the sea is the resultant of a complex and reciprocal interplay between physics and biology (**Figure 6**). The effect on daily, water-column production consequent on changes in irradiance at the sea surface depends, at least, on the processes shown in **Figure 6**. A guide to the scale of the effect (not accounting for spectral dependence) is provided by Equation (18). Spectral effects can be parameterized as a function of chlorophyll concentration (**Figure 2**).

AUTHOR CONTRIBUTIONS

TP, SS, and HB originated the work and organized assembly of the global data bases. GW and TP developed the theoretical background. TJ and SSP made most of the calculations. TP wrote the first draft of the paper. All authors contributed to revision of the paper.

FUNDING

Funding for this work was provided through the European Space Agency (ESA), contracts 4000113690/15/I-LG and 4000110968/14/I-BG; and through the National Centre for Earth Observation (NCEO), grant PR140015. Additional support from the Simons Foundation CBIOMES project is acknowledged.

ACKNOWLEDGMENTS

This work is a contribution to the "PAR for Primary Production" Project and the "Photosynthesis Parameters from Space" Project of the European Space Agency. Support from the Jawaharlal Nehru Science Fellowship to TP is gratefully acknowledged. Additional support was received from the National Centre for Earth Observation (NCEO) of the Natural Environment Research Council of UK. The authors thank the two reviewers for their helpful comments.

REFERENCES

Bird, R. E. (1984). A simple, solar spectral model for direct-normal and diffuse horizontal irradiance. *Sol. Energy* 32, 461–471. doi: 10.1016/0038-092X(84)90260-3

Bird, R. E., and Riordan, C. J. (1986). Simple simple solar spectral model for direct and diffuse irradiance on horizontal and tilted planes at the earth's surface for cloudless atmospheres. *J. Clim. Appl. Meteorol.* 25, 87–97.

Jackson, T., Sathyendranath, S., and Platt, T. (2017). An exact solution for modelling photoacclimation of the carbon-to-chlorophyll ratio in phytoplankton. *Front. Mar. Sci.* 4:283. doi: 10.3389/fmars.2017.00283

Kyewalyanga, M., Platt, T., and Sathyendranath, S. (1992). Ocean primary production calculated by spectral and broad-band models. *Mar. Ecol. Prog. Ser.* 85, 171–185. doi: 10.3354/meps085171

Laufkötter, C., Vogt, M., Gruber, N., Aita-Noguchi, M., Aumont, O., Bopp, L., et al. (2015). Drivers and uncertainties of future global marine primary production in marine ecosystem models. *Biogeosciences* 12, 6955–6984. doi: 10.5194/bg-12-6955-2015

Lohrenz, S. E. (1993). Estimation of primary production by the simulated *in situ* method. *ICES Mar. Sei. Symp.* 197, 159–171.

Lohrenz, S. E., Wiesenburg, D. A., Rein, C. R., Arnone, R. A., Taylor, C. D., Knauer, G. A., et al. (1992). A comparison of *in situ* and simulated *in situ* methods for estimating oceanic primary production. *J. Plankt. Res.* 14, 201–221. doi: 10.1093/plankt/14.2.201

Longhurst, A. R., Sathyendranath, S., Platt, T., and Caverhill, C. M. (1995). An estimate of global primary production in the ocean from satellite radiometer data. *J. Plankton Res.* 17, 1245–1271. doi: 10.1093/plankt/17.6.1245

Mélin, F., and Hoepffner, N. (2004). *Global Marine Primary Production: A Satellite View*. Ispra: Institute for Environment and Sustainability.

Mélin, F., and Hoepffner, N. (2011). *Monitoring Phytoplankton Productivity from Satellite*, Chap. 6. Dartmouth, NS: EU PRESPO and IOCCG.

Morel, A. (1991). Light and marine photosynthesis: a spectral model with geochemical and climatological implications. *Prog. Oceanogr.* 26, 263–306. doi: 10.1016/0079-6611(91)90004-6

Morel, A., Antoine, D., Babin, M., and Dandonneau, Y. (1996). Measured and modeled primary production in the northeast atlantic (eumeli jgofs program): the impact of natural variations in photosynthetic parameters on model predictive skill. *Deep Sea Res. I* 43, 1273–1304. doi: 10.1016/0967-0637(96)00059-3

OBPG (2014a). *Medium Resolution Imaging Spectrometer (MERIS) Photosynthetically Available Radiation Data, 2014 Reprocessing. NASA OB.DAAC*. Greenbelt, MD: NASA Goddard Space Flight Center, Ocean Ecology Laboratory, Ocean Biology Processing Group. doi: 10.5067/ENVISAT/MERIS/L3B/PAR/2014

OBPG (2014b). *Moderate-resolution Imaging Spectroradiometer (MODIS) Aqua Photosynthetically Available Radiation Data; 2014 Reprocessing. NASA OB.DAAC*. Greenbelt, MD: NASA Goddard Space Flight Center, Ocean Ecology Laboratory, Ocean Biology Processing Group. doi: 10.5067/AQUA/MODIS/L3B/PAR/2014

Platt, T., Caverhill, C., and Sathyendranath, S. (1991). Basin-scale estimates of oceanic primary production by remote sensing: the north atlantic. *J. Geophys. Res.* 96, 15147–15159. doi: 10.1029/91JC01118

Platt, T., Denman, K., and Jassby, A. (1977). "Modeling the productivity of phytoplankton," in *The Sea: Ideas and Observations on Progress in the Study of the Seas*, Vol. VI, ed E. D. Goldberg (New York, NY: John Wiley), 807–856.

Platt, T., and Jassby, A. D. (1976). The relationship between photosynthesis and light for natural assemblages of coastal marine phytoplankton. *J. Phycol.* 12, 421–430. doi: 10.1111/j.1529-8817.1976.tb02866.x

Platt, T., and Sathyendranath, S. (1988). Oceanic primary production: estimation by remote sensing at local and regional scales. *Science* 241, 1613–1620. doi: 10.1126/science.241.4873.1613

Platt, T., and Sathyendranath, S. (1991). Biological production models as elements of coupled, atmosphere-ocean models for climate research. *J. Geophys. Res.* 96, 2585–2592. doi: 10.1029/90JC02305

Platt, T., and Sathyendranath, S. (1993). Estimators of primary production for interpretation of remotely sensed data on ocean color. *J. Geophys. Res.* 98, 14561–14576. doi: 10.1029/93JC01001

Platt, T., Sathyendranath, S., Forget, M.-H., White, G. N. III, Caverhill, C., Bouman, H., et al. (2008). Operational mode estimation of primary production at large geographical scales. *Remote Sens. Environ.* 112, 3437–3448. doi: 10.1016/j.rse.2007.11.018

Platt, T., Sathyendranath, S., and Longhurst, A. (1995). Remote sensing of primary production in the ocean: promise and fulfilment. *Philos. Trans. R. Soc. Lond. B* 348, 191–202. doi: 10.1098/rstb.1995.0061

Platt, T., Sathyendranath, S., and Ravindran, P. (1990). Primary production by phytoplankton: analytic solutions for daily rates per unit area of water surface. *Proc. R. Soc. Lond. Ser. B* 241, 101–111. doi: 10.1098/rspb.1990.0072

Platt, T., Sathyendranath, S., White, G. N. III, Fuentes-Yaco, C., Zhai, L., Devred, E., et al. (2009). Diagnostic properties of phytoplankton time series from remote sensing. *Estuar. Coasts* 32, 428–439. doi: 10.1007/s12237-009-9161-0

Sathyendranath, S., Groom, S., Grant, M., Brewin, R., Thompson, A., Chuprin, A., et al. (2016). ESA ocean colour climate change initiative (ocean_colour_cci): Version 1.0 data. *Centre Environ. Data Anal.* doi: 10.5285/E32FEB53-5DB1-44BC-8A09-A6275BA99407

Sathyendranath, S., Longhurst, A., Caverhill, C. M., and Platt, T. (1995). Regionally and seasonally differentiated primary production in the North Atlantic. *Deep Sea Res. I* 42, 1773–1802. doi: 10.1016/0967-0637(95)00059-F

Sathyendranath, S., and Platt, T. (1988). The spectral irradiance field at the surface and in the interior of the ocean: a model for applications in oceanography and remote sensing. *J. Geophys. Res.* 93, 9270–9280. doi: 10.1029/JC093iC08p09270

Sathyendranath, S., and Platt, T. (1989). Computation of aquatic primary production: extended formalism to include effect of angular and spectral distribution of light. *Limnol. Oceanogr.* 34, 188–198. doi: 10.4319/lo.1989.34.1.0188

Sathyendranath, S., Platt, T., Caverhill, C. M., Warnock, R. E., and Lewis, M. R. (1989). Remote sensing of oceanic primary production: computations using a spectral model. *Deep Sea Res. I* 36, 431–453. doi: 10.1016/0198-0149(89)90046-0

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Platt, Sathyendranath, White, Jackson, Saux Picart, and Bouman. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Marine Science

ORIGINAL RESEARCH
published: 11 December 2017
doi: 10.3389/fmars.2017.00378

Intercomparison of Ocean Color Algorithms for Picophytoplankton Carbon in the Ocean

Víctor Martínez-Vicente[1], Hayley Evers-King[1], Shovonlal Roy[2], Tihomir S. Kostadinov[3,4], Glen A. Tarran[1], Jason R. Graff[5], Robert J. W. Brewin[1,6], Giorgio Dall'Olmo[1,6], Tom Jackson[1], Anna E. Hickman[7], Rüdiger Röttgers[8], Hajo Krasemann[8], Emilio Marañón[9], Trevor Platt[1] and Shubha Sathyendranath[1,6]*

[1] Plymouth Marine Laboratory, Plymouth, United Kingdom, [2] Department of Geography and Environmental Sciences, School of Agriculture, Policy and Development, University of Reading, Reading, United Kingdom, [3] Department of Geography and the Environment, University of Richmond, Richmond, VA, United States, [4] Division of Hydrologic Sciences, Desert Research Institute, Reno, NV, United States, [5] Department of Botany and Plant Pathology, Oregon State University, Corvallis, OR, United States, [6] Plymouth Marine Laboratory, National Centre for Earth Observation, Plymouth, United Kingdom, [7] Ocean and Earth Science, National Oceanography Centre Southampton, University of Southampton, Southampton, United Kingdom, [8] Helmholtz-Zentrum Geesthacht, Center for Materials and Coastal Research, Geesthacht, Germany, [9] Departamento de Ecología y Biología Animal, Universidade de Vigo, Vigo, Spain

OPEN ACCESS

Edited by:
Hervé Claustre, Centre National de la Recherche Scientifique (CNRS), France

Reviewed by:
Emmanuel Devred, Fisheries and Oceans Canada, Canada
Severine Alvain, Centre National de la Recherche Scientifique (CNRS), France

***Correspondence:**
Víctor Martínez-Vicente
vmv@pml.ac.uk

Specialty section:
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

Received: *03 April 2017*
Accepted: *10 November 2017*
Published: *11 December 2017*

Citation:
Martínez-Vicente V, Evers-King H, Roy S, Kostadinov TS, Tarran GA, Graff JR, Brewin RJW, Dall'Olmo G, Jackson T, Hickman AE, Röttgers R, Krasemann H, Marañón E, Platt T and Sathyendranath S (2017) Intercomparison of Ocean Color Algorithms for Picophytoplankton Carbon in the Ocean. Front. Mar. Sci. 4:378. doi: 10.3389/fmars.2017.00378

The differences among phytoplankton carbon (C_{phy}) predictions from six ocean color algorithms are investigated by comparison with *in situ* estimates of phytoplankton carbon. The common satellite data used as input for the algorithms is the Ocean Color Climate Change Initiative merged product. The matching *in situ* data are derived from flow cytometric cell counts and per-cell carbon estimates for different types of pico-phytoplankton. This combination of satellite and *in situ* data provides a relatively large matching dataset ($N > 500$), which is independent from most of the algorithms tested and spans almost two orders of magnitude in C_{phy}. Results show that not a single algorithm outperforms any of the other when using all matching data. Concentrating on the oligotrophic regions (Chlorophyll-a concentration, B, less than 0.15 mg Chl m^{-3}), where flow cytometric analysis captures most of the phytoplankton biomass, reveals significant differences in algorithm performance. The bias ranges from −35 to +150% and unbiased root mean squared difference from 5 to 10 mg C m^{-3} among algorithms, with chlorophyll-based algorithms performing better than the rest. The backscattering-based algorithms produce different results at the clearest waters and these differences are discussed in terms of the different algorithms used for optical particle backscattering coefficient (b_{bp}) retrieval.

Keywords: phytoplankton carbon, carbon-to-chlorophyll, ocean color remote sensing, picophytoplankton, flow cytometry, optical water class, algorithm uncertainty

1. INTRODUCTION

One of the standard products from ocean-color remote sensing is the concentration of chlorophyll-a (B) in the surface layers of the ocean, which is an estimation of phytoplankton abundance. This product has proven to be extremely useful for various applications (e.g., Platt and Sathyendranath, 2008). More recently, there has been a growing interest in monitoring the standing stock of phytoplankton in carbon units (CEOS, 2014), in addition to chlorophyll units. There are many reasons for this interest, which include calculation of primary production using carbon-based

models (Behrenfeld et al., 2005; Westberry et al., 2008); estimating phytoplankton loss rates (Zhai et al., 2008, 2010); comparison with estimates of phytoplankton biomass in carbon units from marine ecosystem models (Dutkiewicz et al., 2015); and establishing the budget of the pools of carbon in the ocean (CEOS, 2014), their turnover rates (Casey et al., 2013), and their exchanges with the atmospheric and terrrestrial domains (CEOS, 2014). With increasing appreciation of the different roles of various phytoplankton functional types in the oceanic biogeochemical cycles (Le Quéré et al., 2005), there is a corresponding need to know the pools of carbon associated with the different phytoplankton types, rather than just the total phytoplankton carbon.

A handful of algorithms have been proposed for deriving phytoplankton carbon from satellite data. These include methods based on particle back-scattering coefficient (b_{bp}) at a single wavelength (Behrenfeld et al., 2005; Martínez-Vicente et al., 2013); empirical relationships based on chlorophyll concentration (Sathyendranath et al., 2009; Marañón et al., 2014); and methods based on allometric considerations combined with either the spectral slope of the particle back-scattering spectrum (Kostadinov et al., 2009, 2016) or with the phytoplankton absorption characteristics (Roy et al., 2017). Of these, the method proposed by Martínez-Vicente et al. (2013) dealt with a fraction of the phytoplankton community (diameter $< 20\,\mu m$), whereas those of Behrenfeld et al. (2005), Sathyendranath et al. (2009), and Marañón et al. (2014) dealt with the whole phytoplankton community. The methods based on allometric structure (Kostadinov et al., 2009, 2016; Roy et al., 2017), on the other hand, have the advantage of being able to target the whole of the phytoplankton community, and partition phytoplankton carbon among any user-defined size-intervals. Comparison of these algorithms is not straightforward, because of the differences in approaches used and the products obtained. Furthermore, they have been subjected to varying degrees of validation, with differences in the number of validation points used and in their regional and seasonal coverage. Another difficulty lies with having access to *in situ* data in sufficient quantity and comprehensive enough for algorithm assessment.

Various methods for *in situ* measurements of phytoplankton carbon in the laboratory or in the field have been reviewed by Casey et al. (2013). Some of the *in situ* methods require a proxy measurement, which is then calibrated against phytoplankton carbon. Subsequently, the carbon concentration is inferred from measurements of the proxy, which would typically be easier to measure than the carbon concentration itself. The proxies include adenosine triphosphate (ATP) (Sinclair et al., 1979); the refractive index of phytoplankton cells (Stramski, 1999); and the forward light scatter by phytoplankton cells in a flow cytometer (Casey et al., 2013). Redalje and Laws (1981) used chlorophyll-a labeling and showed that the specific activity of carbon in chlorophyll-a became equivalent to that of total phytoplankton carbon in incubations of 6–12 h, and so chlorophyll-a labeling could be used to infer phytoplankton carbon and growth rates. Graff et al. (2015) used flow cytometer cell sorting (Graff et al., 2012) to measure phytoplankton carbon in sorted samples, thereby avoiding contamination of results by non-pigmented particles. An accepted approach to estimating phytoplankton carbon at sea is to use a flow-cytometer to count phytoplankton cells sorted into different types. Using laboratory-based estimates of carbon per cell and typical (or measured mean) cell diameters for those phytoplankton types, the total carbon is computed by adding the carbon contribution of each phytoplankton cell type. This is obtained by multiplying the number of cells enumerated with the flow cytometer by the carbon per cell (DuRand et al., 2001; Oubelkheir et al., 2005; Martínez-Vicente et al., 2013). Such methods have an upper limit on measured cell size, depending on how the flow-cytometer is set up (typically $D < 50\,\mu m$).

We present in this work a comparison of six different algorithms for estimating phytoplankton carbon from space. The algorithms have been selected as representative of all existing state-of-the-art approaches. The comparisons are based on a newly-compiled, global, flow-cytometric dataset that is used to compute the *in situ* picophytoplankton carbon, matched with satellite data from the same location, and for the same day. The performance of these products is explored in different optical water classes. The comparison is limited to picophytoplankton, because the flow-cytometric database dealt largely with this size class. The objective of the comparison is to learn more about the advantages and limitations of the algorithms, rather than to rank them. We expect that the results will allow a more informed use of phytoplankton carbon products from satellites, for example, when they are compared with model outputs, and serve to identify areas where improvements are needed and potential avenues for achieving them. The analysis also brings to light some of the limitations of the *in situ* database, and highlights areas where progress is needed, to enable better validation of satellite data.

2. METHODOLOGY

2.1. *In Situ* Dataset

As part of this study, more than 12,000 observations of picophytoplankton abundance have been collated from coastal and oceanic regions (**Table 1**), building upon a dataset compiled by the modeling community through MAREDAT (http://maremip.uea.ac.uk/maredat.html) (Buitenhuis et al., 2012). Additional data come from a long-term observation program, the Atlantic Meridional Transect (AMT); as well as recently available data collected independently during AMT-22 and in the Pacific (Graff et al., 2015) and from other regions in the Atlantic ocean (Taylor et al., 2013). The dataset assembled consists of cell counts (in cells per milliliter), from water samples originating between 0 and 200 m depth, and collected in the period between 1997 and 2014, to match satellite observations available. Flow cytometry analysis of the samples provides cell abundances segregated into different types of phytoplankton. At this stage, the database consisted of 12,431 sample entries. Only the picophytoplankton cells ($<2\,\mu m$) were available in the MAREDAT dataset, which were separated into *Prochlorococcus* spp., *Synechococcus* spp. and picoeukaryotic phytoplankton. For consistency, only information on the same phytoplankton types were extracted from the additional data sources (Zubkov et al., 1998; Tarran et al., 2001, 2006; Taylor et al., 2013;

TABLE 1 | Summary of *in situ* data.

Area	Number of *in situ* observations Selected match-ups (Total)	C_{phy} (mg m^{-3}) (Match-ups median ± IQR/2)	References
Pacific and Atlantic Ocean	2 (57)	18.3 ± N.A.	Graff et al., 2015
Atlantic Ocean (AMT)	174 (5,860)	10.6 ± 4.5	Zubkov et al., 1998; Tarran et al., 2001, 2006
Western English Channel (WCO)	0 (1,196)	N.A	Tarran, 2015; Tarran and Bruun, 2015
Atlantic Ocean	48 (224)	11.7 ± 3.5	Taylor et al., 2013
Global Oceans	333 (5,153)	12.1 ± 5.9	MAREDAT (Buitenhuis et al., 2012)
All data	557 (12,490)	11.5 ± 5.3	This study

Graff et al., 2015; Tarran, 2015; Tarran and Bruun, 2015) (see **Table 1**). The carbon concentration (C_{phy}, in mg C m^{-3}) for each phytoplankton group (i) and for each sample (j), $C_{phy}(i,j)$, was calculated as follows:

$$C_{phy}(i,j) = 10^{-6} \times N(i,j)\varepsilon(i) \quad (1)$$

where $N(i,j)$ is cell abundance (cell mL^{-1}) for each of the three phytoplankton types (i = *Prochlorococcus* spp., *Synechococcus* spp. or picoeukaryotic phytoplankton) at sample j; and $\varepsilon(i)$ is cellular carbon per cell (fgC cell^{-1}) for each of the picophytoplankton types. The factor 10^{-6} converts mL to m^3 and fgC to mgC. We used the mean $\varepsilon(i)$ for each phytoplankton type proposed by Buitenhuis et al. (2012): 60 fgC cell^{-1} for *Prochlorococcus* spp., 154 fgC cell^{-1} for *Synechococcus* spp. and 1319 fgC cell^{-1} for picoeukaryotic phytoplankton. These values of $\varepsilon(i)$ are comparable to values from the Bermuda Atlantic Time-series Study (BATS) (Casey et al., 2013), for *Prochlorococcus* spp. and *Synechococcus* spp., whereas picoeukaryotic phytoplankton $\varepsilon(i)$ values are lower than in BATS. The total picophytoplankton carbon concentration per sample j, i.e., $C_{phy}(j)$ is the sum of the contributions from each picophytoplankton type (i.e., $C_{phy}(i,j)$), and will be hereafter referred to as C_{phy} at a given location and depth.

The choice of phytoplankton types included in this computation, as well as the parameters used for the conversion to carbon, matches the modeling community approach as represented in Buitenhuis et al. (2012). The choice of phytoplankton types is such that phytoplankton types with diameter >2 μm are not taken into account. Furthermore, the choice of a mean carbon concentration per cell for each phytoplankton type does not permit accounting for any variations in size or cellular carbon spatially or temporally for each type of phytoplankton. To test our choice of carbon conversion parameters we compared direct measurements of C_{phy} with estimates computed using the conversion factors above. In samples from the AMT-22 ($N = 15$) (Graff et al., 2015), the slope of the regression between direct measurements of C_{phy} and computed C_{phy}, was 0.8 ($r^2 = 0.6$, $p < 0.05$). According to this result, the estimates of picoplankton C_{phy} in our dataset are significantly correlated with direct estimates of phytoplankton carbon, and could be an overestimate of direct observations of C_{phy}, which include nanophytoplankton, although a larger sample is required to support this conclusion.

2.2. *In Situ* and Satellite Match-up Selection

The *in situ* database described above was matched with merged ocean-color satellite data from the Ocean Color Climate Change Initiative (OC-CCI) (Sathyendranath et al., 2012). These merged products were used to maximize the possibility of finding matching *in situ* data as well as to use a set of common inputs to the different algorithms. The OC-CCI version 2 data had a daily sinusoidal projection (binned) and a 4 km spatial resolution. These satellite data were used as inputs for C_{phy} algorithms: total B from OC4v6, b_{bp} from the Quasi-Analytical Algorithm (QAA) v5 (a modification to v4 in Lee et al., 2002, 2007, but that does not include Raman scattering, Westberry et al., 2013; Lee and Huot, 2014). Second, the water class membership (Moore et al., 2001, 2009; Jackson et al., 2017).

The procedure for match-up selection was the same as that used for particulate organic carbon (POC) data (Evers-King et al., 2017). The day of year the *in situ* sample was collected was matched with the same day of year from the merged satellite products. Then all relevant data were extracted from a 3 × 3 pixel set with the sample location at the center. The number of valid data, within the 3 × 3 grid, and mean and standard deviation of the valid points were recorded for each computed C_{phy} product. The 3 × 3 grid was used to identify where sufficient satellite data were available. In this dataset only 11 matched points had 3 valid pixels or less. The C_{phy} algorithms were applied to the central pixel of the satellite matched up data. The match-up process reduced the sample size considerably. Further reduction came from depth-averaging (between 0 and 10 m) the C_{phy} profiles that matched the satellite data, and ignoring the deeper samples, leaving 647 data points. Finally, to remove outliers, the top and bottom 2 percentile were removed from the dataset, leaving $N = 557$ for the analysis. The geographical distribution of match-up database for the picophytoplankton carbon concentration, C_{phy}, is given in **Figure 1**. The match-up dataset usable for the algorithm comparison was only about 5% of the inital data (**Table 1**). It is worth emphasizing that the match-up data set has not been used for the calibration or development of most of the algorithms compared (see section 4).

2.3. Ocean-Color Phytoplankton Carbon Algorithms

The following section describes the six algorithms compared in this exercise. All the phytoplankton algorithms were implemented using as input data the appropriate OC-CCI product for consistency and to isolate the effects of the different algorithms. **Table 2** provides a comparison of the input data and the phytoplankton size range that is included in the

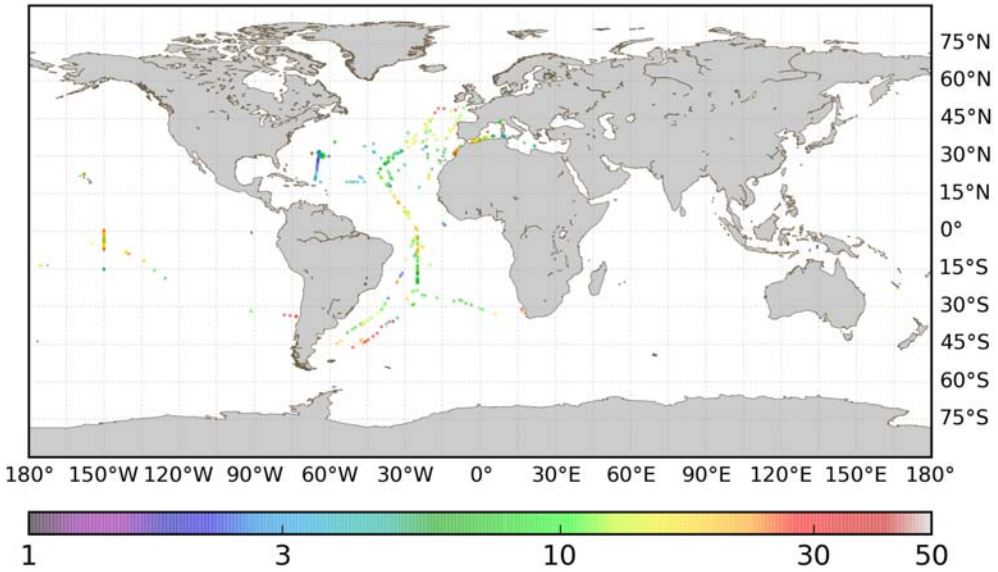


FIGURE 1 | Geographical distribution of the match-up dataset ($N = 557$) used for algorithm testing. Color scale is concentration of picophytoplankton, C_{phy} in mg C m^{-3}.

TABLE 2 | Summary of C_{phy} algorithms main characteristics and median values predicted for the *in situ* match-up database ($N = 557$).

Model	Algorithm type	Input (wavelength)	Output	C_{phy} (Median ± IQR/2)	References
A	Chlorophyll-based	B	All sizes	16.8 ± 6.7	Sathyendranath et al., 2009
B	Chlorophyll-based	B	All sizes	9.2 ± 5.2	Marañón et al., 2014
C	Backscattering-based	b_{bp} (440)	All sizes	18.7 ± 5.2	Behrenfeld et al., 2005
D	Backscattering-based	b_{bp} (470)	D < 2 μ m	20.9 ± 6.4	Martínez-Vicente et al., 2013; modified, see section 2.3.2
E	Allometric conversion	b_{bp} spectral slope from the 490, 510, and 555 nm bands, and b_{bp} at 443 nm	0.5 < D < 2 μ m	5.6 ± 0.8	Kostadinov et al., 2009, 2016
F	Allometric conversion	a_{phy} (676) and B	D < 2 μ m	5.0 ± 2.2	Roy et al., 2017

Chlorophyll concentration, B, in mg Chla m^{-3}; optical particulate backscattering coefficient, b_{bp}, in m^{-1}; phytoplankton absorption coefficients a_{phy}, in m^{-1}; pico phytoplankton carbon concentration, C_{phy}, in mg C m^{-3}.

outputs of each algorithm. These are important characteristics of the algorithms, required for the interpretation of the results. For phytoplankton carbon, C_{phy} in mg C m^{-3}, six products were derived and they are briefly described in this section. According to their common characteristics, they can be grouped into chlorophyll-based, backscattering-based and allometric algorithms.

2.3.1. Chlorophyll-Based Algorithms

This family of algorithms use chlorophyll concentration as an input, B with units of mg Chla m^{-3}. Chlorophyll in this study is obtained from OC-CCI merged dataset with the algorithm OC4v6, which is a band switching algorithm, mainly a fourth-order polynomial relationship between remote sensing reflectance in the blue and green bands. The two algorithms in this group use the same input and have a similar formulation, however, the assumptions made in their construction and hence their definition of C_{phy} are different. Algorithm A (Sathyendranath et al., 2009) was developed from an empirical relationship between *in situ* measurements of total particulate carbon and B. For this model, C_{phy} in Equation (2) below is an upper bound on the total phytoplankton carbon:

$$C_{phy} = 65 \times B^{0.63}. \quad (2)$$

Algorithm B (Marañón et al., 2014) was also developed from an empirical relationship using *in situ* measurements of B, and not originally designed as an algorithm for ocean color applications. However, the estimates of total phytoplankton carbon originated from applying a conversion factor to microscope (counting cells with diameter, $D > 5\,\mu$m) and flow cytometry ($D < 10\,\mu$m) phytoplankton cell counts. This model is formulated in Equation (3) as:

$$C_{phy} = 62 \times B^{0.89}. \quad (3)$$

Because of the difference in their definition of C_{phy}, Algorithm A and Algorithm B have been considered separately in our

analysis. A priori, the expectation is that both chlorophyll-based algorithms, using total chlorophyll concentration as input data, will overestimate the picophytoplankton carbon from our *in situ* match-up dataset, since they are both designed to calculate total phytoplankton carbon, rather than just the picophytoplankton in our dataset. Further, it is also worth noting that the conversion factors used to compute phytoplankton carbon from cell abundance in Algorithm B are different to the ones used in our *in situ* match-up dataset.

2.3.2. Backscattering-Based Algorithms

Some semi-empirical algorithms use the (wavelength dependent) optical particulate backscattering coefficient, b_{bp} in m^{-1}, to estimate C_{phy}. The backscattering coefficient in this study is obtained from the OC-CCI merged dataset by applying the algorithm by the Quasi-Analytical Algorithm (QAA) v5 (a modification to v4 in Lee et al., 2002, 2007). In essence, the QAA first computes $b_{bp}(555)$, from combining remote sensing reflectance at 555 nm with an empirical relationship between remote sensing reflectance ratios and the total absorption coefficient and the backscattering of pure seawater (modeled). Then, to propagate the $b_{bp}(555)$ at other wavelengths, the algorithm uses a band ratio (again blue to green bands) to compute the backscattering spectral slope. The same QAA b_{bp} product is used for both backscattering based C_{phy} algorithms, but at different wavelengths. Algorithm C (Behrenfeld et al., 2005) uses $b_{bp}(443)$ as an input:

$$C_{phy} = 13000 \times (b_{bp}(443) - 0.00035). \quad (4)$$

As this algorithm was developed from MODIS-Aqua (Moderate Resolution Imaging Spectroradiometer) ocean-color data, and 443 nm is a native OC-CCI band, no spectral adjustment is therefore needed. However, it is worth noting that the algorithm was developed originally using the GSM algorithm (Garver and Siegel, 1997; Maritorena et al., 2002; Siegel et al., 2002), but in this test, the b_{bp} input come from the QAA algorithm. The C_{phy} derived with this algorithm includes all the phytoplankton size ranges. Algorithm D (Martínez-Vicente et al., 2013) is another semi-empirical algorithm, developed from the relationship between *in situ* flow cytometry-based carbon and $b_{bp}(470)$, but is included in the comparison with some changes. The first modification was to re-compute the coefficients in the original equation by using the same computation of picoplankton as the one used in this work, which meant ignoring the nanoeukaryotes, cryptophytes and coccolithophorids contributions to the picoplankton carbon and use the same carbon to cell conversion factors as in this study (i.e., those of MAREDAT; Buitenhuis et al., 2013). This re-calculation led to lower (pico)phytoplankton carbon estimates which were, on average, 27% less than the published values of phytoplankton carbon (from pico- and nano-plankton) used in Martínez-Vicente et al. (2013). When the new picophytoplankton C_{phy} estimate was used with the original *in situ* b_{bp} data, the resulting fit was:

$$C_{phy} = 18000 \times (b_{bp}(470) - 43 * 10^{-5}), N = 70. \quad (5)$$

This equation explains considerably less variance in the observed data ($r^2 = 0.4$) than the r^2 of 0.89 reported in the original work. However, it makes the definition of C_{phy} by this model directly comparable to the *in situ* data. The second modification was to adjust the backscattering coefficient wavelength from the available value, 490 to 470 nm. To do so, the spectral slope of the b_{bp} from the OC-CCI data was obtained by doing an ordinary least squares fit to the $\log_{10}$ transformed data and calculated the new $b_{bp}(470)$ needed for Equation (5).

2.3.3. Allometric Type Algorithms

These algorithms belong to a family of algorithms that use optical properties to compute phytoplankton size structure and then convert it into biomass (Mouw et al., 2017). Algorithm E (Kostadinov et al., 2016) retrieves the absolute and fractional phytoplankton carbon biomass in three phytoplankton size classes (or, approximately equivalent − phytoplankton functional types) − picophytoplankton (0.5–2 μm in diameter), nanophytoplankton (2–20 μm) and microphytoplankton (20–50 μm). The algorithm uses retrievals of the particle size distribution (PSD) to estimate particle volume. Note that the PSD is estimated for all particles in suspension in the water. Particle volume is then converted to carbon concentrations using a compilation of existing allometric relationships between size and carbon content of phytoplankton cells (Menden-Deuer and Lessard, 2000). Derived carbon concentration is then divided by 3 to estimate the living phytoplankton carbon fraction. The PSD retrievals themselves are based on a PSD algorithm (Kostadinov et al., 2009), which relates the spectral slope and magnitude of the backscattering coefficient spectrum to the underlying parameters of an assumed power-law PSD, via look-up tables (LUTs) constructed using Mie theory modeling. In the implementation used here, the input backscattering spectrum comes from the standard QAA products of the OC-CCI dataset, which are derived using Lee et al. (2002) algorithm, as summarized above. This is different from the original implementations (Kostadinov et al., 2009, 2016), where the Loisel and Stramski (2000) algorithm was used to retrieve spectral b_{bp}. The PSD parameters retrieved are the power-law slope (ξ) and the scaling parameter (i.e., differential particle number concentration at a reference diameter of 2 μm, N_o, $[m^{-4}]$). Kostadinov et al. (2016) applied an empirical correction to the PSD scaling parameter N_o obtained from the model LUT, to improve absolute phytoplankton carbon concentration estimates.

A further allometry-based method, Algorithm F (Roy et al., 2017), uses chlorophyll concentration and the absorption coefficient of phytoplankton at 676 nm, $a_{ph}(676)$, to compute phytoplankton carbon. In this algorithm, the exponent of the phytoplankton size spectrum (ξ) is first computed from the specific-absorption coefficient of phytoplankton at 676 nm, $a^*_{ph}(676)$ using a method developed by Roy et al. (2013). This algorithm uses as input B from OC4v6 and $a_{ph}(676)$ from QAA, both standard products in the OC-CCI dataset. The estimated exponent of the size spectrum ξ and the allometric relationships between the cellular content of phytoplankton carbon (C_{cell}) and cell volume (V_{cell}) reported by Menden-Deuer and Lessard (2000) are then used to compute the

concentration of phytoplankton carbon (C_{total}, in mg C m^{-3}) contained in the cells within any specified diameter range. To do so, the allometric parameters corresponding to the mixed populations of phytoplankton are derived from the allometric relationships found for individual groups of phytoplankton, as reported in Menden-Deuer and Lessard (2000), by performing linear regression. The derived allometric relationship is used then to compute the magnitude of the carbon-to-chlorophyll ratio (χ), using the derived allometric expressions for the concentrations of chlorophyll and phytoplankton carbon, C_{total}. Finally, phytoplankton carbon for the specified size range is computed as the product of χ and satellite-derived chlorophyll concentration (for more details see Roy et al., 2013, 2017). The allometry-based algorithms E and F were used to compute picophytoplankton concentration within a diameter range 0.2–2.0 μm, which is directly comparable with the size range included in the matching *in situ* database.

In summary, each method is thus based on a satellite measurement that provides underlying variability in the resulting C_{phy} (reflectance ratio or analytically-derived backscatter) which is then combined with selected empirical relationships that scale those measurements to C_{phy} (using either linear or non-linear relationships and sometimes including more than one step, such as via *B*). The strength of this study lies in the use of the OC-CCI satellite dataset as a common source of inputs for all the algorithms, which removes sources of uncertainty from other parts of the satellite processing. A limitation, however, comes from the differences in the definition of the C_{phy} for each algorithm (**Table 2**). It is expected that the algorithms which compute total C_{phy} (i.e., Algorithms A, B, and C) will be most comparable to the *in situ* data when the contribution to the phytoplankton Carbon by nano and picoplankton is not significant.

2.4. Statistical Metrics and Their Contribution to the Study

Ranking of algorithms according to their performance is a classic exercise for the ocean-color community, that has evolved from comparisons of chlorophyll algorithms (O'Reilly et al., 1998) to more complex and comprehensive approaches recently (Brewin et al., 2015; Kostadinov et al., 2017). Typically, a battery of statistical metrics is used to construct an index of overall performance against a set of matched data with *in situ* observations (Brewin et al., 2015). In this exercise, however, we do not use a scoring system to rank algorithms, since one of the aims of this work is to provide an overall idea of the current accuracy of the phytoplankton Carbon product from a group of algorithms. The Kolmogorov-Smirnov test for normality of the *in situ* match-up data showed a significant deviation from normality for $\log_{10}$ transformed and un-transformed data and the residuals ($p < 0.001$). Therefore, statistical metrics that assume normality would be less reliable. For completeness, the statistical tests were computed for $\log_{10}$ transformed data, using parametric tests; and for un-transformed data, using non parametric, rank-based, statistics. Statistical metrics computed were:

- Pearsons correlation coefficient for $\log_{10}$ transformed data, and Spearman's correlation for un-transformed data (r_p and r_s respectively),
- Root mean square differences (RMSD, Ψ),
- Signed average bias (δ),
- Median absolute percentage deviation between predictions and observations, (MAPD in %) was an estimate of bias and precision was estimated as the interquartile range (IQR) of the absolute percentage deviation for the untransformed data.
- Center pattern root mean square differences for $\log_{10}$ transformed and un-transformed data (Δ), and
- Slope and intercept (*S*, *I*) from a Type-II linear regression (Reduced Major Axis) for $\log_{10}$ transformed and un-transformed data.

To provide an indication of the stability of the statistics and to compute confidence intervals on them, bootstrapping (Efron, 1979; Efron and Tibshirani, 1993) with random re-sampling and replacement was used to construct 1,000 different datasets from which confidence intervals were computed for some of the statistical metrics above. These metrics were computed for all the algorithms tested against the match-up dataset as a whole and, in adition, they were also computed after segregating the match-up dataset according to the dominant water class at the central match-up pixel. Because of the nature of the ocean color CCI satellite data and the C_{phy} algorithms, it is expected that algorithm performance will degrade toward more turbid environments (water classes 8 and over). Furthermore, the statistical results per water class provided a measure of dispersion of the phytoplankton Carbon product among algorithms, describing in which optical environments the algorithms show greater agreement. The statistics per water class were used to produce uncertainty maps (RMSD and bias). To generate the uncertainty maps, the optical class memberships at each pixel, and the per-class uncertainty values for each class were used to produce a weighted average uncertainty value for the pixel, with the weighting function being provided by the class membership. This is the method followed by the OC-CCI and described fully in Jackson et al. (2017).

3. RESULTS

3.1. Distribution of *in Situ* Data and Accuracy of Algorithms

The sources and geographic distribution of the *in situ* data, as well as the corresponding median values of the picophytoplankton carbon data are summarized in **Table 1**. The spatial distribution of the match-ups (**Figure 1**) shows their limited coverage of the oceans, with most of the data (71%) located in the Northern Hemisphere and from the Atlantic Ocean. The overall median value of C_{phy} from the match-up database was 11.7 ± 5.3 mg C m^{-3} (median ± IQR/2), with values ranging from 1.7 to 60.2 mg C m^{-3}. As a comparison,chlorophyll concentration (*B*) from the coincident satellite data was 0.12 ± 0.08 mg Chla m^{-3}, ranging from 0.01 to 3.53 mg Chla m^{-3}. The median values of C_{phy} from the algorithms applied to the matching data (**Table 2**) were not significantly different to the C_{phy} *in situ* (Mann-Witney

test, $N = 557$, $p < 0.05$). Relative frequency histograms of these data (**Figure 2**) however, show some bias. The peaks of the histogram of the chlorophyll-based algorithms (Algorithms A and B) were closest to the *in situ*; the backscattering-based models (Algorithms C and D) were double the median of *in situ* and allometric algorithms (Algorithms E and F), were half the *in situ* median value (**Figure 2**). The distribution spread of chlorophyll-based algorithms were about the same or wider than the *in situ* (C.V. between 40 and 56%), with backscattering and allometric algorithms having narrower distributions (C.V. lower than 30%).

Figure 3 shows the results from the algorithm estimates of C_{phy} against *in situ* derived estimates of C_{phy} and the relevant statistical metrics are in **Table 3**. All the algorithms showed different predictions of C_{phy}, but as expected, there were commonalities among models that shared the same formulation. For example, both chlorophyll-based algorithms (**Figures 3A,B**) presented an elongated cloud with a weak but positive (∼0.6) and significant correlation with *in situ* data, and stayed mostly along the 1:1 line, with slopes close to 1; whereas both backscattering-based algorithms (**Figures 3C,D**), had weaker correlations (∼0.4) and, although the slopes where also close to 1, the data cloud does not capture the lower C_{phy} measurements. Contrary to the other two groups of algorithms, the allometric-based do not share a formulation, hence their results (**Figures 3E,F**) differ significantly among them.

The statistical metrics (**Table 3**) provide a range of values among the algorithms tested, showing that there is not a clearly superior performance of a single algorithm on all metrics. The bias (δ) ranged from 3.5 mg C m^{-3} for Algorithm B (Marañón et al., 2014) to 15 mg C m^{-3} for the Algorithm D (Martínez-Vicente et al., 2013) as modified in this work. Chlorophyll and backscattering based algorithms had a positive bias, less than 15 mg C m^{-3}, whereas allometric algorithms had a negative bias, less than 7 mg C m^{-3}. The un-biased RMSD (Δ), which gives an idea of the dispersion of the predictions, ranged from 8.9 mg C m^{-3} for Algorithm E (Kostadinov et al., 2016) to 29 mg C m^{-3} for Algorithm D, (Martínez-Vicente et al., 2013) as modified in this work. The lowest median absolute percentage difference (MAPD), a measure of accuracy, for the untransformed data was for Algorithm B (Marañón et al., 2014), in agreement with the bias indicator for log and non-log transformed data. The inter-quartile range of the MAPD (IQR), a measure of precision of the algorithm, was lowest for Algorithm E (Kostadinov et al., 2016), and coincided with another indication of dispersion (Δ) in the non-log statistics, but differed for the log-transformed data.

Overall, chlorophyll-based algorithms had higher correlation and indicators of lower bias (i.e., δ, MAPD), whereas allometric algorithms had indicators of lower dispersion (i.e., Δ, IQR of MAPD). Between algorithms, there was a factor 4 of difference between the maximum and minimum predictions from the algorithms for all matchups pooled together as a median (i.e., the median of the fractional difference between the minimum and the maximum predictions by the algorithms in each match-up point).

An additional way to assess model performance is to study their emergent properties (de Mora et al., 2016). Here we have compared the *in situ* and the algorithm derived C_{phy} to the chlorophyll concentration, B (standard OC4v6 OC-CCI-product) for the match-ups (**Figure 4** and **Table 4**), to investigate the behavior of the C_{phy}:B ratio. **Figure 4A** displays the positive correlation between the satellite derived B (for the whole of the phytoplankton assemblage) and the *in situ* derived C_{phy} for the picophytoplankton fraction only, over more than two orders of magnitude of chlorophyll concentration. Because of the mismatch between the particle assemblage in B and C_{phy},

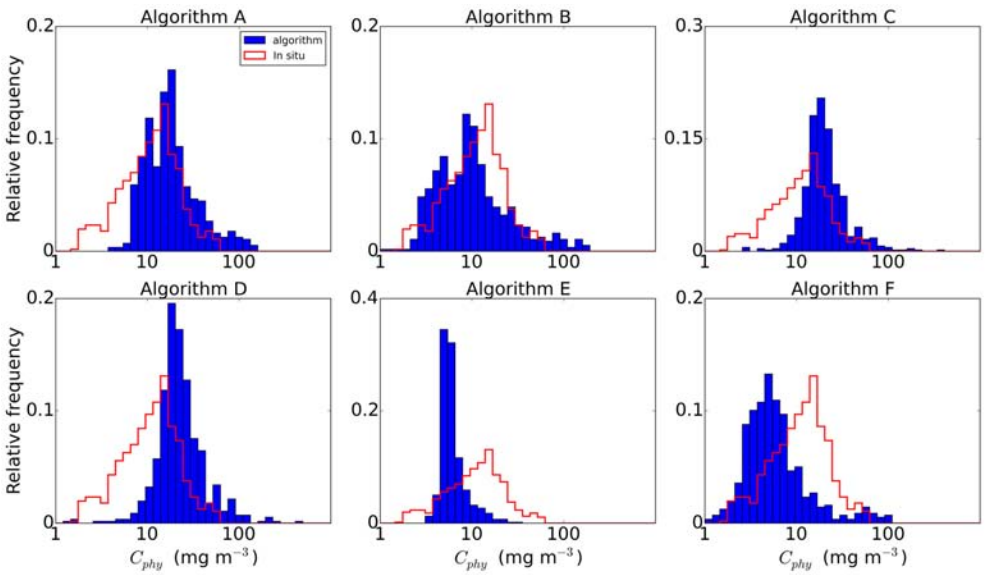


FIGURE 2 | Relative frequency distribution of C_{phy} of the *in situ* data compared with the algorithms outputs. Note change in the scale of y-axis for Algorithms C and E.

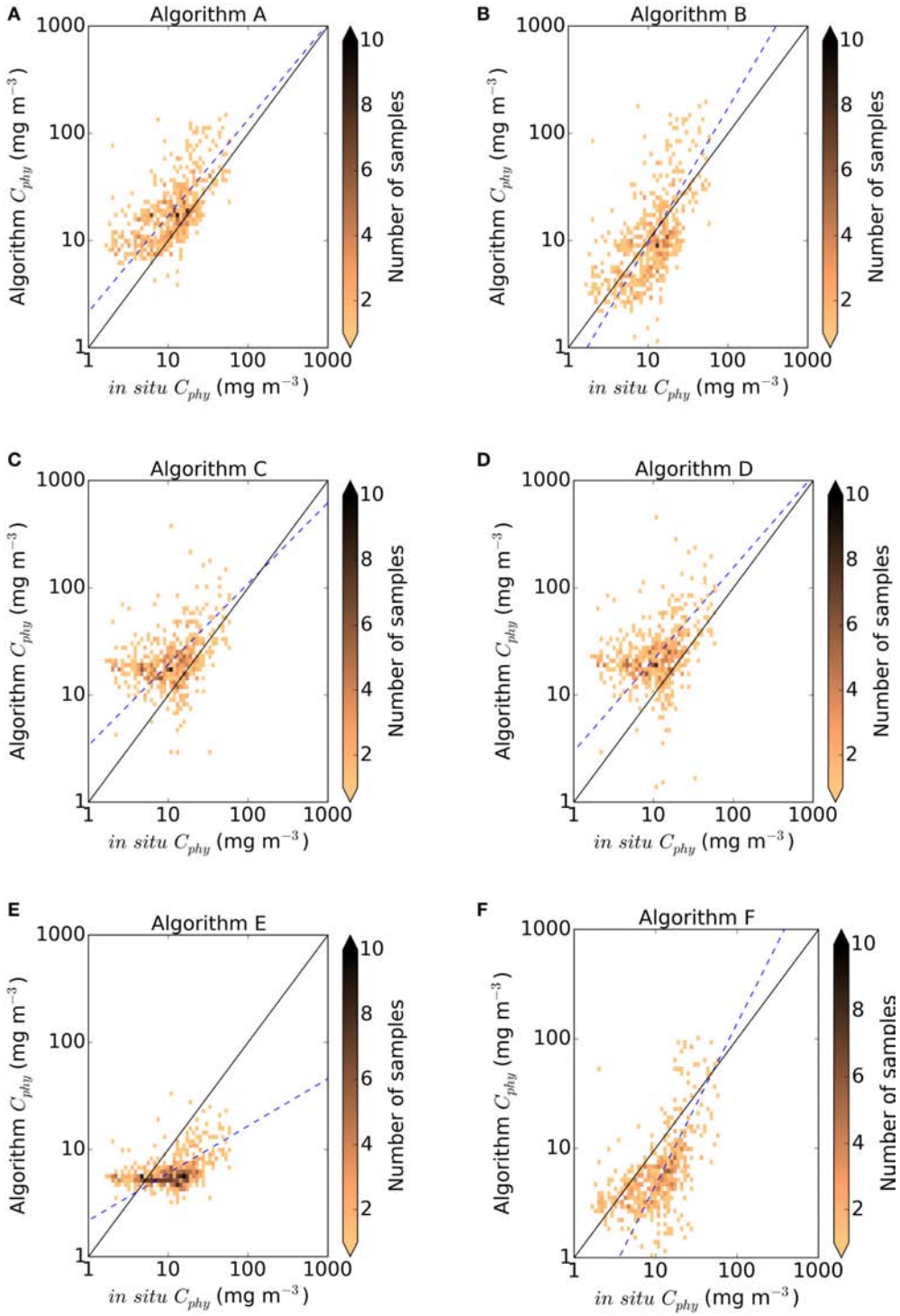


FIGURE 3 | Density plot of *in situ* vs. algorithm C_{phy} from the match-up database ($N = 557$). Black solid line is the 1:1 line and blue dashed line is the Type II linear regression (Reduced Major Axis) fitted to the $\log_{10}$ converted data. **(A)** Sathyendranath et al. (2009), **(B)** Marañón et al. (2014), **(C)** Behrenfeld et al. (2005), **(D)** Martínez-Vicente et al. (2013) modified, **(E)** Kostadinov et al. (2009, 2016), **(F)** Roy et al. (2017).

the overall values reported for the C_{phy}:B ratio are smaller than they would be if the ratio had been derived from B for picoplankton only. However, with a median value of 91 mg C mg Chla^{-1}, the *in situ* C_{phy}-to-satellite-B-ratio falls within or close to observed values in oligotrophic areas. For instance, Sathyendranath et al. (2009) reported average values of this

TABLE 3 | Summary of statistics of algorithm performance (algorithms A–F, columns) for log_{10} and untransformed data ($N = 557$).

Statistic	A	B	C	D	E	F
log_{10} TRANSFORMED DATA						
r_p	**0.61**	**0.61**	0.32	0.31	0.42	0.32
Ψ	0.35 ± 0.02	0.33 ± 0.02	0.43 ± 0.02	0.48 ± 0.02	0.37 ± 0.02	**0.57 ± 0.03**
δ	0.23 ± 0.02	**−0.02 ± 0.03**	0.28 ± 0.03	0.33 ± 0.03	−0.24 ± 0.02	−0.32 ± 0.04
Δ	0.27 ± 0.01	0.32 ± 0.02	0.33 ± 0.02	0.35 ± 0.02	0.29 ± 0.01	**0.47 ± 0.03**
S	0.90 ± 0.05	1.27 ± 0.07	0.75 ± 0.17	**0.86 ± 0.21**	0.44 ± 0.05	1.47 ± 0.44
I	0.34 ± 0.04	**−0.29 ± 0.05**	0.53 ± 0.04	0.48 ± 0.05	0.34 ± 0.02	−0.81 ± 0.10
UNTRANSFORMED DATA						
r_S	**0.58**	**0.54**	0.36	0.37	0.49	0.52
Ψ	19.9 ± 1.37	22.2 ± 1.63	24.6 ± 1.75	32.6 ± 2.34	**11.3 ± 0.67**	13.0 ± 0.83
δ	9.55 ± 1.45	**3.55 ± 1.82**	10.8 ± 1.84	15.0 ± 2.41	−7.04 ± 0.74	−4.76 ± 1.01
Δ	17.4 ± 1.16	21.8 ± 1.54	22.1 ± 1.62	29.0 ± 2.16	**8.86 ± 0.51**	12.1 ± 0.83
S	1.76	2.16	1.77	3.87	2.29	**0.25**
I	2.65	−8.42	**1.57**	−10.9	3.46	3.62
MAPD ± IQR/2	55.1 ± 70.1	**46.1** ± 19.4	78.9 ± 81.4	99.9 ± 98.8	55.1 ± **18.0**	53.2 ± 18.6

Statistics provided have uncertainty estimates (95% confidence interval), from 1,000 bootstrap realizations (See section 2.4). Bold numbers are the best results for each statistic: highest value for r_s and r_p, lowest for Ψ, δ, Δ, I and MAPD, and closest to one for S.

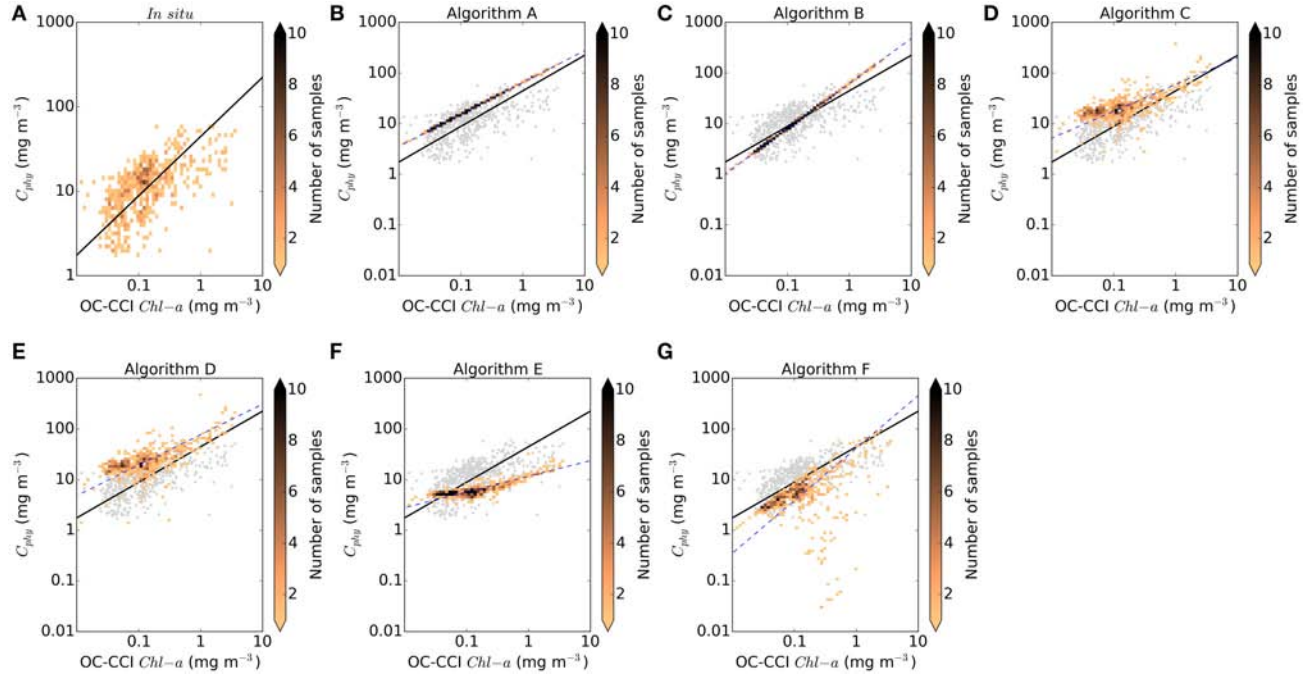


FIGURE 4 | Density plot of covariance between OC-CCI standard chlorophyll product and C_{phy} from the match-up database ($N = 557$). **(A)** *In situ* C_{phy}, **(B)** Sathyendranath et al. (2009), **(C)** Marañón et al. (2014), **(D)** Behrenfeld et al. (2005), **(E)** Martínez-Vicente et al. (2013) modified, **(F)** Kostadinov et al. (2009, 2016), **(G)** Roy et al. (2017). Blue dashed line is the Type II linear regression (Reduced Major Axis) fitted to the $\log_{10}$ converted data. Black solid line is the regression line between chlorophyll and *in situ* for comparison with the regressions by the algorithms. For comparison, data in **(A)** are repeated (gray) in the other panels.

ratio greater than 100 mg C mg Chla^{-1} for prymnesiophytes, cyanobacteria and *Prochlorococcus* sp. Direct observations of pico and nano plankton carbon in the Northern and Southern Atlantic gyres produced carbon-to-chlorophyll ratio estimates, on average, of 106 and 190 mg C mg Chla^{-1}, respectively (Graff et al., 2015). Marañón et al. (2014) also reports values in the range of 80–117 mg C mg Chla^{-1} for oligotrophic regions. Therefore, the C_{phy}:B ratio used as reference in this study compares well with existing observations reported in the literature, despite the mismatch between the phytoplankton assemblages considered. These data are repeated as the background of the other panels in **Figure 4** for reference along with their corresponding regression

TABLE 4 | Summary of statistics for the C_{phy} to B relationships.

Algorithm	Median C_{phy}:B ± IQR/2 (mg C/mg Chla)	r_p	$\alpha \times B + \beta$
In situ	91.3 ± 47.3	0.6	$22.1 \times B + 7.94$
A	143 ± 34.1	N.A.	N.A.
B	78.5 ± 5.4	N.A.	N.A.
C	162 ± 102	0.7	$53.4 \times B + 10.7$
D	180 ± 115	0.7	$70.2 \times B + 10.5$
E	47.9 ± 29.5	0.8	$6.70 \times B + 4.85$
F	52.5 ± 18.6	0.4	$30.5 \times B + 1.03$

C_{phy} was computed from the algorithms and B from the standard OC-CCI chlorophyll product for the match-up dataset (N = 557). Pearson correlation (r_p) was computed for log_{10} converted data. Type II linear regression (Reduced Major Axis) was computed for untransformed data, with dependent variable C_{phy} and independent variable is B.

line. Algorithms A and B are chlorophyll-based, therefore their predictions fall on the line of the equations used, respectively Equations 2, 3 in **Figures 4B,C**. Their predictions are close to the *in situ* data cloud and the resulting median phytoplankton carbon-to-chlorophyll ratio encompass the *in situ* reference value, Algorithm A providing an upper limit, as expected from the assumptions made in its construction. Backscattering-based algorithms (Algorithms C and D, in **Figures 4D,E**, respectively) overestimate the C_{phy}:B reference relationship, specially at the lower concentrations of chlorophyll, and produce median phytoplankton carbon-to-chlorophyll ratio values up to two times greater than the reference. However these algoritms capture some of the variability around the prediction line, which is in the same order of the *in situ* data. Algorithms using inherent optical properties with allometric conversions (Algorithms E and F, in **Figures 4F,G**, respectively) underestimate the reference C_{phy}:B reference relationship, with Algorithm E showing a narrower distribution of data points than Algorithm F.

The *in situ* C_{phy} dataset is representative only of a fraction of the particle population (picophytoplankton). However, its geographical distribution, the median C_{phy} concentrations and carbon-to-chlorophyll ratios derived from this dataset are in agreement with existing observations in oligotrophic oceanic conditions. Taking into account this charcteristic of the dataset, the overall performance of the algorithms was on the low side, with chlorophyll-based algorithms producing slightly lower bias and allometric algorihms slightly lower dispersion. Among algorithms there was a median dispersion of a factor 4 between minimum and maximum predictions. The algorithms were also tested on their ability to produce realistic a C_{phy}:B ratio, which again highlighted great dispersion in predictions among and within algorithms types. Arguably, part of the dispersion in the statistical results may have arised from the fact that the *in situ* C_{phy} dataset is representative only of a fraction of the particle population (i.e., picophytoplankton). Therefore if we limited the study to the optical cases where picophytoplankon is expected to dominate the phytoplankton carbon pool and the chlorophyll content, we would expect an improvement on the results from the algorithms. In the next part of this study we present results obtained from segregating the data into optical water types.

3.2. Algorithm Comparisons for Individual Optical Water Types

An optical water class, in this context, is defined by a mean remote sensing reflectance spectrum representative of particular optical characteristics, i.e., an end-member spectrum. Each extracted satellite pixel coinciding with a match-up *in situ* data has contributions from the different end-members in different proportions. The water class contributing with the largest proportion to the pixel water class membership is classed as belonging to that water class (Jackson et al., 2017). The geographic locations of the match-up points per water class and the number of observations per class are shown in **Figure 5**. There was a good correspondence between the geographic location of the optical water types (note that the classes are numbered such that $i = 1$ is the most oceanic type and $i = 14$ the most coastal type) in **Figure 5A** and the C_{phy} concentrations (**Figure 1**), such that the higher-numbered optical classes tend to be representative of higher concentrations of phytoplankton carbon. The majority of the data (63%), though can be classed as representative of an oligotrophic environment (i 1 to 6, $B < 0.15$ mg Chla m^{-3}). A boxplot summary of the descriptive statistics per optical water class per algorithm is provided (**Figure 6**). **Table 5** summarizes the median C_{phy} values highlighting a steady increase from oceanic to coastal waters, except for $i = 13$, which only has $N = 5$ observations, and is hereafter discarded from the analysis. The magnitude and the increase of the picophytoplankton carbon is in agreement with oceanic and coastal data (Tarran et al., 2006). For instance, using the current carbon conversion parameters on existing abundance data (Tarran and Bruun, 2015), C_{phy} median in the coastal area of the Western English Channel is 12.1 ± 6.1 mg C m^{-3} ($N = 68$).

Algorithm performance per water class was quantified by the signed bias (δ) and the center-pattern RMSD (Δ) (see **Table 5** and **Figure 7**). These statistical indicators of performance improved by limiting the analysis to oligotrophic waters (i 1 to 6) as expected: bias, δ, was similar or lower than those obtained when considering all data, for all the algorithms (section 3.1, **Table 3**). Center-pattern RMSD, Δ an indicator of precision, was, on average, half that when considering all data, indicating decreased noise in the retrieval for all algorithms (for non-log results, section 3.1). Chlorophyll-based algorithms had, on average, similar δ and Δ to the allometric-based algorithms, with back-scattering based algorithms producing higher bias and higher Δ values.

Mesotrophic waters ($0.15 < B < 0.7$ mg Chla m^{-3}, or optical water classes 7–10) comprise 32% of data. With respect to the results obtained for all data available (section 3.1), the chlorophyll based algorithms had an increased δ, whereas Δ was similar. Backscattering-based algorithms had higher uncertainties than chlorophyll-based algorithms in the more turbid waters (high optical class numbers) in the untransformed data (**Figure 7D**). Bias increased for all of the algorithms for these water classes except for Algorithm E, which remained negative and relatively constant at −60%. However, the results for mesotrophic and more turbid waters, should be taken with caution as the use of *in situ* C_{phy} data comprising only picophytoplankton is more problematic in these waters.

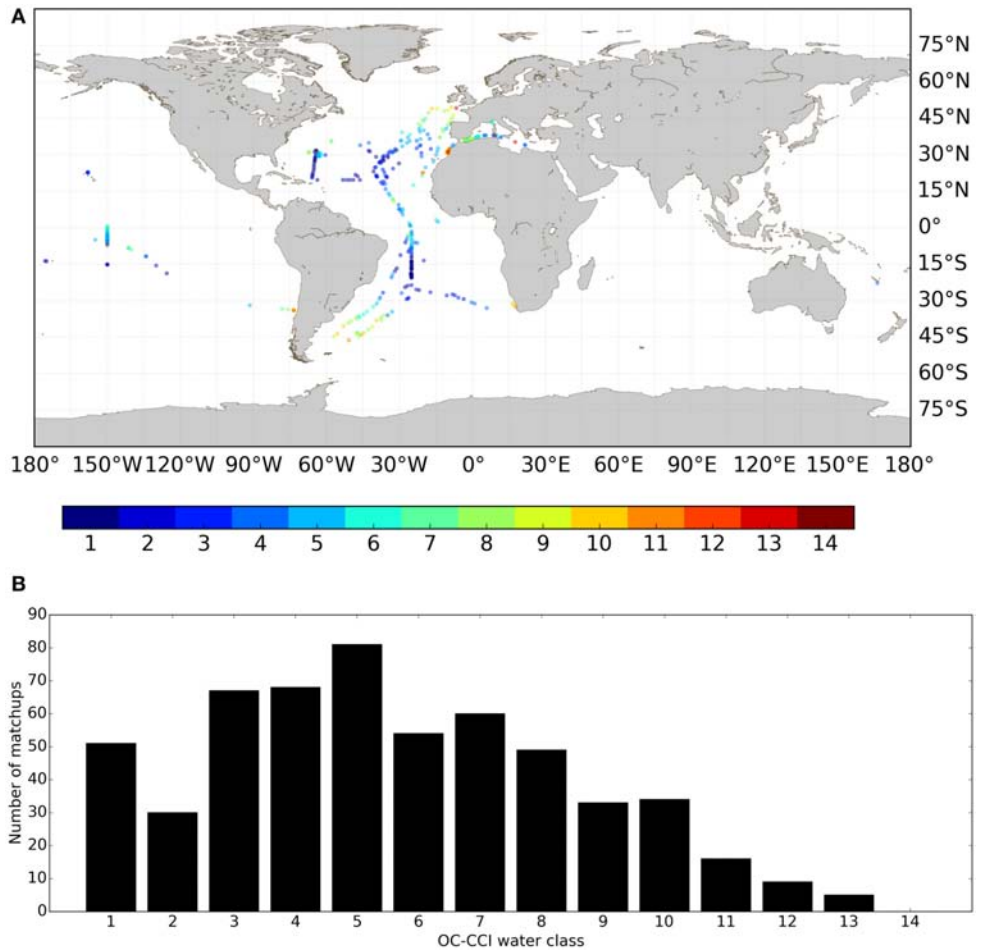


FIGURE 5 | Distribution of the C_{phy} from the match-up dataset per optical water class. **(A)** Geographical distribution. **(B)** Number of data per optical water class.

TABLE 5 | Median C_{phy} from *in situ* matchups and non-log average bias (δ) and un-biased RMSD (Δ) of the different algorithms (algorithms A–F, columns headers) per water class number (i) in mg C m^{-3}, alongside with the number of observations per water class (N_i).

Water class	Median C_{phy} (mg C m^{-3})	N_i	A		B		C		D		E		F	
i			δ	Δ	δ	Δ	δ	Δ	δ	Δ	δ	Δ	δ	Δ
1	4.8	51	2.6	4.0	−2.3	3.6	13.3	7.0	15.5	8.2	−0.2	3.5	−2.8	3.4
2	7.5	30	1.2	3.9	−3.8	3.7	9.1	6.2	10.7	7.2	−2.3	3.8	−4.1	3.9
3	6.4	67	3.2	3.3	−2.5	3.3	10.1	4.2	12.0	4.7	−1.8	3.4	−3.3	3.3
4	10.5	68	3.1	5.7	−3.6	5.5	10.3	10.3	13.1	12.2	−4.9	5.5	−5.3	4.9
5	12.9	81	3.6	6.3	−3.9	6.1	7.1	10.4	9.6	12.3	−7.2	5.9	−6.7	5.2
6	13.6	54	4.2	8.3	−3.7	8.2	5.1	11.8	7.7	13.6	−8.9	8.2	−7.3	6.9
7	13.8	60	6.6	5.7	−1.7	5.5	3.3	8.5	5.5	10.2	−8.8	5.3	−7.3	5.4
8	15.9	49	9.5	8.3	0.6	8.3	4.4	8.8	8.0	9.9	−11.7	10.2	−10.0	8.0
9	20.4	33	18.4	14.5	11.1	16.1	8.9	13.9	16.2	16.8	−16.4	14.3	−8.4	13.6
10	21.1	34	36.0	23.6	33.3	30.2	14.1	16.9	22.9	21.2	−13.0	11.4	4.5	28.8
11	16.7	16	44.3	33.0	45.3	46.8	12.6	24.2	20.6	32.3	−8.8	10.2	−5.0	22.4
12	34.0	9	79.1	15.7	103.2	25.5	99.8	46.2	141	59.5	−14.3	12.0	31.8	30.1
13	6.8	5	19.1	19.2	11.9	21.1	113.1	119.8	144.8	157.4	4.2	12.6	−3.3	8.8

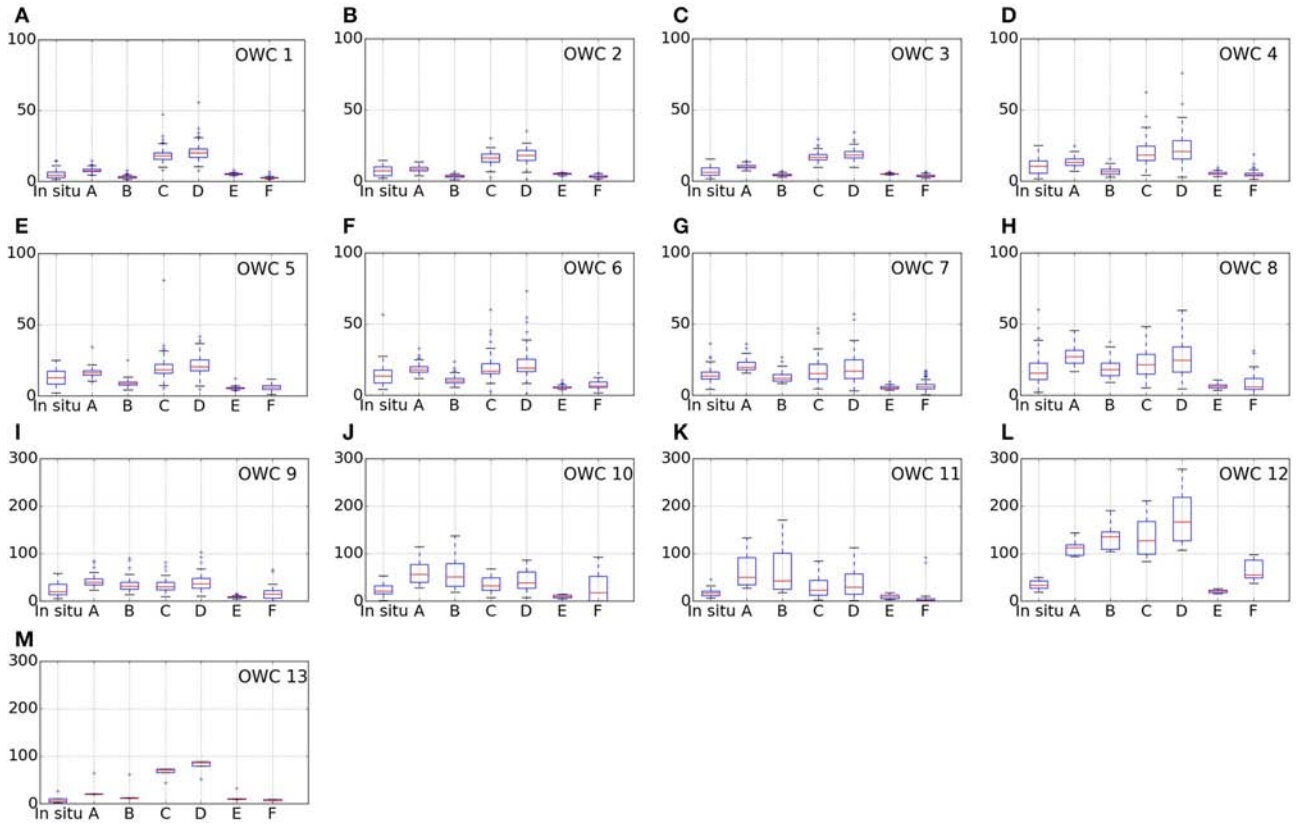


FIGURE 6 | Box-whiskers plots of C_{phy} (in mg C m^{-3}) for the optical water classes (OWC in the figures) 1–13 corresponding to sub-figures **(A–M)** respectively. The output of each algorithm **(A–F)** is compared to the *in situ* measurement for each optical water type. Note change of vertical scale in plots **(I–M)**, corresponding to optical water classes greater than 9. Box, Inter Quartile Range (IQR); red line, median; whiskers, ±1.5×IQR.

The *in situ* median C_{phy}:B ratio by optical water class were also compared to the median C_{phy}:B ratio from the algorithms (**Figure 8**). The *in situ* data (solid gray line and error bars) is repeated as a reference throughout **Figure 8**, showing that for i from 1 to 6 (oligotrophic waters), the range of variation is narrow (106 to 165 mg C mg Chla^{-1}, median 133 mg C mg Chla^{-1}) and error bars are overlapping among the optical water classes 1 to 6. For mesotrophic waters (i from 7 to 10), there was a decreasing tendency of the *in situ* C_{phy}:B ratio.

The analysis of the algorithm predictions of the C_{phy}:B ratio focusses on the optical water classes 1 to 6, where C_{phy} and B are expected to describe the same phytoplankton assemblage. Essentially, the behavior observed for the C_{phy} is also repeated here. Algorithm A was an upper limit to the C_{phy}, it is also an upper limit to C_{phy}:B ratio, which decreases with increasing optical water class number (**Figure 8A**). Algorithm B shows relatively little variation of the median C_{phy}:B ratio for the optical water classes of interest and beyond ($i > 6$). Backscattering-based algorithms, Algorithm C (Behrenfeld et al., 2005) and Algorithm D (Martínez-Vicente et al., 2013), showed also decreasing median C_{phy}:B ratio with increasing water class number, with large overestimations with respect to the *in situ* C_{phy}:B ratio at the clearest waters. The allometric-based algorithms, Algorithm E (Kostadinov et al., 2016) and Algorithm F (Roy et al., 2017), were generally predicting lower than observed C_{phy}, and also predicted lower median C_{phy}:B ratios. However, both algorithms had a decreasing tendency for the median values with increasing turbidity (or water class number, in this study), with Algorithm E being closest to observations for i from 1 to 3.

Finally, **Figure 9** shows an example of the C_{phy} product from algorithms A to D using the OC-CCI monthly product from May 2005. All algorithms reproduce the broad patterns that would be associated with C_{phy} i.e., increased values in high-chlorophyll areas (upwelling sites and coastal regions) and lower concentrations in the gyres, however the salient point of this Figure is the large differences in predictions among the algorithms, as expected from the statistical results.

4. DISCUSSION

4.1. The Picophytoplankton C Match-up Dataset

This study has compiled a large *in situ* database of picophytoplankton carbon, building on a combination of a substantial pre-existing effort by the modeling community (the

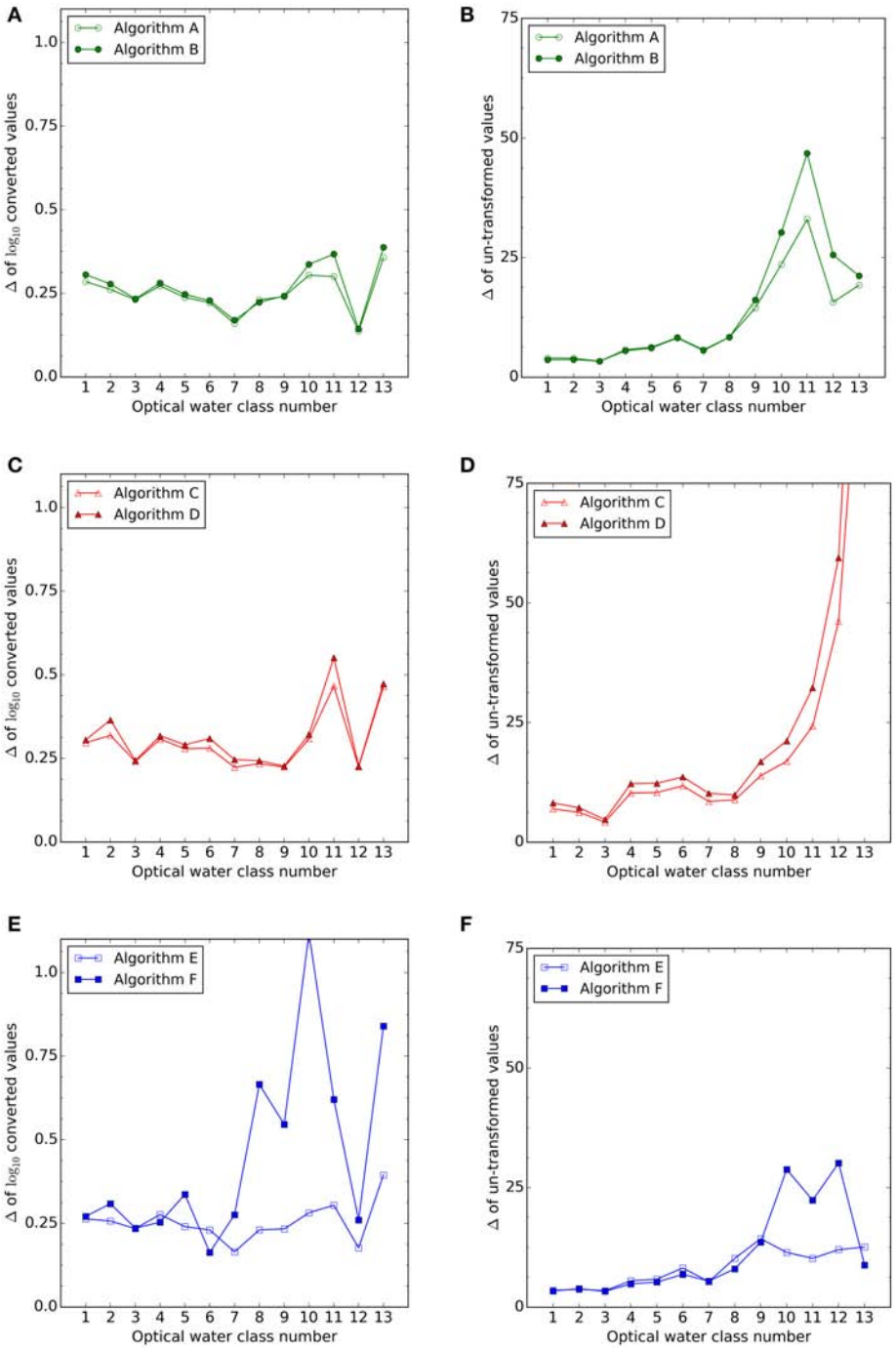


FIGURE 7 | Center pattern root mean square deviation (Δ) for $\log_{10}$ **(A,C,E)** and non-transformed **(B,D,F)** data from the C_{phy} match-up dataset per optical water class. Out of scale from the plot in **(D)** Algorithm C for water class 13 has $\Delta = 120$ mg C m^{-3}; Algorithm D for water class 13 has $\Delta = 157$ mg C m^{-3}. **(E)** Algorithm F for water class 10 has $\Delta = 1.12$ mg C m^{-3}.

MAREDAT data) and long time series observation programmes in the open ocean (Atlantic Meridional Transect, AMT). The ambition is to see this dataset growing with time, as new data are incorporated.

There are a number of advantages and limitations for this dataset with respect to its use for algorithm testing and validation. An advantage is that only a small fraction of the data have been used for the development of any of the algorithms tested here. Algorithms A, B, C, and E are completely independent of the match-up dataset. Algorithm D as implemented by Martínez-Vicente et al. (2013), was developed using a small subset of the new database, $N < 70$, but the subset included nanoplankton.

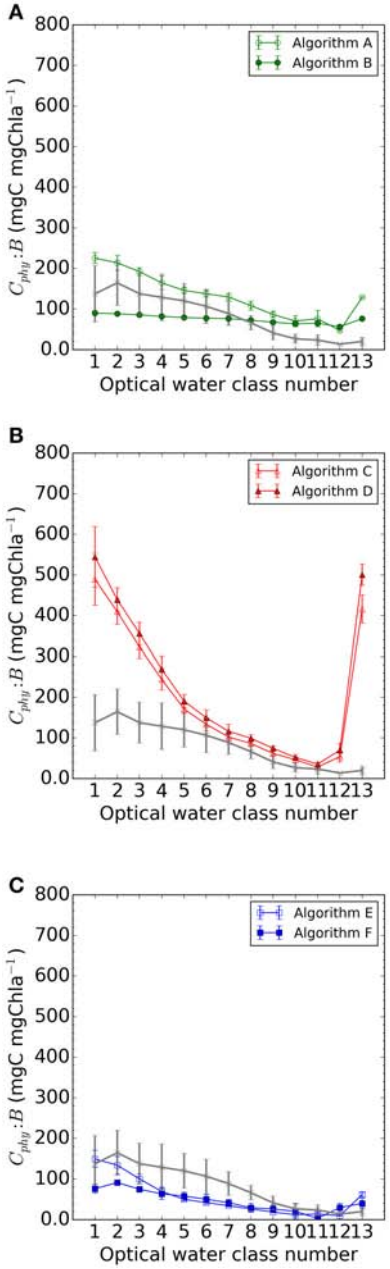


FIGURE 8 | Median C_{phy}:B for each optical water class in the match-up dataset per type of algorithm: **(A)** chlorophyll-based algorithms A and B; **(B)** backscattering-based algorithms C and D and **(C)** allometric algorithms E and F. Error bars are the IQR/2 for the water class. Gray line and error bars in all sub-figures are obtained from the *in situ* dataset.

Algorithm F used MAREDAT for testing algorithm performance, but not for its development. So the data for the validation presented here are mostly independent of the data used in the construction of algorithms. The geographic distribution of the match-up database, though largely limited to the Atlantic, with some additional points from the Pacific, does cover a variety of oceanic regions. It is a purpose-built database for satellite validation studies, and therefore, only an average of the matching data in the top 10 m have been selected for convenience.

However, the dataset also has limitations. One of them is that it is only composed of picophytoplankton: though they are important contributors to the open-ocean phytoplankton biomass, picophytoplankton form a decreasing proportion of the phytoplankton biomass in more productive waters, where larger cells tend to become more important (Marañón et al., 2012; Marañón, 2015). One interesting avenue would be to expand the database with other phytoplankton groups (Buitenhuis et al., 2013; Sal et al., 2013). Nanoplankton can also be counted using flow cytometry, and microphytoplankton groups counted using microscope or automated image processing (Sosik and Olson, 2007; Álvarez et al., 2012), but the relationship between carbon and abundance becomes more variable for larger and more irregularly-shaped phytoplankton (Moberg and Sosik, 2012; Saccà, 2016).

Because the data are confined to the surface layers, they may also be adversely affected by underestimation of *Prochlorococcus* sp. abundance by flow cytometry because of extremely low fluorescence per cell (Partensky et al., 1996; Heywood et al., 2006). Furthermore, by limiting our dataset to the top 10 m, we have precluded the possibility of testing any potential impact that the vertical structure in the first optical depth might have on algorithm performance. Examples exist in the literature where the depth variations in particulate organic carbon been taken into account (Duforêt-Gaurier et al., 2010), and this may be an avenue worth exploring also for phytoplankton carbon algorithms.

Finally, the conversion of abundance to carbon using estimates from the laboratory could cause errors in the computed C_{phy} in the field, if the laboratory estimates do not hold under natural environmental conditions. These factors can vary with physiological state and with depth (Casey et al., 2013) and have been discussed previously (Buitenhuis et al., 2012; Martínez-Vicente et al., 2013). It is important to highlight that in this study we have used indirect estimates of phytoplankton carbon (through cell counts and cell size) only because of the lack of direct measurements. However, methods for direct quantification of C_{phy} have recently become available (Graff et al., 2012; Casey et al., 2013) and the expectation is that there will be more direct C_{phy} data available in the future.

Limiting the study to the optical water classes where the composition of the phytoplankton assemblage is expected to be dominated by picoplankton (i 1 to 6, B up to 0.15 mg Chl m^{-3}) median values of C_{phy} and C_{phy}:B ratio matching the literature (Marañón et al., 2001; Marañón, 2015), were obtained, and the algorithms results showed more stability and less dispersion. However, some algorithms displayed significant differences which are discussed hereafter.

4.2. Algorithm Comparison by Type: Chlorophyll Based, Backscattering Based and Allometric

The algorithms presented here were broadly classified into three classes: chlorophyll-based, back-scattering based and allometric. In the results presented here, it is evident that algorithms based

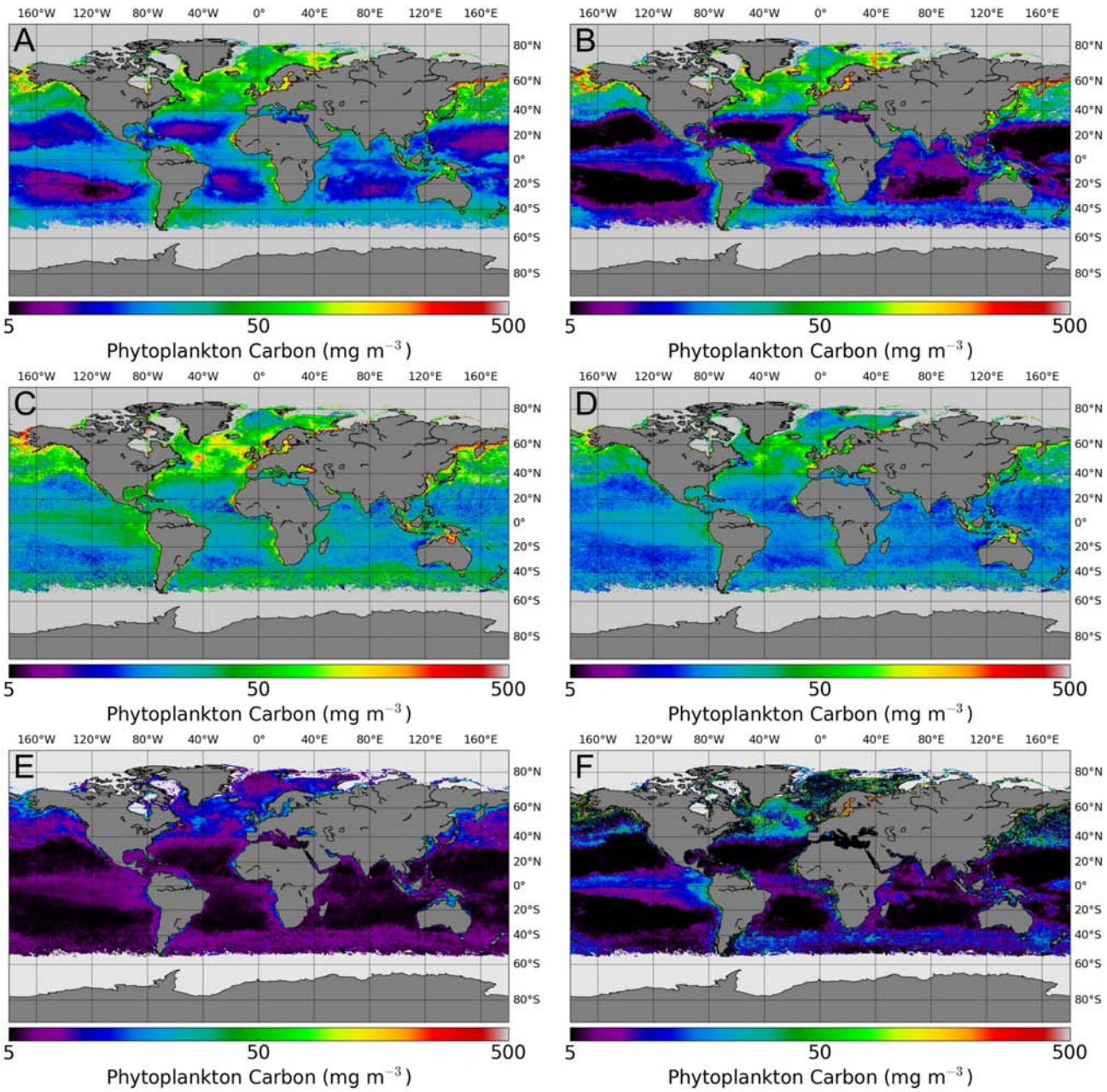


FIGURE 9 | An example of the C_{phy} product for May 2005, estimated using each of the algorithms **(A–F)** applied to the monthly OC-CCI data. Black color in the gyres indicate values close or below 5 mg C m^{-3}, light gray indicates invalid retrieval or unavailable input data.

on similar approaches perform alike. So it is worth examining each of these approaches in some detail, to explain the differences observed in the results for oligotrophic waters.

The two chlorophyll-based algorithms were designed to consider all the phytoplankton groups: Algorithm A (Sathyendranath et al., 2009) provides an upper limit to the phytoplankton contribution to the total particulate carbon pool and Algorithm B (Marañón et al., 2014) is based on phytoplankton carbon computed from flow-cytometric data supplemented with microscopic counts for larger phytoplankton. Yet, when compared with only one fraction of the total phytoplankton pool (the picophytoplankton in this study), the results are similar, with Algorithm A slightly overestimating and Algorithm B slightly underestimating *in situ* C_{phy}. Algorithm A was designed as the upper limit to the phytoplankton carbon, hence this result is as expected. Algorithm B was computed using a similar conversion method between cell count and carbon concentration to the one used in this study, but with different conversion coefficients. We speculate that a possible reason for the difference observed between the predicted C_{phy} by Algorithm

B and the *in situ* C_{phy} could be found in the differences in conversion between cell counts and carbon. Production of datasets with consistent conversion factors would help eliminate this source of discrepancy.

Differences in results are greater with the backscattering-based algorithm, which produced overestimation of C_{phy} and greater dispersion statistics than the other algorithms. Because it has been verified *in situ* that b_{bp} increases with pico and nano plankton carbon (Martínez-Vicente et al., 2013; Graff et al., 2015), the degradation of results may be linked to the backscattering input from the algorithm. Algorithm C equation was derived using the GSM algorithm for obtaining a relationship between the backscattering coefficient and the *B*. It has been found *in situ*, that at low *B* there may be an overestimation of the b_{bp} by using this satellite-derived relationship (Huot et al., 2008; Antoine et al., 2011; Brewin et al., 2012), yet *in situ* data also have shown overestimation of backscattering by the QAA algorithm (Behrenfeld et al., 2013). It has been found recently, that Raman scattering plays a role in the discrepancy of the retrievals of the backscattering coefficient in very clear waters from satellite (Westberry et al., 2013), causing an overestimation in backscattering. This effect of Raman scattering has been only identified recently (Lee and Huot, 2014), and was not incorporated in the QAA version used for the production of the OC-CCI version 2 data used for this study, but can be analytically corrected (McKinna et al., 2016), and future versions of the OC-CCI dataset will address this issue.

While validation of the OC-CCI chlorophyll product has been performed intensively in Atlantic oligotrophic waters and showed low error statistics (Brewin et al., 2016), investigations to improve the validation of b_{bp} and to understand better the relationships of b_{bp} with particles in oligotrophic areas are still required (Brewin et al., 2015). At a fundamental level, backscattering is dependent on the refractive index of the cells, their abundance and size, whose interplay is not yet fully understood. But in addition to phytoplankton, other particles (e.g., detritus) are known to contribute to the backscattering signal, and their variability relative to phytoplankton could potentially contribute to the discrepancies observed in this study.

Algorithm E used in its original formulation an inversion model to retrieve the backscattering spectral slope (Loisel et al., 2006) that is different from the one used in this comparison as input (i.e., QAAv5). The QAA used here invoked a band ratio to solve for the backscattering spectral slope, and this may account for at least part of the observed tighter coupling between B and Algorithm E. Algorithm E provides consistently low Δ values across the water classes, although it underestimates the *in situ* data systematically. This may be pointing to a need to re-adjust the size scaling parameter (like N_o), the empirical correction for which is now based on a rather limited *in situ* validation data set (Kostadinov et al., 2016). Ultimately, a better optical closure is needed between modeled and observed backscattering spectra, and a better understanding of the underlying particle assemblages, their refractive indices, and their relative contributions to the backscattering coeffient. It remains to be validated if at more productive waters, the prediction of low C_{phy} (from Algorithm E picophytoplankton fraction) remains valid, whereas the current *in situ* dataset showed an increase (**Table 5**).

Algorithm F was similar to algorithm E in all water classes except for water classes 8 to 11, where the uncertainty is almost double compared with the rest of the algorithms, pointing perhaps to a vulnerability to uncertainties in the $a_{ph}(676)$ retrieval in these waters. More accurate estimates of phytoplankton carbon by Algorithm F would possibly require improving the retrievals of the input $a_{ph}(676)$ values, especially in high-chlorophyll waters.

5. CONCLUSIONS AND FUTURE WORK

Further work is required to extend the *in situ* dataset to include additional phytoplantkon sizes to evaluate if uncertainties can be reduced in the product by including larger phytoplankton to capture phytoplankton dynamics at wider scales. Despite the limitations of the *in situ* data used, it has been shown that where chlorophyll concentrations are less than 0.15 mg Chla m^{-3}, chlorophyll-based algorithms provide the best estimates of C_{phy}, allometric-based algorithms consistently underestimate C_{phy} and backscattering-based algorithms, can produce large overestimations of C_{phy}, at least for the particular case of back-scattering data, provided using the QAA algorithm, as implemented in OC-CCI version 2.0. To improve back-scattering products from satellites, fundamental optical work on explaining the relationship between the b_{bp} and particles in oligotrophic areas is still needed. Satellite-based phytoplankton carbon product, once validated to a level that meets user requirements, and *in situ* datasets, similar to the one presented here, will be useful for validation of marine ecosystem and biogeochemical models at a wider scale (Dutkiewicz et al., 2015).

DATA AVAILABLE

The C_{phy} data, computed from *in situ* phytoplankton counts, and the matching C_{phy} data, computed from different algorithms using OC-CCI version 2.0 satellite data, can be obtained from http://www.zenodo.org with doi: 10.5281/zenodo.1067229.

AUTHOR CONTRIBUTIONS

VM-V: Lead writer, collation of *in situ* database, statistical analysis and production of code for match-up evaluation. HE-K: Production of code for satellite processing and match-up extraction, and writer. RB: Provision of code for statistical analysis based on OC-CCI methodology. TJ: Provision of code for calculation of per optical water class uncertainties, based on the OC-CCI methodology. GD: Provision of *in situ* data. AH: Perspectives on ecosystem modeling. SR: algorithm provider, input in statistical analysis/interpretation. RR and HK: Review of the *in situ* database. TK and EM: Algorithm providers. JG: Data provider. GT: Data provider. TP: Project leader, scientific advice. SS: Co-lead development of the concept, work plan and

guidance. All authors reviewed and provided comments on the draft manuscript.

FUNDING

Several authors (VM-V, SS, TP, SR, HE-K, GD, AH, RR, and HK) were supported by POCO, which is a project funded by the European Space Agency (ESA) under the program of Science Exploitation of Operational Missions (SEOM) following Contract: 4000113692/15/I-LG. This study is a contribution to the international IMBeR project and was supported by the UK Natural Environment Research Council National Capability funding to Plymouth Marine Laboratory and the National Oceanography Centre, Southampton. This is contribution number 319 of the AMT programme. TK was supported on NASA grant # NNX13AC92G and by the Division of Hydrologic Sciences, Desert Research Institute. JG was supported through NASA Grant # NNX10AT70G. Thanks to ESA for sponsoring this publication.

ACKNOWLEDGMENTS

The authors would like to thank Peter Regner at ESA for his support and management of the POCO project. The authors would like to thank Oliver Fisher for his contributions to the project during his student internship. The authors would like to thank the participants of the ESA sponsored Color and Light in the Ocean (CLEO) Workshop for constructive discussions and comments. Also thanks to the providers of the Ocean Color Climate Change Initiative dataset, Version 2.0, European Space Agency, available online at http://www.esa-oceancolour-cci.org and to British Oceanographic Data Centre (BODC) for the provision of the archived AMT and WCO flow cytometry data. VM-V thanks Prof. A. Bracher for directing us to additional data sources and Dr. L. Polimene for useful discussions on the variations of the C_{phy}:B ratio. We thank the two reviewers for their time and comments, which helped to improve significantly the quality of the work.

REFERENCES

Álvarez, E., López-Urrutia, A., and Nogueira, E. (2012). Improvement of plankton biovolume estimates derived from image-based automatic sampling devices: application to FlowCAM. *J. Plankt. Res.* 34, 454–469. doi: 10.1093/plankt/fbs017

Antoine, D., Siegel, D. A., Kostadinov, T. S., Maritorena, S., Nelson, N. B., Gentili, B., et al. (2011). Variability in optical particle backscattering in contrasting bio-optical oceanic regimes. *Limnol. Oceanogr.* 56, 955–973. doi: 10.4319/lo.2011.56.3.0955

Behrenfeld, M. J., Boss, E., Siegel, D. A., and Shea, D. M. (2005). Carbon-based ocean productivity and phytoplankton physiology from space. *Glob. Biogeochem. Cycles* 19:GB1006. doi: 10.1029/2004GB002299

Behrenfeld, M. J., Hu, Y., Hostetler, C. A., Dall'Olmo, G., Rodier, S. D., Hair, J. W., et al. (2013). Space-based lidar measurements of global carbon stocks. *Geophys. Res. Lett.* 40, 4355–4360. doi: 10.1002/grl.50816

Brewin, R. J. W., Dall'Olmo, G., Pardo, S., van Dongen-Vogels, V., and Boss, E. S. (2016). Underway spectrophotometry along the atlantic meridional transect reveals high performance in satellite chlorophyll retrievals. *Remote Sens. Environ.* 183, 82–97. doi: 10.1016/j.rse.2016.05.005

Brewin, R. J. W., Dall'Olmo, G., Sathyendranath, S., and Hardman-Mountford, N. J. (2012). Particle backscattering as a function of chlorophyll and phytoplankton size structure in the open-ocean. *Opt. Exp.* 20, 17632–17652. doi: 10.1364/OE.20.017632

Brewin, R. J. W., Sathyendranath, S., Müller, D., Brockmann, C., Deschamps, P.-Y., Devred, E., et al. (2015). The ocean colour climate change initiative: III. A round-robin comparison on in-water bio-optical algorithms. *Remote Sens. Environ.* 162, 271–294. doi: 10.1016/j.rse.2013.09.016

Buitenhuis, E. T., Li, W. K. W., Vaulot, D., Lomas, M. W., Landry, M. R., Partensky, F., et al. (2012). Picophytoplankton biomass distribution in the global ocean. *Earth Syst. Sci. Data* 4, 37–46. doi: 10.5194/essd-4-37-2012

Buitenhuis, E. T., Vogt, M., Moriarty, R., Bednaršek, N., Doney, S. C., Leblanc, K., et al. (2013). Maredat: towards a world atlas of marine ecosystem data. *Earth Syst. Sci. Data* 5, 227–239. doi: 10.5194/essd-5-227-2013

Casey, J. R., Aucan, J. P., Goldberg, S. R., and Lomas, M. W. (2013). Changes in partitioning of carbon amongst photosynthetic pico- and nano-plankton groups in the Sargasso Sea in response to changes in the North Atlantic Oscillation. *Deep Sea Res. II Top. Stud. Oceanogr.* 93, 58–70. doi: 10.1016/j.dsr2.2013.02.002

CEOS (2014). *Ceos Strategy for Carbon Observations from Space. The Comitee on Earth Observation Satellites (ceos) Response to the Group on Earth Observation (geo) Carbon Strategy*. Technical report, JAXA and IA Corp.

De Mora, L., Butenschön, M., and Allen, J. I. (2016). The assessment of a global marine ecosystem model on the basis of emergent properties and ecosystem function: a case study with ersem. *Geosci. Model Dev.* 9, 59–76. doi: 10.5194/gmd-9-59-2016

Duforêt-Gaurier, L., Loisel, H., Dessailly, D., Nordkvist, K., and Alvain, S. (2010). Estimates of particulate organic carbon over the euphotic depth from *in situ* measurements. Application to satellite data over the global ocean. *Deep Sea Res. I* 57, 351–367. doi: 10.1016/j.dsr.2009.12.007

DuRand, M., Olson, R., and Chisholm, S. (2001). Phytoplankton population dynamics at the Bermuda Atlantic Time-series station in the Sargasso Sea. *Deep Sea Res. II* 48, 1983–2003. doi: 10.1016/S0967-0645(00)00166-1

Dutkiewicz, S., Hickman, A. E., Jahn, O., Gregg, W. W., Mouw, C. B., and Follows, M. J. (2015). Capturing optically important constituents and properties in a marine biogeochemical and ecosystem model. *Biogeosciences* 12, 4447–4481. doi: 10.5194/bg-12-4447-2015

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Ann. Stat.* 7, 1–26. doi: 10.1214/aos/1176344552

Efron, B., and Tibshirani, R. J. (1993). *An Introduction to the Bootstrap*. Boca Raton, FL: Chapman and Hall/CRC.

Evers-King, H., Martinez-Vicente, V., Brewin, R. J. W., Dall'Olmo, G., Hickman, A. E., Jackson, T., et al. (2017). Validation and intercomparison of ocean color algorithms for estimating particulate organic carbon in the oceans. *Front. Mar. Sci.* 4:251. doi: 10.3389/fmars.2017.00251

Garver, S. A., and Siegel, D. A. (1997). Inherent optical property inversion of ocean color spectra and its biogeochemical interpretation .1. Time series from the sargasso sea. *J. Geophys. Res. Oceans* 102, 18607–18625. doi: 10.1029/96JC03243

Graff, J. R., Milligan, A. J., and Behrenfeld, M. J. (2012). The measurement of phytoplankton biomass using flow-cytometric sorting and elemental analysis of carbon. *Limnol. Oceanogr. Methods* 10, 910–920. doi: 10.4319/lom.2012.10.910

Graff, J. R., Westberry, T. K., Milligan, A. J., Brown, M. B., Dall'Olmo, G., Dongen-Vogels, V. V., et al. (2015). Analytical phytoplankton carbon measurements spanning diverse ecosystems. *Deep Sea Res. I Oceanogr. Res. Papers* 102, 16–25. doi: 10.1016/j.dsr.2015.04.006

Heywood, J. L., Zubkov, M. V., Tarran, G. A., Fuchs, B., and Holligan, P. (2006). Prokaryoplankton standing stocks in oligotrophic gyre and equatorial provinces of the atlantic ocean: evaluation of inter-annual variability. *Deep Sea Res. II* 53, 1530–1547. doi: 10.1016/j.dsr2.2006.05.005

Huot, Y., Morel, A., Twardowski, M., Stramski, D., and Reynolds, R. A. (2008). Particle optical backscattering along chlorophyll gradient in the upper layer of the eastern south pacific ocean. *Biogeosciences* 5, 495–507. doi: 10.5194/bg-5-495-2008

Jackson, T., Sathyendranath, S., and Mélin, F. (2017). An improved optical classification scheme for the ocean colour essential climate variable and its applications. *Remote Sens. Environ.* 203, 152–161. doi: 10.1016/j.rse.2017.03.036

Kostadinov, T. S., Cabré, A., Vedantham, H., Marinov, I., Bracher, A., Brewin, R. J. W., et al. (2017). Inter-comparison of phytoplankton functional type phenology metrics derived from ocean color algorithms and earth system models. *Remote Sens. Environ.* 190, 162–177. doi: 10.1016/j.rse.2016.11.014

Kostadinov, T. S., Milutinović, S., Marinov, I., and Cabré, A. (2016). Carbon-based phytoplankton size classes retrieved via ocean color estimates of the particle size distribution. *Ocean Sci.* 12, 561–575. doi: 10.5194/os-12-561-2016

Kostadinov, T. S., Siegel, D. A., and Maritorena, S. (2009). Retrieval of the particle size distribution from satellite ocean color observations. *J. Geophys. Res.* 114:C09015. doi: 10.1029/2009JC005303

Le Quéré, C., Harrison, S. P., Prentice, I. C., Buitenhuis, E. T., Aumont, O., Bopp, L., et al. (2005). Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models. *Glob. Change Biol.* 11, 2016–2040.

Lee, Z., and Huot, Y. (2014). On the non-closure of particle backscattering coefficient in oligotrophic oceans. *Opt. Exp.* 22, 29223–29233. doi: 10.1364/OE.22.029223

Lee, Z., Weidemann, A., Kindle, J., Arnone, R., Carder, K. L., and Davis, C. (2007). Euphotic zone depth: its derivation and implication to ocean-color remote sensing. *J. Geophys. Res. Oceans* 112:C03009. doi: 10.1029/2006JC003802

Lee, Z. P., Carder, K. L., and Arnone, R. A. (2002). Deriving inherent optical properties from water color: a multiband quasi-analytical algorithm for optically deep waters. *Appl. Opt.* 41, 5755–5772. doi: 10.1364/AO.41.005755

Loisel, H., Nicolas, J., Sciandra, A., Stramski, D., and Poteau, A. (2006). Spectral dependency of optical backscattering by marine particles from satellite remote sensing of the global ocean. *J. Geophys. Res. Oceans* 111:C09024.

Loisel, H., and Stramski, D. (2000). Estimation of the inherent optical properties of natural waters from the irradiance attenuation coefficient and reflectance in the presence of raman scattering. *Appl. Opt.* 39, 3001–3011. doi: 10.1364/AO.39.003001

Marañón, E., Cermeño, P., Huete-Ortega, M., López-Sandoval, D. C., Mouriño-Carballido, B., and Rodríguez-Ramos, T. (2014). Resource supply overrides temperature as a controlling factor of marine phytoplankton growth. *PLoS ONE* 9:e99312. doi: 10.1371/journal.pone.0099312

Marañón, E., Cermeño, P., Latasa, M., and Tadonléké, R. D. (2012). Temperature, resources, and phytoplankton size structure in the ocean. *Limnol. Oceanogr.* 57, 1266–1278. doi: 10.4319/lo.2012.57.5.1266

Marañón, E. (2015). Cell size as a key determinant of phytoplankton metabolism and community structure. *Ann. Rev. Mar. Sci.* 7, 241–264. doi: 10.1146/annurev-marine-010814-015955

Marañón, E., Holligan, P., Barciela, R., González, N., Mouriño, B., Pazó, M. J., et al. (2001). Patterns of phytoplankton size structure and productivity in contrasting open-ocean environments. *Mar. Ecol. Prog. Series* 216, 43–56. doi: 10.3354/meps216043

Maritorena, S., Siegel, D. A., and Peterson, A. R. (2002). Optimization of a semianalytical ocean color model for global-scale applications. *Appl. Opt.* 41, 2705–2714. doi: 10.1364/AO.41.002705

Martínez-Vicente, V., Dall'Olmo, G., Tarran, G., Boss, E., and Sathyendranath, S. (2013). Optical backscattering is correlated with phytoplankton carbon across the atlantic ocean. *Geophys. Res. Lett.* 40, 1–5.

McKinna, L. I. W., Werdell, P. J., and Proctor, C. W. (2016). Implementation of an analytical raman scattering correction for satellite ocean-color processing. *Opt. Exp.* 24, A1123–A1137. doi: 10.1364/OE.24.0A1123

Menden-Deuer, S., and Lessard, E. J. (2000). Carbon to volume relationships for dinoflagellates, diatoms, and other protist plankton. *Limnol. Oceanogr.* 45, 569–579. doi: 10.4319/lo.2000.45.3.0569

Moberg, E. A., and Sosik, H. M. (2012). Distance maps to estimate cell volume from two-dimensional plankton images. *Limnol. Oceanogr.* 10, 278–288. doi: 10.4319/lom.2012.10.278

Moore, T. S., Campbell, J. W., and Dowell, M. D. (2009). A class-based approach to characterizing and mapping the uncertainty of the modis ocean chlorophyll product. *Remote Sens. Environ.* 113, 2424–2430. doi: 10.1016/j.rse.2009.07.016

Moore, T. S., Campbell, J. W., and Hui, F. (2001). A fuzzy logic classification scheme for selecting and blending satellite ocean color algorithms. *IEEE Trans. Geosci. Remote Sens.* 39, 1764–1776. doi: 10.1109/36.942555

Mouw, C. B., Hardman-Mountford, N. J., Alvain, S., Bracher, A., Brewin, R. J. W., Bricaud, A., et al. (2017). A consumer's guide to satellite remote sensing of multiple phytoplankton groups in the global ocean. *Front. Mar. Sci.* 4:41. doi: 10.3389/fmars.2017.00041

O'Reilly, J. E., Maritorena, S., Mitchell, B. G., Siegel, D. A., Carder, K. L., Garver, S. A., et al. (1998). Ocean color chlorophyll algorithms for seawifs. *J. Geophys. Res. Oceans* 103, 24937–24953.

Oubelkheir, K., Claustre, H., Sciandra, A., and Babin, M. (2005). Bio-optical and biogeochemical properties of different trophic regimes in oceanic waters. *Limnol. Oceanogr.* 50, 1795–1809. doi: 10.4319/lo.2005.50.6.1795

Partensky, F., Blanchot, J., Lantoine, F., Neveux, J., and Marie, D. (1996). Vertical structure of picophytoplankton at different trophic sites of the tropical northeastern atlantic ocean. *Deep Sea Res. I Oceanogr. Res. Papers* 43, 1191–1213. doi: 10.1016/0967-0637(96)00056-8

Platt, T., and Sathyendranath, S. (2008). Ecological indicators for the pelagic zone of the ocean from remote sensing. *Remote Sens. Environ.* 112, 3426–3436. doi: 10.1016/j.rse.2007.10.016

Redalje, D. G., and Laws, E. A. (1981). A new method for estimating phytoplankton growth rates and carbon biomass. *Mar. Biol.* 62, 73–79. doi: 10.1007/BF00396953

Roy, S., Sathyendranath, S., Bouman, H., and Platt, T. (2013). The global distribution of phytoplankton size spectrum and size classes from their light-absorption spectra derived from satellite data. *Remote Sens. Environ.* 139, 185–197. doi: 10.1016/j.rse.2013.08.004

Roy, S., Sathyendranath, S., and Platt, T. (2017). Size-partitioned phytoplankton carbon and carbon-to-chlorophyll ratio from ocean colour by an absorption-based bio-optical algorithm. *Remote Sens. Environ.* 194, 177–189. doi: 10.1016/j.rse.2017.02.015

Saccà, A. (2016). A simple yet accurate method for the estimation of the biovolume of planktonic microorganisms. *PLoS ONE* 11:e0151955. doi: 10.1371/journal.pone.0151955

Sal, S., López-Urrutia, A., Irigoien, X., Harbour, D. S., and Harris, R. P. (2013). Marine microplankton diversity database. *Ecology* 94, 1658–1658. doi: 10.1890/13-0236.1

Sathyendranath, S., Brewin, B., Mueller, D., Doerffer, R., Krasemann, H., Mélin, F., et al. (2012). "Ocean colour climate change initiative: approach and initial results," in *2012 IEEE International Geoscience and Remote Sensing Symposium* (Munich), 2024–2027.

Sathyendranath, S., Stuart, V., Nair, A., Oka, K., Nakane, T., Bouman, H., et al. (2009). Carbon-to-chlorophyll ratio and growth rate of phytoplankton in the sea. *Mar. Ecol. Prog. Series* 383, 73–84. doi: 10.3354/meps07998

Siegel, D. A., Maritorena, S., Nelson, N. B., Hansell, D. A., and Lorenzi-Kayser, M. (2002). Global distribution and dynamics of colored dissolved and detrital organic materials. *J. Geophys. Res. Oceans* 107:3228. doi: 10.1029/2001JC000965

Sinclair, M., Keighan, E., and Jones, J. (1979). Atp as a measure of living phytoplankton carbon in estuaries. *J. Fish. Res. Board Canada* 36, 180–186. doi: 10.1139/f79-028

Sosik, H. M., and Olson, R. J. (2007). Automated taxonomic classification of phytoplankton sampled with imaging-in-flow cytometry. *Limnol. Oceanogr. Methods* 5, 204–216. doi: 10.4319/lom.2007.5.204

Stramski, D. (1999). Refractive index of planktonic cells as a measure of cellular carbon and chlorophyll a content. *Deep Sea Res. I Oceanogr. Res. Papers* 46, 335–351. doi: 10.1016/S0967-0637(98)00065-X

Tarran, G. (2015). *Abundance of Phytoplankton, Heterotrophic Nanoflagellates and Bacteria through the Water Column at Time Series Station L4 in the Western English Channel, 2007–2014*. British Oceanographic Data Centre - Natural Environment Research Council, UK. doi: 10.5285/116a5210-2d58-5ceb-e053-6c86abc0d420

Tarran, G. A., and Bruun, J. T. (2015). Nanoplankton and picoplankton in the Western English Channel: abundance and seasonality from 2007–2013. *Prog. Oceanogr.* 137(Pt B), 446–455. doi: 10.1016/j.pocean.2015.04.024

Tarran, G. A., Heywood, J. L., and Zubkov, M. V. (2006). Latitudinal changes in the standing stocks of nano- and picoeukariotic phytoplankton in the Atlantic Ocean. *Deep Sea Res. II* 53, 1516–1529. doi: 10.1016/j.dsr2.2006.05.004

Tarran, G. A., Zubkov, M. V., Sleigh, M. A., Burkill, P. H., and Yallop, M. (2001). Microbial comunity structure and standing stocks in the NE Atlantic in June and July of 1996. *Deep Sea Res. II* 48, 963–985. doi: 10.1016/S0967-0645(00)00104-1

Taylor, B. B., Taylor, M. H., Dinter, T., and Bracher, A. (2013). Estimation of relative phycoerythrin concentrations from hyperspectral underwater radiance measurements—a statistical approach. *J. Geophys. Res. Oceans* 118, 2948–2960. doi: 10.1002/jgrc.20201

Westberry, T., Behrenfeld, M. J., Siegel, D. A., and Boss, E. (2008). Carbon-based primary productivity modeling with vertically resolved photoacclimation. *Glob. Biogeochem. Cycles* 22:GB2024. doi: 10.1029/2007GB003078

Westberry, T. K., Boss, E., and Lee, Z. (2013). Influence of raman scattering on ocean color inversion models. *Appl. Opt.* 52, 5552–5561. doi: 10.1364/AO.52.005552

Zhai, L., Platt, T., Tang, C., Dowd, M., Sathyendranath, S., and Forget, M.-H. (2008). Estimation of phytoplankton loss rate by remote sensing. *Geophys. Res. Lett.* 35:L23606. doi: 10.1029/2008GL035666

Zhai, L., Platt, T., Tang, C., Sathyendranath, S., Fuentes-Yaco, C., Devred, E., et al. (2010). Seasonal and geographic variations in phytoplankton losses from the mixed layer on the northwest atlantic shelf. *J. Mar. Syst.* 80, 36–46. doi: 10.1016/j.jmarsys.2009.09.005

Zubkov, M. V., Sleigh, M. A., Tarran, G. A., Burkill, P. H., and Leakey, J. G. (1998). Picoplanktonic community structure on an Atlantic transect from 50 N to 50 S. *Deep Sea Res. I* 45, 1339–1355. doi: 10.1016/S0967-0637(98)00015-6

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer SA and handling Editor declared their shared affiliation.

Copyright © 2017 Martínez-Vicente, Evers-King, Roy, Kostadinov, Tarran, Graff, Brewin, Dall'Olmo, Jackson, Hickman, Röttgers, Krasemann, Marañón, Platt and Sathyendranath. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Marine Science

ORIGINAL RESEARCH
published: 12 December 2017
doi: 10.3389/fmars.2017.00386

Comparison of Seasonal Cycles of Phytoplankton Chlorophyll, Aerosols, Winds and Sea-Surface Temperature off Somalia

Muhammad Shafeeque[1,2], *Shubha Sathyendranath*[3]*, *Grinson George*[1], *Alungal N. Balchand*[2] *and Trevor Platt*[1,4]

[1] Fishery Resources Assessment Division, Central Marine Fisheries Research Institute, Kochi, India, [2] School of Marine Sciences, Cochin University of Science and Technology, Kochi, India, [3] National Centre for Earth Observation, Plymouth Marine Laboratory, Plymouth, United Kingdom, [4] Plymouth Marine Laboratory, Plymouth, United Kingdom

OPEN ACCESS

Edited by:
Catherine Jeandel,
Centre National de la Recherche Scientifique (CNRS), France

Reviewed by:
Hubert Loisel,
Université du Littoral Côte d'Opale, France
David Antoine,
Curtin University, Australia

***Correspondence:**
Shubha Sathyendranath
ssat@pml.ac.uk

Specialty section:
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

Received: 30 March 2017
Accepted: 15 November 2017
Published: 12 December 2017

Citation:
Shafeeque M, Sathyendranath S, George G, Balchand AN and Platt T (2017) Comparison of Seasonal Cycles of Phytoplankton Chlorophyll, Aerosols, Winds and Sea-Surface Temperature off Somalia. Front. Mar. Sci. 4:386. doi: 10.3389/fmars.2017.00386

In climate research, an important task is to characterize the relationships between Essential Climate Variables (ECVs). Here, satellite-derived data sets have been used to examine the seasonal cycle of phytoplankton (chlorophyll concentration) in the waters off Somalia, and its relationship to aerosols, winds and Sea Surface Temperature (SST). Chlorophyll-a (Chl-a) concentration, Aerosol Optical Thickness (AOT), Ångström Exponent (AE), Dust Optical Thickness (DOT), SST and sea-surface wind data for a 16-year period were assembled from various sources. The data were used to explore whether there is evidence to show that dust aerosols enhance Chl-a concentration in the study area. The Cross Correlation Function (CCF) showed highest positive correlation ($r^2 = 0.3$) in the western Arabian Sea when AOT led Chl-a by 1–2 time steps (here, 1 time step is 8 days). A $2 \times 2^\circ$ box off Somalia was selected for further investigations. The correlations of alongshore wind speed, Ekman Mass Transport (EMT) and SST with Chl-a were higher than that of AOT, for a lag of 8 days. When all four variables were considered together in a multiple linear regression, the increase in r^2 associated with the AOT is only about 0.02, a consequence of covariance among AOT, SST, EMT and alongshore wind speed. The AOT data show presence of dust aerosols most frequently during the summer monsoon season (June–September). When the analyses were repeated for the dust aerosol events, the correlations were generally lower, but still significant. Again, the inclusion of DOT in the multiple linear regression increased the correlation coefficient by only 2%, indicating minor enhancement in Chl-a concentration. Interestingly, during summer monsoon season, there is a higher probability of finding more instances of positive changes in Chl-a after one time step, regardless of whether there is dust aerosol or not. On the other hand, during the winter monsoon season (November–December) and rest of the year, the probability of Chl-a enhancement is higher when dust aerosol is present than when it is absent. The phase relationship in the 8-day climatologies of Chl-a and AOT (derived from NASA's SeaWiFS and MODIS-A ocean colour processing chain) showed that AOT

led Chl-a for most of the summer monsoon season, except when Chl-a was very high, during which time, Chl-a led AOT. The phase shift in the Chl-a and AOT climatological relationship at the Chl-a peak was not observed when AOT from Aerosol Climate Change Initiative (Aerosol-CCI) was used.

Keywords: essential climate variables, aerosol optical thickness, Ångström exponent, chlorophyll-a, ocean colour climate change initiative, climate change, remote sensing, dust aerosols

1. INTRODUCTION

Phytoplankton, Sea-Surface Temperature (SST), sea-surface winds and aerosols are all Essential Climate Variables (ECVs) identified by the Global Climate Observation System (GCOS, 2011) as being worthy of sustained global observations at high spatial resolution and over long time scales, to aid studies of Earth's climate and climate change. As we strive to understand how the Earth system might respond holistically to climate change, it is important to explore not only the behavior of individual ECVs, but also their inter-relationships and the feedbacks between them. In the western Arabian Sea, the relationships between phytoplankton, winds and SST are better understood than that between phytoplankton and aerosols.

Yet, there are known functional links between marine aerosols and phytoplankton. For example, dust aerosols, transported by winds over the ocean, can be an important source of micronutrients such as iron, essential for phytoplankton growth (Duce and Tindale, 1991; Martin et al., 1991, 1994; Prospero et al., 2002; Cropp et al., 2005; Jickells et al., 2005; Mahowald et al., 2005; Meskhidze et al., 2005; Gallisai et al., 2014), with the proviso that not all the iron contained in dust particles is usable by phytoplankton. Winds over the ocean are also responsible for the formation of aerosols through generation of sea salt sprays (O'Dowd et al., 1997; Smirnov et al., 2003; Satheesh et al., 2006; Mulcahy et al., 2008; Glantz et al., 2009; Huang et al., 2010; Meskhidze and Nenes, 2010) and the same winds also mix the surface layer of the ocean, dictating the entrainment of nutrients from the deeper waters into the surface layer and controlling the average light available for phytoplankton growth in the layer. In addition to sea salt sprays, biological particles (for example, fragments of phytoplankton) contained in sea spray can also aid aerosol formation (Leck and Bigg, 2005; Facchini et al., 2008; Hawkins and Russell, 2010; Quinn and Bates, 2011). Feedback mechanisms (both positive and negative) have been proposed between dimethyl sulphide in the atmosphere of phytoplanktonic origin and the Earth's radiation budget, *via* aerosols (Charlson et al., 1987; Lovelock, 2006).

Positive (Martin et al., 1994; Jickells et al., 2005; Patra et al., 2007; Banerjee and Prasanna Kumar, 2014) and negative (Mallet et al., 2009; Paytan et al., 2009; Jordi et al., 2012) correlations between marine aerosols and phytoplankton concentration have been reported for different parts of the world ocean. Some studies have also identified regions where no relationship exists between the two (Cropp et al., 2005; Gallisai et al., 2014). Possible explanations for the positive correlations include the fertilizing role of iron contained in dust aerosols, or phytoplankton themselves, acting as a source of marine aerosols. Negative correlations might arise from high winds causing production of wind-spray aerosols, while at the same time forming deep mixed layers that may be able to support only low concentrations of phytoplankton, because of low average light levels available in the layer.

Satellite-based measurements provide a valuable tool for studies of aerosols and phytoplankton. Aerosol Optical Thickness (AOT), amenable to remote sensing, is an often-used measure of aerosol concentration. The Ångström Exponent (AE), which defines the wavelength dependence of AOT, is indicative of the type of aerosols present, and is also available through remote sensing. Dust Optical Thickness (DOT) can be inferred from AOT and the AE. Satellite data have been used to track dust aerosols for thousands of kilometers away from their source (Myhre et al., 2005). Likewise, ocean colour measured from space provides information on the concentration of chlorophyll-a (Chl-a), which is a major photosynthetic pigment contained in phytoplankton. Furthermore, estimates of winds (speed and direction) and SST, essential for understanding phytoplankton dynamics, are also available through remote sensing. An advantage of remote sensing is that it provides data at large scales and over many years, allowing studies of time-series at multiple locations in a systematic manner. But some caution should be exercised when using ocean colour derived Chl-a concentration, AOT and AE. Sometimes they are all produced from the same processing chain, and one might argue that, in the extreme case, any relationships observed between the three are purely artifacts of the processing algorithm. Furthermore, the effects of clouds on satellite retrievals are significant and sometimes lead to biases by overestimation or underestimation of aerosol data, particularly for dust aerosols (Levy et al., 2007; Torres et al., 2007; Baddock et al., 2009; Kahn et al., 2010). However, some authors have used cloud-screening techniques to reduce such errors (Kaufman et al., 2005). Therefore, the processing chain issues should be verified to arrive at conclusive results.

In this paper, we examine the relationships of Chl-a with winds, SST, AOT and dust aerosols in the western Arabian Sea, at a selected site off Somalia. The region is characterized by a high dynamic range in Chl-a values that vary seasonally, in response to the reversing wind patterns and associated upwelling (Prasanna Kumar et al., 2001; Schott and McCreary, 2001; Schott et al., 2002; Shankar et al., 2002; Wiggert et al., 2005; Lévy et al., 2007; Wiggert and Murtugudde, 2007; Prakash et al., 2012). Diverse physical forcings of both oceanic and atmospheric origins drive biological production off Somalia region. During summer monsoon season, the Somalia coastal region is characterized by strong upwelling with high primary productivity due to the swift Somali current caused by strong

south-westerlies along the coast (Smith and Codispoti, 1980; Schott, 1983; Hitchcock and Olson, 1992; Brock et al., 1994; Schott et al., 2002; deCastro et al., 2016). The anti-cyclonic eddies associated with the Somali current during the same season further enhance production by transporting and mixing upwelled water (Fischer et al., 1996; McCreary et al., 1996; Schott et al., 1997; Koning et al., 2001; Schott et al., 2002; Santos et al., 2015). The consequent nutrient enrichment in the mixed layer of the ocean leads to high phytoplankton production during summer monsoon season (Banse, 1987; Owens et al., 1993). Because of its proximity to the Arabian Peninsula, the region also receives seasonally-varying dust deposition (Pease et al., 1998; Li and Ramanathan, 2002; Prospero et al., 2002; Léon and Legrand, 2003; Zhu et al., 2007; Prasanna Kumar et al., 2010). Thus, the same winds that transport dust aerosols to the western Arabian Sea during the summer monsoon season also induce upwelling, favoring phytoplankton blooms. Hence the relationship between Chl-a and aerosols in this region would be incomplete, unless we examined the effect of winds on phytoplankton dynamics as well. Here, we use 16 years of satellite data (1998–2013) to make a systematic study of the relationship of Chl-a with AOT, winds and SST in the waters off Somalia.

2. MATERIALS AND METHODS

2.1. Data

Level-3 8-day composite Aerosol Optical Thickness (AOT) at 865 nm and Ångström Exponent (AE) from Sea-viewing Wide Field-of-view Sensor (SeaWiFS) during January 1998–December 2010 and Moderate Resolution Imaging Spectro-radiometer (MODIS) Aqua during January 2011–December 2013 downloaded from National Aeronautics and Space Administration's (NASA's) ocean colour website (https://oceancolor.gsfc.nasa.gov) were used in this work. The AOT and AE data from NASA are referred to here as NASA-AOT and NASA-AE respectively. The daily AOT at 550 nm and AE data from European Space Agency's (ESA's) Aerosol Climate Change Initiative (Aerosol-CCI) programme (de Leeuw et al., 2015; Popp et al., 2016, see also http://www.esa-aerosol-cci.org) were also used in this study, as an independent source of aerosol data, unconnected with ocean colour atmospheric correction routines. The AOT and AE data from the Aerosol-CCI website are referred to here as CCI-AOT and CCI-AE respectively. The data are available at 1° spatial resolution for the period from January 1998–December 2010.

The relationship between the AOT (τ) at any given wavelength λ_0 and that at any other wavelength λ depends on the AE (α) through the equation:

$$\left(\frac{\tau_\lambda}{\tau_{\lambda_0}}\right) = \left(\frac{\lambda}{\lambda_0}\right)^{-\alpha}. \quad (1)$$

In principle, if the optical thickness at one wavelength and the AE are known, the optical thickness can be computed at any other wavelength using Equation (1).

Chlorophyll-a (Chl-a) concentration, for the period January 1998–December 2013, was obtained from ESA's Ocean Colour-Climate Change Initiative (OC-CCI) website (Sathyendranath et al., 2016, see also https://www.oceancolour.org). One of the major reasons for the choice of the Chl-a data was the improved coverage provided by the OC-CCI data in the Arabian Sea, especially during the summer monsoon season. The 8-day composite AOT data from SeaWiFS are available at only 9 km resolution, so we used MODIS Aqua data at the same resolution (9 km) even though they are available at 4 km resolution. The Chl-a concentration from OC-CCI (version-2), which is available at 4 km resolution, was also re-gridded to 9 km resolution. Since the CCI-AOT data are available at 1° spatial resolution, the Chl-a concentration from OC-CCI was also re-gridded to 1° resolution to analyse the correlation between them. The daily value of AOT at 865 nm was calculated from daily CCI-AOT at 550 nm and CCI-AE using Equation (1). The data were merged to genereate 8-day composites and extracted for the region off Somalia. The daily 1° gridded Sea Surface Temperature (SST) data were obtained for the period January 1998–December 2013 from Woods Hole Oceanographic Institute's (WHOI's) objectively-analyzed air-sea heat fluxes available at Asia-Pacific Data-Research Centre (APDRC) website (http://apdrc.soest.hawaii.edu). The SST anomaly has been calculated using these data after merging into 8-day composites. In addition, the daily NCEP/NCAR reanalysis U-wind (zonal velocity) and V-wind (meridional velocity) data with 2.5 × 2.5° spatial resolution at 10 m above the sea surface were obtained for the same period from their official website (https://www.esrl.noaa.gov/psd). The data have been merged to generate 8-day composites and used to derive the south westerly wind component along the Somalia coast. All the above mentioned information is summarized in **Table 1**.

2.2. Methods

The methods used in this study are shown schematically in **Figure 1**, and described below.

2.2.1. Correlation between Chl-a and AOT in the Arabian Sea

Correlation between Chl-a and AOT concentration for the 1998–2013 period over the Arabian Sea was studied using the 8-day composites. The results showed areas of both positive and negative correlation. The western Arabian Sea showed strong positive correlation. A 2 × 2° box (54–56° E longitude and 10–12° N latitude) off Somalia coast, with high positive correlation, was chosen for further analyses.

2.2.2. CCF Analysis and Lagged Correlation

We studied the lags in the correlation between Chl-a and AOT using Cross Correlation Function (CCF). CCF analysis produces cross correlations in which the observations of one time series are correlated with the observations of another time series at different lags and leads, to identify the variables which are leading or lagging indicators of other variables. The basic premise is that, if the relationships between the variables were merely a processing artifact, the correlations would peak at zero lag. In instances where phytoplankton might be contributing biological material for aerosol formation, the correlation would

TABLE 1 | Summary of data sets analyzed in the study.

Variable	Sensor/provider	Units	Spatial resolution	Temporal extent	Source
Chlorophyll-a (Chl-a)	Ocean Colour Climate Change Initiative (OC-CCI)	mg m^{-3}	4 km	1998–2013	https://www.oceancolour.org
Aerosol Optical Thickness (AOT)	Sea-viewing Wide Field-of-view Sensor (SeaWiFS) and Moderate Resolution Imaging Spectro-radiometer Aqua (MODIS-Aqua)	Dimensionless	9 km	1998–2013	https://oceancolor.gsfc.nasa.gov
	Aerosol Climate Change Initiative (Aerosol-CCI)	Dimensionless	1°	1998–2010	http://www.esa-aerosol-cci.org
Ångström Exponent (α)	SeaWiFS and MODIS Aqua	Dimensionless	9 km	1998–2013	https://oceancolor.gsfc.nasa.org
	Aerosol-CCI	Dimensionless	1°	1998–2010	http://www.esa-aerosol-cci.org
Sea Surface Temperature (SST)	Woods Hole Oceanographic Institute (WHOI)	°C	1°	1998–2013	http://apdrc.soest.hawaii.edu
U & V wind components	NCEP NCAR	ms^{-1}	2.5°	1998–2013	https://www.esrl.noaa.gov/

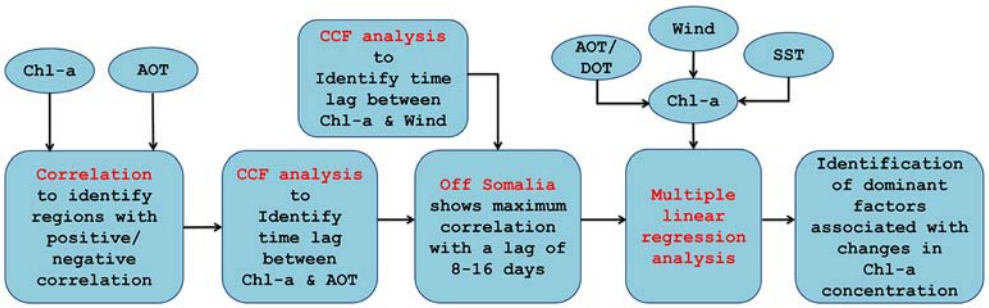


FIGURE 1 | Schematic diagramme showing the methods and analyses used in this work.

be maximum when AOT lagged behind Chl-a concentration. On the other hand, if the oceans were fertilized by aerosols, then Chl-a would lag behind AOT.

The CCF analysis was also carried out between Chl-a concentration and alongshore component of wind speed. If wind-induced upwelling were a causative factor for the increment in Chl-a concentration in the Somalia coast, then we anticipate that the correlation between them would peak when Chl-a lagged behind wind (because of the finite time it takes for phytoplankton to bloom in response to the nutrients brought to the surface by upwelling). Though the alongshore wind speed over the Somalia coast is a fairly good indicator of upwelling strength, we have calculated the Ekman Mass Transport (EMT) as an upwelling index for the analysis. Since a surface signature of upwelling is a decrease of SST in the upwelling zone, we have also taken SST as another proxy for upwelling.

2.2.3. Ekman Mass Transport

For the Somalia region, the alongshore component of the wind stress is favorable for upwelling during summer monsoon season. A positive value for the EMT represents upwelling along the coast of Somalia. The alongshore wind stress for Somalia coast was calculated by the bulk aerodynamic formula from Koracin et al. (2004) as shown in Equation (2):

$$\tau_y = \rho_a \times C_d \times w \times v. \quad (2)$$

where τ_y is the alongshore wind stress; ρ_a is the density of air, which was taken to be 1.2 kg/m^3; w is the magnitude of the wind speed; v is the alongshore component of wind speed in m/s; and C_d is the nonlinear drag coefficient based on Large and Pond (1981) and Trenberth et al. (1990) for low wind speeds. So, the EMT along the Somalia coast can be calculated using Equation (3):

$$M_{ev} = \frac{\tau_y}{f}. \quad (3)$$

where, M_{ev} is mass transport by the alongshore wind, f is the Coriolis parameter ($2 \times \Omega \times \sin\phi$), Ω is the angular frequency of the Earth and ϕ is the latitude.

Multiple linear regression analysis with Chl-a as dependent variable and NASA-AOT (or DOT), alongshore wind speed, EMT and SST as independent variables was carried out. We used 8-day composites with lags of 1–2 time steps for this analysis (these lags correspond to the maximum correlation between Chl-a and NASA-AOT data). The analysis was repeated by replacing NASA-AOT with CCI-AOT (or DOT) with a lag of 3 time steps, corresponding to the maximum correlation between Chl-a and CCI-AOT. We have also calculated the 8-day climatologies of all these variables, and plotted against time of year, to study their phase relationships.

2.2.4. Derivation of Dust Optical Thickness (DOT)

The desert dust transported by winds over the ocean contains micronutrients such as iron, which can regulate phytoplankton

activity (Martin et al., 1994; Lenes et al., 2001; Muhs et al., 2007; Donaghay et al., 2015). Since the seasonal monsoon winds bring large quantities of iron-containing dust aerosols to the study area (Li and Ramanathan, 2002; Banerjee and Prasanna Kumar, 2014), we investigated the effect of dust aerosols on Chl-a concentrations. The AE, which is often used as a qualitative indicator of aerosol particle size, and AOT, which indicates the aerosol load, can be used to differentiate dust aerosol from other types of aerosol. Generally, a higher value of AE ($\alpha > 1$) is indicative of fine, submicron aerosols, whereas lower values ($\alpha < 1$) are representative of coarse, super-micron particles (Kaufman, 1993; Gobbi et al., 2007; Yoon et al., 2012). The AOT values are lower for fine aerosols and higher for coarse aerosols. In the literature, different criteria have been proposed to identify dust aerosols at different locations: for example, $\alpha < 0.6$ (Dubovik et al., 2002; Brindley et al., 2015), $\alpha < 0.8$ (Eck et al., 2005; Che et al., 2013), $\alpha < 1$ (Eck et al., 1999; Schuster et al., 2006; Papaynannis et al., 2007; Yoon et al., 2012; Valenzuela et al., 2014; Zu et al., 2014; Pakszys et al., 2015) and $\alpha < 1.4$ (Gobbi et al., 2007; Pereira et al., 2011; Shinozuka et al., 2011); similarly, AOT > 0.11 (Toledano et al., 2007; Balarabe et al., 2016), AOT > 0.2 (Salinas et al., 2009; Pakszys et al., 2015) and AOT > 0.25 (Guleria et al., 2012) have been recommended to identify dust aerosols. After considering all these studies, we have adopted the ranges of AOT and AE for off Somalia as follows: AE less than 1, and AOT at 440 nm (τ_{440}) greater than 0.2 (i.e., $\alpha < 1$ and $\tau_{440} > 0.2$) are designated as DOT or dust aerosols. The NASA-AOT at 865 nm (τ_{865}) and NASA-AE were used to calculate NASA-AOT at 440 nm (τ_{440}), using Equation (1). Similarly, CCI-AOT at 550 nm (τ_{550}) and CCI-AE were used to calculate CCI-AOT at 440 nm (τ_{440}), using Equation (1).

3. RESULTS

3.1. Relationship between Chl-a and AOT in the Arabian Sea

The correlation between Chl-a and NASA-AOT using 8-day time series from 1998 to 2013 data for the Arabian Sea is mapped in **Figure 2**. The results are based on data for all the seasons rather than for specific seasons as in Patra et al. (2007) or in Banerjee and Prasanna Kumar (2014). The western Arabian Sea exhibits high positive correlations, whereas the south eastern Arabian Sea shows low to moderate positive correlations. There are also regions (south central) where no statistically-significant correlation is evident and extensive regions (north-central and north-eastern) of significant negative correlations. The region off Somalia shows high positive correlation between Chl-a and NASA-AOT and it is located along the path of winds carrying dust aerosols emanating from South Asia, South-West Asia, North Africa (Sahara) and the eastern Horn of Africa (Pease et al., 1998; Ginoux et al., 2001; Goudie and Middleton, 2001; Prospero et al., 2002; Léon and Legrand, 2003). Although there are several studies (Banzon et al., 2004; Kayetha et al., 2007; Patra et al., 2007; Singh et al., 2008; Nezlin et al., 2010; Banerjee and Prasanna Kumar, 2014) that have examined the relationship between Chl-a and AOT in various parts of the Arabian Sea, the region off Somalia has not yet been explored in detail, and it is the region selected for our investigation.

3.2. Climatologies of Chl-a, Aerosols, Winds and SST off Somailia

The 16-year 8-day climatological seasonal cycles of Chl-a concentration, NASA-AOT, CCI-AOT, SST and along-shore wind speed are shown in **Figure 3A**, for the selected study area off Somalia. SST data are reported as anomalies from 8-day average. When the aerosols are identified as dust aerosols, they are indicated in the plot using black and purple filled circles. Out of 46 observations involved in both AOT data sets, for the 8-day climatology, 24 observations were dust aerosols for CCI-AOT data whereas 14 were identified as dust aerosols for NASA-AOT data. It was found that the CCI-AOT data showed the presence of dust aerosols not only during the summer monsoon season, but also during the winter monsoon season. The corresponding climatological wind vectors are shown in **Figure 3B**. During the first 100 days of the year, winds are north easterly, the wind speed decreasing with time. These conditions are unfavorable for upwelling off Somalia. During this period, SST increases steadily by some 3°C. At the same time, the Chl-a concentrations decrease, and AOT also remains low. After this, the winds reverse direction and intensify, resulting in upwelling (indicated by decreasing SST) that favors phytoplankton growth. We note that the initial response of phytoplankton to the intense south westerly winds is a decrease in concentration, perhaps a consequence of the phytoplankton being mixed into deeper layers. After this, the Chl-a increases, with a lag of a couple of time steps behind the increasing wind speed. Both AOT and Chl-a reach their respective maxima during the summer monsoon season.

In **Figure 3A**, the NASA-AOT and Chl-a reach their respective maxima during the summer monsoon season. Although the seasonality of CCI-AOT is more or less similar to that of NASA-AOT, the occurrence of peak values is different. The maximum value for CCI-AOT occurred during early summer monsoon season (Day of Year, DoY 170) while the Chl-a is still increasing, whereas the NASA-AOT peak occurred at DoY 224 during the waning phase of summer monsoon season and after the Chl-a peaks at DoY 216. An interesting feature in the figure is that, towards the peak of the summer monsoon (around DoY 180), when Chl-a concentration reaches $\approx 0.8\,\mathrm{mg\,m^{-3}}$, there is a brief period when Chl-a continues to increase and leads NASA-AOT by up to 3 time steps until DoY $\approx$ 220. However, this feature was not found in CCI-AOT data. Just before the Chl-a peak is reached, the wind speed starts to drop, followed by Chl-a and NASA-AOT, until all variables reach minima toward DoY 300, at which point the wind direction again reverses. SST starts to increase when the south-westerly winds drop, reaching a secondary peak at around DoY 310.

The seasonal patterns are consistent with the known geography of the area. However, there is a tantalizing suggestion in **Figure 3** that when the winds speed are at their highest, and Chl-a levels are high, the NASA-AOT concentrations may be enhanced by maritime aerosols, in addition to the dust aerosols,

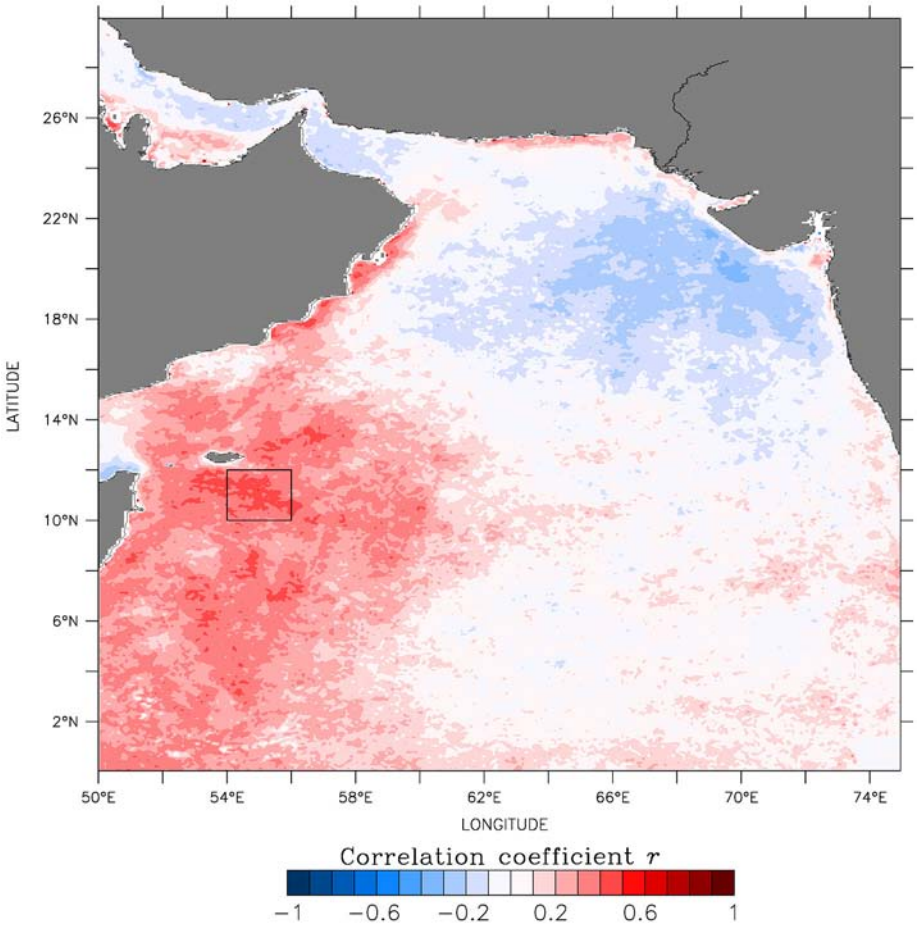


FIGURE 2 | Correlation map between 8-day composited Chl-a and NASA-AOT during 1998–2013 for the Arabian Sea. Pale to dark red shading and pale to dark blue shading represent positive and negative correlation respectively. The black square box off Somalia shows the study area selected for further analyses.

and that some of these aerosols may have a biological origin, as indicated by Chl-a leading NASA-AOT during this period. However, this observation is not supported by CCI-AOT, and in the absence of additional information, it would be premature to conclude that such is the case. But it would be a point worthy of further investigation.

3.3. The Relationship between Chl-a, AOT and DOT off Somalia

Since **Figure 3** indicates that there is a lag in the relationships between Chl-a and the other variables studied here, further analysis has been made for the $2 \times 2^{\circ}$ box using Cross Correlation Function (CCF) between Chl-a and AOT. The result (**Figure 4A**) shows that the highest significant positive correlation ($r = 0.55$) between Chl-a and NASA-AOT in the study region occurred for Chl-a lagging NASA-AOT by 1 to 2 time steps (1 time step is 8 days). The CCF analysis was also carried out between CCI-AOT and Chl-a and shows a significant positive correlation. Further, the maximum correlation ($r = 0.54$) occurred when Chl-a lagged behind CCI-AOT by 3 time steps (**Figure 4B**). So the analysis using CCI-AOT data confirmed the results obtained using NASA-AOT on the existence of a significant correlation between Chl-a and AOT in the region off Somalia, the magnitude of the correlation and also the sign of the lag.

The relationship between Chl-a and AOT (or DOT) with lag of 8 days is explored further in **Figure 5** using NASA-AOT (or DOT). Scatter plot between Chl-a and AOT is shown in **Figure 5A**, with the fitted curve and the r value of 0.55 for the fit, consistent with the CCF. However, we recognize that the relationship of maritime and dust aerosols with Chl-a would be functionally different (for example, we do not anticipate that maritime aerosols could fertilize the oceans, whereas it would be plausible with dust aerosols). Dust aerosols are present more frequently during the summer monsoon season because of the favorable wind from adjacent land masses, compared with other seasons. Out of 736 observations over 16 years, around 203 observations were identified as dust aerosols. **Figure 5B** shows the relationship between Chl-a and DOT. We see that there is a general tendency for Chl-a to increase with DOT.

We checked further whether the presence of dust aerosols enhances the Chl-a concentration in the subsequent time steps by calculating the difference in Chl-a (ΔChl-a) in 1 time step after

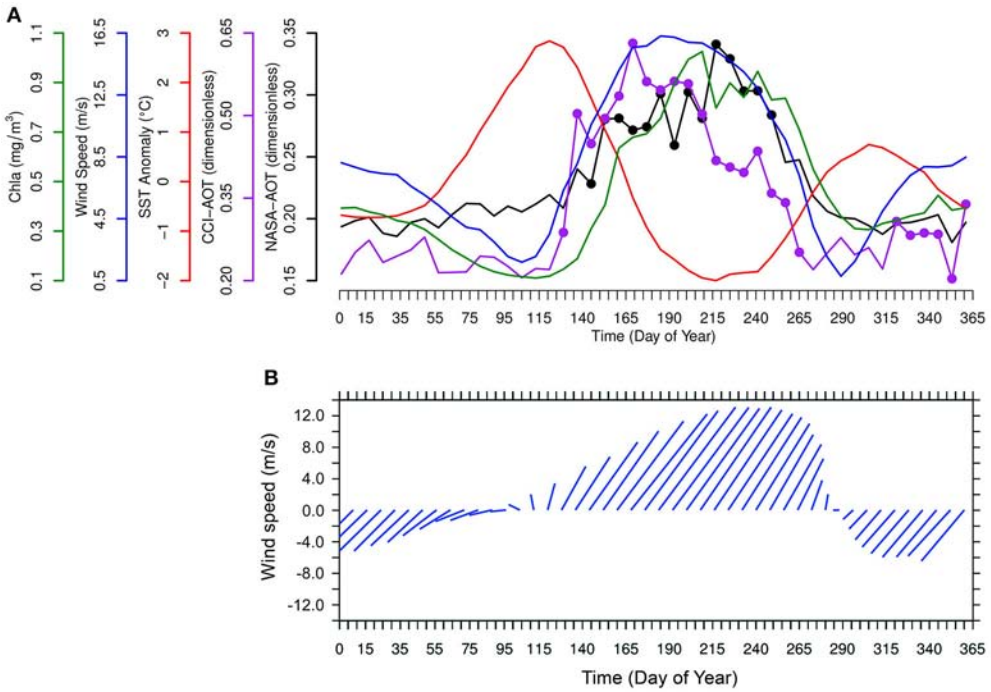


FIGURE 3 | Time series of 8-day climatology, **(A)** for Chl-a, NASA-AOT, CCI-AOT, SST anomaly and Wind speed and **(B)** for the wind vectors. The black and purple coloured dots in the NASA-AOT and CCI-AOT graph denote the presence of dust aerosols.

a dust event, and plotting it against NASA-AOT (**Figure 5C**). The presence of more positive ΔChl-a following high aerosol events would be indicative of a positive effect of aerosols on phytoplankton concentration. The data (**Figure 5C**) show no obvious relationship between aerosols and ΔChl-a either for all aerosols taken together or for dust aerosol events (circles in red colour) by themselves. However, for all the DOT events considered by themselves, the frequency of ΔChl-a is slightly skewed toward positive numbers, with some 114 values being positive out of 203 events (see histogram of ΔChl-a, **Figure 5D**). So the probability that Chl-a enhancement is associated with the presence of dust aerosols throughout the year is 56% (114 out of 203), compared with 238 out of 532 in the absence of dust aerosols (45%). The higher number of positive ΔChl-a observations is significant ($p < 0.05$) according to a binomial test. For the non-dust events, there is a higher number of negative values (294) compared with positive values (238) of ΔChl-a. These results are summarized in **Table 2**.

We supplemented these calculations after splitting the data according to monsoon (summer monsoon) and non-monsoon seasons, recognizing the differences in oceanographic and meteorological conditions during these two parts of the year (**Table 2**). Out of 224 observations during the summer monsoon season, 140 are dust aerosol events and 84 are non-dust events. Within these 140 dust events, the number of positive ΔChl-a values is 82 (59%), compared with 58 (41%) negative values. However, for non-dust events during this season, we also find more positive ΔChl-a values (58 events, or 69%) than negative ones (26 events, or 31%). For the non-monsoon season, out of 511 total observations, 448 are non-dust aerosol events and 63 are dust events. Within these non-dust observations, there is higher number of negative values (268, or 60%) when compared with positive values (180 or 40%). But during dust events, the number of positive observations is slightly higher, with 32 (51%) positive values compared with 31 (49%) negative ones. We conclude from all of the above that the probability of Chl-a enhancement during the summer monsoon season does not depend much on the presence or absence of dust aerosols. In other words, during the summer monsoon season, there is a higher probability of finding positive ΔChl-a values, regardless of whether there is a dust event or not. On the other hand, during the rest of the year, the probability of chlorophyll enhancement is a little higher during dust events than during non-dust events.

The analysis was also repeated for CCI-AOT data to verify the above results and is presented in **Table 3**. For this dataset, the probability of Chl-a enhancement is again more in the presence of dust aerosols when the whole year is considered, at 53% (134 out of 251), compared with 142 out of 344 in the absence of dust aerosols (41%). Out of 208 observations during the summer monsoon season, 154 are dust aerosol events and 54 are non-dust events. Within these 154 dust events, the number of positive ΔChl-a values is 83 (54%), compared with 71 (46%) negative values. However, for non-dust events during this season, we also find more positive ΔChl-a values (31, or 57%) than negative ones (23, or 43%). The results from winter monsoon season indicate

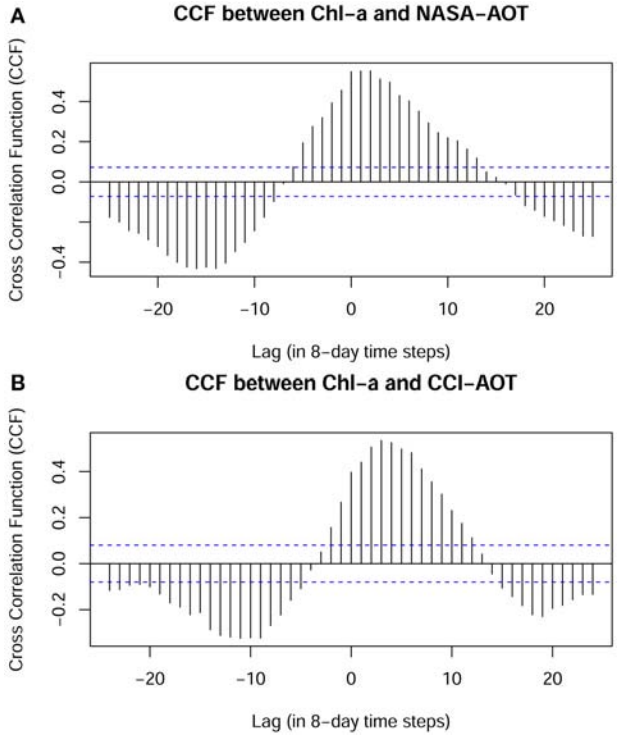


FIGURE 4 | Cross Correlation Function between **(A)** Chl-a and NASA-AOT **(B)** Chl-a and CCI-AOT for study area off Somalia.Time step is 8 days.

that, though the dust aerosol events are fewer in number (49) compared with non-dust (93) within 142 observations, there were more positive ΔChl-a values (28, or 57%) than negative values (21, or 43%) when dust aerosols were present in the region. But, during the absence of dust aerosols, there is a higher number of negative values (48, or 52%) compared with positive values (45 or 48%).

Thus both NASA-AOT and CCI-AOT lead to the conclusion that the probability of Chl-a enhancement during the summer monsoon season does not depend on the presence or absence of dust aerosols. In other words, during the summer monsoon season, there is a higher probability of finding positive ΔChl-a values, regardless of whether there is a dust event or not. On the other hand, during the winter monsoon season and rest of the year, the probability that dust events may be associated with chlorophyll enhancement is higher than that during non-dust periods.

3.4. Relationship of Chl-a with Winds, SST, AOT, and DOT

To elucidate further the relationship between Chl-a and environmental conditions, we next examined the CCF between Chl-a and alongshore wind speed, since it is known that the alongshore winds determine upwelling, and hence influence phytoplankton dynamics in the area (Goes et al., 2005; Gregg et al., 2005; Wiggert et al., 2005; Prasanna Kumar et al., 2010); (see also **Figure 3**). The result (**Figure 6**) shows, similar to the CCF between Chl-a and NASA-AOT, that the correlation peaks with a lag of 1-2 time steps, with wind speed leading Chl-a, but with a higher correlation coefficient ($r = 0.69$, $p < 0.05$).

Since the correlation coefficients of Chl-a with both aerosols and wind speed peak with a lag of 1–2 time steps, we chose a lag of 1 time step, for a linear step-wise multiple regression study with Chl-a as dependent variable, and NASA-AOT (or DOT), Ekman Mass Transport (EMT), alongshore wind speed and SST as independent variables. The upwelling indices, the wind speed and EMT both show more or less similar correlation with Chl-a. So, we excluded the EMT from the multiple linear regression analysis (but the results from the multiple linear regression including EMT are presented as Table S1). When the correlations with each of the independent variables are considered individually, the highest r^2 values were found for alongshore wind speed ($r^2 = 0.47$) for the ensemble of year-round data, with the corresponding r^2 dropping to 0.17 when dust aerosol events are considered separately (140 dust events during the summer monsoon, and 63 outside of it, totalling 203), followed by SST ($r^2 = 0.33$ and $r^2 = 0.20$ for the same two cases respectively), and then by NASA-AOT ($r^2 = 0.30$ and $r^2 = 0.08$ for the corresponding cases). From the results of pair-wise regression analysis, we see that the addition of NASA-AOT (or DOT) as an independent variable, in addition to wind speed, increases r^2 values by a modest 0.02. With all three variables taken together as independent variables, the explained variance (r^2) is 0.52 for all data, and 0.25 for DOT events (**Table 4**). The results for a lag of 2 time steps (not shown) are similar to those for lag of 1 time step, but with lower correlation coefficients.

The multiple regression analysis was also repeated for CCI-AOT data with Chl-a as dependent variable and CCI-AOT (or DOT), alongshore wind speed and SST as independent variables (**Table 5**). Since the correlation coefficients of Chl-a with CCI-AOT peak with a lag of 3 time steps, we chose a lag of 3 time steps for this analysis. When the correlations with each of the independent variables are considered individually, the highest r^2 values were again found for alongshore wind speed ($r^2 = 0.49$) for year-round data and $r^2 = 0.38$ for dust aerosol events (154 dust events during the summer monsoon, and 97 outside of it, totalling 251) considered separately, followed by SST ($r^2 = 0.32$ and $r^2 = 0.15$ for the same two cases respectively), and then by CCI-AOT ($r^2 = 0.29$ and $r^2 = 0.09$ for the corresponding cases). From the results of pair-wise regression analysis, the addition of CCI-AOT (or DOT) on wind speed as independent variables did not make any improvement in the r^2 value of 0.49. However, a small increase in r^2 value by 0.06 or 0.08 was found when adding CCI-AOT (or DOT) respectively to SST.

4. DISCUSSION

4.1. The Satellite Data Used

Much of the interpretation of results for the region off Somalia depends on the quality of the satellite data used for the analysis, especially during the summer monsoon season, since this is a highly dynamic season, with high winds, high AOT and high Chl-a concentrations. The OC-CCI Chl-a dataset (Sathyendranath et al., 2016) was selected because of the significantly-improved

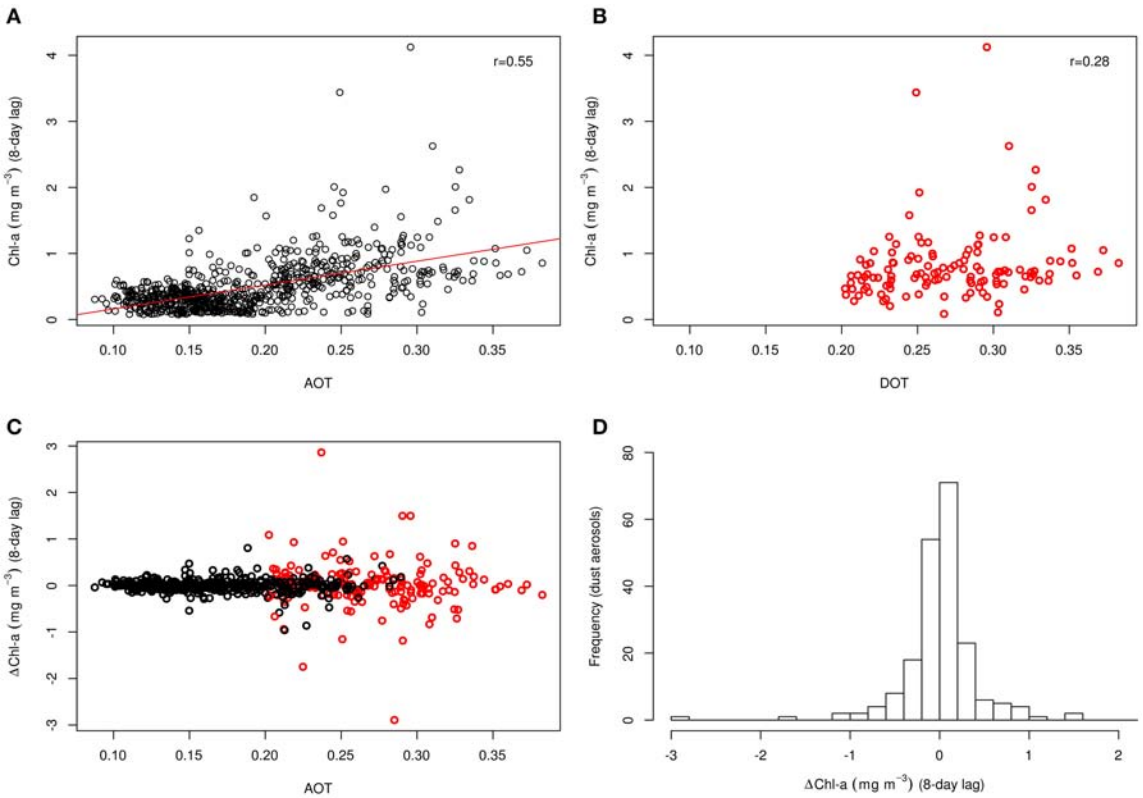


FIGURE 5 | (A) The scatter plot between 8-day composite AOT (NASA-AOT) and Chl-a with lag of one time steps, for the study area. The straight line in red colour indicates the regression equation and the correlation coefficient (*r*) is shown in the top right side. **(B)** Scatter plot between DOT and Chl-a with lag of one time steps (subset of all the data points in **(A)**. **(C)** Scatter plot between AOT and ΔChl-a with lag of one time step. The circles in red colour indicate dust aerosols. **(D)** Histogram for ΔChl-a during the presence of dust aerosols with lag of one time steps.

TABLE 2 | The number of observations with enhancements in Chl-a (+ve ΔChl-a) or reductions in Chl-a (−ve ΔChl-a), for all data, and for the summer monsoon, for the non-monsoon and sorted according to whether the aerosols were identified as dust or not (here, dust aerosols were derived from NASA-AOT data).

Dust + non-dust, All seasons		Dust, All seasons		Non-dust, All seasons		Dust + non-dust, Summer monsoon season		Dust, Summer monsoon season		Non-dust, Summer monsoon season		Dust + non-dust, Non-monsoon season		Dust, Non-monsoon season		Non-dust, Non-monsoon season	
TOTAL NUMBER OF OBSERVATIONS																	
735		203		532		224		140		84		511		63		448	
NUMBER OF POSITIVE AND NEGATIVE ΔChl-a VALUES																	
+ve	−ve	+ve	−ve	+ve	−ve	+ve	−ve	+ve	−ve	+ve	−ve	+ve	−ve	+ve	−ve	+ve	−ve
352	383	114	89	238	294	140	84	82	58	58	26	212	299	32	31	180	268
PERCENTAGE OF POSITIVE AND NEGATIVE ΔChl-a VALUES																	
48	52	56	44	45	55	62	38	59	41	69	31	41	59	51	49	40	60

seasonal coverage that the data provide in the study area, compared with other datasets, especially during the summer monsoon season. But it is important to reassure ourselves that the data are of sufficient quality for the analysis presented. Though the OC-CCI data have been validated using a global dataset as part of the project, and also for the neighboring Red Sea (Brewin et al., 2015) and the Gulf of Aden (Gittings et al., 2016), we do not have *in situ* data from off Somalia region for local validation. However, the data are reassuring in some respects: the first one is that, if the relationship between Chl-a and AOT were an artifact of the processing, then one would anticipate that the relationship would peak at zero lag. In fact, we see

TABLE 3 | The number of observations with enhancements in Chl-a (+ve ΔChl-a) or reductions in Chl-a (−ve ΔChl-a), for all data, and for the summer monsoon, for the winter monsoon and sorted according to whether the aerosols were identified as dust or not (here, dust aerosols were derived from CCI-AOT data).

Dust + non-dust, All seasons		Dust, All seasons		Non-dust, All seasons		Dust + non-dust, Summer monsoon season		Dust, Summer monsoon season		Non-dust, Summer monsoon season		Dust + non-dust, Winter monsoon season		Dust, Winter monsoon season		Non-dust, Winter monsoon season	
TOTAL NUMBER OF OBSERVATIONS																	
595		251		344		208		154		54		142		49		93	
NUMBER OF POSITIVE AND NEGATIVE ΔChl-a VALUES																	
+ve	−ve	+ve	−ve	+ve	−ve	+ve	−ve	+ve	−ve	+ve	−ve	+ve	−ve	+ve	−ve	+ve	−ve
276	319	134	117	142	202	114	94	83	71	31	23	73	69	28	21	45	48
PERCENTAGE OF POSITIVE AND NEGATIVE ΔChl-a VALUES																	
46	54	53	47	41	59	55	45	54	46	57	43	51	49	57	43	48	52

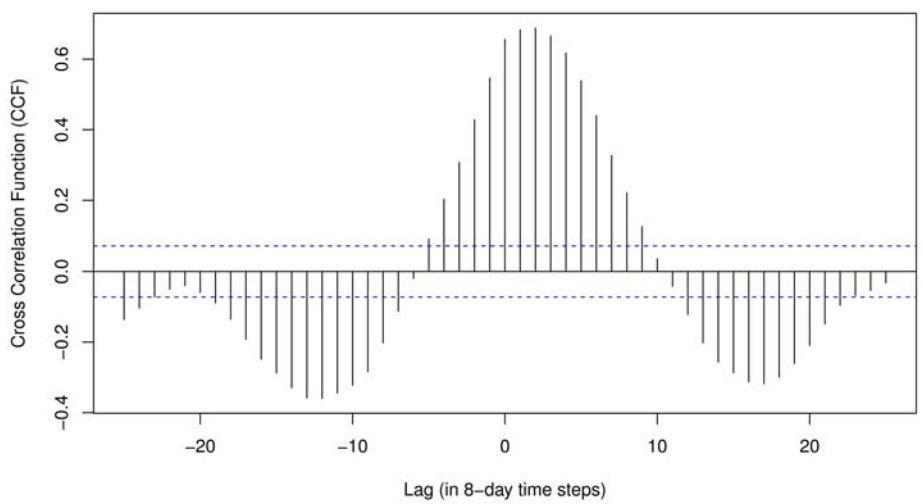


FIGURE 6 | Cross Correlation Function between Chl-a and alongshore wind speed for study area off Somalia. Time step is 8 days.

that, typically, the maximum correlation occurred with a lag, suggesting a functional relationship between the two variables, rather than an artifact. The second is that the seasonal patterns in Chl-a are consistent with the known oceanography of the area, and appear as a consequence of the seasonal changes in the oceanographic conditions, as indicated by the winds and SST. The AOT and AE data from both NASA and CCI also show seasonal changes with high AOT values and low AE during summer monsoon season and vice versa for the rest of the year.

We have used aerosol data from the NASA ocean colour web site, partly to reassure ourselves that the aerosol and Chl-a products that are outputs of the same processing chain do not show inter-dependencies associated with the assumptions that underlie the processing. In the OC-CCI processing version-2 used here, SeaWiFS and MODIS-Aqua data were processed using NASA's SeaDAS software, consistent with the processing chain that generated the aerosol products at the NASA ocean colour website. That the analysis presented here has indicated that the patterns in Chl-a and in the aerosol properties are consistent with the known oceanography of the study area, and that the correlations vary with region (**Figure 2**) as oceanographic and meteorological conditions change, lends some confidence to the quality of the data, in the absence of direct validation data. To further substantiate the application of satellite data to studies of relationship between aerosol and phytoplankton, Aerosol-CCI data sets were also subjected to identical analysis and the data confirmed our findings.

4.2. Aerosols and Phytoplankton in the Western Arabian Sea off Somalia

There have been a few previous studies that dealt with the influence of aerosols on phytoplankton dynamics in the Arabian Sea. A recent study (Banerjee and Prasanna Kumar, 2014) has shown that episodic dust storms could generate phytoplankton

TABLE 4 | Results of multiple linear regression analysis with Chl-a as the dependent variable and NASA-AOT or DOT, alongshore wind speed and SST as independent variables for 1 time step lag.

NASA-AOT	DOT	Wind	SST	*N*	r^2	Adj. r^2	*r*
+	−	+	+	735	0.52	0.52	0.72
+	−	−	−	735	0.30	0.30	0.55
−	−	+	−	735	0.47	0.47	0.69
−	−	−	+	735	0.33	0.33	0.57
+	−	+	−	735	0.49	0.49	0.70
+	−	−	+	735	0.46	0.46	0.68
−	−	+	+	735	0.50	0.49	0.71
−	+	+	+	203	0.25	0.24	0.50
−	+	−	−	203	0.08	0.07	0.28
−	−	+	−	203	0.17	0.17	0.41
−	−	−	+	203	0.20	0.19	0.45
−	+	+	−	203	0.19	0.18	0.44
−	+	−	+	203	0.22	0.21	0.47
−	−	+	+	203	0.24	0.24	0.49

Number of observations (N), r^2, adjusted r^2 and r are shown, for each of the analyses and they are statistically significant ($p < 0.05$). The first set of calculations with 735 observations is for the whole year. The second set, with 203 observations, is for the dust aerosol events. Plus signs indicate variables that were used, and minus signs indicate variables that were excluded in each analysis.

TABLE 5 | Results of multiple linear regression analysis with Chl-a as the dependent variable and CCI-AOT or DOT, alongshore wind speed and SST as independent variables for 3 time step lag.

CCI-AOT	DOT	Wind	SST	*N*	r^2	Adj. r^2	*r*
+	−	+	+	595	0.49	0.49	0.70
+	−	−	−	595	0.29	0.29	0.54
−	−	+	−	595	0.49	0.49	0.70
−	−	−	+	595	0.32	0.32	0.57
+	−	+	−	595	0.49	0.49	0.70
+	−	−	+	595	0.38	0.38	0.62
−	−	+	+	595	0.49	0.49	0.70
−	+	+	+	251	0.39	0.38	0.62
−	+	−	−	251	0.09	0.09	0.30
−	−	+	−	251	0.38	0.38	0.62
−	−	−	+	251	0.15	0.15	0.39
−	+	+	−	251	0.38	0.38	0.62
−	+	−	+	251	0.23	0.22	0.48
−	−	+	+	251	0.38	0.38	0.62

Number of observations (N), r^2, adjusted r^2 and r are shown, for each of the analyses and they are statistically significant ($p < 0.05$). The first set of calculations with 595 observations is for the whole year. The second set, with 251 observations, is for the dust aerosol events. Plus signs indicate variables that were used, and minus signs indicate variables that were excluded in each analysis.

blooms in the central Arabian Sea during the winter monsoon. Nezlin et al. (2010) reported a correlation between Chl-a and aerosols when studying inter-annual variations in the Persian Gulf area. Prasanna Kumar et al. (2010) reported an increasing trend in phytoplankton in the central Arabian Sea during winter months of 1997–2007, and attributed it to increasing supply of iron by dust aerosols. Singh et al. (2008) studied a series of dust storms in the northern Arabian Sea during a 3-year period, and reported chlorophyll enhancement within 1–4 days of dust events, but also pointed out other mechanisms that might be responsible for the relationship observed.

Our results for the western Arabian Sea off Somalia indicate only a possible minor role for dust aerosols enhancing Chl-a concentration during the summer monsoon, supplementing the major role of alongshore winds inducing upwelling favorable for phytoplankton growth. The upwelling component of winds off Somalia during summer monsoon season appears to be far stronger than the classic eastern coastal upwelling zones in the world ocean (Bakun et al., 1998). In the data used here, the wind speed was greater than 15 m/s during summer monsoon season over the Somalia coast. Recently, deCastro et al. (2016) studied the evolution of Somali coastal upwelling under future warming scenarios using models. When the intensity of Somali coastal upwelling during summer monsoon season was projected for the twenty first century, the trends showed that changes in coastal upwelling were mainly related to the wind-induced Ekman transport. Further, our findings are consistent with those of Gallisai et al. (2014) for the Mediterranean: they concluded that the main driver of phytoplankton dynamics is the supply of nutrients from the deep water to the surface layers through vertical mixing. However, the results of the multiple regression presented here do not necessarily imply that the effect of aerosols on Chl-a is only 2%, but only that, because AOT covaries with the other variables, especially wind speed, it is difficult to disentangle their individual effects on Chl-a concentration. Perhaps more interesting is the possibility that the effect of dust events on Chl-a enhancement might be a little stronger during the winter monsoon season and rest of the year than during the summer monsoon season (**Tables 2**, **3**), consistent with the results of Prasanna Kumar et al. (2010) for the central Arabian Sea during winter monsoon season. The direction of the winds during the winter monsoon would suggest an origin in the Asian subcontinent for these dust aerosols, rather than the Arabian peninsula.

We used the cross correlation function to study the phase relationship between aerosol (AOT) and phytoplankton (Chl-a) dynamics. The correlation between the two variables peaked at a lag of 1–2 time steps, with AOT leading. However, since a similar lag was found in the CCF between Chl-a and alongshore winds, it is difficult to attribute a causal relationship to the aerosols by themselves. The phase relationship also throws light on whether or not the biological particles might be enhancing the production of aerosols in the study area. If such events were commonplace, then one would expect that Chl-a enhancement might occur prior to increase in aerosol concentration. The CCF results do not support this in general, but the climatologies of the studied variables (**Figure 3A**) do show that there is a reversal in the phase relationship for a brief period, with Chl-a leading NASA-AOT when Chl-a concentration approaches its peak during the summer monsoon season. However, this result

is not confirmed by CCI-AOT data. Thus, conclusive evidence for biological enhancement of aerosols remains elusive. The intriguing result with the NASA-AOT certainly merits further investigation.

5. CONCLUDING REMARKS

Essential Climate Variables, or ECVs, are our sentinels for observation of climate change. However, to understand climate change, it is not sufficient to study individual ECVs in isolation. Instead, it is also important to study how they interact with each other, and to understand how these interactions might change in the future. Of the marine ECVs, Chl-a concentration is the only biological ECV that is currently amenable to routine observations by remote sensing.

In this paper, we have examined one piece of the puzzle, by studying how the variability of Chl-a in the western Arabian Sea is related to those in three other ECVs: aerosols, winds and SST, focussing more on aerosol- Chl-a interactions, using 16 years of satellite data. What emerges is a complex pattern of relationships, in an area where many ECVs co-vary with each other. While it is difficult to elucidate causal relationships from simple correlations, the phase relationships between the variables can throw some light on the underlying causes.

A question that had to be addressed first, when using satellite data for the analysis, was whether there were artifacts in the patterns in Chl-a, introduced by the atmospheric correction process, which depends to some extent on aerosol optical properties. The correlation between Chl-a and NASA-AOT (a by-product of ocean colour processing) peaking with a lag provided reassurance on this point, since the peak should have been observed at zero lag had processing artifacts been the cause of the correlation. This point was reinforced by repeating the analysis with data from Aerosol-CCI products, which are derived independently of the ocean colour processing chain.

Though the NASA aerosol properties and the CCI aerosol properties are generally consistent with each other, there is a significant phase shift in the time when they peak during the summer monsoon season. The underlying causes for this difference deserve to be investigated further, but fall outside the scope of this paper. In the Somali region, under upwelling regimes, the Chl-a concentration is strongly correlated with wind. Analysis of Ekman Mass Transport supports the hypothesis that wind-induced upwelling is the underlying cause of the high correlation between wind and Chl-a. According to the linear multiple regression analysis, aerosols have a modest effect on Chl-a, at best, with a lag of one to two time steps during this period. An unexpected outcome from this study is related to the importance of dust aerosols in stimulating Chl-a enhancement during the winter monsoon season, suggesting that the abundance of dust aerosols might enhance Chl-a in the absence of wind-induced upwelling.

AUTHOR CONTRIBUTIONS

MS carried out all the data analyses and produced the figures. MS and SS wrote the manuscript. SS conceived the scientific plan, with help from TP. TP provided scientific advice, and led the project. GG contributed to the planning and discussions, and along with AB, provided supervision.

ACKNOWLEDGMENTS

The authors acknowledge Department of Science and Technology (DST), India for the Jawaharlal Nehru Science Fellowship (JNSF) awarded to TP. The authors thank the Director, CMFRI, Kochi for all support and encouragement. This work is a contribution to the Ocean Colour Climate Change Initiative of the European Space Agency, and to the activities of the National Centre for Earth Observations of UK. We also thank the two reviewers for their helpful comments, which have improved the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2017.00386/full#supplementary-material

REFERENCES

Baddock, M., Bullard, J., and Bryant, R. (2009). Dust source identification using MODIS: a comparison of techniques applied to the Lake Eyre Basin, Australia. *Remote Sens. Environ.* 113, 1511–1528. doi: 10.1016/j.rse.2009.03.002

Bakun, A., Roy, C., and Lluch-Cota, S. (1998). *Coastal Upwelling and Other Processes Regulating Ecosystem Productivity and Fish Production in the Western Indian Ocean: Large Marine Ecosystems of the Indian Ocean: Assessment, Sustainability, and Management*. Malden, MA: Blackwell Science.

Balarabe, M., Abdullah, K., and Nawawi, M. (2016). Seasonal variations of aerosol optical properties and identification of different aerosol types based on AERONET data over sub-Sahara West-Africa. *Atmos. Clim. Sci.* 6, 13–28. doi: 10.4236/acs.2016.61002

Banerjee, P., and Prasanna Kumar, S. (2014). Dust-induced episodic phytoplankton blooms in the Arabian Sea during winter monsoon. *J. Geophys. Res. Oceans* 119, 7123–7138. doi: 10.1002/2014JC010304

Banse, K. (1987). Seasonality of phytoplankton chlorophyll in the central and northern Arabian Sea. *Deep Sea Res. A* 34, 713–723. doi: 10.1016/0198-0149(87)90032-X

Banzon, V. F., Evans, R. E., Gordon, H. R., and Chomko, R. M. (2004). SeaWiFS observations of the Arabian Sea southwest monsoon bloom for the year 2000. *Deep Sea Res. Part II* 51, 189–208. doi: 10.1016/j.dsr2.2003.10.004

Brewin, R., Raitsos, D., Dall'Olmo, G., Zarokanellos, N., Jackson, T., Racault, M.-F., et al. (2015). Regional ocean-colour chlorophyll algorithms for the Red Sea. *Remote Sens. Environ.* 165, 64–85. doi: 10.1016/j.rse.2015.04.024

Brindley, H., Osipov, S., Bantges, R., Smirnov, A., Banks, J., Levy, R., et al. (2015). An assessment of the quality of aerosol retrievals over the Red Sea and evaluation of the climatological cloud-free dust direct radiative effect in the region. *J. Geophys. Res. Atmos.* 120, 10862–10878. doi: 10.1002/2015JD023282

Brock, J., Sathyendranath, S., and Platt, T. (1994). A model study of seasonal mixed-layer primary production in the arabian sea. *Proc. Indian Acad. Sci. (Earth Planet Sci.)* 103, 163–176.

Charlson, R. J., Lovelock, J. E., Andreae, M. O., and Warren, S. G. (1987). Oceanic phytoplankton, atmospheric sulphur, cloud albedo and climate. *Nature* 326, 655–661. doi: 10.1038/326655a0

Che, H., Wang, Y., Sun, J., Zhang, X., Zhang, X., Zhang, X., et al. (2013). Variation of aerosol optical properties over the taklimakan desert in China. *Aerosol Air Qual. Res.* 13, 777–785. doi: 10.4209/aaqr.2012.07.0200

Cropp, R. A., Gabric, A. J., McTainsh, G. H., Braddock, R. D., and Tindale, N. (2005). Coupling between ocean biota and atmospheric aerosols: dust, dimethylsulphide, or artifact? *Global Biogeochem. Cycles* 19:GB4002. doi: 10.1029/2004GB002436

deCastro, M., Sousa, M. C., Santos, F., Dias, J. M., and M., G.-G. (2016). How will Somali coastal upwelling evolve under future warming scenarios? *Sci. Rep.* 6:30137. doi: 10.1038/srep30137

de Leeuw, G., Holzer-Popp, T., Bevan, S., Davies, W. H., Descloitres, J., and Grainger, R. G. (2015). Evaluation of seven European aerosol optical depth retrieval algorithms for climate analysis. *Remote Sens. Environ.* 162, 295–315. doi: 10.1016/j.rse.2013.04.023

Donaghay, P. L., Liss, P. S., Duce, R. A., Kester, D. R., Hanson, A. K., Villareal, T., et al. (2015). The role of episodic atmospheric nutrient inputs in the chemical and biological ecosystems. *Oceanography* 4, 62–70. doi: 10.5670/oceanog.1991.04

Dubovik, O., Holben, B. N., Eck, T. F., Smirnov, A., Kaufman, Y. J., King, M. D. M., et al. (2002). Variability of absorption and optical properties of key aerosol types observed in worldwide locations. *J. Atmos. Sci.* 59, 590–608. doi: 10.1175/1520-0469(2002)059<0590:VOAAOP>2.0.CO;2

Duce, R. A., and Tindale, N. W. (1991). Atmospheric transport of iron and its deposition in the ocean. *Limnol. Oceanogr.* 36, 1715–1726. doi: 10.4319/lo.1991.36.8.1715

Eck, T., Holben, B., Dubovik, O., Smirnov, A., Goloub, P., Chen, H., et al. (2005). Columnar aerosol optical properties at AERONET sites in central eastern Asia and aerosol transport to the tropical mid- Pacific. *J. Geophys. Res. Atmos.* 110:D06202. doi: 10.1029/2004JD005274

Eck, T., Holben, B., Reid, J., Dubovik, O., Smirnov, A., O'Neill, N., et al. (1999). Wavelength dependence of the optical depth of biomass burning, urban and desert dust aerosols. *J. Geophys. Res. Atmos.* 104, 31333–31349. doi: 10.1029/1999JD900923

Facchini, M. C., Rinaldi, M., Decesari, S., Carbone, C., Finessi, E., Mircea, M., et al. (2008). Primary submicron marine aerosol dominated by insoluble organic colloids and aggregates. *Geophys. Res. Lett.* 35:L17814. doi: 10.1029/2008GL034210

Fischer, J., Schott, F., and Stramma, L. (1996). Current and transports of the Great Whirl-Socotra Gyre system during the summer monsoon, August 1993. *J. Geophys. Res. Oceans* 101, 3573–3587. doi: 10.1029/95JC03617

Gallisai, R., Peters, F., Volpe, G., Basart, S., and Baldasano, J. M. (2014). Saharan dust deposition may affect phytoplankton growth in the Mediterranean sea at ecological time scales. *PLoS ONE* 9:e110762. doi: 10.1371/journal.pone.0110762

GCOS (2011). *Systematic Observation Requirements from Satellite-based Data Products for Climate 2011 Update. Supplemental Details to the Satellite-based Component of the "Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC*. Technical report, No. 154, World Meteorological Organisation (WMO), Geneva.

Ginoux, P., Chin, M., Tegen, I., Prospero, J., Holben, B., Dubovik, O., et al. (2001). Sources and global distributions of dust aerosols simulated with the GOCART model. *J. Geophys. Res. Atmos.* 106, 20255–20273. doi: 10.1029/2000JD000053

Gittings, J., Raitsos, D., Racault, M.-F., Brewin, R., Pradhan, Y., Sathyendranath, S., et al. (2016). Seasonal phytoplankton blooms in the Gulf of Aden revealed by remote sensing. *Remote Sens. Environ.* 189, 56–66. doi: 10.1016/j.rse.2016.10.043

Glantz, P., Nilsson, E. D., and von Hoyningen-Huene, W. (2009). Estimating a relationship between aerosol optical thickness and surface wind speed over the ocean. *Atmos. Res.* 92, 58–68. doi: 10.1016/j.atmosres.2008.08.010

Gobbi, G. P., Kaufman, Y. J., Koren, I., and Eck, T. F. (2007). Classification of aerosol properties derived from AERONET direct sun data. *Atmos. Chem. Phys.* 7, 453–458. doi: 10.5194/acp-7-453-2007

Goes, J. I., Thoppil, P. G., Gomes, Hdo, R., and Fasullo, J. T. (2005). Warming of the Eurasian landmass is making the Arabian Sea more productive. *Science* 308, 545–547. doi: 10.1126/science.1106610

Goudie, A. S., and Middleton, N. J. (2001). Saharan dust storms: nature and consequences. *Earth Sci. Rev.* 56, 179–204. doi: 10.1016/S0012-8252(01)00067-8

Gregg, W. W., Casey, N. W., and McClain, C. R. (2005). Recent trends in global ocean chlorophyll. *Geophys. Res. Lett.* 32:L03606. doi: 10.1029/2004GL021808

Guleria, R. P., Kuniyal, J., Sharma, N., and Dhyani, P. P. (2012). Seasonal variability in aerosol optical and physical characteristics estimated using the application of the Ångström formula over Mohal in the northwestern Himalaya, India. *J. Earth Syst. Sci.* 121, 697–710. doi: 10.1007/s12040-012-0194-6

Hawkins, L. N., and Russell, L. M. (2010). Polysaccharides, proteins, and phytoplankton fragments: Four chemically distinct types of marine primary organic aerosol classified by single particle spectromicroscopy. *Adv. Meteorol.* 2010, 1–14. doi: 10.1155/2010/612132

Hitchcock, G. L., and Olson, D. B. (1992). *NE and SW Monsoon Conditions Along the Somali Coast during 1987*. Oxford Publishing Co.

Huang, H., Thomas, G. E., and Grainger, R. G. (2010). Relationship between wind speed and aerosol optical depth over remote ocean. *Atmos. Chem. Phys.* 10, 5943–5950. doi: 10.5194/acp-10-5943-2010

Jickells, T. D., An, Z. S., Andersen, K. K., Baker, A. R., Bergametti, G., Brooks, N., et al. (2005). Global iron connections between desert dust, ocean biogeochemistry, and climate. *Science* 308, 67–71. doi: 10.1126/science.1105959

Jordi, A., Basterretxea, G., Tovar-Sánchez, A., Alastuey, A., and Querol, X. (2012). Copper aerosols inhibit phytoplankton growth in the Mediterranean Sea. *Proc. Natl. Acad. Sci. U.S.A.* 109, 21246–21249. doi: 10.1073/pnas.1207567110

Kahn, R. A., Gaitley, B. J., Garay, M. J., Diner, D. J., Eck, T. F., Smirnov, A., et al. (2010). Multiangle Imaging Spectro Radiometer global aerosol product assessment by comparison with the Aerosol Robotic Network. *J. Geophys. Res. Atmos.* 115:D23209. doi: 10.1029/2010JD014601

Kaufman, Y. J. (1993). Aerosol optical thickness and atmospheric path radiance. *J. Geophys. Res. Atmos.* 98, 2677–2692. doi: 10.1029/92JD02427

Kaufman, Y. J., Remer, L. A., Tanrè, D., Li, R.-R., Kleidman, R., Mattoo, S., et al. (2005). A critical examination of the residual cloud contamination and diurnal sampling effects on MODIS estimates of aerosol over ocean. *IEEE Trans. Geosci. Remote Sens.* 43, 2886–2897. doi: 10.1109/TGRS.2005.858430

Kayetha, V. K., Senthil Kumar, J., Prasad, A. K., Cervone, G., and Singh, R. P. (2007). Effect of dust storm on ocean color and snow parameters. *J. Indian Soc. Remote Sens.* 35, 1–9. doi: 10.1007/BF02991828

Koning, E., van Iperen, J. M., van Raaphorst, W., Helder, W., Brummer, G.-J. A., and van Weering, C. E. (2001). Selective preservation of upwelling-indicating diatoms in sediments off Somalia, NW Indian Ocean. *Deep Sea Res. I* 48, 2473–2495. doi: 10.1016/S0967-0637(01)00019-X

Koracin, D., Dorman, C., and Dever, E. (2004). Coastal perturbations of marine-layer winds, wind stress, wind stress curl along California and Baja California in June 1999. *J. Phys. Ocean.* 34, 1152–1173. doi: 10.1175/1520-0485(2004)034<1152:CPOMWW>2.0.CO;2

Large, W., and Pond, S. (1981). Open ocean momentum flux measurements in moderate to strong winds. *J. Phys. Ocean.* 11, 324–336.

Leck, C., and Bigg, E. K. (2005). Biogenic particles in the surface microlayer and overlaying atmosphere in the central Arctic Ocean during summer. *Tellus* 57, 305–316. doi: 10.3402/tellusb.v57i4.16546

Lenes, J. M., Darrow, B. P., Cattrall, C., Heil, C. A., Callahan, M., Vargo, G. A., et al. (2001). Iron fertilization and the Trichodesmium response on the West Florida shelf. *Limnol. Oceanogr.* 46, 1261–1277. doi: 10.4319/lo.2001.46.6.1261

Léon, J.-F., and Legrand, M. (2003). Mineral dust sources in the surroundings of the north Indian Ocean. *Geophys. Res. Lett.* 30, 42–1–42–4. doi: 10.1029/2002GL016690

Lévy, M., Shankar, D., André, J. M., Shenoi, S. S. C., Durand, F., and de Boyer Montégut, C. (2007). Basin-wide seasonal evolution of the Indian Ocean's phytoplankton blooms. *J. Geophys. Res. Ocean* 112:C12014. doi: 10.1029/2007JC004090

Levy, R., Remer, L., Mattoo, S., Vermote, E., and Kaufman, Y. (2007). Second-generation operational algorithm: retrieval of aerosol properties over land from inversion of Moderate Resolution Imaging Spectroradiometer spectral reflectance. *J. Geophys. Res. Atmos.* 112:D13211. doi: 10.1029/2006JD007811

Li, F., and Ramanathan, V. (2002). Winter to summer monsoon variation of aerosol optical depth over the tropical Indian Ocean. *J. Geophys. Res. Atmos.* 107, AAC 2–1–AAC 2–13. doi: 10.1029/2001JD000949

Lovelock, J. (2006). *The Revenge of Gaia: Earth's Climate Crisis and the Fate of Humanity*. New York, NY: Basic Books.

Mahowald, N. M., Baker, A. R., Bergametti, G., Brooks, N., Duce, R. A., Jickells, T. D., et al. (2005). Atmospheric global dust cycle and iron inputs to the ocean. *Global Biogeochem. Cycles* 19:GB4025. doi: 10.1029/2004GB002402

Mallet, M., Chami, M., Gentili, B., Sempéré, R., and Dubuisson, P. (2009). Impact of sea-surface dust radiative forcing on the oceanic primary production: a 1D modeling approach applied to the West African coastal waters. *Geophys. Res. Lett.* 36:L15828. doi: 10.1029/2009GL039053

Martin, J. H., Coale, K. H., Johnson, K. S., Fitzwater, S. E., Gordon, R. M., Tanner, S. J., et al. (1994). Testing the iron hypothesis in ecosystems of the equatorial Pacific Ocean. *Nature* 371, 123–129. doi: 10.1038/371123a0

Martin, J. H., Gordon, R. M., and Fitzwater, S. E. (1991). The case for iron. *Limnol. Oceanogr.* 36, 1793–1802. doi: 10.4319/lo.1991.36.8.1793

McCreary, J. P., Han, W., Shankar, D., and Shetye, S. R. (1996). Dynamics of the east India coastal current: 2. Numerical solutions. *J. Geophys. Res. Oceans* 101, 13993–14010. doi: 10.1029/96JC00560

Meskhidze, N., Chameides, W. L., and Nenes, A. (2005). Dust and pollution: a recipe for enhanced ocean fertilization? *J. Geophys. Res. Atmos.* 110:D03301. doi: 10.1029/2004JD005082

Meskhidze, N., and Nenes, A. (2010). Effects of ocean ecosystem on marine aerosol-cloud interaction. *Adv. Meteorol.* 2010:239808. doi: 10.1155/2010/239808

Muhs, D. R., Budahn, J. R., Prospero, J. M., and Carey, S. N. (2007). Geochemical evidence for African dust inputs to soils of western Atlantic islands: barbados, the Bahamas and Florida. *J. Geophys. Res. Earth Surface* 112:F02009. doi: 10.1029/2005JF000445

Mulcahy, J. P., O'Dowd, C. D., Jennings, S. G., and Ceburnis, D. (2008). Significant enhancement of aerosol optical depth in marine air under high wind conditions. *Geophys. Res. Lett.* 35:L16810. doi: 10.1029/2008GL034303

Myhre, G., Stordal, F., Johnsrud, M., Diner, D. J., Geogdzhayev, I. V., Haywood, J. M., et al. (2005). Intercomparison of satellite retrieved aerosol optical depth over ocean during the period September 1997 to December 2000. *Atmos. Chem. Phys.* 5, 1697–1719. doi: 10.5194/acp-5-1697-2005

Nezlin, N. P., Polikarpov, I. G., Al-Yamani, F. Y., Subba Rao, D. V., and Ignatov, A. M. (2010). Satellite monitoring of climatic factors regulating phytoplankton variability in the Arabian (Persian) Gulf. *J. Mar. Syst.* 82, 47–60. doi: 10.1016/j.jmarsys.2010.03.003

O'Dowd, C. D., Smith, M. H., Consterdine, I. E., and Lowe, J. A. (1997). Marine aerosol, sea-salt, and the marine sulphur cycle: a short review. *Atmos. Environ.* 31, 73–80. doi: 10.1016/S1352-2310(96)00106-9

Owens, N. J. P., Burkill, P. H., Mantoura, R. F. C., Woodward, E. M. W., Bellan, I. E., Aiken, J., et al. (1993). Size-fractionated primary production and nitrogen assimilation in the northwestern Indian Ocean. *Deep Sea Res. II* 40, 697–709. doi: 10.1016/0967-0645(93)90053-P

Pakszys, P., Zielinski, T., Markowicz, K., Petelski, T., Makuch, P., Lisok, J., et al. (2015). *Annual Changes of Aerosol Optical Depth and Ångström Exponent over Spitsbergen*. Switzerland: Springer.

Papaynannis, A., Zhang, H. Q., Amiridis, V., Ju, H. B., Chourdakis, G., Georgussis, G., et al. (2007). Extraordinary dust event over Beijing, China, during April 2006: Lidar, Sun photometric, satellite observations and model validation. *Geophys. Res. Lett.* 34:L07806. doi: 10.1029/2006GL029125

Patra, P. K., Kumar, M. D., Mahowald, N., and Sarma, V. V. S. S. (2007). Atmospheric deposition and surface stratification as controls of contrasting chlorophyll abundance in the North Indian Ocean. *J. Geophys. Res. Ocean* 112:C05029. doi: 10.1029/2006JC003885

Paytan, A., Mackeya, K. R., Chena, Y., Lima, I. D., Doney, S. C., Mahowald, N., et al. (2009). Toxicity of atmospheric aerosols on marine phytoplankton. *Proc. Natl. Acad. Sci. U.S.A.* 106, 4601–4605. doi: 10.1073/pnas.0811486106

Pease, P. P., Chakerian, V. P. T., and Tindale, N. W. (1998). Aerosols over the Arabian Sea: geochemistry and source areas for aeolian desert dust. *J. Arid Environ.* 39, 477–496. doi: 10.1006/jare.1997.0368

Pereira, S. N., Wagner, F., , and Silva, A. M. (2011). Seven years of measurements of aerosol scattering properties, near the surface, in the southwestern Iberia Peninsula. *Atmos. Chem. Phys.* 11, 17–29. doi: 10.5194/acp-11-17-2011

Popp, T., de Leeuw, G., Bingen, C., Brühl, C., Capelle, V., Chedin, A., et al. (2016). Development, production and evaluation of aerosol climate data records from European satellite observations (Aerosol-CCI). *Remote Sens.* 8:421. doi: 10.3390/rs8050421

Prakash, P., Prakash, S., Rahaman, H., Ravichandran, M., and Nayak, S. (2012). Is the trend in chlorophyll-a in the Arabian Sea decreasing? *Geophys. Res. Lett.* 39:L23605. doi: 10.1029/2012GL054187

Prasanna Kumar, S., Madhupratap, M., Dileep Kumar, M., Muraleedharan, P. M., De Souza, S. N., Gauns, M., et al. (2001). High biological productivity in the central Arabian Sea during the summer monsoon driven by Ekman pumping and lateral advection. *Curr. Sci.* 81, 1633–1638.

Prasanna Kumar, S., Roshin, R. P., Narvekar, J., Dinesh Kumar, P. K., and Vivekanandan, E. (2010). What drives the increased phytoplankton biomass in the Arabian Sea? *Curr. Sci.* 99, 101–106.

Prospero, J. M., Ginoux, P., Torres, O., Nicholson, S. E., and Gill, T. E. (2002). Environmental characterization of global sources of atmospheric soil dust identified with the NIMBUS 7 Total Ozone Mapping Spectrometer (TOMS) absorbing aerosol product. *Rev. Geophys.* 40, 2–1–2–31. doi: 10.1029/2000RG000095

Quinn, P. K., and Bates, T. S. (2011). The case against climate regulation via oceanic phytoplankton sulphur emissions. *Nature* 480, 51–56. doi: 10.1038/nature10580

Salinas, S. V., Chew, B. N., and Liew, S. C. (2009). Retrievals of aerosol optical depth and Ångström exponent from ground-based Sun-photometer data of Singapore. *Appl. Opt.* 48, 1473–1484. doi: 10.1364/AO.48.001473

Santos, F., Gomez-Gesteira, M., deCastro, M., and Días, J. M. (2015). A dipole-like SST trend in the Somalia region during the monsoon season. *J. Geophys. Res. Oceans* 120, 597–607. doi: 10.1002/2014JC010319

Satheesh, S. K., Srinivasan, J., and Moorthy, K. K. (2006). Contribution of sea-salt to aerosol optical depth over the Arabian Sea derived from MODIS observations. *Geophys. Res. Lett.* 33:L03809. doi: 10.1029/2005GL024856

Sathyendranath, S., Groom, S., Grant, M., Brewin, R. J. W., Thompson, A., Chuprin, A., et al. (2016). *ESA Ocean Colour Climate Change Initiative (Ocean-Colour-CCI): Version 2.0 Data*. Centre for Environmental Data Analysis.

Schott, F. (1983). Monsoon response of the Somali Current and associated upwelling. *Prog. Oceanogr.* 12, 357–381. doi: 10.1016/0079-6611(83)90014-9

Schott, F., Dengler, M., and Schoenefeldt, R. (2002). The shallow overturning circulation of the Indian Ocean. *Prog. Oceanogr.* 53, 57–103. doi: 10.1016/S0079-6611(02)00039-3

Schott, F., Fischer, J., Garternicht, U., and Quadfasel, D. (1997). Summer monsoon response of the Northern Somali Current, 1995. *Geophys. Res. Lett.* 24, 2565–2568. doi: 10.1029/97GL00888

Schott, F., and McCreary, J. P. (2001). The monsoon circulation of the Indian Ocean. *Prog. Oceanogr.* 51, 1–123. doi: 10.1016/S0079-6611(01)00083-0

Schuster, G. L., Dubovik, O., and Holben, B. N. (2006). Ångström exponent and bimodal aerosol size distributions. *J. Geophys. Res. Atmos.* 111:D07207. doi: 10.1029/2005JD006328

Shankar, D., Vinayachandran, P. N. N., Unnikrishnan, A. S. S., and Shetye, S. R. (2002). The monsoon currents in the north Indian Ocean. *Prog. Oceanogr.* 52, 63–120. doi: 10.1016/S0079-6611(02)00024-1

Shinozuka, Y., Redemann, J., Livingston, J. M., Russell, P. B., Clarke, A. D., Howell, S. G., et al. (2011). Airborne observation of aerosol optical depth during ARC-TAS: vertical profiles, inter-comparison and fine-mode fraction. *Atmos. Chem. Phys.* 11, 3673–3688. doi: 10.5194/acp-11-3673-2011

Singh, R. P., Prasad, A. K., Kayetha, V. K., and Kafatos, M. (2008). Enhancement of oceanic parameters associated with dust storms using satellite data. *J. Geophys. Res. Oceans* 113:C11008. doi: 10.1029/2008JC004815

Smirnov, A., Holben, B. N., Eck, T. F., Dubovik, O., and Slutsker, I. (2003). Effect of wind speed on columnar aerosol optical properties at Midway Island. *J. Geophys. Res. Atmos.* 108, AAC 15–1–AAC 15–8. doi: 10.1029/2003JD003879

Smith, S. L., and Codispoti, L. A. (1980). Southwest monsoon of 1979: chemical and biological response of Somali coastal waters. *Science* 209, 597–600. doi: 10.1126/science.209.4456.597

Toledano, C., Cachorro, V., Benjon, A., de Frutos, A., Sorribas, M., de la Morena, B., et al. (2007). Aerosol optical depth and Ångström exponent climatology at El arenosillo AERONET site (Huelva, Spain). *Q. J. R. Meteorol. Soc.* 133, 795–807. doi: 10.1002/qj.54

Torres, O., Tanskanen, A., Veiheiman, B., Ahn, C., Braak, R., Bhartia, P., et al. (2007). Aerosols and surface UV products from Ozone Monitoring

Instrument observations: an overview. *J. Geophys. Res. Atmos.* 112:D24S47. doi: 10.1029/2007JD008809

Trenberth, K. E., Large, W. G., and Olson, J. G. (1990). The mean annual cycle in global ocean wind stress. *J. Phys. Oceanogr.* 20, 1742–1760.

Valenzuela, A., Olmo, F., Lyamani, H., Antón, M., Titos, G., Cazorla, A., et al. (2014). Aerosol scattering and absorption Ångström exponents as indicators of dust and dust-free days over Granada (Spain). *Atmos. Res.* 154, 1–13. doi: 10.1016/j.atmosres.2014.10.015

Wiggert, J. D., Hood, R. R., Banse, K., and Kindle, J. C. (2005). Monsoon-driven biogeochemical processes in the Arabian Sea. *Prog. Oceanogr.* 65, 176–213. doi: 10.1016/j.pocean.2005.03.008

Wiggert, J. D., and Murtugudde, R. G. (2007). The sensitivity of the southwest monsoon phytoplankton bloom to variations in aeolian iron deposition over the Arabian Sea. *J. Geophys. Res. Oceans* 112:C05005. doi: 10.1029/2006JC003514

Yoon, J., von Hoyningen-Huene, W., Kokhanovsky, A. A., Vountas, M., and Burrows, J. P. (2012). Trend analysis of aerosol optical thickness and Ångström exponent derived from the global AERONET spectral observations. *Atmos. Meas. Tech.* 5, 1271–1299. doi: 10.5194/amt-5-1271-2012

Zhu, A., Ramanathan, V., Li, F., and Kim, D. (2007). Dust plumes over the Pacific, Indian, and Atlantic oceans: climatology and radiative impact. *J. Geophys. Res. Atmos.* 112:D16208. doi: 10.1029/2007JD008427

Zu, J., Che, H. Z., Xia, X. G., Chen, H. B., Goloub, P., and Zhang, W. X. (2014). Column- integrated aerosol optical and physical properties at regional background atmosphere in North China Plain. *Atmos. Environ.* 84, 54–60. doi: 10.1016/j.atmosenv.2013.11.019

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Shafeeque, Sathyendranath, George, Balchand and Platt. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Marine Science

ORIGINAL RESEARCH
published: 14 February 2018
doi: 10.3389/fmars.2018.00003

Satellite Radiation Products for Ocean Biology and Biogeochemistry: Needs, State-of-the-Art, Gaps, Development Priorities, and Opportunities

Robert Frouin[1], Didier Ramon[2], Emmanuel Boss[3], Dominique Jolivet[2], Mathieu Compiègne[2], Jing Tan[1], Heather Bouman[4], Thomas Jackson[5], Bryan Franz[6], Trevor Platt[5] and Shubha Sathyendranath[5]*

[1] Scripps Institution of Oceanography, La Jolla, CA, United States, [2] HYGEOS, Euratechnologies, Lille, France, [3] School of Marine Sciences, University of Maine, Orono, ME, United States, [4] Department of Earth Sciences, University of Oxford, Oxford, United Kingdom, [5] Plymouth Marine Laboratory, Plymouth, United Kingdom, [6] NASA Goddard Space Flight Center, Greenbelt, MD, United States

OPEN ACCESS

Edited by:
Laura Lorenzoni, University of South Florida, United States

Reviewed by:
Andrew Clive Banks, National Physical Laboratory, United Kingdom
Giuseppe Zibordi, Joint Research Centre, Italy

***Correspondence:**
Robert Frouin
rfrouin@ucsd.edu

Specialty section:
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

Received: *24 April 2017*
Accepted: *10 January 2018*
Published: *14 February 2018*

Citation:
Frouin R, Ramon D, Boss E, Jolivet D, Compiègne M, Tan J, Bouman H, Jackson T, Franz B, Platt T and Sathyendranath S (2018) Satellite Radiation Products for Ocean Biology and Biogeochemistry: Needs, State-of-the-Art, Gaps, Development Priorities, and Opportunities. Front. Mar. Sci. 5:3. doi: 10.3389/fmars.2018.00003

Knowing the spatial and temporal distribution of the underwater light field, i.e., the spectral and angular structure of the radiant intensity at any point in the water column, is essential to understanding the biogeochemical processes that control the composition and evolution of aquatic ecosystems and their impact on climate and reaction to climate change. At present, only a few properties are reliably retrieved from space, either directly or via water-leaving radiance. Existing satellite products are limited to planar photosynthetically available radiation (PAR) and ultraviolet (UV) irradiance above the surface and diffuse attenuation coefficient. Examples of operational products are provided, and their advantages and drawbacks are examined. The usefulness and convenience of these products notwithstanding, there is a need, as expressed by the user community, for other products, i.e., sub-surface planar and scalar fluxes, average cosine, spectral fluxes (UV to visible), diurnal fluxes, absorbed fraction of PAR by live algae (APAR), surface albedo, vertical attenuation, and heating rate, and for associating uncertainties to any product on a pixel-by-pixel basis. Methodologies to obtain the new products are qualitatively discussed in view of most recent scientific knowledge and current and future satellite missions, and specific algorithms are presented for some new products, namely sub-surface fluxes and average cosine. A strategy and roadmap (short, medium, and long term) for usage and development priorities is provided, taking into account needs and readiness level. Combining observations from satellites overpassing at different times and geostationary satellites should be pursued to improve the quality of daily-integrated radiation fields, and products should be generated without gaps to provide boundary conditions for general circulation and biogeochemical models. Examples of new products, i.e., daily scalar PAR below the surface, daily average cosine for PAR, and sub-surface spectral scalar fluxes are presented. A procedure to estimate algorithm uncertainties in the total uncertainty budget for above-surface daily PAR, based on radiative simulations for expected situations, is described. In the future, space-borne

lidars with ocean profiling capability offer the best hope for improving our knowledge of sub-surface fields. To maximize temporal coverage, space agencies should consider placing ocean-color instruments in L1 orbit, where the sunlit part of the Earth can be frequently observed.

Keywords: photosynthetically available radiation, average cosine, attenuation coefficient, ocean color, remote sensing

INTRODUCTION

From the point of view of biology and chemistry, solar radiation in the photosynthetically active range (roughly 400–700 nm), referred to as PAR, controls the growth of aquatic plants (e.g., Ryther, 1956; Platt et al., 1977; Kirk, 1994; Falkowski and Raven, 1997). It ultimately regulates the composition and dynamics of marine ecosystems. Solar radiation in the ultraviolet (UV) (280–400 nm), by damaging cellular constituents, may stress phytoplankton and inhibit their growth (e.g., Cullen and Neale, 1994; Häder et al., 2011). UV light, via photo-oxidation of colored dissolved organic matter (CDOM), may increase the bioavailability of nutrients (Sulzberger and Durisch-Kaiser, 2009). In the process, absorption by CDOM is reduced, increasing light penetration. Knowing the distribution (spectral, spatial, and temporal) of UV and visible solar radiation in the upper ocean is critical to understanding biogeochemical cycles of carbon, nutrients, and oxygen, and to addressing climate and global change issues, such as the fate of anthropogenic atmospheric carbon dioxide (CO_2).

From the point of view of physics, sunlight absorbed by phytoplankton and other water constituents (CDOM, mineral particles, etc.) heats the upper ocean and distributes heat horizontally and vertically, affecting mixed-layer dynamics and oceanic circulation (e.g., Nakamoto et al., 2000, 2001; Ballabrera-Poy et al., 2007). These changes in turn influence atmospheric temperature and circulation, with remote effects (Miller et al., 2003; Shell et al., 2003). Solar radiation diffusely reflected by the ocean also affects the outgoing radiative flux from the planet (planetary albedo), with climate consequences (Frouin and Iacobellis, 2002). In order to make predictions for future conditions, we need to get some idea of how the phytoplankton concentrations and optical properties will evolve with changing conditions. Many processes and feedbacks in which solar radiation absorption plays a role are involved and difficult to untangle, and a large fraction of the uncertainties in projections of future climate is associated with physical-biological interactions (Friedlingstein et al., 2006).

This article reviews operational satellite radiation products, the user needs and gaps, and it provides a scientific roadmap for the use of, and priorities for improving existing products and developing new products, i.e., for closing the gaps in ocean biology and biogeochemistry, including studies of biological-physical interactions and feedbacks. This roadmap emerged from the presentations (oral and poster) and discussions during the Color and Light in the Ocean from Earth Observation (CLEO) workshop at ESRIN, Frascati, Italy on 6–8 September 2016. The following questions are addressed: (1) Do existing shortwave downward flux products meet the requirements of the dynamics and bio-geochemical communities? What can be done to serve better the needs of the user community in general and the modeling community in particular? (2) What additional products should be added to the processing streams to increase their usefulness? What should be the characteristics of these products in terms of temporal, spatial, and spectral resolution, spectral range, and accuracy? (3) What are the needs in terms of harmonization between sensors, methodologies, ancillary data and radiative transfer tools?

Current Products

The underwater light field is defined at any point in space by the spectral radiance ($W/m^2/sr/nm$) from all directions. Useful properties can be derived from radiance, i.e., planar and scalar irradiance, average cosine, reflectance, and vertical attenuation coefficient (see, e.g., Mobley, 1994 or Kirk, 1994 for definitions). Only a few of these properties are presently inferred reliably and operationally from space, namely daily above-surface downward planar irradiance integrated from 400 to 700 nm (known as "daily PAR product"), above-surface planar spectral UV irradiance at noon, spectral reflectance of the water body, and diffuse attenuation coefficient at 490 nm (derived from water reflectance). These products are generated for each ocean color mission individually (note that UV products are not available in standard ocean-color missions).

Examples of level-3 daily, weekly, and monthly above-surface PAR products at 9 km resolution from MODIS-Aqua data are displayed in **Figure 1**. Dates are March 22, March 22–29, and March 1–31, 2010, respectively. The NASA Ocean Color Biology Group (OBPG) in Greenbelt, Maryland generates and archives these products operationally. Typical uncertainty (RMS) is ± 6.5 (19%), ± 4.2 (12%), and ± 2.6 $Em^{-2}d^{-1}$ (7%) for daily, weekly, and monthly estimates (Frouin et al., 2012). The daily maps exhibit missing values, especially at low latitudes, due to the limited spatial coverage of the instruments, but the weekly and a fortiori monthly maps are completely filled, except at latitudes where the Sun zenith angle at the time of satellite overpass is above 75 degrees, since the data is discarded in the OBPG ocean-color processing line. The weekly and monthly PAR fields exhibit similar patterns, but the monthly product is smoother (lower variability at small scales), which is expected when the averaging period is longer.

The corresponding maps of level-3 daily, weekly, and monthly diffuse attenuation coefficient (K_d) at 490 nm, also produced by the OBPG, are displayed in **Figure 2**. RMS uncertainty on log-transformed instantaneous estimates is about ± 0.1 (e.g., Morel et al., 2007). Since K_d is only retrieved in clear sky conditions, the

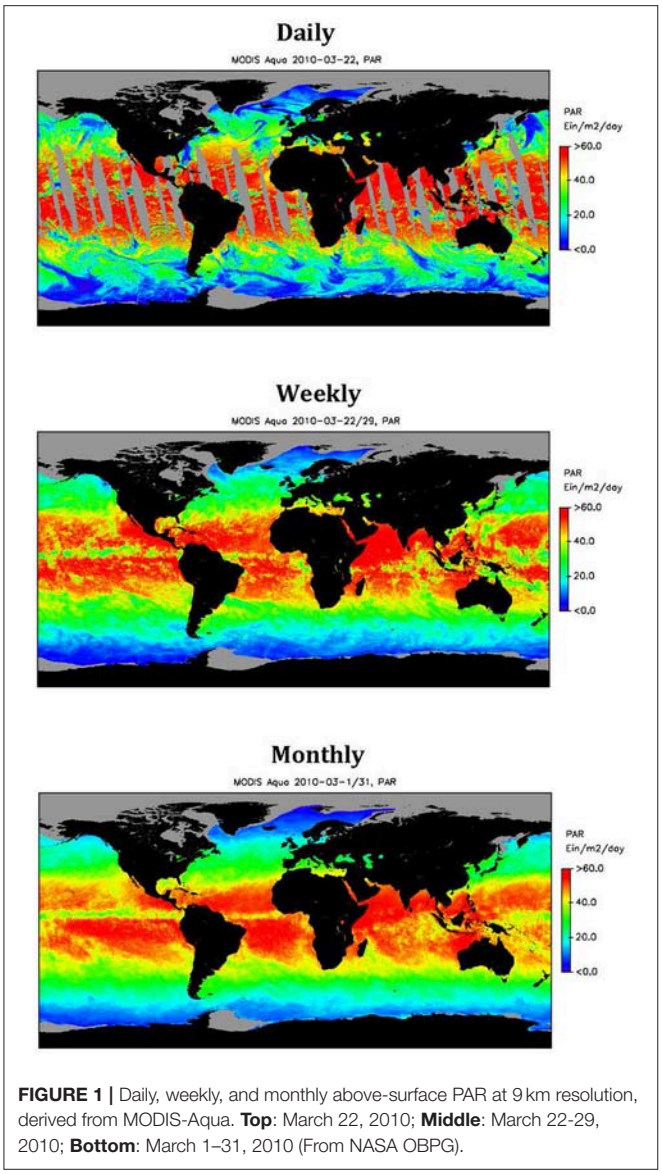


FIGURE 1 | Daily, weekly, and monthly above-surface PAR at 9 km resolution, derived from MODIS-Aqua. **Top**: March 22, 2010; **Middle**: March 22-29, 2010; **Bottom**: March 1–31, 2010 (From NASA OBPG).

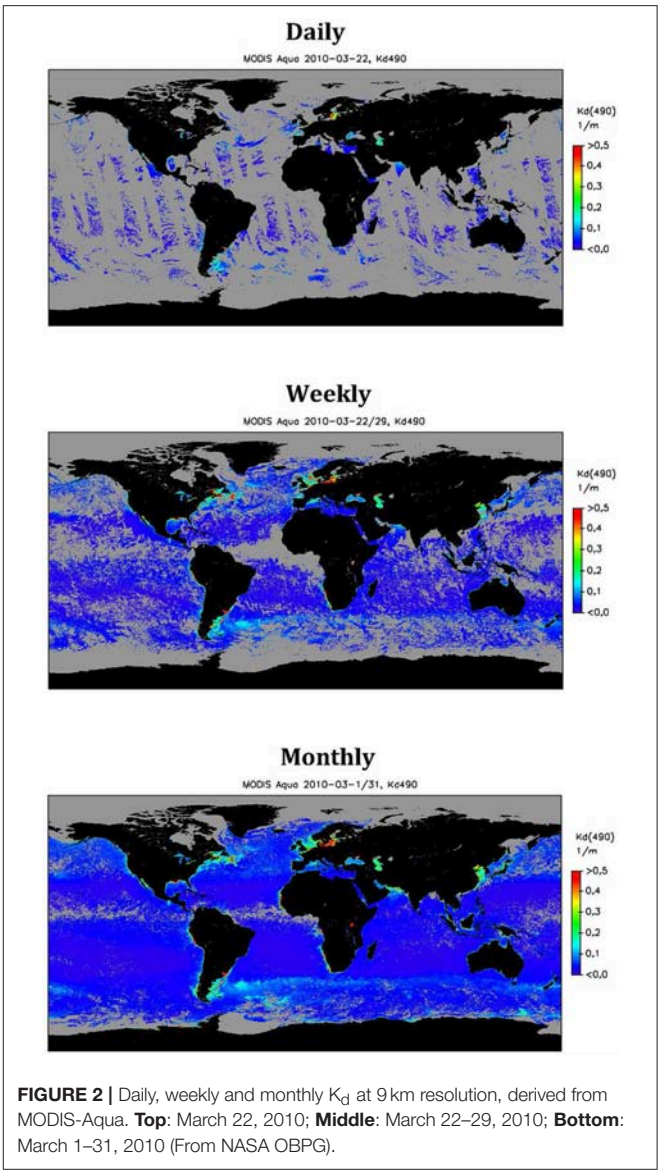


FIGURE 2 | Daily, weekly and monthly K_d at 9 km resolution, derived from MODIS-Aqua. **Top**: March 22, 2010; **Middle**: March 22–29, 2010; **Bottom**: March 1–31, 2010 (From NASA OBPG).

daily map has many gaps (only about 10–15 % of the observed pixels typically pass through the strict glint and cloud filters), and the weekly product show many areas with no information. On a monthly time scale, information is still missing at low latitudes. The spatial gaps limit considerably the utility of the K_d retrievals for propagating light below the surface. In the open ocean, global coverage every 3–5 days is necessary to resolve variability associated with seasonal biological phenomena such as phytoplankton blooms. In coastal waters, wind forcing create "events" (e.g., upwelling) that occur every 2–10 days, and 1-day coverage is the requirement for resolving the event time scale.

Figure 3 provides examples of level 3 daily, weekly, and monthly maps of noon surface UV irradiance at 1° resolution and 324 nm from OMI-Aura. Irradiance at 305, 310, and 380 nm as well as erythemally weighted daily dose and erythemal dose rate are also available. Dates are the same as those of **Figure 1**. The data are routinely processed at the OMI Science Investigator-led Processing System (SIPS) Facility in Greenbelt, Maryland, and are archived at the NASA Goddard Earth Sciences Data and Information Services Center (GES DISC). Overall uncertainty of UV irradiance estimates ranges from ±5 to over ±30%, depending on atmospheric conditions and geolocation (e.g., Arola et al., 2009). The UV irradiance and PAR fields have similar spatial coverage (**Figures 1**, **3**), with missing values due to instrument swath on a daily time scale, except that UV irradiance estimates are obtained at high latitudes. Variability patterns are also similar, since UV irradiance variability is also governed by Sun zenith angle and cloudiness, although ozone absorption plays a bigger role in modulating the surface values (but gradients of total ozone content remain mostly latitudinal). Spatial resolution is coarser than for the MODIS products (OMI sensor footprint is

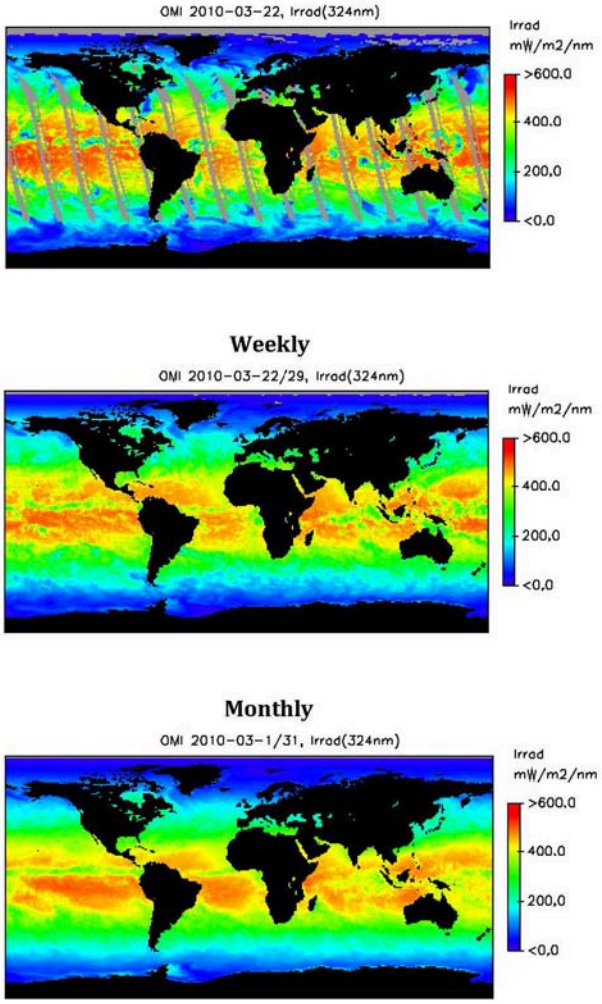


FIGURE 3 | Daily, weekly and monthly noon above-surface UV irradiance at 1° resolution, derived from OMI-Aura. **Top**: March 22, 2010; **Middle**: March 22–29, 2010; **Bottom**: March 1-31, 2010. (From NASA GES DISC).

$13 \times 24\ km^2$), , and no daily-averaged values (only daily values at noon or overpass time) are generated.

The situation regarding spatial coverage is summarized in **Figure 4**, which displays the percentage of the ocean surface covered by PAR, K_d, and UV irradiance products on daily, weekly, and monthly time scales (imagery of **Figures 1–3**). In the equatorial region, percent coverage is 65, 80, and 5% for daily PAR, UV irradiance at 324 nm, and K_d, respectively. It is increased to 100% for weekly and monthly PAR and UV irradiance, and to 30 and 70% for weekly and monthly K_d. In middle latitude regions, daily coverage is almost 100% for PAR, about 80% for UV irradiance, and 15–20% for K_d. Weekly and monthly coverage is 100% for PAR and UV irradiance, and reaches 70–75 and 100% in sub-tropical regions for K_d. At high latitudes (>70°), monthly coverage is <40% for PAR and K_d. This lack of coverage is limiting in view of the large productivity of high latitude marine ecosystems (Southern ocean, Arctic ocean). Furthermore, monthly K_d products may only contain estimates during a few days, and therefore may not represent accurately actual monthly values in dynamic regions (IOCCG, 2015).

Multiple satellites can improve the daily ocean coverage, especially for K_d (can only be retrieved in clear sky conditions). For example, three instruments of MODIS type, flying in a constellation on satellite orbits differing by the mean anomaly (angular distance from pericenter), would increase the daily spatial coverage of water reflectance, therefore K_d, from 15 to 25% over 1 day and from 40 to 60% over 4 days (Gregg et al., 1998). **Figure 5** shows, for March 22, 2010, the increase in daily ocean coverage obtained by combining estimates from MODIS-Aqua (overpass at 13:30 local time) and –Terra (overpass at 10:30 local time) instead of using MODIS-Aqua only. For PAR, the increase is from 60 to 80% in equatorial regions, and complete coverage is reached at sub-tropical latitudes. For K_d, the ocean coverage is more than doubled at most latitudes (e.g., 35–50%, instead of 15–20% in the sub-tropics). Combining PAR estimates from instruments orbiting at different times not only increases spatial coverage, but perhaps more importantly, also takes into account cloud diurnal variability, yielding more accurate estimates.

Satellite instruments in geostationary orbit, by observing the same target multiple times during the day (e.g., every 30 min for GOCI onboard COMS and 10 min for AHI onboard Hiwamari-8) offer an efficient way to account for diurnal changes in cloudiness in daily PAR products. **Figure 6** displays daily PAR imagery obtained with GOCI data acquired on April 5, 2011 at 3:16 GMT and at 00:16, 01:16, 02:16, 0.3:16, 04:16, 05:16, 06:16, and 07:16 GMT. In clear-sky regions (north of Japan), the values are close using one or 8 observations, which is expected since the governing parameter is the Sun zenith angle. In cloudy regions (South of Japan), the PAR spatial field is smoother and the range of values smaller since cloudiness changes are accounted for in the daily average. The lowest value is about 12 $Em^{-2}d^{-1}$ with 8 observations instead of 5 $Em^{-2}d^{-1}$ with one observation. The two types of estimates compare well, with a bias of 0.11 $Em^{-2}d^{-1}$ (0.2%), i.e., slightly higher values using one observation, and a root-mean-squared (RMS) difference of 5.92 $Em^{-2}d^{-1}$ (13.5%), largely due to differences in cloudy situations. **Figure 7** displays an example of AHI daily PAR product (July 20, 2011) and the corresponding MODIS-Aqua product. Observations every 10 min were used to estimate daily PAR from AHI data. The patterns of variability are similar in both products, but as for GOCI, the AHI PAR imagery is smoother and contains less extreme values. The AHI product, unlike the MODIS product, does not exhibit spatial gaps. In terms of comparison statistics, the AHI values are lower by 1.38 $Em^{-2}d^{-1}$ (4.8%) on average, and the RMS difference is 6.46 $Em^{-2}d^{-1}$ (22.7%).

In summary, existing satellite products generally do not cover the global open oceans (e.g., retrievals limited to Sun zenith angles < 75°), except for UV irradiance, and they do not provide information below sea ice, where significant blooms may develop (e.g., Arrigo et al., 2012). In addition, cloud diurnal variability

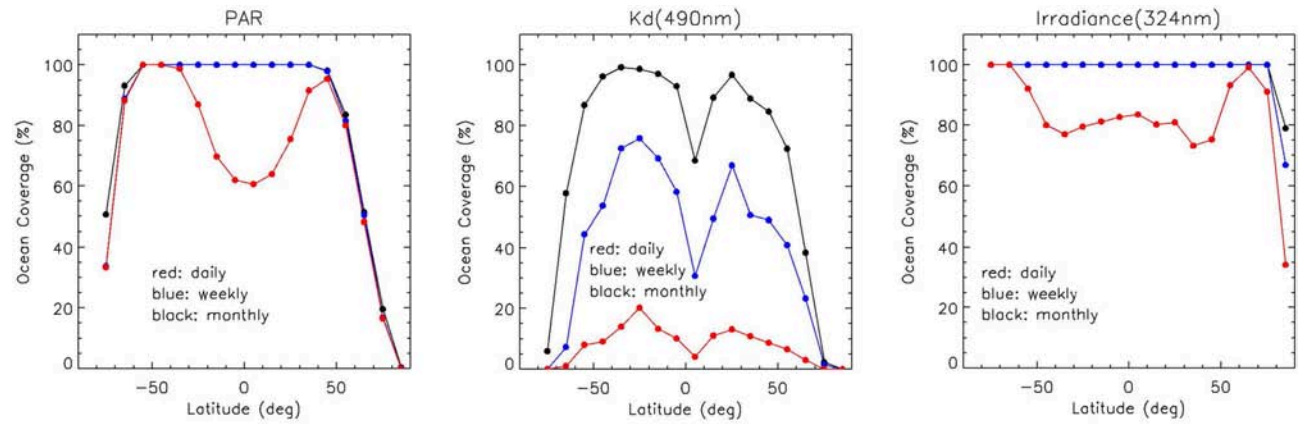


FIGURE 4 | Percentage of the ocean surface covered by **(Left)** PAR, **(Middle)** K_d, and **(Right)** noon irradiance products on daily, weekly, and monthly time scales, i.e., March 22, March 22–29, and March 1–31, 2010. Situation is dramatic for K_d, which is only retrieved in clear-sky conditions (<15% coverage at most latitudes).

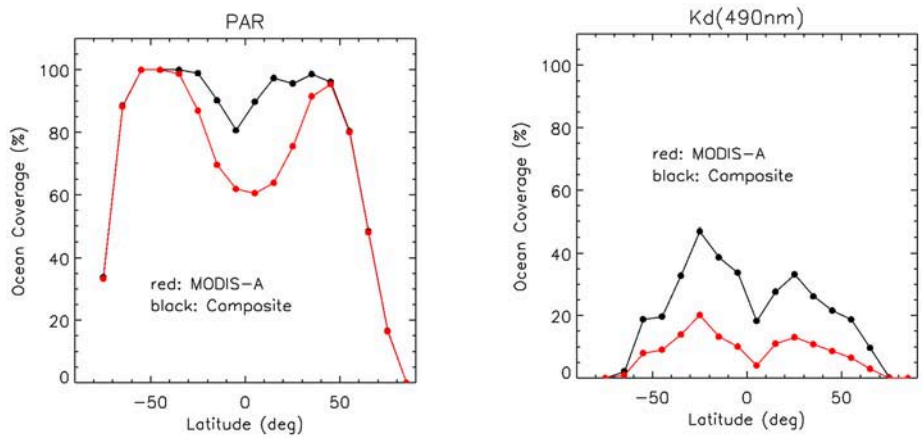


FIGURE 5 | Percentage of ocean coverage by daily PAR and K_d products when using MODIS-Aqua data only (red curves) and combining MODIS-Aqua and -Terra data (black curves). Date is March 22, 2010. Coverage is almost total (>80%) in low to middle latitude regions for PAR **(left)** and at least doubled at most latitudes for K_d **(right)**.

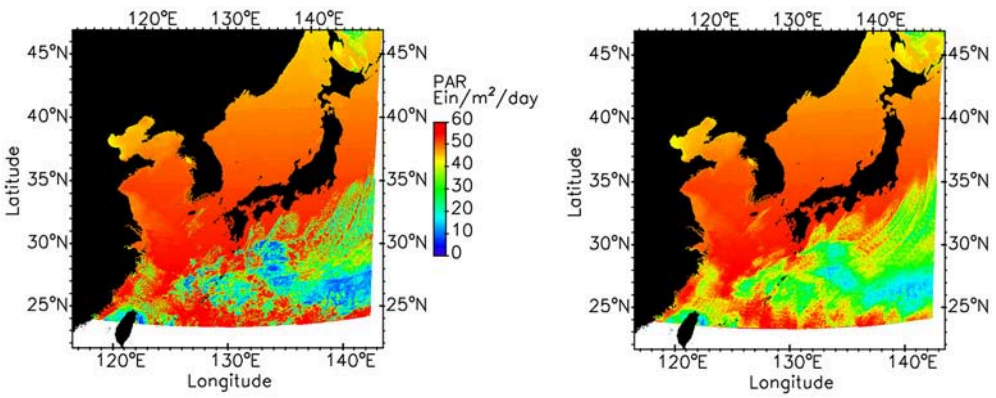


FIGURE 6 | Daily above-surface PAR distribution for April 5, 2011 obtained from 8 hourly GOCI observations during the day **(right)** and a single GOCI observation at 03:16 GMT **(left)**. In the stormy region South of Japan, the range of PAR values is smaller when using 8 observations due to diurnal changes in cloudiness. Reproduced with permission from Springer Nature (After Frouin and McPherson, 2012).

is not accounted for in daily PAR calculations when using polar orbiting satellites. Geostationary satellites observing frequently during the day account properly for cloudiness changes, but the drawback is a decreased spatial resolution at high latitudes and potentially large uncertainties for slanted viewing geometries. Propagation of surface radiation to depth currently assumes that the ocean is homogenous, neglecting potentially important effects of stratification on the absorption of solar radiation. In sum, our view of the underwater light field from space is limited. Nevertheless, the operational radiation products have been used to address a variety of topics related to aquatic photosynthesis, for example biosphere productivity during an El Niño transition (Behrenfeld et al., 2001), phytoplankton class-specific productivity (Uitz et al., 2010), chlorophyll and carbon-based ocean productivity modeling (Behrenfeld et al., 2005; Platt et al., 2008), climate-driven trends in productivity (Behrenfeld et al., 2006; Kahru et al., 2009; Henson et al., 2010), and inter-comparison of productivity algorithms (Carr et al., 2006; Lee et al., 2015). They have also been used to check the stability of CERES measurements (Loeb et al., 2006).

Users Needs

User needs vary widely in terms of products, spectral, spatial, and temporal resolution, and acceptable uncertainties, depending on the scientific or societal subject of interest. Applications requiring knowledge of radiomeric quantities and apparent properties (radiance, irradiance, average cosine, attenuation coefficients) are multiple and diverse, including phytoplankton phenology, carbon inventory, heat budget and ocean dynamics, fisheries and ecosystem management, toxic algal blooms, and eutrophication (see National Research Council, 2011 for a comprehensive list). Observational and uncertainty requirements (satellite products) generally range from 1 h to 1 day, 0.1 to 50 km, and ±5 to ±20% for PAR and 0.1 to 10 km, 1 to 7 days, and ±10 to ±25% for spectral K_d (Malenovsky and Schaepman, 2011), but higher resolution may be needed in some cases, for example rapidly changing phenomena occurring in small water bodies. For applications that need analyzing long-term records (e.g., associated with climate), the products need to be sensor independent, consistent, and continuous across satellite missions.

The satellite products should be defined unambiguously and completely, they should be easily accessible, and they should have associated ATBDs with detailed protocols including description of all ancillary data used and their sources. For example, defining a PAR product merely as downward quantum flux at the surface in the 400–700 nm spectral range is insufficient. One needs to precise whether the flux is just above or just below the surface, whether it is instantaneous or time-averaged (e.g., over 24 h for "daily PAR"), and to indicate spatial resolution. Advantages and limitations should be specified (e.g., product not valid at Sun zenith angles above 75°, or over sea ice, or in the presence of Sun glint), and uncertainties assessed, preferentially provided on a pixel-by-pixel basis, to make sure that observed changes are interpreted correctly. Computer codes used to derive the products should be available to users with proper documentation, as well as standardized processing tools (e.g., open-source toolboxes).

A survey about satellite PAR observations was conducted in 2015 by Plymouth Marine Laboratory (PML), requesting feedback from the user community on adequacy of available products, including importance, usage, and accuracy, and additional features one would like to see. **Figure 8** summarizes the results about desired PAR attributes and new products and acceptable PAR uncertainties. Most respondents answered it was very or extremely important to have uncertainties associated with PAR products. About 50% of the respondents indicated that ±10–25% uncertainty was acceptable, and about 40% wanted uncertainty better than ±10%. A substantial majority of

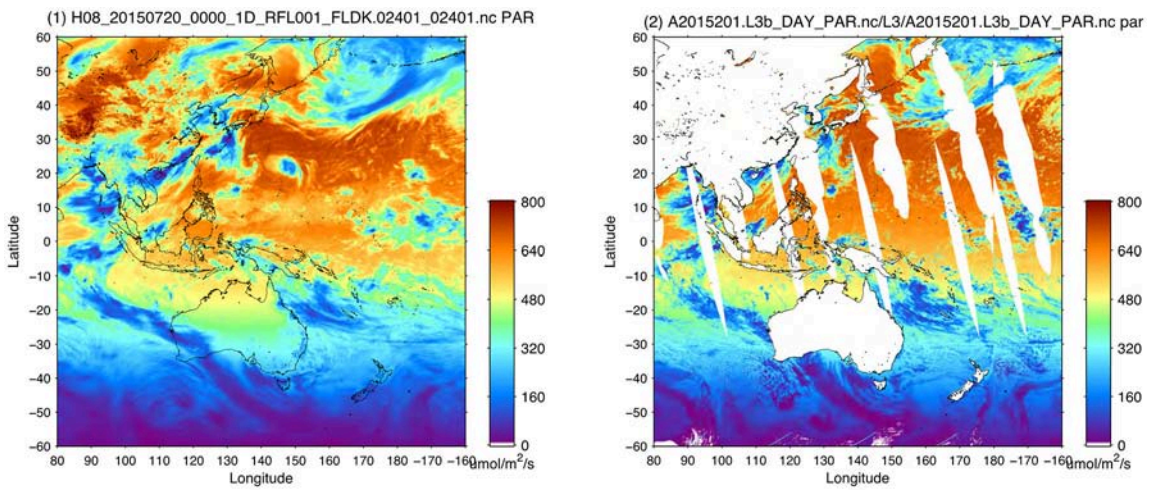


FIGURE 7 | Daily above-surface PAR for July 20, 2011 obtained from AHI (observations every 10 min) and MODIS-A data at 9 km resolution (**left** and **right**, respectively). Patterns of spatial variability are similar for the two sensors, but the AHI PAR field is smoother, contains less extreme values, and does not exhibit spatial gaps (Courtesy of H. Murakami, Japanese Aerospace Exploration Agency).

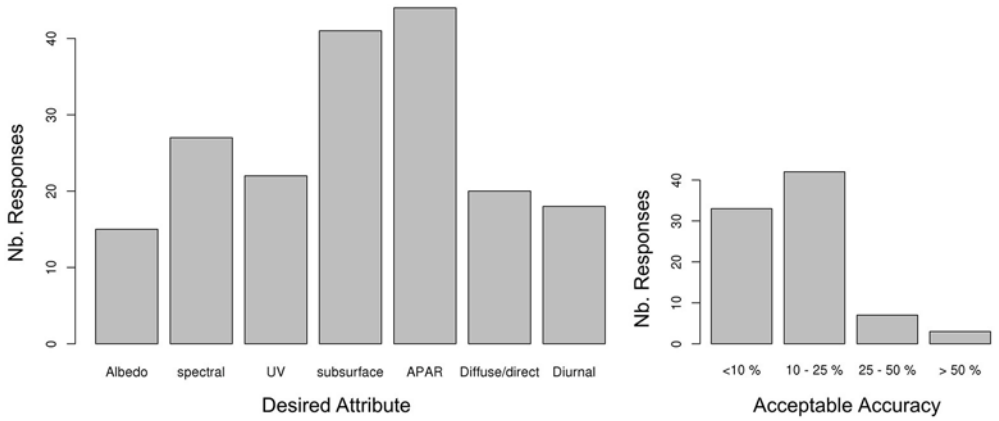


FIGURE 8 | Results of PML user requirement questionnaire for satellite-derived PAR **(Left)**. Histogram of desired products and attributes **(Right)**. Histogram of acceptable uncertainty for PAR. (Based on data from 63 respondents).

respondents ranked PAR below the surface and the fraction of PAR absorbed by phytoplankton about equally as the top two new products to generate. Based on this survey and in view of the extensive list of research and societal applications, current and potential, a list of required (new) products was compiled, including products that may be challenging to generate from space:

- *Sub-surface planar and scalar irradiance (as opposed to above the surface).*
- *Fraction of PAR absorbed by phytoplankton (APAR).*
- *Diffuse fraction of total irradiance (average cosine of light field just below the surface).*
- *Spectral planar and scalar irradiance.*
- *Spectral diffuse attenuation coefficient for downward irradiance.*
- *Surface albedo (ratio of planar upward irradiance to downward planar irradiance just above the ocean surface).*
- *UV-A, UV-B scalar irradiance (with photon and energy units).*
- *Products without gaps (in space/time) to provide boundary conditions to models.*
- *Upper-ocean heating profile.*
- *Diurnal distribution of PAR and its attenuation.*
- *Averaged mixed-layer PAR.*
- *Under-ice light fields.*

Gap Analysis

Some new products (such as below surface planar and scalar PAR) can be easily implemented while others require development (e.g., APAR, vertical attenuation of PAR). For some products (e.g., under-ice light fields), the readiness level, in terms of methodology, is low. The state-of-the-art, however, is such that the strategy to obtain the new products described above is known and summarized below.

Sub-surface planar PAR and scalar PAR depend essentially on the sunlight transmission across the air-water interface, therefore the angular distribution of radiance incident at the surface and surface roughness (Mobley and Boss, 2012). The 24 h-averaged quantities, as well as the average cosine for total light, can be parameterized as a function of latitude and daily cloud factor (i.e., the ratio of actual PAR and clear-sky PAR) and wind speed. This may require look-up tables for clear sky and overcast quantities. Details about procedures are provided in section Examples of New Products, where examples of sub-surface products are presented. Note that in the OBPG approach to estimating above-surface daily planar PAR, it is actually easier (more direct) to compute the downward flux below the surface. This "penetrative" flux is obtained by subtracting from the incident extraterrestrial solar irradiance the reflected flux and the flux absorbed by the surface/atmosphere system. The "penetrative" flux is then corrected by $1/(1-A_s)$, where A_s is the surface albedo, to yield the incident flux onto the surface (e.g., Frouin et al., 2012). This second step introduces uncertainty, but allows for extensive evaluation at PAR measuring sites, an activity that cannot be performed easily for sub-surface fluxes (lack of data and difficulty to measure sub-surface fluxes accurately).

The above variables, integrated over the PAR spectral range, can be calculated without difficulty for the spectral bands of the ocean color sensors (i.e., visible to near infrared), in which cloud absorption is negligible. Providing the information at regular spectral intervals (e.g., every 5 or 10 nm), a requirement of some primary production models (e.g., Sathyendranath et al., 1989; Antoine et al., 1996), is straightforward since cloud optical properties (extinction coefficient, asymmetry factor) are similar, i.e., cloud albedo can be assumed constant in the entire spectral range, and the coupling between molecules and cloud droplets/crystals is relatively small (i.e., fairly unique relation between cloud transmittance at different wavelengths). In spectral regions of strong gaseous absorption, however, uncertainties may be introduced due to the coupling between absorption and scattering processes (depends on the unknown vertical distribution of the absorbers and scatterers). Extending the calculations to the ultraviolet (e.g., UV-A, UV-B) from measurements in the visible is more complicated, but definitely feasible. The complication is not due to cloud optical properties

(they remain similar to those in the visible), but to the coupling between molecules and cloud droplets/crystals, which is effective. In other words, the relation between cloud transmittance (in the presence of molecules and aerosols) in the ultraviolet and visible depends on the type of clouds and their location in the vertical. This indicates that de-coupling the clear atmosphere from clouds, as it is done in the OBPG PAR algorithm, may introduce significant errors in the estimation of ultraviolet irradiance. Suitable modeling of the relation between atmospheric transmittance at visible and ultraviolet wavelengths is therefore required, which can be accomplished via radiation transfer calculations to various levels of accuracy depending on user needs (may require additional information on clouds).

From the spectral downward irradiance at the surface, minimally affected by photons reflected by the surface and backscattered by the water body (spherical albedo of the atmosphere is small, i.e., about 0.15 in the visible) and the spectral upward irradiance at the surface, one may compute the spectral surface albedo. The computation, however, can only be done in clear-sky conditions since water optical properties are necessary but generally not retrieved in cloudy conditions from ocean-color sensors. Estimating the spectral upward irradiance at the surface requires retrieving spectral water reflectance (i.e., the signal backscattered by the water body), which is routinely achieved by satellite project offices, and modeling the bidirectional reflectance function of the surface (e.g., using Cox and Munk, 1954) and the water body (e.g., using Morel and Gentili, 1996 or Park and Ruddick, 2005).

Although daily averaged quantities are often used (therefore required) in applications, instantaneous quantities (i.e., determined at time of satellite overpass) can be easily provided. In fact instantaneous products, unlike 24 h-averaged products, do not require assumptions about diurnal changes in atmospheric and oceanic properties. Diurnal variability can be described well using sensors onboard geostationary satellites, such as GOCI and AHI, see section Current Products, all the more as these sensors have ocean-color capabilities. One limitation, however, is the reduced spatial resolution at high latitudes, and managing data from different instruments, which may be operated by different space agencies, to achieve global coverage. Observations from the Earth-Sun the Lagrangian-1 (L1) point, 1.5 million kilometers from Earth, such as those made by the EPIC camera onboard DISCOVR (1–2 h temporal resolution, 21 km spatial resolution), provide a suitable alternative to several sensors on geostationary orbit. Because of the high orbit, spatial resolution at high latitudes is much less an issue. Diurnal information may also be obtained from a constellation of Sun-synchronous instruments with local overpass times spread during the day, for example MODIS-Terra at 10:30 am, SeaWiFS-SeaStar observing at noon, and MODIS-Aqua observing at 1:30 pm. From a unique instrument in Sun-synchronous orbit, ancillary data about variability of clouds and aerosols is necessary, for example Modern-Era Retrospective analysis for Research and Applications version 2 (MERRA-2) products, available at a 1/2 × 2/3 degree every hour for the day of the satellite observation.

Propagating fluxes vertically below the surface requires knowledge of the vertical profile of diffuse attenuation coefficient, K_d. From space, this can only be achieved in clear sky conditions with passive optical sensors, since K_d is deduced from water reflectance (empirical algorithms) or inherent optical properties (e.g., IOCCG, 2006). The K_d estimates are actually weighted averages over one optical depth (from which most of the photons from the water body reaching the satellite sensor originate); no vertical information is obtained. In first approximation, one may propagate light below the first optical depth (shallower than the depth of the euphotic zone) by assuming no vertical variation in spectral diffuse attenuation. However, this is generally inaccurate, as many oceanic regions (e.g., oligotrophic provinces) exhibit maximum chlorophyll concentration well below the first optical depth. One has therefore to rely on statistical relations between concentrations of oceanic constituents at the surface and below (e.g., Morel and Berthon, 1989) to estimate the K_d depth profile, or to use outputs of predictive coupled physical-biogeochemical numerical models. In the future, with the advent of space-borne polarization lidars such as CALIOP onboard CALIPSO or the Aerosol/Cloud/Ecosystems (ACE) lidar (being designed), one will be able to profile K_d in both clear and cloudy conditions, day and night, up to 3 optical depths in the green (532 nm) at a vertical resolution of 3 to 30 m (e.g., Lu et al., 2014; Behrenfeld et al., 2016). The sub-surface fluxes and the vertical profile of diffuse attenuation coefficient give access to fluxes at the bottom (important for studies of shallow coastal ecosystems), average fluxes in the mixed layer (requires knowledge of mixed-layer depth, e.g., from ocean circulation models), and the upper-ocean heating rate profile. Knowing mixed-layer fluxes and vertical heat distribution is especially useful to characterize the role of solar penetration and biological-physical interactions on ocean circulation and climate (Olhmann et al., 1996; Shell et al., 2003; see section Introduction).

In a homogeneous ocean, APAR depends on the ratio of the spectral absorption coefficient by live phytoplankton, a_{ph}, and total absorption (water, yellow substances, non-algal particles), a_{tot}, in the PAR spectral range and the spectral planar irradiance just below the surface. In a vertically heterogeneous ocean, the vertical distribution of those quantities plays a role, as well as the vertical distribution of the diffuse attenuation coefficient for downward irradiance, K_d. Computing APAR from space, therefore, requires estimates of spectral planar irradiance below the surface, and vertical profiles of a_{ph}, a_{tot}, and K_d. Sub-surface spectral irradiance and its vertical attenuation can be obtained as discussed above, and absorption coefficients using various techniques (e.g., IOCCG, 2006). Some of these variables are difficult to retrieve with good accuracy, in particular a_{ph} (requires partitioning a_{tot} into its components), and vertical information in the euphotic zone is generally not directly available (except from future space-borne lidars, see above). Consequently, uncertainties on APAR computations based on satellite estimates of individual variables may be large. Since APAR strongly depends on the spectral ratio of sub-surface reflectance, $R(0^-)$ and pure seawater reflectance, $R_w(0^-)$, one may envision approximating APAR by a linear combination of $R(0^-)/R_w(0^-)$ in the PAR spectral range (Frouin et al., 2014). Since APAR is expressed linearly in water reflectance, the method is applicable to average values of water reflectance (e.g., spatially

averaged), which may reduce the impact of water reflectance noise on the APAR estimate.

Primary production under sea ice is considerable, as evidenced from in situ measurements of phytoplankton concentrations and suggested by numerical model simulations. In the Arctic Ocean, it may constitute more than 30% of the total production (Popova et al., 2010). As a consequence, the seasonal cycle is shifted, with maximum primary production occurring in July, not in August-September (case of open waters). There is a need, for primary production modeling and studies of under-ice phytoplankton blooms, to determine light fields under the ice (Laliberté et al., 2016). This is quite difficult to realize from space, since knowledge of sunlight transmission though sea ice is required, and this parameter is quite variable depending on ice type and thickness, and the presence of snow and melt ponds. Transmission models are now becoming available (Arndt and Nicolaus, 2014), and they can make use of sea-ice thickness and age, snow depth, and melt pond fraction observations from microwave and optical sensors (SIRAL on Cryosat-2, AMSR-2 on CGOM-W1, ATLAS on IceSta-2, MODIS on Terra and Aqua). One issue is separating the contribution of clouds to the planetary albedo, in order to access the surface albedo. Algorithms are not mature for generating operationally shortwave fluxes under sea ice.

The possible methodologies, difficulties, opportunities, and readiness level for developing and creating the new products were discussed above. The following recommendations regarding these products are made:

1) Spectral fields should be provided at the sensor resolution with protocols (and codes) describing how to interpolate and extrapolate for obtaining other spectral distributions (e.g., 1 nm irradiance field from a multi-spectral sensor).
2) Vertical propagation of products requires an appropriate attenuation coefficient from which other products can be derived (e.g., euphotic depth, mixed-layer depth, isolume depth). Clear guidelines on how to produce the derived products using the attenuation should be provided.
3) Horizontal/temporal gap filling is necessary for certain applications (e.g., ecological forecasting). This can be done using merged products across sensors and/or interpolation schemes (using known de-correlation scales or models). Many techniques are available (e.g., Pottier et al., 2006; Alvera-Azcarate et al., 2007; Krasnopolsk et al., 2016).
4) Products should have associated uncertainties that are consistent with those obtained when validating estimates against in situ measurements, taking into account uncertainties in the in situ data. This requires a calibration/validation program. The product protocol should provide a description of how the uncertainty was derived. It is desirable to provide a per-pixel uncertainty. It is recognized that the level of effort to obtain a very accurate uncertainty estimate can be very large and therefore some trade-offs may need to be done, in consultation with user requirements.
5) Data access should be tailored to users need. For example, modelers will use multithreading with simultaneous access to associated error fields. In contrast, the Earth observation community will, typically, want to access data using FTP. Most users do not care about the satellite mission from which a product was derived, but rather care about the products being continuous in time and consistent across missions.
6) Cross-agency efforts should be made to homogenize their respective products so it is easy for users to use these products (e.g., the definition of a PAR product should be the same). For climate relevant products, it is critical to merge them (and de-bias) across missions so that models to not experience secular jumps as they assimilate such data.

EXAMPLES OF NEW PRODUCTS

Par Simulator

Radiative Transfer Code

To develop new shortwave radiation products from satellite data (such as those discussed above) and assess accuracies, one needs to simulate the TOA radiance measured by a given sensor and the variable to retrieve (e.g., planar or scalar spectral irradiance below the surface, average cosine). For this, we use the Speed up Monte-Carlo Advanced Radiative Transfer using GPU (SMART-G) radiative transfer code (Ramon et al., 2017). This code, based on the Monte-Carlo method, is fast and massively parallel. It computes the complete light field (i.e., radiance, including polarization) in the ocean and atmosphere.

The code simulates the transfer/propagation of solar radiation in a 1-dimensional coupled ocean-atmosphere system with a wavy interface. It accounts for absorption and scattering by molecules, aerosols, and hydrosols, and Fresnel reflection/refraction at the interface. Polarization properties of the various atmospheric and oceanic constituents and the surface are explicitly considered. Inelastic processes (i.e., Raman scattering, fluorescence) are omitted in the current version. The ocean can be infinitely deep or bounded by a reflective bottom at finite depth. The computations are made in either plane-parallel or spherical geometry. The four components of the Stokes vector can be obtained at any wavelength of the solar spectrum and any level of the coupled ocean-atmosphere system.

Gaseous absorption is treated either by correlated k-distribution (Kato et al., 1999) or an equivalent like REPTRAN (Gasteiger et al., 2014; Emde et al., 2016). In the multispectral mode, each photon is assigned a wavelength. All optical properties of the medium are pre-calculated for these wavelengths. The spectrum is computed in one pass but is under the influence of Monte-Carlo noise. For higher spectral resolution computations without spectral noise, for example in order to handle line-by-line (LBL) gas absorption, the ALIS method (Emde et al., 2011) is also implemented in SMART-G. This allows for the calculation of spectra by tracing the photon paths once for all wavelengths (by absorption/scattering decoupling).

Simulations

In the Monte-Carlo code, photon packets are carrying planar irradiance (Wm^{-2}) perpendicular to their direction of propagation. Therefore downward or upward spectral planar irradiances are obtained by performing the weighted sum of

irradiances crossing a unit area of horizontal surface located above or below the air-sea interface. For spherical irradiances, each photon is assigned an additional weight of $1/\cos(\theta_i)$ where θ_i is the photon's zenith angle arriving on the detector. The runs were optimized for the calculation of spectral daily fluxes above and below the ocean surface. For 1 day, one run was executed by injecting a large number of photons whose wavelength was chosen randomly as well as the injection angles at TOA and in the atmosphere depending on the hour of the day. Typical runtime is 1 s to reach ±1% uncertainty for clear sky conditions and for 1 day. It becomes 30 s for a totally overcast situation with a cloud optical thickness of 50.

The aerosol and cloud optical properties are taken from the OPAC database (Hess et al., 1998) and distributed within the libradtran software package (www.libradtran.org). Aerosols are supposed to be spherical. The clouds are supposed to be composed of liquid water droplets with a varying effective radius. Rayleigh optical depth is computed according to Bodhaine et al. (1999). The gaseous absorption is parameterized as a correlated k-distribution with spectral intervals of 10 nm from line-by-line calculations using the Py4Cats code (Schreier and Gimeno Garcia, 2013) and HITRAN 2012 (Rothman et al., 2013) absorption parameters for H_2O and O_2. Ozone and NO_2 smooth absorption coefficient are taken from Bogumil et al. (2003). The wind-roughened sea surface is modeled as an ensemble of uncorrelated facets with a slope distribution (Cox and Munk, 1954) and a uniform azimuth distribution. The reflection and transmission coefficients are then computed using Fresnel's laws. Ocean bulk optical properties correspond to Case-I waters with a chlorophyll-a concentration of 0.5 g.m^{-3}. The ocean's bottom is black and is located at a depth of 50 m. The ocean phase function is represented by a Fournier-Forand function with a varying truncation angle.

An example of outputs of the SMART-G code in the context of PAR simulations (both spectral irradiance and reflectance) is given in **Figure 9** for various levels in the atmosphere-ocean system (top, just above the surface, just below the surface, and at the black bottom of the water column) and several atmospheric conditions (clear and cloudy situations with different Sun zenith angles). The TOA reflectance in the MERIS bands is also indicated. The various graphs show the importance of Sun zenith angle and cloud optical thickness in controlling the downward PAR above and below the surface. They also illustrate the usefulness of the SMART-G tool for algorithm development.

For the year 2011, and for 14 latitudes ranging from −64.5 to 64.5, the various spectral irradiances listed in the Annex were computed at 11 times regularly distributed throughout the day. The typical spectral TOA reflectance for a MERIS-like instrument measuring at 10:30 local time was also computed. The atmospheric content was changing according to the MERRA-2, hourly, gridded datasets of water vapor and ozone contents, aerosol optical depths of black carbon, dust, organic carbon, sea salt, and sulfates aerosols at 550 nm, and cloud optical thicknesses. Individual clear sky and overcast plane-parallel radiative transfer calculations were then mixed using the Independent Pixel Approximation (IPA) using MERRA-2 cloud cover variable as the mixing value.

The following quantities were then computed: $PAR_o^{0-}(t)$ and $PAR_d^{0+}(t)$ (to describe diurnal variation of sub-surface fluxes), $\langle PAR_d^{0+}\rangle_{24h}$ (the main product), $\langle PAR_o^{0-}\rangle_{24h}$ (the key product for primary production and photo-chemical processes), $\langle PAR_o^{0-}\rangle_{24h}(\lambda_i)$ (spectral scalar flux, as requested by modelers, at a resolution of 10 nm), and $\langle \mu_0^{0-}\rangle_{24h}$ (to characterize the angular structure of the light field). Definitions are provided in the Annex.

Algorithm for $\langle PAR_o^{0-}\rangle_{24h}$ and $\langle \mu_0^{0-}\rangle_{24h}$

Rationale

We define a Cloud Factor (CF) as the deviation from a pure clear sky daily averaged PAR above the surface. It is a "measure" of the influence of clouds on the daily PAR:

$$\langle CF\rangle_{24h} = \frac{\langle PAR_d^{0+}\rangle_{24h}}{\langle PAR_d^{0+}\rangle_{24h}^{clear}}$$

One typical day of simulations is displayed in **Figure 10** for a latitude of 55.5°N. The Sun zenith angle and the length of the day mainly drive the diurnal cycle of both clear sky PAR and partly cloudy PAR. For that particular day, the influence of cloudiness is important (and somehow stable) reducing the daily PAR by a factor $\langle CF\rangle_{24h} = 0.34$. The average cosine μ_0^{0-} is very stable for cloudy conditions, with a value around 0.82 throughout the day. A clear sky has a variable μ_0^{0-} culminating at noon. In both cases the spectral variation of μ_0^{0-}, not shown here, is weak. The 24 h-averaged value of μ_0^{0-} should be representative of the direction of propagation of the largest fraction of the daily PAR. That is why we define the 24 h-averaged $\langle \mu_0^{0-}\rangle_{24h}$ as the ratio of the 24 h-averaged net and scalar PAR (and not the 24 h average of their instantaneous ratios):

$$\langle \mu_0^{0-}\rangle_{24h} = \frac{\langle E_{net}^{0-}\rangle_{24h}}{\langle E_o^{0-}\rangle_{24h}}$$

The normalized spectral PAR, obtained by dividing the spectral PAR by its value at 675 nm, is defined in the same manner as:

$$\left\langle \widetilde{PAR_o^{0-}} \right\rangle_{24h}(\lambda_i) = \frac{\langle PAR_o^{0-}\rangle_{24h}(\lambda_i)}{\langle PAR_o^{0-}\rangle_{24h}(675\text{nm})}$$

It gives the spectral shape of the PAR and it is very stable and close to the TOA solar irradiance spectrum in most cases. The spectral shape of the clear sky PAR is slightly influenced by the mean Sun zenith angle because the absorption of ozone in Chappuis bands becomes more and more effective. The spectral dependence of the cloudy PAR is smaller. It increases in the blue part of the spectrum as cloud influence increases.

Following the approach of Mobley and Boss (2012), 24 h-averaged secondary radiative quantities may be obtained from a reduced set of parameters, the most important ones being the location and date which control the day length and mean Sun zenith angle, then the influence of the clouds which is between null (clear sky) and maximum (100% cloud cover), and finally the wind speed. The chlorophyll content

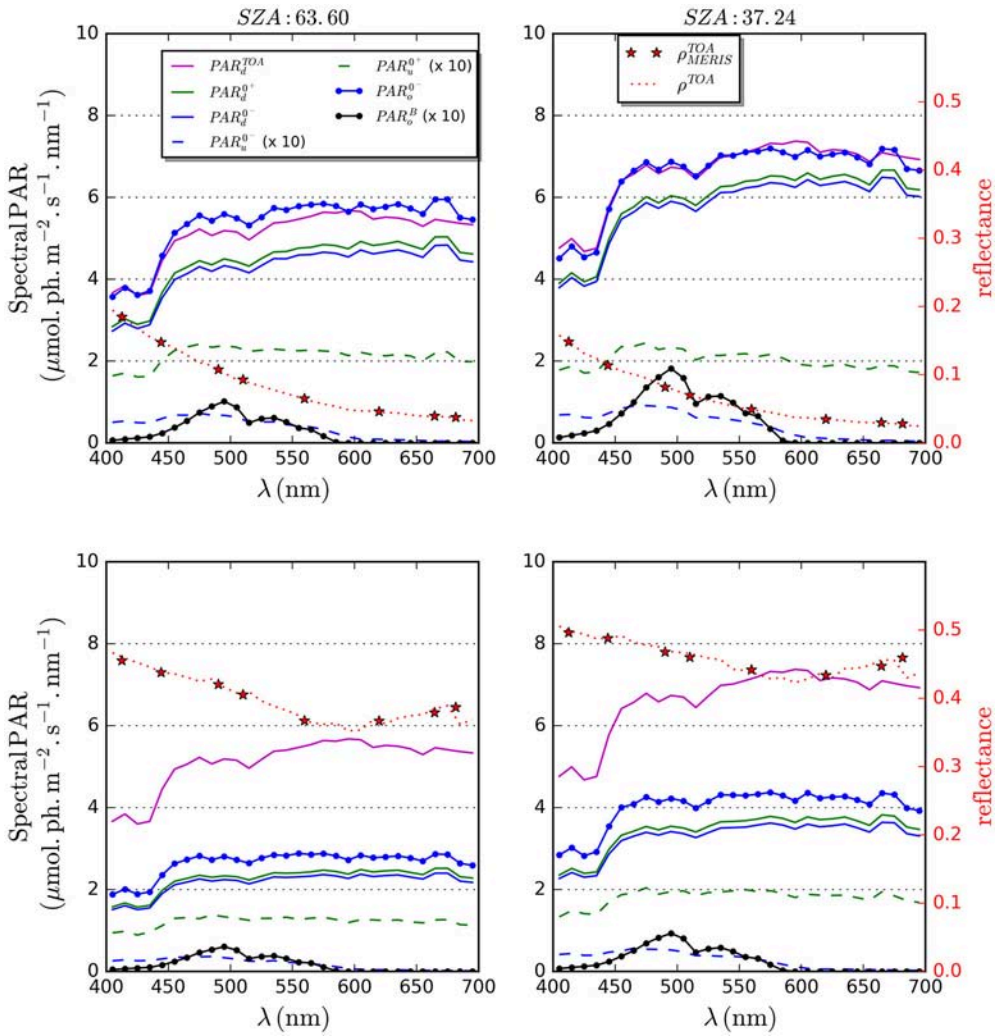


FIGURE 9 | Examples of outputs of the SMART-G RTC code for 4 situations. On the left axis is reported the "spectral PAR" in units of light quanta per unit time, unit area and unit wavelength for planar and spherical upwelling ("u") or downwelling irradiances ("d") at several levels in the ocean-atmosphere coupled system: at TOA, just above (0^+) or below (0^-) the surface, or at the black bottom (superscript B) of the ocean located here at a depth of 50 m. On the right axis is plotted the TOA spectral reflectance at the same resolution, and at the center of MERIS wavebands. The top two graphs are for a clear-sky situation for two Sun zenith angles (SZA) and the bottom two graphs a liquid water cloud, located between 2 and 4 km with droplets of effective radius = 11 mm whose of optical thickness 10 is added. The clear atmosphere model is the US62 standard atmosphere, with maritime polluted aerosols as described in the OPAC database with an AOT at 550 nm equal to 0.1. The ozone column is 300 DU and the precipitable water quantity is 2 g/cm^2.

of the water is of minor importance for calculating the scalar and net PAR just below the surface. **Figure 11** displays the coefficients to be applied to $\langle PAR_d^{0+} \rangle_{24\,h}$ in order to obtain $\langle PAR_o^{0-} \rangle_{24h}$, or $\langle PAR_d^{0-} \rangle_{24h}$, and also shows how $\langle \mu_0^{0-} \rangle_{24h}$ and $\langle \widetilde{PAR_o^{0-}} \rangle_{24h}(\lambda_i)$ vary for various latitudes and wind speeds and for the two extreme cases, i.e., clear and totally overcast. **Figure 12** displays the result of the application of the coefficients described above for estimating the scalar PAR below the surface. We processed one full year of global simulated data (2011, see section PAR Simulator) for 3 wind speeds (0, 7, and 15 m.s^{-1}) and checked the quality of the regression vs. the "actual" (or prescribed) values, i.e., the values obtained by running the Monte Carlo code with the various input variables (MERRA-2 hourly data, see above). These values provide the reference in calculating algorithm performance statistics. The regression is excellent with a residual bias and a R.M.S. difference of about 0.5 mol.ph m^{-2}.day^{-1}. Wind speed or cloudiness do not apparently impact the quality of the regression.

Figures 13, **14** depict the normalized spectral PAR and average cosine for the year 2011, for a latitude of 55.5°N and a wind speed of 7 m.s^{-1}. For both parameters, the clear sky and totally overcast situations are also plotted. They constitute an envelope wihin which the actual values are included, and the deviation from the clear sky value seems proportional to the actual cloud factor $\langle CF \rangle_{24h}$. This suggests a method to derive $\langle \widetilde{PAR_o^{0-}} \rangle_{24h}(\lambda_i)$

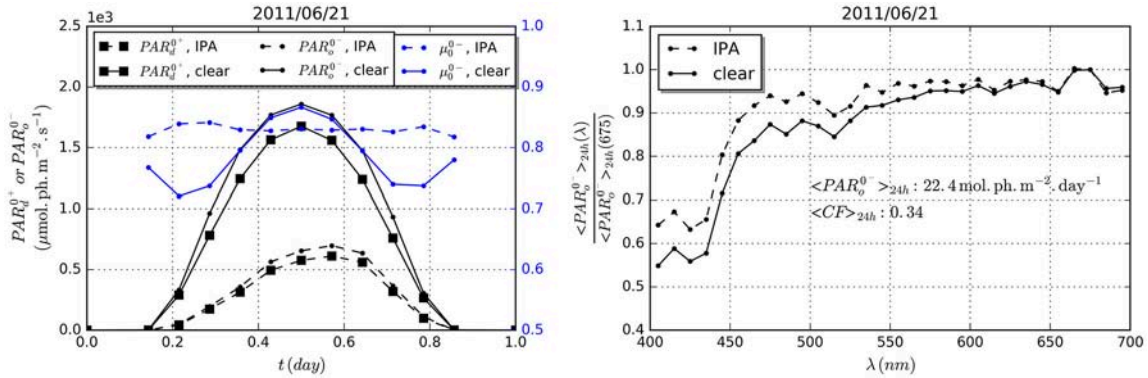


FIGURE 10 | (Left plot) Diurnal variation of the downward above-surface PAR (squares) and scalar below-surface PAR (dots) for the 21st of June and for a latitude of 55.5° N and a wind speed of 0 m/s. On the right vertical axis is reported the mean cosine below the surface. Two simulations are shown: (i) one for clear sky (solid lines) and (ii) one for "real" sky with cloud cover and cloud optical thickness as given by MERRA-2 and mixed with the clear-sky simulations using the Independent Pixel Approximation (IPA). **(Right plot)** Normalized spectral PAR at 675 nm for both clear and IPA situations.

and $\langle \mu_0^{0-} \rangle_{24h}$ from look-up tables of the clear sky and overcast situations and from an estimation of the actual cloud factor.

Equations

We propose to use the observed cloud factor $\langle \mathrm{CF} \rangle^{\mathrm{obs}}$ as a proxy of $\langle CF \rangle_{24h}$ and then linearly interpolate between clear sky and overcast look-up tables as a function of $\langle \mathrm{CF} \rangle^{\mathrm{obs}}$. We have:

$$\langle \mathrm{CF} \rangle^{\mathrm{obs}} = \frac{\langle \mathrm{PAR}_{\mathrm{d}}^{0+} \rangle_{24\,h}^{\mathrm{obs}}}{\langle \mathrm{PAR}_{\mathrm{d}}^{0+} \rangle_{24\,h}^{\mathrm{clear}}}$$

$$S1 = \frac{\left\langle \widetilde{PAR_o^{0-}} \right\rangle_{24h}^{\mathrm{clear}} (\lambda_i) - \left\langle \widetilde{PAR_o^{0-}} \right\rangle_{24h}^{\mathrm{overcast}} (\lambda_i)}{1 - \langle CF \rangle_{24h}^{\mathrm{overcast}}}$$

$$\langle \widetilde{PAR_o^{0-}} \rangle_{24h} (\lambda_i) = S1 \,.\, \left(\langle CF \rangle^{obs} - \langle CF \rangle_{24h}^{overcast} \right) + \langle \widetilde{PAR_o^{0-}} \rangle_{24h}^{\mathrm{overcast}} (\lambda_i)$$

$$S2 = \frac{\langle \mu_0^{0-} \rangle_{24h}^{\mathrm{clear}} - \langle \mu_0^{0-} \rangle_{24h}^{\mathrm{overcast}}}{1 - \langle CF \rangle_{24h}^{overcast}}$$

$$\langle \mu_0^{0-} \rangle_{24h} = S2 \,.\, \left(\langle CF \rangle^{obs} - \langle CF \rangle_{24h}^{\mathrm{overcast}} \right) + \langle \mu_0^{0-} \rangle_{24h}^{\mathrm{overcast}}$$

The error will be large when cloudiness changes a lot during the day and thus we may suspect that $\langle \mathrm{CF} \rangle^{\mathrm{obs}}$ deviates substantially from $\langle CF \rangle_{24h}$ and when the values of the parameters in clear or totally cloudy conditions are substantially different.

Look- Up Tables for Clear Sky and Overcast Situations

Models

The clear sky model is based on the AFGL US 62 standard atmosphere with a surface pressure of 1012.15 hPa, an O_3 vertically integrated content of 300 DU, and a H_2O vertically integrated content of 2 g.cm^{-2}. The aerosol model is the maritime clean model from the OPAC database with an AOT of 0.1 at 550 nm. The air-sea interface is a wind-roughened surface. The ocean bulk optical properties correspond to Case-I waters with a chlorophyll a concentration of 0.5 g.m^{-3}. The ocean bottom is black and is located at a depth of 50 m. For the totally overcast model we added a permanent cloud layer between 5 and 10 km consisting of water droplets with $r_{\mathrm{eff}} = 11\,\mu$m and a cloud optical thickness of 50. The aerosols, hydrosols and cloud phase matrices are computed at 550 nm and are assumed spectrally invariant between 400 and 700 nm.

Computations

For latitudes between −90° and 90° by step of 10°, for every 30 days along the year, and for 3 wind speeds: 0, 7 and 15 m.s^{-1}, we computed the following quantities:

$$\langle CF \rangle_{24h}^{overcast}, \langle \widetilde{PAR_o^{0-}} \rangle_{24h}^{\mathrm{overcast}} (\lambda_i), \langle \widetilde{PAR_o^{0-}} \rangle_{24h}^{\mathrm{clear}} (\lambda_i), \langle \mu_0^{0-} \rangle_{24h}^{\mathrm{overcast}}, \text{ and } \langle \mu_0^{0-} \rangle_{24h}^{\mathrm{clear}}.$$

First Results

For the MERIS sensor, first examples of the new parameters (daily and monthly global products for May 15 and May 1–31, 2011) are displayed in **Figures 15–17**. Data at high latitudes were masked using ESA CCI Sea Ice Concentration products v2.0 (cci.esa.int). The scalar PAR below the surface (**Figure 15**) follows the planar PAR above the surface (not shown here), but the values are somewhat higher, as expected. The spatial coverage of the daily MERIS product is less than for the MODIS products (**Figure 1**), because of the narrower swath of MERIS and the glitter mask (more glint in the MERIS imagery). The average cosine product (**Figure 16**) is smoother because it is influenced mainly by the average solar elevation. The latitudinal variation is important with the highest values (0.85) in the tropics whatever the cloudiness. At high latitudes the contrast between clear and cloudy sky conditions becomes more marked with an average cosine closer to the tropics values in cloudy sky conditions and the lowest values (0.65) under clear skies. For monthly averages, the mixture between cloudy and clear skies tends to further

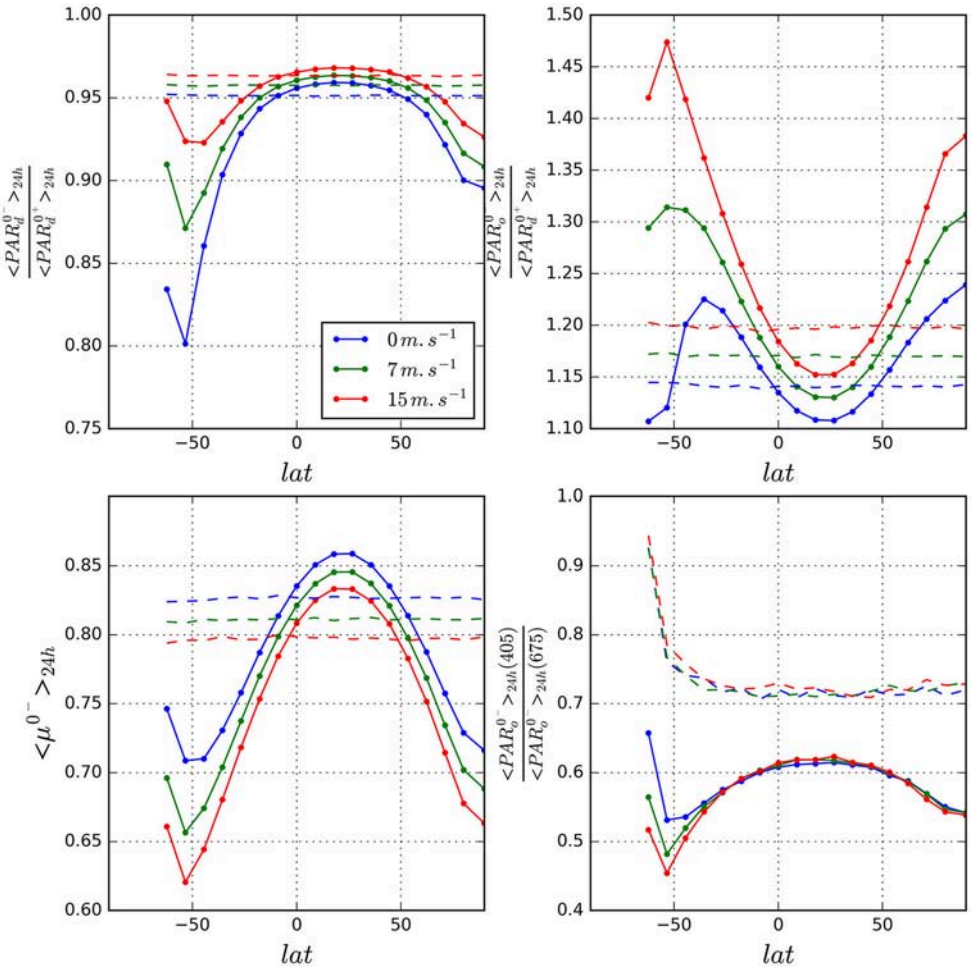


FIGURE 11 | Simulations of 24 h averaged secondary radiative products for clear-sky (solid lines) and totally overcast situations (dashed lines, cloud optical thickness = 50, constant along the day) and for several wind speed as a function of latitude. **(Top left)** Ratio of planar PAR below and above surface. **(Top right)** Ratio of scalar PAR below the surface to planar PAR above surface. **(Bottom left)** Average cosine below the surface. **(Bottom right)** Ratio of spectral scalar PAR below the surface at 405 and 765 nm. The date is 21st of June.

smooth the product, which becomes a simple function of latitude. The normalized spectral PAR at 405 nm (**Figure 17**), like the mean cosine, exhibits a contrast between clear and cloudy sky situations that is increasing toward the high latitudes. However on a monthly time scale the latitudinal gradient is very weak and the dispersion of the product is also very weak with a mean of 0.63 and a standard deviation of 0.024 over the global ocean.

Algorithm Uncertainties for $\langle PAR_d^{0+} \rangle_{24h}$

Associating uncertainties to the satellite radiation products, preferentially on a pixel-by-pixel basis, is obligatory to quantify their quality. This is important to ensure that variability or trends detected in scientific analyses are of geophysical nature, i.e., that the data are interpreted properly in view of their strengths and limitations, and to merge different data sets. This is also essential for data assimilation, a primary application of the products (large uncertainties will have little impact on model runs, small uncertainties will constrain the model to behave like the data). Expressing uncertainties requires modeling the measurement, identifying all possible error sources (e.g., noise in the input variables, imperfect/incomplete mathematical model), and determining the combined uncertainty, as described in JGCM-100 (2008) and subsequent publications.

In the following, algorithm uncertainties associated with $\langle PAR_d^{0+} \rangle_{24h}$ are considered, i.e., those due to model approximations and parameter errors (e.g., decoupling effects of clouds and clear atmosphere, neglecting diurnal variability of clouds, using aerosol climatology) assuming that the input variables (TOA reflectance at wavelengths in the PAR spectral range) are known perfectly. A procedure is provided to estimate and provide, for each pixel of a product, this uncertainty component of the total uncertainty budget, which is expected to dominate. The uncertainty characterization has been done using an extended simulation dataset covering the 2003–2012 time period still using 1 hourly MERRA-2 input data. The large number of data points allows one to sample well the atmospheric variability and in particular many variations of daytime nebulosity, for all latitudes.

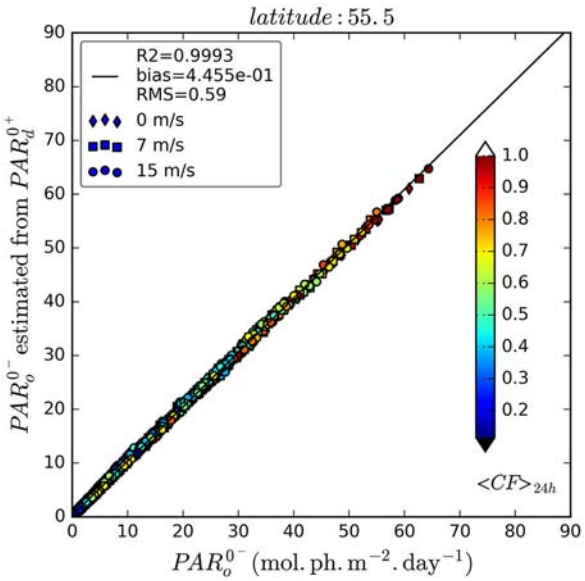


FIGURE 12 | Estimated scalar PAR below the surface from the planar PAR above surface and the clear sky "transmission factors" shown in **Figure 11** for all the days of 2011 and 3 wind speeds. The color of the symbols is a function of the cloud factor.

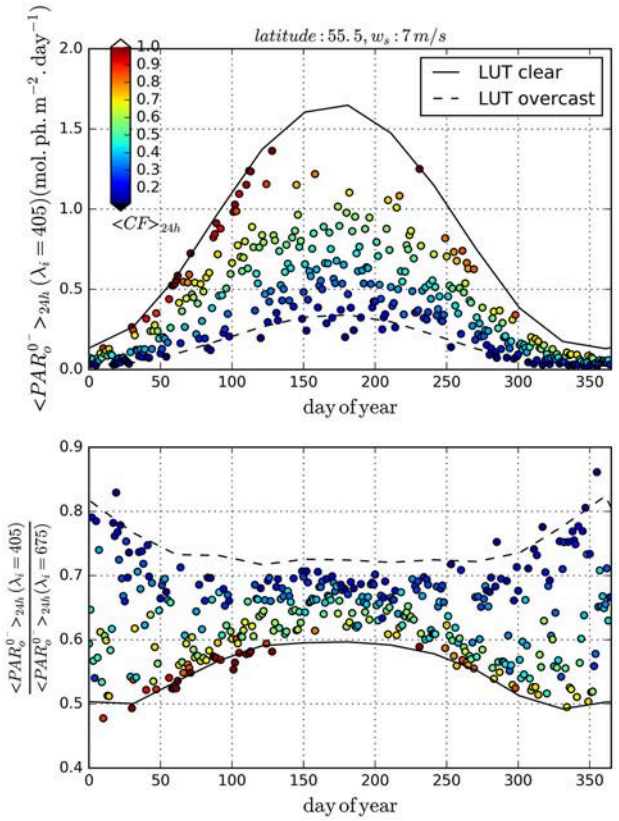


FIGURE 13 | (Top) Scalar PAR below the surface integrated in a narrow spectral band (400–410 nm) for all the days of 2011 (dots), for a particular latitude and wind speed, as well as predicted values for clear sky (solid line) and totally overcast situations (dashed lines). **(Bottom)** Same as top but normalized by the band integrated scalar PAR below the surface between 670 and 680 nm. The color of the dots is a function of the cloud factor.

Figure 18 displays the result of the uncertainty analysis for the whole dataset. The bias and the standard deviation of the daily PAR estimates are plotted as a function of the clear-sky PAR (which depends itself mainly on latitude and date) and the actual cloud factor for that day. The Monte Carlo calculations, assumed accurate, provide the reference. This is justified in view of the bias and standard deviation values. The bias exhibits a slight dependence on the clear sky PAR, suggesting an overestimation reaching 2.5 E.m^{-2}.day^{-1} (∼4%) for the maximum clear sky PAR values, and a slight underestimation 1 E.m^{-2}.day^{-1} (∼6.5%) when the clear sky PAR reaches a low value of 15 E.m^{-2}.day^{-1}. When looking at the dependence upon $\langle CF \rangle_{24h}$, a slight overestimation exists between 1 and 1.5 E.m^{-2}.day^{-1}. This overestimation is independent of the cloudiness of the day, with the exception of totally clear days ($\langle CF \rangle_{24h} = 1$), for which the bias drops to 0.5 E.m^{-2}.day^{-1}. The bias is also quite small for totally overcast situations. A bias correction can be considered, based on the clear-sky PAR value.

The standard deviation (SD) is peaked toward intermediate cloud factors, where the risk of deviation of cloudiness at the time of the satellite measurement and the mean cloudiness of the day is maximum. When the clear sky PAR is large and $\langle CF \rangle_{24h}$ is about 0.5, SD reaches 8 E.m^{-2}.day^{-1} (∼11%). But SD drops to 1 E.m^{-2}.day^{-1} (1.5%) for clear sky situations ($\langle CF \rangle_{24h} = 1$). SD is drastically reduced when dealing with monthly PAR estimates. Whatever the cloudiness, SD is lower than 2 E.m^{-2}.day^{-1}. The main feature is that SD seems to be proportional to the clear-sky PAR value, and thus we can consider associating an uncertainty to each pixel of the product from an estimate of the clear-sky value, in addition to the cloud factor. This model uncertainty component could be extended to situations with multiple satellite measurements per day, as it is often the case for high latitudes with polar orbiting sensors like MERIS or VIIRS. In that case we anticipate a reduction of the standard deviation of the daily PAR product. For a complete per-pixel uncertainty budget, the uncertainty associated with TOA reflectance noise (radiometric and due to vicarious calibration) should be included, which may require evaluating the sensitivity of the daily PAR to the TOA reflectance and the covariance of the input reflectance in the various spectral bands (since the measurements are correlated).

SUMMARY AND RECOMMENDATIONS

Studying and understanding the chemical, physical, geological, and biological processes that govern the composition of the marine environment requires knowledge of the underwater light field. Ideally, one wants to figure out and monitor the spectral and angular structure of the radiant intensity at any point in the water body. From space, by means of remote sensing, only limited information about the radiative properties of a water body can be obtained, but the advantage is global and repetitive coverage. Existing satellite products are restricted to planar PAR

and UV irradiance above the surface, and diffuse attenuation coefficient (average in the upper layer). These products, despite their drawbacks (e.g., no information at high Sun zenith angles and diurnal variability poorly described for PAR, no retrieval in cloudy conditions for the diffuse attenuation coefficient), have been useful to many studies of aquatic photosynthesis, heat budget, and chemical effects of light. There is a need, however, for other products, i.e., sub-surface planar and scalar fluxes, average cosine, spectral fluxes (UV to visible), diurnal fluxes, APAR, surface albedo, vertical attenuation, and heating rate, and for associating uncertainties to any product on a pixel-by-pixel basis. Possible approaches and methodologies to generate these new products, in view of state-of-the-art knowledge and current and planned satellite missions, were discussed, including difficulties, assumptions, and readiness level. A strategy and a roadmap, with development priorities and opportunities to obtain the new products, could be established. Examples of new products, i.e., daily scalar PAR below the surface, daily average cosine for PAR, and spectral scalar fluxes below the surface, and their algorithms, were presented. A statistical way to estimate uncertainties for each pixel, based on radiative transfer simulations for expected clear and cloudy situations, was proposed.

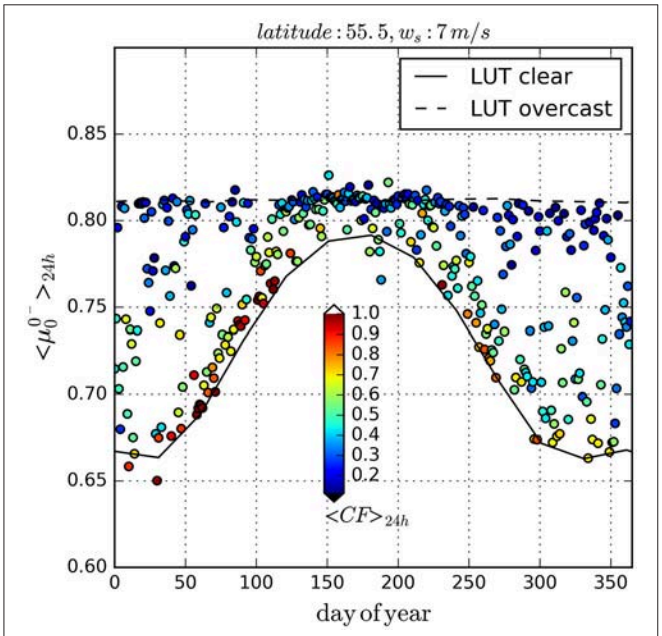


FIGURE 14 | Mean cosine PAR for all the days of 2011 (dots), for a particular latitude and wind speed, as well as predicted values for clear sky (solid line) and totally overcast situations (dashed lines). The color of the dots is a function of the cloud factor.

In the short term, the focus should be on improving/completing existing products from satellite ocean-color sensors, i.e., extending calculations of above-surface fluxes to high Sun zenith angles and accounting, at least statistically, for diurnal variability of clouds. One should also work on how to compute similar products from different missions and their likely uncertainties, to obtain platform-independent global products. On the other hand, with existing scientific knowledge, the following new products (see sections Current Products and Users Needs) should be derived from past and current missions: planar and scalar PAR below the surface, average cosine for PAR below the surface, diffuse attenuation coefficient for PAR, and respective spectral quantities (visible to near infrared), spectral albedo, and APAR. These products should be in the form of daily averages, except for diffuse attenuation coefficient. New processing lines will require links with other ocean and atmosphere products and/or ancillary data (e.g., reanalysis products from observations and models, such as MERRA-2). A calibration/validation program should be planned to evaluate the new products and their uncertainty over a representative set of conditions.

In the medium term (longer-term effort, not lower priority), with specific efforts and new avenues in algorithm development, the aim should be to generate the following products, not only from past and current missions, but also from future missions: spectral fluxes below the surface and diffuse attenuation in the UV (or integrated over UV-A and UV-B ranges), in both photon and energy units, especially using TROPOMI on Sentinel 5P/5, average mixed-layer PAR (will require mixed-layer depth fields from Argo-assimilated circulation models), upper-ocean heating profile, and under-ice light fields (in conjunction with cryosphere missions and modeling). Diurnal-cycle resolving measurements, combining different satellites over-passing at different times with geostationary satellites, should be pursued

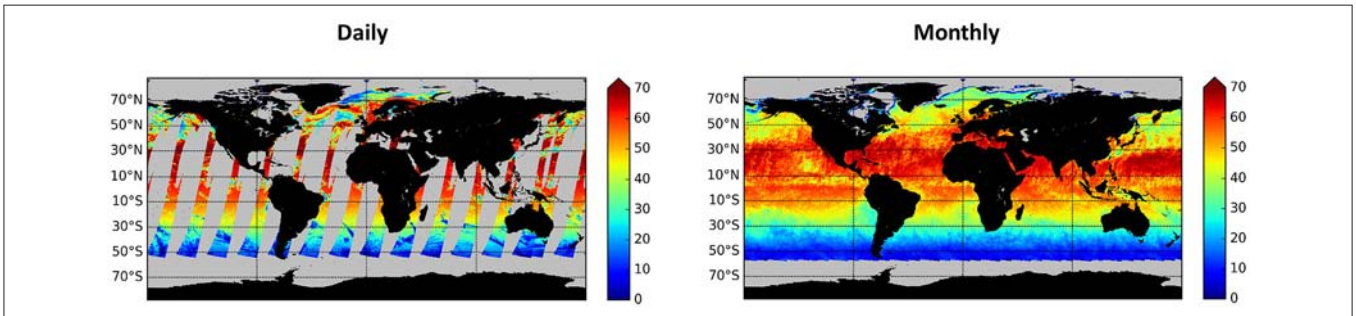


FIGURE 15 | Daily, and monthly scalar PAR below surface (unit: E/m^2/day), 0.1° spatial resolution derived from MERIS. **Top Left**: May 15, 2011; **Top Right**: May 1–31, 2011.

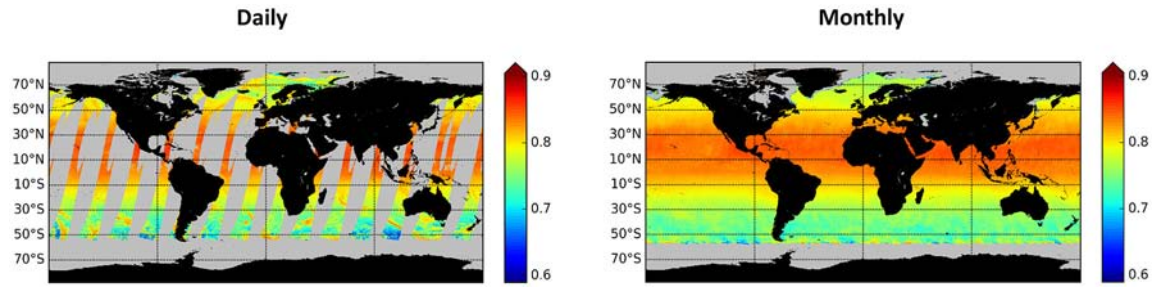


FIGURE 16 | Same as **Figure 15** but for the average cosine below the surface.

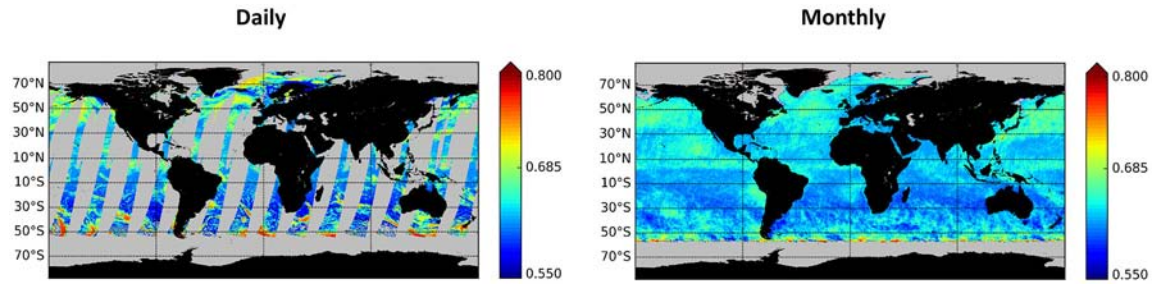


FIGURE 17 | Same as **Figure 15** but for the ratio of spectral scalar PAR below the surface at 405 and 675 nm.

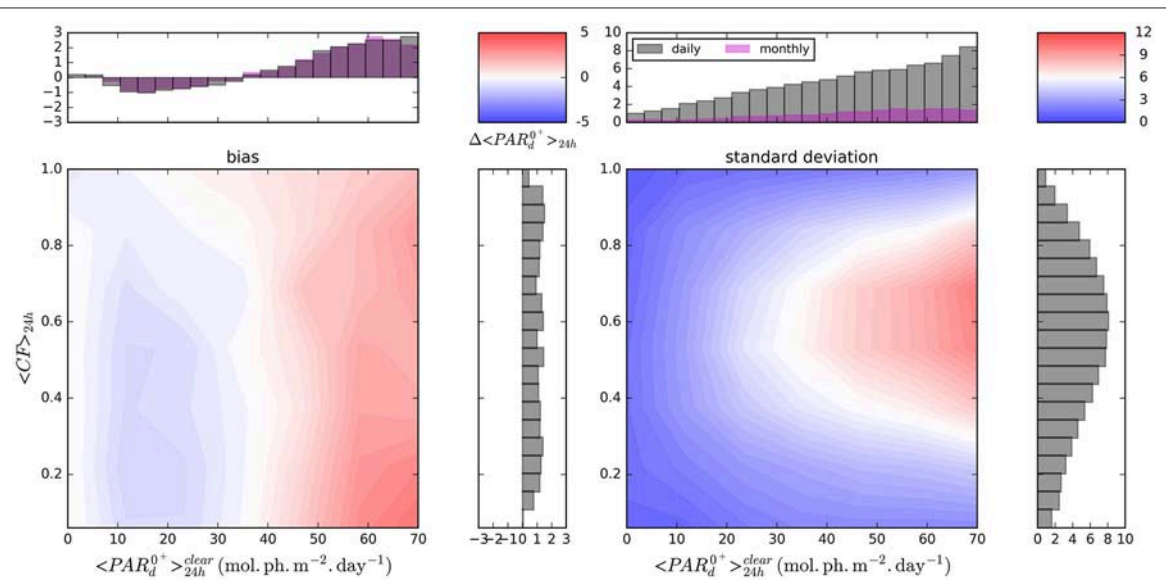


FIGURE 18 | Error budget of the daily "MERIS" PAR product above the surface (estimated PAR–actual PAR) from simulations for the period 2003–2012 using 1-hourly resolved MERRA-2 input data. For each day, a set of MERIS spectral reflectance data is simulated for a typical observation at 10:30 UT local time and several viewing geometries (nadir and 20° view zenith angle (VZA) with relative azimuth of 0, 90, and 180°, with sun glint avoidance). **(Left)** 2D plot of the error bias as a function of the clear sky daily PAR above surface (x axis) and cloud factor (y axis), along with 1D (marginal) distribution. **(Right)** Same as left but for the error standard deviation. For the 1D marginal distribution as a function of the clear-sky PAR is also reported the monthly PAR error bias and standard deviation in magenta.

to describe hourly changes in radiation fields and improve daily-integrated values (e.g., due to clouds). The products should also be generated without gaps (applying gap-filling techniques) to provide boundary conditions for general circulation models. At this stage, evaluation of the products and their uncertainty should be ongoing on a continuous basis.

In the long term (future vision), for significant improvement of sub-surface light fields, space lidars could be used (e.g., the CNES MESCAL), as they can resolve the vertical distribution of material in the ocean (while all the products described above, derived from passive optical sensors, assume a homogeneous upper ocean). This is particularly critical in high-latitude regions (near the ice) and near land. Approaches using hyper-spectrally resolved sensors such as SCHIMACHY to retrieve the availability of light in the ocean (depth-integrated scalar irradiance) from the vibrational Raman scattering effect of water molecules (Dinter et al., 2015) should be explored. Finally, satellites missions with instruments in L1 orbit (such as the NASA EPIC onboard DISCOVR) would offer the opportunity to continuously observe the sun-lit part of the ocean, maximizing the temporal coverage. Space agencies should consider exploiting using such orbit for an ocean-color satellite mission.

AUTHOR CONTRIBUTIONS

All the authors contributed to the conception and organization of the study, to the analysis of the various products and results, and to the writing of the findings. RF, DR, and EB arranged the contents of the various sections, and wrote the first version of the manuscript. HB provided insight into light and biogeochemical processes. JT and BF carried out the work related to current MODIS and OMI products. DR, DJ, MC, and RF developed algorithms for new products and devised ways to associate uncertainties. DR, DJ, and MC generated examples of new products. TJ, TP, and SS organized the survey about satellite PAR observations and compiled and synthesized user responses. All the authors critically revised the manuscript and helped to bring about its final version.

FUNDING

The National Aeronautics and Space Administration (NASA) and the European Space Agency (ESA) provided funding for the study under Grant NNX14AL91G and Contract 40001 1 3690/1 4/I-LG, respectively.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the NASA Ocean Biology Processing Group (OBPG) and the OMI Science Investigator-led Processing System (SIPS) Facility at Goddard Space Flight center, Greenbelt, Maryland for generating and maintaining ocean-color and radiation products used in the study, and Mr. John McPherson, from Scripps Institution of Oceanography and Mr. Monstreleet and Mr. Laurent Wandrebeck from HYGEOS for providing technical support. The survey about PAR products, which involved human participants, did not require ethics approval as per institutional and national guidelines. The participants consented to release the results by virtue of survey completion.

REFERENCES

Alvera-Azcarate, A., Barth, J.-H., Beckers, M., and Weisberg, R. (2007). Multivariate reconstruction of missing data in sea surface temperature, chlorophyll, and wind satellite fields. *J. Geophys. Res.* 112:C03008. doi: 10.1029/2006JC003660

Antoine, D., André, J. M., and Morel, A. (1996). Oceanic primary production: 2. Estimation at global scale from satellite (coastal zone color scanner) chlorophyll, *Glob. Biogeochem. Cycles* 10, 57–69. doi: 10.1029/95GB02832

Arndt, S., and Nicolaus, M. (2014). Seasonal cycle and long-term trend of solar energy fluxes through Arctic sea ice. *Cryosphere* 8, 2219–2233, doi: 10.5194/tc-8-2219-2014

Arola, A., Kazadzis, S., Lindfors, A., Krotkov, N., Kujanpää, J., Tamminen, J et al. (2009). A new approach to correct for absorbing aerosols in OMI UV. *Geophys. Res. Lett*. 36:L22805, doi: 10.1029/2009GL041137

Arrigo, K. R., Perovich, D. K., Pickart, R. S., Brown, Z. W., van Dijken, G. L., et al. (2012). Massive phytoplankton blooms under Arctic sea ice. *Science* 336:1408. doi: 10.1126/science.1215065

Ballabrera-Poy, J., Murtugudde R., Zhang, R.-H., and Busalacchi., A. J. (2007). Coupled ocean–atmosphere response to seasonal modulation of ocean color: impact on interannual climate simulations in the Tropical Pacific. *Am. Meteorol. Soc.* 20, 353–374. doi: 10.1175/JCLI3958.1

Behrenfeld, M., Hostetler, C., Sarmiento, J., in collaboration with the, N. A. S. A., Ocean Biology and Biogeochemistry Advance Planning Committee (2016). *A Global Ocean Science Target for the Coming Decade*. Washington, DC: Response to the 2017 NRC Decadal Survey Request for Information.

Behrenfeld, M. J., O'Malley, R. T., Siegel, D. A., McClain, C. R., Sarmiento, J. L., Feldman, G. C., et al. (2006). Climate-driven trends in contemporary ocean productivity. *Nature* 444, 752–755. doi: 10.1038/nature05317

Behrenfeld, M. J., Boss, E., Siegel, D. A., and Shea, D. M. (2005). Carbon-based ocean productivity and phytoplankton physiology from space. *Glob. Biogeochem. Cycles* 19:GB1006. doi: 10.1029/2004GB002299

Behrenfeld, M. J., Randerson, J., McClain, C., Feldman, G., Los, S., Tucker, C., et al. (2001). Biospheric primary production during an ENSO transition. *Science* 291, 2594–2597. doi: 10.1126/science.1055071

Bodhaine, B. A., Wood, N. B., Dutton, E. G., and Slusser, J. R. (1999). On Rayleigh optical depth calculations. *J. Atmos. Oceanic Tech.* 16, 1854–1861.

Bogumil, K., Orphal, J., Homann, T., Voigt, S., Spietz, P., Fleischmann, O. C., et al. (2003). Measurements of molecular absorption spectra with the SCIAMACHY pre-flight model: instrument characterization and reference data for atmospheric remote sensing in the 230–2380 nm region. *J. Photochem. Photobio.* 157, 167–184. doi: 10.1016/S1010-6030(03)00062-5

Carr, M.-E., Friedrichs, M. A. M., Schmeltz, M., Noguchi Aita, M., Antoine, D., Arrigo, K. R., et al. (2006). A comparison of global estimates of marine primary production from ocean color. *Deep Sea Res. II* 53, 741–770. doi: 10.1016/j.dsr2.2006.01.028

Cox, C., and Munk, W. (1954). Measurement of the roughness of the sea surface from photographs of the suns glitter. *J. Opt. Soc. Am.* 44, 838–850. doi: 10.1364/JOSA.44.000838

Cullen, J. J., and Neale, P. J. (1994). Ultraviolet radiation, ozone depletion, and marine photosynthesis. *Photosynth. Res.* 39, 303–320. doi: 10.1007/BF00014589

Dinter, T., Rozanov, V. V., Burrows, J. P., and Bracher, A. (2015). Retrieving the availability of light in the ocean utilising spectral signatures of vibrational Raman scattering in hyper-spectral satellite measurements. *Ocean Sci.* 11, 373–389. doi: 10.5194/os-11-373-2015

Emde, C., Buras-Schnell, R., Kylling, A., Mayer, B., Gasteiger, J., Hamann, U., et al. (2016). The libRadtran software package for radiative transfer calculations (version 2.0.1). *Geosci. Model Dev.* 9, 1647–1672. doi: 10.5194/gmd-9-1647-2016

Emde, C., Buras, R., and Mayer, B. (2011). ALIS: an efficient method to compute high spectral resolution polarized solar radiances using the Monte Carlo approach. *J. Quant. Spec. Rad. Transfer.* 112, 1622–1631. doi: 10.1016/j.jqsrt.2011.03.018

Falkowski, P. G., and Raven, J. A. (1997). *Aquatic Photosynthesis*. Oxford: Blackwell.

Friedlingstein, P. P., Cox, R., Betts, L., Bopp, W., von Bloh, V., Brovkin, V., et al. (2006). Climate–carbon cycle feedback analysis: results from the C4MIP model intercomparison. *Am. Meteorol. Soc.* 19, 3337–3353. doi: 10.1175/JCLI 3800.1

Frouin, R., and Iacobellis, S. F. (2002). Influence of phytoplankton on the global radiation budget. *J. Geophys. Res.* 107, ACL 5-1–ACL 5-10. doi: 10.1029/2001JD000562

Frouin, R. J., McPherson, J., Ueyoshi, K., and Franz, B. A. (2012). "A time series of photosynthetically available radiation at the ocean surface from SeaWiFS and MODIS data," in *Proceedings SPIE* (Bellingham), 8525.

Frouin, R. J., Ruddorff, N. M., and Kampel, M. (2014). "Estimating solar radiation absorbed by live phytoplankton from satellite ocean-color data," in *Proceedings SPIE* (Bellingham), 9261.

Frouin, R., and McPherson, J. (2012). Estimating photosynthetically available radiation at the ocean surface from GOCI data. *Ocean Sci. J.* 47, 313–321. doi: 10.1007/s12601-012-0030-6

Gasteiger, J., Emde, C., Mayer, B., Buras, R., Buehler, S. A., and Lemke, O. (2014). Representative wavelengths absorption parameterization applied to satellite channels and spectral bands. *J. Quant. Spec. Rad. Transfer* 148, 99–115. doi: 10.1016/j.jqsrt.2014.06.024

Morel, A., and Gentili, B. (1996). Diffuse reflectance of oceanic waters. III. Implication of bidirectionality for the remote-sensing problem. *Appl. Opt.* 35, 4850–4862. doi: 10.1364/AO.35.004850

Gregg, W. W., Esaias, W. E., Feldman, G. C., Frouin, R., Hooker, S. B., McClain, C. R., et al. (1998). Coverage opportunities for global ocean color in a multi mission era. *IEEE Trans. Geos. Rem. Sen.* 36, 1620–1628. doi: 10.1109/36.718865

Häder, D. P., Helbling, E. W., Williamson, C. E., and Worrest, R. C. (2011). Effects of UV radiation on aquatic ecosystems and interactions with climate change, *Photochem. Photobiol. Sci.* 10, 242–260. doi: 10.1039/c0pp90036b

Henson, S. A., Sarmiento, J. L., Dunne, J. P., Bopp, L., Lima, I., Doney, S. C., et al. (2010). Detection of anthropogenic climate change in satellite records of ocean chlorophyll and productivity. *Biogeosciences* 7, 621–640. doi: 10.5194/bg-7-621-2010

Hess, M., Koepke, P., and Schult, I. (1998). Optical properties of aerosols and clouds: the software package OPAC. *Bull. Am. Met. Soc.* 79, 831–844. doi: 10.1175/1520-0477(1998)079<0831:OPOAAC>2.0.CO;2

IOCCG (2006). *Remote Sensing of Inherent Optical Properties: Fundamentals, Tests of Algorithms, and Applications.* Report No. 5 of the International Ocean-Colour Coordinating Group, ed Z. Lee (Dartmouth, NS: IOCCG), 126.

IOCCG (2015). *Ocean Colour Remote Sensing of Polar Regions*. Reports of the International Ocean-Colour Coordinating Group, eds M. Babin, M.-H. Forget (Dartmouth, NS:: IOCCG), 129.

JGCM-100 (2008). *Evaluation of Measurement Data—Guide to the Expression of Uncertainty in Measurement, Document Produced By Working Group 1 of the Joint Committee for Guides in Metrology (Jcgm/Wg 1),* Published by the JCGM member organizations (BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and *OIML)*, 134.

Kahru, M., Kudela, R., Manzano_Sarabia, M., and Mitchell, B. G. (2009). Trends in primary production in the California current detected with satellite data. *Geophys. Res. Lett.* 114:C02004. doi: 10.1029/2008JC004979

Kato, S., Ackerman, T. P., Mather, J. H., and Clothiaux, E. E. (1999). The k-distribution method and correlated-k approximation for a shortwave radiative transfer model. *J. Quant. Spec. Rad. Transfer* 62, 109–121. doi: 10.1016/S0022-4073(98)00075-2

Kirk, J. T. O. (1994). *Light and Photosynthesis in Aquatic Ecosystems*. Bristol: Cambridge University Press

Krasnopolsk, V., Nadiga, S., Mehra, A., Bayler, E., and Behringer, D. (2016). Neural networks technique for filling gaps in satellite measurements: application to ocean color observations. *Comput. Intell. Neurosci.* 2016:6156513. doi: 10.1155/2016/6156513

Laliberté, J., Bélanger, S., and Frouin, R. (2016). Evaluation of satellite-based algorithms to estimate photosynthetically available radiation (PAR) reaching the ocean surface at high northern latitudes, *Remote Sens. Environ.* 184, 199–211. doi: 10.1016/j.rse.2016.06.014

Lee, Y. J., Matrai, P. A., Friedrichs, M. A., Saba, V. S., Antoine, D., Ardyna, M., et al. (2015). An assessment of phytoplankton primary productivity in the Arctic Ocean from satellite ocean color/*in situ* chlorophyll-a based models. *J. Geophys. Res.* 120, 6508–6541, doi: 10.1002/2015JC011018

Loeb, N. G., Wielicki, B., Wenying, S. U, Loukachine, K., Sun, W., Wong, T., et al. (2006). Multi-instrument comparison of top-of-atmosphere reflected solar radiation, *Am. Meteorol. Soc.* 20, 575–591. doi: 10.1175/JCLI4018.1

Lu, X., Hu, Y., Trepte, C., Zeng, S., and Churnside, J. H. (2014). Ocean subsurface studies with the CALIPSO spaceborne lidar. *J. Geophys. Res.* 119, 4305–4317. doi: 10.1002/2014JC009970

Malenovsky, Z., and Schaepman, M. (2011). *Preliminary Scientific Needs for Ocean Sentinel-1,−2, and−3 Products, SEN4SCI (Sentinels for Science)*. Workshop on Assessing Product Requirements for the Scientific Exploitation of the SENTINEL Missions, ESA-ESRIN, (Frascati).

Miller, A. J., Alexander, M. A., Boer, G. J., Chai, F., Denman, K., and Timmermann, A. (2003). Potential feedbacks between Pacific Ocean ecosystems and Inter-decadal climate variations. *Bull. Am. Meteor. Soc.* 84, 617–633. doi: 10.1175/BAMS-84-5-617

Mobley, C. D. (1994). *Light and Water: Radiative Transfer in Natural Waters.* San Diego, CA: Academic Press, 592.

Mobley, C. D., and Boss, E. S. (2012). Improved irradiances for use in ocean heating, primary production, and photo-oxidation calculations *Appl. Opt.* 51, 6549–6560. doi: 10.1364/AO.51.006549

Morel, A., Huot, Y., Gentili, B., Werdell, P. J., Hooker, S. B., and Franz, B. A. (2007). Examining the consistency of products derived from various ocean color sensors in open ocean (Case 1) waters in the perspective of a multi-sensor approach. *Remote Sens. Environ.* 111, 69–88. doi: 10.1016/j.rse.2007.03.012

Morel, A., and Berthon, J.- F. (1989). Surface pigments, algal biomass profiles, and potential-production of the euphotic layer: relationships reinvestigated in view of remote-sensing applications. *Limnol. Oceanogr.* 34, 1545–1562. doi: 10.4319/lo.1989.34.8.1545

Nakamoto, S., Prasana Kumar, S., Oberhuber, J. M., Muneyama, K., and Frouin, R. (2000). Chlorophyll control of sea surface temperature in the Arabian Sea in a mixed layer isopycnal general circulation model. *Geophys. Res. Lett.* 27, 747–750. doi: 10.1029/1999GL002371

Nakamoto, S., Prasanna Kumar, S., Oberhuber, J. M., Ishizaka, J., Muneyama, K., and Frouin, R. (2001). Response of the equatorial Pacific to chlorophyll pigments in a mixed layer isopycnal ocean general circulation model. *Geophys. Res. Lett.* 28, 2021–2024. doi: 10.1029/2000GL012494

Olhmann, J. C., Siegel, D. A., and Gautier, C. (1996). Ocean mixed layer radiant heating and solar penetration: a global analysis. *Am. Meteorol. Soc.* 9, 2265–2280. doi: 10.1175/1520-0442(1996)009<2265:OMLRHA>2.0.CO;2

Park, Y.-J., and Ruddick, K. (2005). Model of remote-sensing reflectance including bidirectional effects for case 1 and case 2 waters. *Appl. Opt.* 44, 1236–1249. doi: 10.1364/AO.44.001236

Platt, T., Denman, K. L., and Jassby, A. D. (1977). "Modeling the productivity of phytoplankton," in *The Sea: Ideas and Observations on Progress in the Study of the Seas*, ed E. D. Goldberg (New York, NY: John Wiley), 807–856.

Platt, T., Sathyendranath, S., Forget, M.-H., White, G. N., III., Caverhill, C., Bouman, H., et al. (2008). Operational estimation of primary production at large geographical scales. *Remote Sens. Environ.* 112, 3437–3448. doi: 10.1016/j.rse.2007.11.018

Popova, E. E., Yool, A., Coward, A. C., Aksenov, Y. K., Alderson, S. G., and Anderson, T. R. (2010). Control of primary production in the Arctic by nutrients and light: insights from a high resolution ocean general circulation model. *Biogeosciences Discuss.* 7, 5557–5620. doi: 10.5194/bgd-7-5557-2010

Pottier, C., Garçon, V., Larnicol, G., Sudre, J., Schaeffer, P., and Le Traon, P.-Y. (2006). Merging SeaWiFS and MODIS/Aqua ocean color data in North and Equatorial Atlantic using weighted averaging and objective analysis. *IEEE Trans. Geosci. Rem. Sen.* 44, 3436–3451. doi: 10.1109/TGRS.2006.878441

Ramon, D., Steinmetz, F., Compiegne, M., and Frouin, R. (2017)."Accurate radiation transfer tool for ocean colour," in *Poster Presented at International Ocean Colour Science Symposium* (San Francisco, CA). Available online at: https://www.eposters.net/poster/

National Research Council (2011). *Assessing the Requirements for Sustained Ocean Color Research and Operations.* Washington, DC: The National Academies Press

Rothman, L. S., Gordon, I. Babikov, E. Y., Hill, C., Kochanovac, R. V., Tan, Y., et al. (2013). The HITRAN 2012 molecular spectroscopic database. *J. Quant. Spec. Rad. Transfer* 130, 4–50. doi: 10.1016/j.jqsrt.2013.07.002

Ryther, J. H. (1956). Photosynthesis in the ocean as a function of light intensity. *Limnol. Oceanogr.* 1, 61–70. doi: 10.4319/lo.1956.1.1.0061

Sathyendranath, S., Platt, T., Caverhill, C. M., Warnock, R. E., and Lewis, M. R. (1989). Remote sensing of oceanic primary production: computations using a spectral model. *Deep-Sea Res.* 36, 431–453. doi: 10.1016/0198-0149(89)90046-0

Schreier, F., and Gimeno Garcia, S. (2013). "Py4CAtS – Python Tools for Line-by-Line Modelling of Atmospheric Radiative Transfer," in *Radiation Processes in the Atmosphere and Ocean (IRS 2012): Proceedings of the International Radiation Symposium (IRC/IAMAS), Vol. 1531 of AIP Conference Proceedings*, eds C. F. Robert and F. Jürgen (Berlin: American Institute of Physics), 123–126.

Shell, K., Frouin, R., Nakamoto, S., and R. C., Somerville, J. (2003). Atmospheric response to solar radiation absorbed by phytoplankton. *J. Geophys. Res.* 108:4445. doi: 10.1029/2003JD003440

Sulzberger, B., and Durisch-Kaiser, E. (2009). Chemical characterization of dissolved organic matter (DOM): a prerequisite for understanding UV-induced changes of DOM absorption properties and bioavailability. *Aquat. Sci.* 71, 104–126. doi: 10.1007/s00027-008-8082-5

Uitz, J., Claustre, H., Gentili, B., and Stramski, D. (2010). Phytoplankton class-specific primary production in the world's oceans: seasonal and interannual variability from satellite observations. *Glob. Biogeochem. Cycles* 24:GB3016. doi: 10.1029/2009GB003680

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Frouin, Ramon, Boss, Jolivet, Compiègne, Tan, Bouman, Jackson, Franz, Platt and Sathyendranath. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

ANNEX: NOTATIONS AND DEFINITIONS

Irradiances (W.m^{-2}.nm^{-1})

-Spectral scalar irradiance

$$E_o^{0\pm}(\lambda,t) = \int_{\phi=0}^{2\pi}\int_{\theta=-\pi/2}^{\pi/2} L^{0\pm}(\lambda,t,\theta,\phi)\sin(\theta)\,d\theta\,d\phi$$

where $L^{0\pm}$ is the radiance just above or just below the surface (superscript $0+$ or $0-$), λ is wavelength, t is time, and θ and ϕ are zenith and azimuth angles.

-Spectral downward planar irradiance

$$E_d^{0\pm}(\lambda,t) = \int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi/2} \cos(\theta)\,L^{0\pm}(\lambda,t,\theta,\phi)\sin(\theta)\,d\theta\,d\phi$$

-Spectral upward planar irradiance

$$E_u^{0\pm}(\lambda,t) = \int_{\phi=0}^{2\pi}\int_{\theta=-\pi/2}^{0} \cos(\theta)\,L^{0\pm}(\lambda,t,\theta,\phi)\sin(\theta)\,d\theta\,d\phi$$

-Spectral net irradiance

$$E_{net}^{0\pm}(\lambda,t) = E_d^{0\pm}(\lambda,t) - E_u^{0\pm}(\lambda,t)$$

Band-Integrated Irradiances (W.m^{-2})

-PAR-integrated scalar irradiance

$$E_o^{0\pm}(t) = \int_{\lambda=400nm}^{700nm} E_o^{0\pm}(t,\lambda)\,d\lambda$$

-PAR-integrated downward planar irradiance

$$E_d^{0\pm}(t) = \int_{\lambda=400nm}^{700nm} E_d^{0\pm}(t,\lambda)\,d\lambda$$

Band-Integrated Quanta Fluxes (mol.ph.m^{-2}.s^{-1})

-PAR-integrated scalar quanta flux

$$PAR_o^{0\pm}(t) = \int_{\lambda=400nm}^{700nm} \frac{E_o^{0\pm}(t,\lambda)\cdot\lambda}{hcN_A}d\lambda$$

-PAR-integrated downward planar quanta flux

$$PAR_d^{0\pm}(t) = \int_{\lambda=400nm}^{700nm} \frac{E_d^{0\pm}(t,\lambda)\cdot\lambda}{hcN_A}d\lambda$$

-Narrowband-integrated scalar quanta flux

$$PAR_o^{0\pm}(t,\lambda_i) = \int_{\lambda_i-\Delta\lambda/2}^{\lambda_i+\Delta\lambda/2} \frac{E_o^{0\pm}(t,\lambda)\cdot\lambda}{hcN_A}d\lambda$$

where λ_i is the central wavelength of any spectral interval $\Delta\lambda$ in the PAR spectral range, h is Plank's constant, c is the speed of light, and N_A is the Avogadro number.

Average Cosines (Dimensionless)

-Spectral average cosine

$$\mu_0^{0\pm}(\lambda,t) = \frac{E_{net}^{0\pm}(\lambda,t)}{E_o^{0\pm}(\lambda,t)}$$

-PAR-integrated average cosine

$$\mu_0^{0\pm}(t) = \frac{E_{net}^{0\pm}(t)}{E_o^{0\pm}(t)}$$

24h-Averaged Quantities

-Daily scalar PAR-integrated quanta flux (mol.ph.m^{-2}.day^{-1})

$$\left\langle PAR_o^{0\pm}\right\rangle_{24h} = N_{\text{sec}}\cdot\int_0^{1day} PAR_o^{0\pm}(t)dt$$

where $N_{sec} = 86400$ s.day^{-1} and t is expressed in day.

-Daily downward planar PAR-integrated quanta flux (mol.ph.m^{-2}.day^{-1})

$$\left\langle PAR_d^{0\pm}\right\rangle_{24h} = N_{\text{sec}}\cdot\int_0^{1day} PAR_d^{0\pm}(t)dt$$

-Daily PAR-integrated average cosine (dimensionless)

$$\left\langle \mu_0^{0\pm}\right\rangle_{24h} = \frac{\left\langle E_{net}^{0\pm}\right\rangle_{24h}}{\left\langle E_o^{0\pm}\right\rangle_{24h}}$$

frontiers in Marine Science

ORIGINAL RESEARCH
published: 20 March 2018
doi: 10.3389/fmars.2018.00087

Optical Classification of the Coastal Waters of the Northern Indian Ocean

S. Monolisha[1], *Trevor Platt*[1,2], *Shubha Sathyendranath*[3], *J. Jayasankar*[1], *Grinson George*[1]* *and Thomas Jackson*[2]

[1] *Fishery Resources Assessment Division, Central Marine Fisheries Research Institute, Kochi, India,* [2] *Plymouth Marine Laboratory, Plymouth, United Kingdom,* [3] *Plymouth Marine Laboratory, National Centre for Earth Observation, Plymouth, United Kingdom*

Coastal waters are optically diverse; studying their optical characteristics is an important application of satellite oceanography. In coastal ecosystems of the northern Indian Ocean, optical diversity has been little studied, except for the global analysis by Mélin and Vantrepotte (2015). This paper is a contribution toward identification and characterization of optical classes in the coastal regions of the northern Indian Ocean. The study identified eight optical classes using the monthly climatological datasets of remote sensing reflectance for the 1998–2013 period from the Ocean Color Climate Change Initiative (OC-CCI, www.oceancolour.org). The optical classification we adopted uses the fuzzy logic method, based on Moore et al. (2009). The seasonal variations of the eight resultant optical classes of the coastal waters of the northern Indian Ocean were explored. From the mean reflectance spectral signals obtained, it appears that classes 1–6 belong to Case-1 waters and classes 7 and 8 correspond to Case-2 waters. Classes 1 to 2 appear in deeper oligotrophic waters; classes 3–6 are present in intermediate depths; classes 7 and 8 are mostly found within inshore eutrophic regions with high chlorophyll concentrations, sediments from river plumes and land runoffs. The optical variability between seasons (the summer and winter monsoon and the intermonsoon seasons) are influenced by variations in physical forcing, such as surface winds, ocean currents, precipitation, and sediment influx from rivers and land runoff. Optical diversity index ranged from around 0.3 to 1.36. High diversity indices ranging between 1 and 1.36 were found in areas dominated by classes 1–4, whereas low diversity indices 0.3 occurred in areas where classes 7 and 8 dominated. The variations in the dominant optical classes are shown to be related to changes in chlorophyll concentration and suspended sediment load, as indicated by remote sensing reflectance at 670 nm. On the other hand, optical diversity appears to be high in zones of transition between dominant optical classes.

Keywords: coastal ecosystems, satellite ocean color, classification, remote sensing reflectance, ecosystem management

OPEN ACCESS

Edited by:
Hervé Claustre, Centre National de la Recherche Scientifique (CNRS), France

Reviewed by:
Pierre Gernez, University of Nantes, France
Aneesh Lotliker, Indian National Centre for Ocean Information Services, India

***Correspondence:**
Grinson George
grinsongeorge@gmail.com

Specialty section:
This article was submitted to Ocean Observation, a section of the journal Frontiers in Marine Science

Received: *30 March 2017*
Accepted: *02 March 2018*
Published: *20 March 2018*

Citation:
Monolisha S, Platt T, Sathyendranath S, Jayasankar J, George G and Jackson T (2018) Optical Classification of the Coastal Waters of the Northern Indian Ocean. Front. Mar. Sci. 5:87. doi: 10.3389/fmars.2018.00087

1. INTRODUCTION

In an ocean under rapid modification by climate change, the boundaries between marine ecological provinces will move, but in ways that are difficult to predict (Karl et al., 1995; Platt and Sathyendranath, 1999). However, there is a premium on knowing the large-scale structure of the ocean ecosystem as it changes through time, in other words on developing and maintaining a

biogeography of the ocean basin. Conventional biogeography relies on collecting and identifying individual specimens through samplings and survey techniques. At large geographical scales, it is a costly and time-consuming procedure to make even a single survey at one time point; making serial surveys to detect possible changes may be prohibitive on the grounds of expense. An alternative approach would be to use data streams from sensors carried on satellites in Earth orbit. Such data have the advantages of high-resolution at the ocean surface, high frequency of coverage, cost-effectiveness and synoptic coverage (Platt and Sathyendranath, 1999, 2008). Potentially, their use could yield a different kind of biogeography, based on data free from the limitations of coarse resolution in space and time. Visible spectral radiometry of the ocean provides a data stream that is particularly useful for ecosystem analysis: the visible spectrum carries information on the pigments and size of phytoplankton cells, as well as on the optical properties of the other constituents (such as suspended sediments and colored or chromophoric dissolved organic matter) in the surface waters of the ocean (Guzman et al., 1995; Babin et al., 2003; Dowell and Platt, 2009; Garaba and Zielinski, 2013). Mélin and Vantrepotte (2015) have pioneered the classification of coastal waters at global scale using annually-averaged fields of optical radiances from satellite data.

The Northern Indian Ocean is landlocked toward the north and bifurcates into two intra-continental seas: the Arabian Sea and the Bay of Bengal. Seasonally reversing monsoons and reversal of ocean currents are the major distinguishing features of the Indian Ocean basin (Shetye, 1998; Qasim, 1999). The monsoonal cycle, including southwest or summer monsoon and northeast or winter monsoon, determines the climate of the region. Southwest monsoon is the continuation of the southern hemisphere trade winds that bring monsoon rains and floods to the Asian landmass (Tomczak and Godfrey, 2001). Northeast monsoon is characterized by high pressure over the Asian land mass and northeasterly winds over the tropics and northern subtropics (Shetye and Shenoi, 1988). A strong coastal upwelling occurs along the western coast during the southwest monsoon season, whereas during the northeast monsoon season, cold continental winds cause convective mixing and winter cooling along the north Indian coast (Tomczak and Godfrey, 2001). Other oceanographic features of interest in this region include the Indian Ocean warm pool (Vinayachandran and Shetye, 1991) and monsoon depressions and cyclones (Schott and McCreary, 2001; Schott et al., 2009).

In the Northern Indian Ocean, biogeographical analysis has so far been restricted to what can be found using conventional methods (Krishnamurthy et al., 1978; Schills and Wilson, 2006; Obura, 2012, 2016; Jeffries et al., 2015). Few notable studies on global ocean biogeographic partitions using satellite datasets include: Longhurst province classification (Longhurst, 1998), based on regional oceanography of major oceanic basins, and a global database of chlorophyll profiles; and the 56 biogeochemical provinces proposed by Reygondeau et al. (2013) using the datasets of Sea Surface Temperature (SST), Chlorophyll and Sea Surface Salinity (SSS). The current study region includes at least parts of four provinces proposed by Longhurst (1998): the Red Sea and Persian gulf province (REDS), Northwest Arabian Sea upwelling province (ARAB), Western India coastal province (INDW), and Eastern India coastal province (INDE). Studies on biogeographic partitioning of the Indian Ocean region using remotely-sensed datasets are relatively few. Here, we follow the lead of Mélin and Vantrepotte (2015) through a detailed implementation of their optical remote-sensing method to the Indian Ocean region. We extend the temporal resolution to reveal seasonal changes in the optical classification of the coastal waters of the region. We interpret the results in the context of the seasonally-reversing wind and ocean current system that is the unique oceanographic characteristic of the region.

2. DATA AND METHODS

2.1. Study Area

Northern Indian Ocean is subdivided by landmasses into the Arabian Sea in the west and the Bay of Bengal in the east and it opens into the equatorial Indian Ocean to the south. The Bay of Bengal coast is shared among India, Bangladesh, Myanmar, Sri Lanka, and the western part of Thailand. The Arabian Sea coast is shared among India, Yemen, Oman, Iran, Pakistan, Sri Lanka, Maldives, and Somalia. The area of interest is the coastal waters of the northern Indian Ocean within the 2,000 m isobath (**Figure 1**) (extending from 0 to 30° N latitude and 50 to 100° E longitude). Rather than using a more shallow depth (100–200 m) as the outer limit of the coastal zone, we have opted to use the 2,000 m isobath for the outer limit. This was to explore whether optical signatures of offshore waters appeared close to shore, and vice versa. In this choice, we were guided by Antony et al. (2002) who suggested that the offshore influence of coastal waters could extend as far out as 400 km from the shore. This region is well-known for the alternate upwelling and downwelling processes occurring during the contrasting seasons of southwest and northeast monsoons.

Here, we use satellite remote-sensing reflectances (R_{rs}) at six wavelengths (412, 443, 490, 510, 555, and 670 nm) to identify optically-distinct regions of the coast. Figure S1 provides a schematic diagram of the methods used in the current study.

2.2. Satellite Dataset

Remote sensing reflectance (R_{rs}) of six wavelengths and Chlorophyll datasets were obtained from Version 2 of the Ocean Color Climate Change Initiative (OC-CCI, see www.oceancolour.org) (Sathyendranath et al., 2016, 2017) with spatial resolution of 4 km. Chlorophyll concentration was calculated from the remote-sensing reflectance, using the National Aeronautics and Space Administration (NASA) Ocean Color Chlorophyll Version 4 (OC4) algorithm (O'Reilly et al., 1998). This algorithm performed best in an algorithm comparison carried out as part of OC-CCI activities (Brewin et al., 2015). The OC-CCI satellite datasets affords superior coverage for the area of interest, compared with previously available data. These data products are band-shifted, bias-corrected and merged data archives obtained from three sensors: Sea-WIFS (Sea-Viewing Wide Field-of-View Sensor), MODIS-Aqua (Moderate Resolution Imaging Spectro-radiometer of the Aqua earth Observing System), and MERIS (Medium Resolution

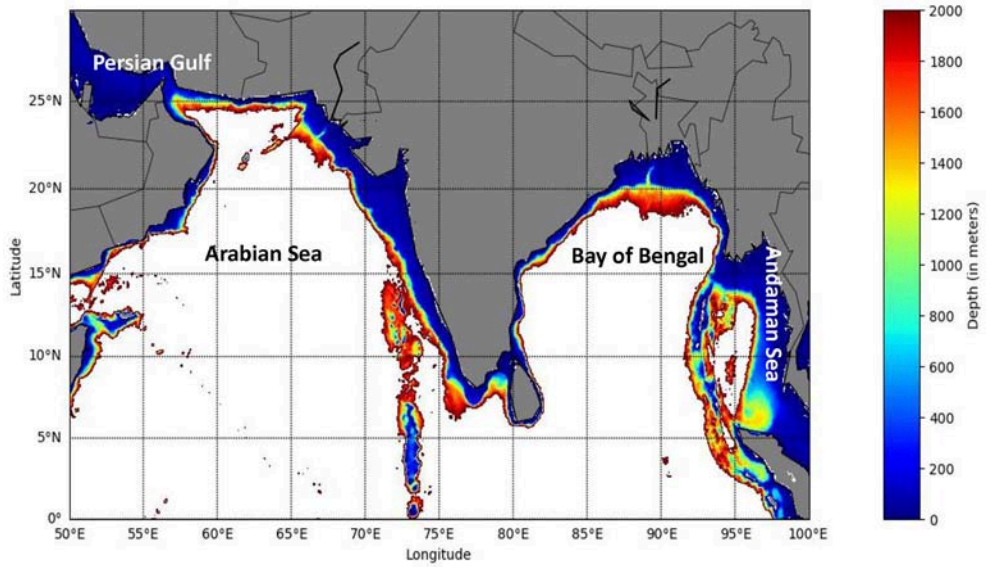


FIGURE 1 | Northern Indian Ocean region showing 2,000 m and shallower isobaths.

Imaging Spectrometer). OC-CCI datasets were validated using the *in-situ* datasets from Teledyne/Webb APEX-Argo floats deployed in the Arabian Sea (Roxy et al., 2016). The OC-CCI dataset are limited to the six SeaWiFS wavebands in the visible. We recognize the limitations of the bandset that were identified by Mélin and Vantrepotte (2015) for coastal optical classification. Therefore, the analyses and interpretation are restricted to the optical differences that are amenable to identification by the available dataset. All grid points of the selected region (depth range of 0–2,000 m) were used in the classification: grid points outside the 2,000 m depth range were excluded. Isobaths of the region were taken from the General Bathymetric Chart of the Oceans (GEBCO) 1-min gridded data set (**Figure 1**).

2.3. Normalization of Dataset

The remote-sensing reflectances (R_{rs}) at six wavelengths (412, 443, 490, 510, 555, and 670 nm) for the years 1998–2013 were used for the study. The remote-sensing reflectance values were skewed in their distribution and to minimize skewness, each R_{rs} spectrum was transformed to its $\log_{10}$ values. They were then normalized by its integral from λ_1 (412 nm) to λ_2 (670 nm), where λ is the wavelength.

$$\vec{x} = \log_{10} R_{rs}(\lambda) / \int_{\lambda_1}^{\lambda_2} \log_{10} R_{rs}(\lambda) d\lambda, \quad (1)$$

where $(\vec{x})$ (in units of nm^{-1}) indicates the normalized spectrum. The denominator was computed by trapezoidal integration. The normalization allows analysis of changes in the shape of the R_{rs} spectra, rather than in their magnitudes. Typically, changes in the shape of the spectra would be more affected by the composition of the materials present in the water, whereas the magnitude of the spectra is likely to be more indicative of the concentration of the substances, especially of highly-scattering substances. In this work, the vector $\vec{x}_j$ of six log-transformed and normalized reflectance values from a particular location and time (pixel, here indexed by subscript j) is referred to as an object. The total number of objects in a classification is N. Notation and Definitions used in this study are presented in **Table 1**.

2.4. Fuzzy C Mean Algorithm

Fuzzy classification evolved from classical set theory. The classical clustering approach determines whether an object is a member or non-member of a given set of any system. Only these two options are possible. In contrast, fuzzy logic allows that an object may have partial memberships in more than one set. The classification algorithms based on fuzzy logic are often used in classifying data from natural systems. The method allows for overlap between boundaries of particular classes or sets, and recognizes that more than one class may be represented at a particular location at any given time.

The membership F_{ij} of a cluster i in the object j is given by $(1 - Q(\vec{Z}_{ij}^2))$ where $\vec{Z}_{ij}$ is the Mahalanobis distance given by $(\vec{x}_j - \vec{M}_i)/\vec{S}_i$ where $\vec{M}_i$ is the mean, $\vec{S}_i$ is the standard deviation and Q is a cumulative χ^2 distribution (Zadeh, 1965).

In this study, the log-transformed, normalized reflectance spectra $(\vec{x})$ were analyzed using the Fuzzy C-mean (FCM) algorithm. Our implementation of fuzzy C mean classification follows Moore et al. (2001). It calculates the centres of each class or cluster and the percentage membership of each class in the data at each pixel. The FCM algorithm also uses several validity functions to assess the optimal number of clusters to be chosen for the classification (Bezdek, 1973; Rezaee, 2010).

TABLE 1 | Definition and Notations used in this study.

Notations	Definitions
C_i	*i*th cluster
c_i	Center of cluster C_i
$d(c_i, c_k)$	Distance between center of clusters *i* and *k* (i ≠ k)
$d(\vec{x}_j, c_i)$	Distance between object $\vec{x}_j$ and center of cluster C_i
$F_{i,j}$	Membership of object *j* in cluster *i*
$F^*_{i,j}$	Membership of object *j* in cluster *i*, normalized by the total membership in all clusters
H_j	Shannon diversity of the object *j*
i	Cluster index, $i = \{1,\ldots n_C\}$
j	Index for object $\vec{x}, j = \{1, \ldots., N\}$
$\vec{M}_i$	Mean of cluster C_i
N	Total number of objects, $\vec{x}$
n_C	Number of clusters
n_i	No. of objects in cluster C_i
P	Partition Coefficient
Q	Cumulative χ^2 square distribution
R_{rs}	Remote-sensing reflectance
$\vec{S}_i$	Standard deviation of Cluster C_i
X	Xie-Beni Index
$\vec{x}_j$	An object, defined as the vector consisting of six remote sensing reflectances R_{rs} after log transformation and normalization (equation 1) for a given pixel *j*
$\vec{Z}_{ij}$	Mahalanobis distance between the object $\vec{x}_j$ and the center C_i of cluster *i*

2.5. Optimal Cluster Validity Functions

Cluster validity function is a statistical measure used to select the optimal number of clusters in the classification (too many clusters would imply that individual clusters resemble each other; too few would imply that all possible cases are not covered). We have used two methods: 1. Xie-beni index and 2. Partition co-efficient. These two methods are used only for selecting the optimal cluster number to run the fuzzy C-means classification. Cluster validity methods are statistical functions that determine the performance of a clustering procedure. Criteria of merit for a clustering method include the distance between clusters (separation) and the distribution of points around a cluster (compactness) (Deborah et al., 2010). We can rely on multiple validity functions to aid selection of the optimal cluster number. The principal strategy used is to cluster the data over a range of cluster numbers (n_c) and evaluate each clustering result with each validity function (Moore et al., 2009).

The Partition Coefficient and the Xie-Beni index are cluster validity methods designed specifically for use with fuzzy algorithms. These two methods are preferred to aid selection of the optimal number of clusters in fuzzy classification (Halkidi et al., 2001).

2.6. Xie-Beni Index

The Xie-Beni index X is one of the measures used to determine the best cluster number for the fuzzy classification of a particular dataset. This index depends on the geometric properties of the dataset and the membership matrix. To calculate X, we need to calculate two quantities: the sum over all clusters of the mean squared distance of each data object from the centre c_i of cluster C_i; and the square of the minimum distance between two cluster centers (Xie and Beni, 1991). The ratio of these two quantities is the Xie-Beni index:

$$X = \Big[\sum_i \sum_{\vec{x}\in C_i} d^2(\vec{x}, c_i)\Big] \Big/ \Big[N. \min_{i,j\neq i}\{d^2(c_i, c_k)\}\Big]. \quad (2)$$

The smallest value of the index indicates the best cluster number (Halkidi et al., 2001; Zhao et al., 2009).

2.7. Partition Coefficient

The Partition Coefficient P is a validity function that uses the membership values (F_{ij}) to provide the optimal cluster number. It measures the amount of overlap between clusters. It is defined as the ratio of the sum of squares of the membership matrix elements of all the clusters to the total number of objects.

$$P = \frac{1}{N}\sum_{j=1}^{N}\sum_{i=1}^{n_c}(F_{ij}^2). \quad (3)$$

The index values lie in the range $[1/n_c, 1]$, where n_c is the number of clusters. The closer the value is to one the better the data are classified. The cluster number with a maximum partition coefficient is said to be the best cluster number to choose for classification (Bezdek, 1973; Bezdek et al., 1984).

2.8. Optical Diversity

Optical diversity is an indicator of the overall variability in optical constituents at a given space and time. Optical diversity, (H_j) is defined here, following Mélin and Vantrepotte (2015), by analogy with the Shannon Diversity Index (Shannon, 2001),

$$H_j = -\sum_{i=n}^{n_c}(F^*_{i,j})\ln(F^*_{i,j}), \quad (4)$$

where $F^*_{i,j}$ is the normalized membership of the optical classes and n_c is the number of classes represented. The membership $F_{i,j}$ was normalized by the integral of $F_{i,j}$ over all optical classes to obtain $F^*_{i,j}$:

$$F^*_{i,j} = (F_{i,j})/\Big(\sum_{i=1}^{n_c} F_{i,j}\Big). \quad (5)$$

3. RESULTS AND DISCUSSION

3.1. Selection of the Optimal Class Number

The Xie-Beni Index and the Partition Coefficient were calculated for monthly climatologies of $\vec{x}$ for the study area, computed from the OC-CCI monthly R_{rs} climatologies, which are based on years 1998–2013. Monthly values allowed study of seasonal variations in the distribution of optical classes in this region, which is known for its pronounced seasonality. Climatologies

were selected to minimize the effect of outliers through averaging, and also to improve the coverage and reduce gaps in data. Climatological data also provide a baseline against which trends in anomalies can be studied at a future date. Both indices varied between months, and the Xie-Beni index often showed a broad minimum, whereas the Partition Coefficient often showed a broad maximum, such that selection of the optical class number was not straightforward. Nevertheless, eight emerged as the optimal number. To aid the selection of optimal class number further, we also studied the maps of cumulative membership (sum of the memberships of all the classes) calculated using class numbers n_c from 5 to 15. The maps were studied for evidence of over-classification (large areas where the cumulative membership was >1) and under-classification (large areas where cumulative membership was <1). This study also showed that n_c = 8 gave the best compromise, with low numbers of both under-classified and over-classified pixels. Therefore, finally, eight classes were selected as the optimal cluster number for all the analyses presented here.

3.2. Identification of the Optical Classes

The mean spectra $\vec{M}_j$ of the eight selected optical classes are shown in **Figure 2**. Optical class 1 is characterized by a maximum in the blue, with the signal decreasing progressively toward the red, indicative of clear oceanic waters. With increasing class number, the signal decreased steadily at the shortest wavelength (412 nm), and the maximum shifted toward longer wavelengths: the maximum is at 490 nm for class 6 and 555 nm for class 8. Conversely, class 1 has the minimum value in the red at 670 nm, whereas classes 7 and 8 have the highest values in the red. The values of the mean spectra at the six SeaWiFs wavelengths for each of the classes and their corresponding covariance matrix are provided in Tables S1, S2. It is useful to assess how these optical classes relate to Case-1 and Case-2 waters as defined by Morel and Prieur (1977) and Prieur and Sathyendranath (1981). From the shapes of the spectra, it appears that classes 1–6 are representative of Case-1 waters and classes 6–8 of turbid Case-2 waters.

The distributions of the dominant classes of representative months of the four seasons are shown in **Figure 3**. The mean (**Figure 2**) and covariance values of the optical classes were then used to classify the waters of the study area for all the months of the year, using the climatological satellite R_{rs} data as inputs, after log-transformation and normalization to obtain $\vec{x}$. Seasonal cycles used in description of the optical classes are: 1. southwest monsoon or summer monsoon (June–September), 2. northeast monsoon or winter monsoon (December–March), 3. spring intermonsoon (April–May), 4. autumn (fall) intermonsoon (October–November).

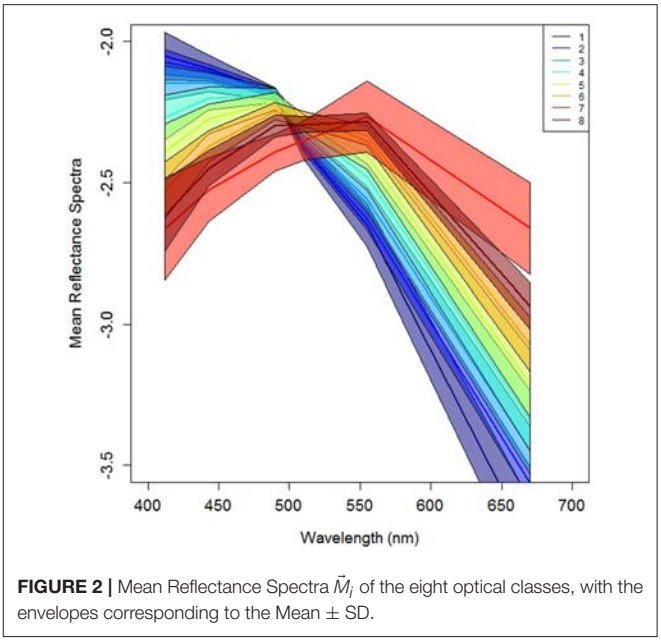


FIGURE 2 | Mean Reflectance Spectra $\vec{M}_j$ of the eight optical classes, with the envelopes corresponding to the Mean ± SD.

3.3. Spatio-Temporal Variations of Optical Classes in the Northern Indian Ocean

3.3.1. Classes 1 and 2

Optical classes 1 and 2 vary strongly with season. They occur along with class 3 over deeper waters (>200 m). During the southwest monsoon season (June–September), these classes represent very few pixels in deeper waters near the Andaman Sea. In the intermonsoon period (October–November) and northeast monsoon (December–March), classes 1 and 2 are present in the deeper waters along the Southwest coast, West Bengal, and Andaman Sea. The Eastern India coastal current carries the low salinity waters (Optical classes 1 and 2) of the Bay of Bengal to the southwest coast of India during the months of February and March in the winter monsoon season. In the months of March and April, the nearshore waters of the Somalia coast and Gulf of Aden are represented by optical classes 1 and 2, extending to the Gulf of Yemen and Oman. The region 17–20° N, 69–72° E (deeper waters) was also characterized by classes 1, 2, and 3 during the transition period (April–May). The Chlorophyll concentration corresponding to these classes ranged from 0 to 0.2 mg m^{-3} and the optical diversity index fell in the range from 1 to 1.3.

3.3.2. Classes 3 and 4

Optical classes 3 and 4 occurs in the isobaths of 100–2,000 m (shallow to deeper depths). These classes show irregular boundaries in the offshore waters during southwest monsoon season along west and east coasts of India, extending to the deeper waters of Andaman Sea. The classes are found near Gulf of Aden and Oman waters only in June, i.e., during the onset of southwest monsoon. In the autumn intermonsoon (October–November), the classes were distributed over the shallow depths (0–500 m) along the near-shore waters of Gulf of Yemen, Oman, Arabian Sea, and Bay of Bengal. In the northeast monsoon, these classes occurred around the islands off Somalia coast, Gulf of Yemen, west coast, and east coast of India. The Gulf of Oman waters flowing toward the Arabian Sea represents class 4 in May (spring intermonsoon). Chlorophyll concentration corresponding to these classes fell in the range of 0.5–0.75 mg m^{-3} and the diversity index ranged from 0.3 to 0.9.

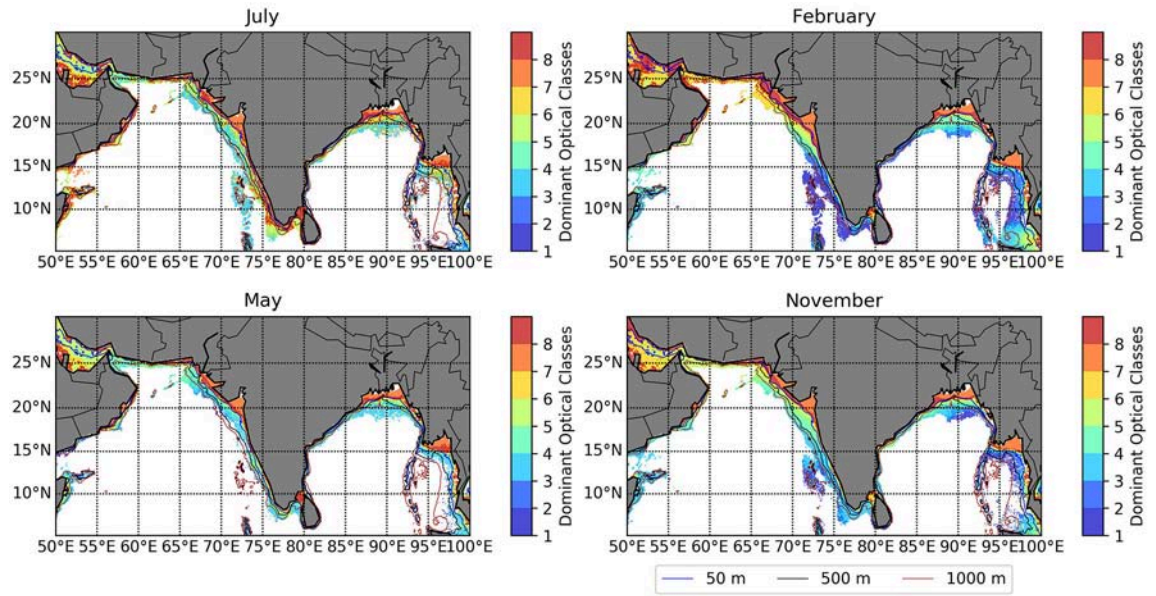


FIGURE 3 | Monthly climatologies of dominant classes (classes of maximum memberships) in the coastal waters of the Northern Indian Ocean. **(Top left)** July, representative of southwest monsoon season. **(Top right)** January, representative of northeast monsoon season. **(Bottom left)** May, representative of Spring intermonsoon season. **(Bottom right)** November, representative of Fall intermonsoon season. All monthly climatologies are calculated for the years 1998–2013.

3.3.3. Class 5

This optical class dominates in regions with isobaths of <1,000 m but >200 m. During the onset of southwest monsoon, class 5 is prominent in the inner Persian Gulf, Strait of Hormuz, Gulf of Oman, Somalia, west and east coasts of India. At the end of southwest monsoon season and onset of the fall intermonsoon period, this class is distributed throughout the coastline in the depth range 0–500 m. This trend persists until the month of January (northeast monsoon) along the entire coastline. In the spring intermonsoon this class is present toward the Persian Gulf, Gulf of Oman flowing into the Northwest coast and further extending toward the east coast of India including the Andaman Sea. This class has chlorophyll levels ranging from 0.75 to 1 mg m^{-3} and the diversity index falls between 0.2 and 0.8.

3.3.4. Classes 6–8

Classes 6–8 dominate in the regions with depths <200 m. Class 6 is dominant in the Persian Gulf characterized by high dense saline waters in all the seasons. Chlorophyll concentration of regions with class 6 varied from 1 to 1.5 mg m^{-3}. Classes 7 and 8 are present in the inner shelf regions with shallower depths influenced by boundary currents and river influx. These classes appear in the near-shore waters off the Somalia coast, Gulf of Oman, Inner Gulf of Kutch and Khambhat, Inner Ganges shelf, and Irrawady river basin near Andaman Sea. Local wind-driven circulation brings in the waters of optical classes 7 and 8 from the major river deltas and minor rivers. The influxes from rivers are seasonally variable and rain-fed according to changing precipitation. In the northeast monsoon, waters belonging to classes 7 and 8 flow toward the Strait of Hormuz and into the Arabian Sea in the months of December to April under the influence of strong northwest wind during winter monsoon (Hunter, 1983), turning the Gulf of Oman waters into classes 7 and 8 in February. These classes do not show major variations in the transition periods. The chlorophyll concentration in the regions with classes 7 and 8 was high, ranging from 2 to 2.5 mg m^{-3}. The diversity index of the classes 6–8 were low, around 0.3.

3.4. Optical Diversity Index

The previous section describes the distribution of the dominant classes, but contains no information on contributions to the optical signal from non-dominant classes. The optical diversity index, which depends on membership of all classes represented in a pixel provides complementary information on the extent to which non-dominant classes are contributing to the signal. If a single class contributed to the optical signal of a pixel, then the optical diversity would be a minimum of 0.26 in our classification with 8 classes represented. On the other hand, if all classes contributed equally, then the optical diversity index would reach a maximum of 2.08.

The optical diversity index H (Equations 4 and 5) was calculated for all months to study the seasonal and regional variations in optical diversity. The optical diversity index H (**Figure 4**) fell mostly between 0.3 and 1.36. Regardless of season, higher H values (1–1.36) were found in deeper waters off the south west coast of India (6–15° N), around Lakshadweep and Maldive Islands in the Arabian Sea, around Andaman and Nicobar Islands in the Bay of Bengal and off Myanmar. The highest H values are found in these locations during the winter or northeast monsoon season (December–March),

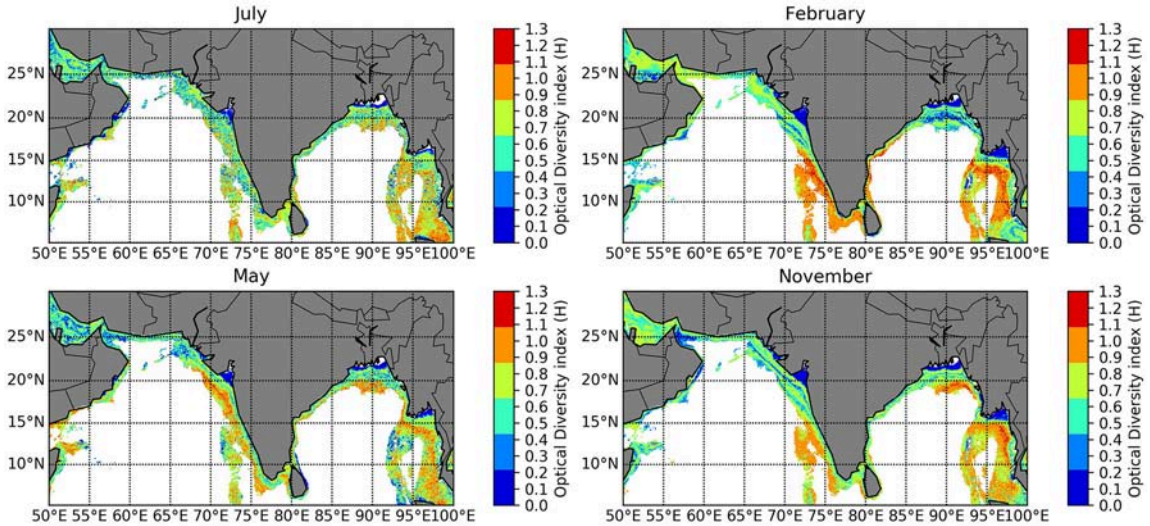


FIGURE 4 | Monthly climatologies of Optical diversity index in the coastal waters of the Northern Indian Ocean. **(Top left)** July, representative of southwest monsoon season. **(Top right)** January, representative of northeast monsoon season. **(Bottom left)** May, representative of Spring intermonsoon season. **(Bottom right)** November, representative of Fall intermonsoon season. All monthly climatologies are calculated for the years 1998–2013.

when the areas covered by high H values were also more extensive. High diversity indices were also found in the fall intermonsoon and spring intermonsoon periods. During the summer monsoon season (June–September), the diversity index lay mostly in the 0.5–1 range along the entire study region with some pixels having index >1 appearing in the deeper waters. High diversity indices (1–1.36) occur along the productive upwelling areas, the transition zones between coast and open ocean, oligotrophic waters and regions with the influence of boundary currents. Low diversity indices occur in the regions of most turbid waters, regions of high river water influx and inland waters.

3.5. Optical Classes, Optical Diversity, and Chlorophyll Concentration

In coastal waters, we know that the optical remote-sensing reflectance spectra are affected not only by chlorophyll concentration, but also by suspended sediment load. The seasonal variability of the chlorophyll concentration for the representative months is shown in **Figure 5**. The remote-sensing reflectance value at 670 nm is often taken to be a measure of suspended sediment load. Therefore, in **Figure 6**, we have plotted the dominant optical class as a function of chlorophyll-a concentration and $R_{rs}(670)$, for the monthly climatology of February, as an example. Only well-classified pixels (cumulative class membership >0.5) are plotted. We see a gradual progression in the optical classes 1–8, with increasing chlorophyll concentration and increasing R_{rs} values, clearly indicating that the optical classification is affected by both chlorophyll concentration and suspended sediment load. Since the chlorophyll concentration in Version 2 of OC-CCI was calculated with a single, global algorithm, we can discount the possibility that the relationship seen in **Figure 6** is emerging from the use of different algorithms for different optical classes. On the other hand, it is worth discussing whether a single chlorophyll algorithm would work equally well in all optical classes in the coastal waters of the northern Indian Ocean. Tilstone et al. (2011) reported that there was a good agreement between OC4v6 and another algorithm (OC5) in open-ocean and coastal waters with chlorophyll concentration up to 2 mg m^{-3} for the Arabian Sea and the Bay of Bengal. In the current study, the classes 7 and 8 had chlorophyll concentrations ranging from 1.5 to 2.5 mg m^{-3}, quite close to conditions discussed by Tilstone et al. (2011), so that we can assume that OC4 algorithm was suitable for even these high-turbid classes. Nevertheless, it would be interesting, in a future study, to explore the advantages of using algorithms designed for coastal waters (e.g., Le et al., 2013; Loisel et al., 2017; Tilstone et al., 2017).

A similar plot (**Figure 7**) for the optical diversity index H reveals a more complex pattern, with very high and very low indices appearing in close juxtaposition to each other in clear waters (where both chlorophyll concentration and R_{rs} values are low). Strands of high values of the index also appear for higher values of chlorophyll concentration and R_{rs}. The differences in the distribution of H values, compared with **Figure 6** for the optical classification, suggest that optical diversity H perhaps tends to be high during transition between optical classes.

3.6. Comparison of Regional Optical Classes With Results of a Global Classification

The question remains whether the regional classification presented here yields results similar to those found in the global classification of Mélin and Vantrepotte (2015). The first difference we note is that the regional classification yielded

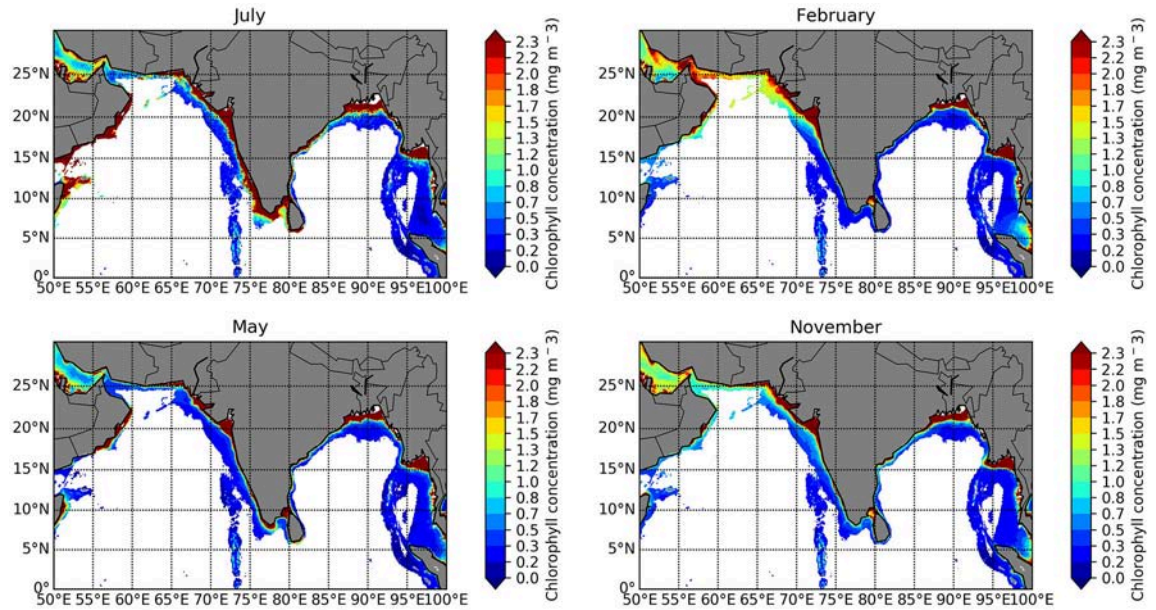


FIGURE 5 | Monthly climatologies of chlorophyll concentrations in the coastal waters of the Northern Indian Ocean. **(Top left)** July, representative of southwest monsoon season. **(Top right)** January, representative of northeast monsoon season. **(Bottom left)** May, representative of Spring intermonsoon season. **(Bottom right)** November, representative of Fall intermonsoon season. All monthly climatalogies are calculated for the years 1998–2013.

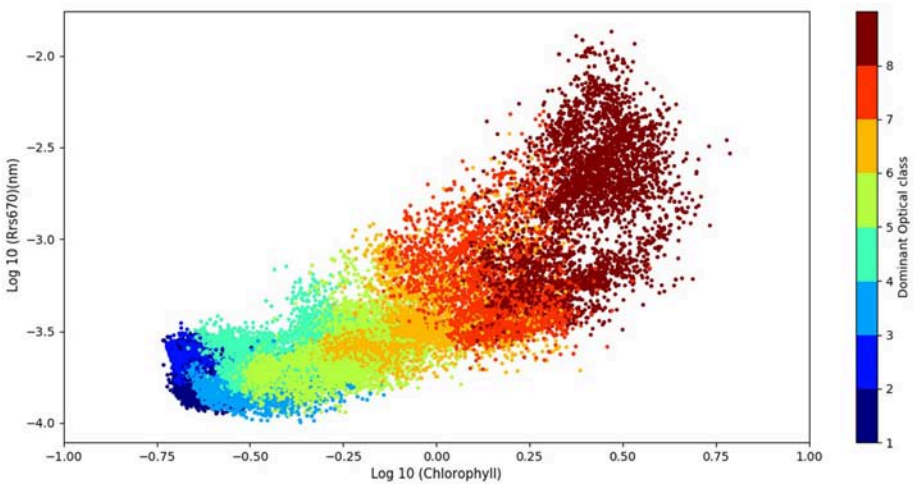


FIGURE 6 | Relationship between chlorophyll concentration and R_{rs}(670). Climatological data for February are shown as an example. Only well-classified pixels (cumulative membership >0.5) are plotted here. The dominant optical classes are identified using different colors.

only eight classes, whereas the global classification produced 16 distinct classes. The log-normalized mean reflectance spectra (**Figure 2**) of our optical classes 1–4 are similar to those of classes 10–16 of Mélin and Vantrepotte (2015). The spectral characteristics of their classes 8–16 of Mélin and Vantrepotte (2015) are typical of clear waters, and are similar to those of our classes 1 and 2. We also see similarities between the optical signatures of our classes 6, 7, and 8 and the classes 1–7 of Mélin and Vantrepotte (2015), and both these sets are typical of highly turbid waters, with mineral particles and dissolved organic matter (Vantrepotte et al., 2012). For the optical diversity, the values of H from this study lie in the range of 0–1.3 which was lower than the range (0–3) reported by Mélin and Vantrepotte (2015) globally. These differences in values of optical diversity are associated, by definition, with the differences in the number of classes, which has a direct impact on the values of optical diversity (Mélin and Vantrepotte, 2015). It is important to note therefore, that the values of optical diversity reported here are not directly comparable with those of Mélin and Vantrepotte (2015). Similarly, the differences in the class numbers have to be

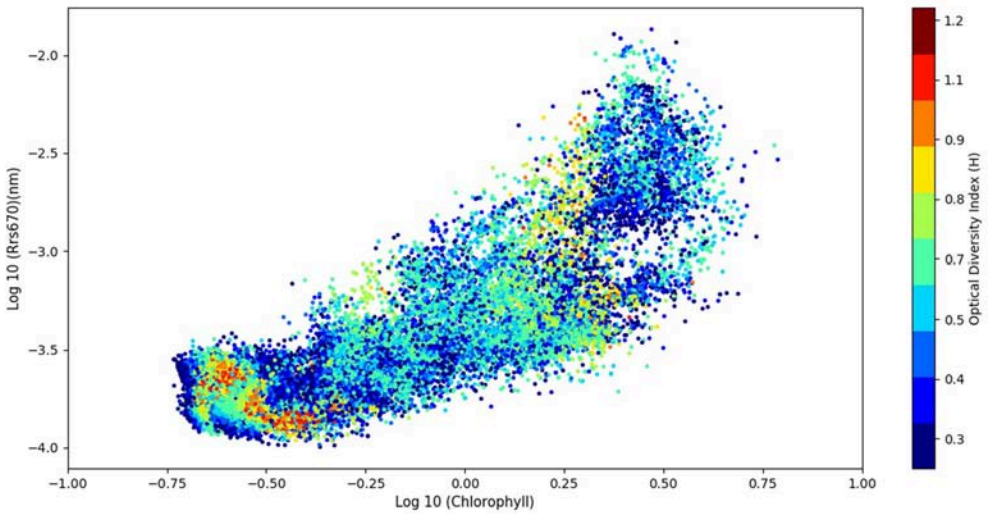


FIGURE 7 | Relationship between chlorophyll concentration and R_{rs}(670). Climatological data for February are shown as an example. Only well-classified pixels (cumulative membership >0.5) are plotted here. The optical diversity indices are identified using different colors.

accounted for, when comparing our results with those of Mélin and Vantrepotte (2015).

CONCLUDING REMARKS

We have implemented an optical classification using a log-transformed, normalized, remote-sensing reflectance (R_{rs}) datasets, with spatial resolution of 4 km. In this study, eight optical classes were obtained in the coastal waters of the northern Indian Ocean. Seasonally-reversing monsoons are a defining oceanographic characteristic of the Indian Ocean. Here, we have discussed variations in optical classes with reference to the southwest and northeast monsoon seasons of the study region. The distribution pattern of optical classes in the study region showed major variations between seasons. An example is the presence of optical classes 1, 2, 3, and 4 in the latitudes (0–18° N) during December–March, whereas they were not found in the months of June–September. The influence of class 5 in intermediate coastal waters is consistent in all the regions with fewer variations in each month. Class 6 is also a minimal contributor to the coastal waters of India, restricted the Persian Gulf in northeast monsoon season. These patterns show that in southwest monsoon season, the optical constituents of the coastline are affected mainly by precipitation and river water intrusion; this condition is not prevalent in northeast monsoon season. The regional distribution of dominant optical classes, and how they are related to physical and biological oceanographic features and processes, is presented in the Table S3.

We have also used the memberships of different optical classes in a given pixel, to study optical diversity within a pixel. Both the dominant optical class and the optical diversity index appear to be related to the chlorophyll concentration and the remote-sensing reflectance at 670 nm (used here as an index of suspended sediment load), but in quite different ways. Whereas, the dominant optical classes transition in a systematic matter from classes 1 to 8 with increasing concentration of chlorophyll and increasing R_{rs} 670, the diversity index appears to be high in areas of transition between optical classes. We also see that the diversity index was high in clear waters around coral islands and in deeper waters away from the shore. Since it is well-known that biological diversity tends to be high when chlorophyll concentration is high, these results suggest that optical diversity indices might run counter to biological diversity. This suggestion can only be verified when data on phytoplankton diversity in the study area become available on a systematic, and extensive basis. But once such relationships are established, optical diversity and optical classification would pave the way for mapping biological diversity at large scales, using remote sensing.

We opted for the Ocean Color Climate Change Initiative products for the study, because of the long time series of data available, which would facilitate extension of our work to study trends and inter-annual variability, and also because of the better spatial coverage, especially during the monsoon season. However, the dataset is limited to six SeaWiFS wavebands in the visible, which was dictated by the historical sensor capabilities. The number of wavebands available also determined the extent to which optical diversity could be explored. No doubt as better-resolution data become available over long time scales from missions such as Sentinel 3, which carries the Ocean and Land Color Instrument (OLCI) sensor with 10 wavebands in the visible domain, it would become possible to investigate optical classes and optical diversity with higher spectral resolution, which may reveal additional optical classes which were not captured in the present analysis.

The optical classification presented in this work enables us to study the seasonal dynamics in the bio-optical characteristics of the coastal waters of the Northern Indian Ocean, and

how they are related to the physical and biological processes. Spatio-temporal variations of the eight optical classes under the influence of seasonally reversing monsoons were profound. This study will aid as a first step for investigations of the inter-annual variations in distribution of optical classes and their shifts in response to changing climatic conditions such as *El Niño* and *La Niña* events.

AUTHOR CONTRIBUTIONS

SM: analyzed the datasets, ran the algorithms, and worked on interpretation of the results; TP and SS: implemented and led the study; JJ: assisted with statistical measures used in the study; GG: contributed to the ideas of physical and biological oceanography of the study region; TJ: provided the computational codes for the work. All the authors contributed to the text.

ACKNOWLEDGMENTS

The authors thank the Department of Science and Technology, Science and Engineering Research Board, India for the Jawaharlal Nehru Science Fellowship to TP. The Director, Central Marine Fisheries Research Institute, India is acknowledged for the support and facilities to pursue this work. The study also contributes to the Ocean Color Climate change Initiative (OC-CCI) of the European Space Agency. This is also a contribution to the activities of the National Centre for Earth Observations of the Natural Environment Research Council of the United Kingdom. We thank Dr. Tim Moore, Research Assistant Professor, University of New Hampshire, for providing the classification algorithm. We also thank Dr. Mike Grant, Plymouth Marine Laboratory, United Kingdom, Mr. Devi Varaprasad and Mr. Abhinav, National Remote Sensing Center, India, Mr. Tarun Joseph, Cochin University of Science and Technology, India and Mr. Muhammad Shafeeque, Central Marine Fisheries Research Institute, India for their extended support in programming and computation.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2018.00087/full#supplementary-material

REFERENCES

Antony, M. K., Swamy, G. N., and Somayajulu, Y. K. (2002). Offshore limit of coastal ocean variability identified from hydrography and altimeter data in the eastern Arabian Sea. *Cont. Shelf Res.* 22, 2525–2536. doi: 10.1016/S0278-4343(02)00100-0

Babin, M., Stramski, D., Ferrari, M. G., Claustre, H., Bricaud, A., Obolensky, G., et al. (2003). Variations in the light absorption coefficients of phytoplankton, nonalgal particles, and dissolved organic matter in coastal waters around Europe. *J. Geophys. Res.* 108:C7. doi: 10.1029/2001JC000882

Bezdek, J. C. (1973). Cluster validity with fuzzy sets. *J. Cybern.* 3, 58–73.

Bezdek, J. C., Ehrlich, R., and Full, W. (1984). FCM: the fuzzy c-means clustering algorithm. *Comput. Geosci.* 10, 191–203.

Brewin, R. J. W., Sathyendranath, S., Müller, D., Brockmann, C., Deschamps, P.-Y., Devred, E., et al. (2015). The ocean colour climate change initiative: III. A round-robin comparison on in-water bio-optical algorithms. *Remote Sens. Environ.* 162, 271–294. doi: 10.1016/j.rse.2013.09.016

Deborah, J. L., Baskaran, R., and Kannan, A. (2010). A survey on internal validity measure for cluster validation. *Int. J. Comput. Sci. Eng. Surv.* 1, 85–102. doi: 10.5121/ijcses.2010.1207

Dowell, M., and Platt, T. (2009). *Partition of the Ocean into Ecological Provinces: Role of Ocean-Colour Radiometry*. International Ocean-Colour Coordinating Group.

Garaba, P. S., and Zielinski, O. (2013). Comparison of remote sensing reflectance from above-water and in-water measurements west of Greenland, Labrador Sea, Denmark Strait, and west of Iceland. *Opt. Express* 21, 15938–15950. doi: 10.1364/OE.21.015938

Guzman, J. B., Claustre, H., and Marty, J. C. (1995). Specific phytoplankton signatures and their relationship to hydrographic conditions in the coastal northwestern Mediterranean Sea. *Mar. Ecol. Prog. Ser.* 124, 247–258.

Halkidi, M., Batistakis, Y., and Vazirgiannis, M. (2001). On cluster validation techniques. *J. Intell. Inform. Syst.* 17, 107–145.

Hunter, J. R. (1983). *Aspects of the Dynamics of the Residual Circulation of the Arabian Gulf*. Anglesey: Plenum Press.

Jeffries, C. T., Ostrowski, M., Williams, B. R., Xie, C., Jensen, R. M., Grzymski, J. J., et al. (2015). Spatially extensive microbial biogeography of the Indian Ocean provides insights into the unique community structure of a pristine coral atoll. *Sci. Rep.* 5:15383. doi: 10.1038/srep15383

Karl, D. M., Leteller, R., Hebel, D., Tupas, L., Dore, J., Christian, J., and Winn, C. (1995). Ecosystem changes in the Pacific subtropical gyre attributed to the 1991-1992 *El Niño*. *Nature* 373, 230–234.

Krishnamurthy, K., Naidu, W. D., and Santhanam, R. (1978). Plankton biogeography of the Indian Ocean. *Int. Rev. Gesamten Hydrobiol.* 63, 721–726.

Le, C., Hu, C., Cannizzaro, J., English, D., Muller-Karger, F., and Lee, Z. (2013). Evaluation of chlorophyll-a remote sensing algorithms for an optically complex estuary. *Remote Sens. Environ.* 129, 75–89. doi: 10.1016/j.rse.2012.11.001

Loisel, H., Vantrepotte, V., Ouillon, S., Ngoc, D. D., Herrmann, M., Tran, V., et al. (2017). Assessment and analysis of the chlorophyll-a concentration variability over the Vietnamese coastal waters from the MERIS ocean color sensor (2002–2012). *Remote Sens. Environ.* 190, 217–232. doi: 10.1016/j.rse.2016.12.016

Longhurst, A. R. (1998). *Ecological Geography of the Sea*. London: Academic Press.

Mélin, F., and Vantrepotte, V. (2015). How optically diverse is the coastal ocean?. *Remote Sens. Environ.* 160, 235–251. doi: 10.1016/j.rse.2015.01.023

Moore, T. S., Campbell, J. W., and Dowell, M. D. (2009). A class-based approach to characterizing and mapping the uncertainty of the MODIS ocean chlorophyll product. *Remote Sens. Environ.* 113, 2424–2430. doi: 10.1016/j.rse.2009.07.016

Moore, T. S., Campbell, J. W., and Feng, H. (2001). A fuzzy logic classification scheme for selecting and blending satellite ocean color algorithms. *Remote Sens. Environ.* 39, 1764–1776. doi: 10.1109/36.942555

Morel, A., and Prieur, L. (1977). Analysis of variations in ocean color. *Limnol. Oceanogr.* 22, 709–722.

Obura, O. D. (2012). The diversity and biogeography of Western Indian Ocean Reef-Building Corals. *PLoS ONE* 9:e45013. doi: 10.1371/journal.pone.0045013

Obura, O. D. (2016). An Indian Ocean centre of origin revisited: Palaeogene and Neogene influences defining a biogeographic realm. *J. Biogeogr.* 43, 229–242. doi: 10.1111/jbi.12656

O'Reilly, E. J., Maritorena, S., Mitchell, G. B., Siegel, A. D., Carder, K. L., Garver, S. A., et al. (1998). Ocean color chlorophyll algorithms for SeaWiFS. *J. Geophys. Res.* 103, 937–953.

Platt, T., and Sathyendranath, S. (1999). Spatial structure of pelagic ecosystem processes in the global ocean. *Ecosystems* 2, 384–394.

Platt, T., and Sathyendranath, S. (2008). Ecological indicators for the pelagic zone of the ocean from remote sensing. *Remote Sens. Environ.* 112, 3426–3436. doi: 10.1016/j.rse.2007.10.016

Prieur, L., and Sathyendranath, S. (1981). An optical classification of coastal and oceanic waters based on the specific spectral absorption curves of phytoplankton pigments, dissolved organic matter, and other particulate materials. *Limnol. Oceanogr.* 26, 671–689.

Qasim, S. Z. (1999). Some unique characteristics of the Indian Ocean. *Qatar Univ. Sci. J.* 6, 111–116.

Reygondeau, G., Longhurst, A., Martinez, E., Beaugrand, G., Antoine, D., and Maury, O. (2013). Dynamic biogeochemical provinces in the global ocean. *Global Biogeochem. Cycles* 27, 1–13. doi: 10.1002/gbc.20089

Rezaee, B. (2010). A cluster validity index for fuzzy clustering. *Fuzzy Sets Syst.* 161, 3014–3025.

Roxy, M. K., Modi, A., Murtugudde, R., Valsala, V., Panickas, S. S., Kumar, P., et al. (2016). A reduction in marine primary productivity driven by rapid warming over the tropical Indian Ocean. *Geophs. Res. Lett.* 43, 826–833. doi: 10.1002/2015GL066979

Sathyendranath, S., Brewin, R. J. W., Jackson, T., Melin, F., and Platt, T. (2017). Ocean-colour products for climate-change studies: what are their ideal characteristics. *Remote Sens Environ.* 203, 125–138. doi: 10.1016/j.rse.2017.04.017

Sathyendranath, S., Groom, S., Grant, M., Brewin, R. J. W., Thompson, A., et al. (2016). *ESA Ocean Colour Climate Change Initiative (Ocean-Colour-CCI): Version 2.0 Data.* Centre for Environmental Data Analysis.

Schills, T., and Wilson, S. (2006). Temperature threshold as a biogeographic barrier in northern Indian Ocean macroalgae. *J. Phycol.* 42, 749–756. doi: 10.1111/j.1529-8817.2006.00242.x

Schott, F. A., and McCreary, J. P. (2001). The monsoon circulation of the Indian Ocean. *Progr. Oceanogr.* 51, 1–123. doi: 10.1016/S0079-6611(01)00083-0

Schott, F. A., Xie, S. P., and McCreary Jr., J. P. (2009). Indian Ocean circulation and climate variability. *Rev. Geophys.* 47, 1–46. doi: 10.1029/2007RG000245

Shannon, C. E. (2001). *A Mathematical Theory of Communication*. New York, NY: ACM SIGMOBILE Mobile Computing and Communications.

Shetye, S., and Shenoi, S. S. C. (1988). Seasonal cycle of surface circulation in the coastal north indian ocean. *Proc. Indian Acad. Sci.* 97, 53–62.

Shetye, S. R. (1998). West India coastal current and Lakshadweep high/low. *Indian Acad. Sci. Limnol. Oceanogr.* 23, 637–651.

Tilstone, G., Angel-Benavides, I. M., Pradhan, Y., Shutler, J. D., Groom, S., and Sathyendranath, S. (2011). An assessment of chlorophyll-a algorithms available for SeaWiFS in coastal and open areas of the Bay of Bengal and Arabian Sea. *Remote Sens. Environ.* 115, 2277–2291. doi: 10.1016/j.rse.2011.04.028

Tilstone, G., Mallor-Hoya, S., Gohin, F., Couto, A. B., Sá, C., Goela, P., et al. (2017). Which ocean colour algorithm for MERIS in NorthWest European waters? *Remote Sens. Environ.* 189, 132–151. doi: 10.1016/j.rse.2016.11.012

Tomczak, M., and Godfrey, J. S. (2001). *Regional Oceanography: An Introduction*. Oxford, UK: Elsevier Science Ltd.

Vantrepotte, V., Loisel, H., Dessailly, D., and Meriaux, X. (2012). Optical classification of contrasted coastal waters. *Remote Sens. Environ.* 123, 306–323. doi: 10.1016/j.rse.2012.03.004

Vinayachandran, P. N., and Shetye, S. R. (1991). The warm pool in the Indian Ocean. *Proc. Indian Acad. Sci.* 100, 165–175.

Xie, X. L., and Beni, G. (1991). A validity measure for fuzzy clustering. *Pattern Anal. Mach. Intell.* 13, 841–847.

Zadeh, L. (1965). Fuzzy sets. *Inf. Control* 8, 338–353.

Zhao, Q., Xu, M., and Franti, P. (2009). Sum-of-squares based cluster validity index and significance analysis. *Pattern Anal. Mach. Intell.* 5495, 313–322.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Monolisha, Platt, Sathyendranath, Jayasankar, George and Jackson. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: info@frontiersin.org | +41 21 510 17 00

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership