frontiers | Research Topics

Advances in ocean data assimilation : Methodologies, forecasting and reanalysis

Edited by
Shiqiu Peng, Yasumasa Miyazawa, Youmin Tang,
Jinyu Sheng and Zhijin Li

Published in
Frontiers in Marine Science

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 978-2-8325-4060-2
DOI 10.3389/978-2-8325-4060-2

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: frontiersin.org/about/contact

Advances in ocean data assimilation : Methodologies, forecasting and reanalysis

Topic editors

Shiqiu Peng — State Key Laboratory of Tropical Oceanography, South China Sea Institute of Oceanology, Chinese Academy of Sciences, China

Yasumasa Miyazawa — Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Japan

Youmin Tang — University of Northern British Columbia Canada, Canada

Jinyu Sheng — Dalhousie University, Canada

Zhijin Li — JEFRESSE, UCLA, United States

Citation

Peng, S., Miyazawa, Y., Tang, Y., Sheng, J., Li, Z., eds. (2023). *Advances in ocean data assimilation : Methodologies, forecasting and reanalysis*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-4060-2

Table of contents

05 Rapid reconstruction of temperature and salinity fields based on machine learning and the assimilation application
Zhihui Chen, Pinqiang Wang, Senliang Bao and Weimin Zhang

21 Advancing parameter estimation with Characteristic Finite Difference Method (CFDM) for a marine ecosystem model by assimilating satellite observations: Spatial distributions
Minjie Xu, Yongzhi Liu, Zihan Zhao, Kai Fu and Xianqing Lv

39 Long-term mean circulation in the Japan Sea as reproduced by multiple eddy-resolving ocean circulation models
Haodi Wang, Kaijun Ren, Ma Lina, Chen Yu, Wen Chen and Wen Zhang

64 Variational parameter estimation in a two-equation turbulence model: A case study with a 3D primitive-equation ocean model
Yan Hu, Xuefeng Zhang and Wei Li

81 A vortex-implanted initialization scheme for the mesoscale eddy prediction: Idealized experiments
Yuhang Zhu, Shiqiu Peng and Yineng Li

96 Application of interpolation methodology with dynamical constraint to the suspended particulate matter in the Liaodong Bay
Haoyi Geng, Haohong Peng, Jie Shi, Xinyan Mao and Xianqing Lv

114 An operational improvement of A-4DEnVar and its application to the estimation of the spatially varying bottom friction coefficients of the M_2 constituent in the Bohai and Yellow seas
Kangzhuang Liang, Wei Li, Guijun Han, Xuan Wang and Xiaobin Qiu

129 Monthly extended ocean predictions based on a convolutional neural network *via* the transfer learning method
Yonglan Miao, Xuefeng Zhang, Yunbo Li, Lianxin Zhang and Dianjun Zhang

142 On the control of spatial and temporal oceanic scales by existing and future observing systems: An observing system simulation experiment approach
Florent Gasparin, Jean-Michael Lellouche, Sophie E. Cravatte, Giovanni Ruggiero, B. Rohith, Pierre Yves Le Traon and Elisabeth Rémy

156 Effects of spatial bottom friction parameterization scheme on the tidal dynamics in the macrotidal East China Seas
Suhui Qian, Yunfei Du, Zilu Wei, Jicai Zhang, Jun Cheng, Daosheng Wang and Ya Ping Wang

172 USFP: An unbalanced severe typhoon formation prediction framework based on transfer learning
Xiaotian Pan, Xiang Wang, Chengwu Zhao, Jianping Wu, Huizan Wang, Senzhang Wang and Sihao Chen

188 A hybrid deep learning model for predicting the Kuroshio path south of Japan
Xiaobo Wu, Guijun Han, Wei Li, Zenghua Ji, Lige Cao and Wanqiu Dong

198 Evaluation of the sensitivity on mesoscale eddy associated with the sea surface height anomaly forecasting in the Kuroshio Extension
Lin Jiang, Wansuo Duan, Hui Wang, Hailong Liu and Lingjiang Tao

208 Fusion of ocean data from multiple sources using deep learning: Utilizing sea temperature as an example
Mingqing Wang, Danni Wang, Yanfei Xiang, Yishuang Liang, Ruixue Xia, Jinkun Yang, Fanghua Xu and Xiaomeng Huang

223 Impact of assimilating repeated subsurface temperature transects on state estimates of a western boundary current
David E. Gwyther, Moninya Roughan, Colette Kerry and Shane R. Keating

239 Impact of satellite and regional *in-situ* profile data assimilation on a high-resolution ocean prediction system in the Northwest Pacific
Inseong Chang, Young Ho Kim, Hyunkeun Jin, Young-Gyu Park, Gyundo Pak and You-Soon Chang

253 Ensemble-based data assimilation for predictable zones and application for non-linear deep-water waves
Wataru Fujimoto and Kinya Ishibashi

265 An assessment of the LICOM Forecast System under the IVTT class4 framework
Weipeng Zheng, Pengfei Lin, Hailong Liu, Yihua Luan, Jinfeng Ma, Huier Mo and Juan Liu

275 Global ocean reanalysis CORA2 and its inter comparison with a set of other reanalysis products
Hongli Fu, Bo Dan, Zhigang Gao, Xinrong Wu, Guofang Chao, Lianxin Zhang, Yinquan Zhang, Kexiu Liu, Xiaoshuang Zhang and Wei Li

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 26 September 2022
DOI 10.3389/fmars.2022.985048

Check for updates

OPEN ACCESS

EDITED BY
Zhijin Li,
Jefresse, UCLA, United States

REVIEWED BY
Liying Wan,
National Marine Environmental Forecasting Center, China
Brian Powell,
University of Hawaii, United States

*CORRESPONDENCE
Weimin Zhang
weiminzhang@nudt.edu.cn

†These authors have contributed equally to this work and share first authorship

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

RECEIVED 03 July 2022
ACCEPTED 08 September 2022
PUBLISHED 26 September 2022

CITATION
Chen Z, Wang P, Bao S and Zhang W (2022) Rapid reconstruction of temperature and salinity fields based on machine learning and the assimilation application.
Front. Mar. Sci. 9:985048.
doi: 10.3389/fmars.2022.985048

COPYRIGHT
© 2022 Chen, Wang, Bao and Zhang. 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(s) 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.

Rapid reconstruction of temperature and salinity fields based on machine learning and the assimilation application

Zhihui Chen†, Pinqiang Wang†, Senliang Bao and Weimin Zhang*

College of Meteorology and Oceanography, National University of Defense Technology, Changsha, China

Satellite observations play important roles in ocean operational forecasting systems, however, the direct assimilation of satellite observations cannot provide sufficient constraints on the model underwater structure. This study adopted the indirect assimilation method. First, we created a 3D temperature and salinity reconstruction model that took into account the advantage of the nonlinear regression of the generalized regression neural network with the fruit fly optimization (abbreviated as FOAGRNN). Compared with the reanalysis product and the WOA13 climatology data, the synthetic T/S (temperature and salinity) profiles had sufficient accuracy and could better describe the characteristics of mesoscale eddies. Then, the synthetic T/S profiles were assimilated into the Regional Ocean Model System (ROMS) using the Incremental Strong constraint 4D Variational (I4D-Var) data assimilation algorithm. The quantitative and qualitative analysis results indicated that compared with the direct assimilation of satellite observations, the root mean square errors (RMSEs) of temperature and salinity were reduced by 26.0% and 23.1% respectively by assimilating the synthetic T/S profiles. Furthermore, this method significantly improved the simulation effect of the model underwater structure, especially in the 300 m to 500 m water layer. Compared with the National Marine Data Center's real-time analysis data, the machine learning-based assimilation system demonstrated a significant advantage in the simulation of underwater salinity structure, while showing a similar performance in the simulation of underwater temperature structure.

KEYWORDS

machine learning, satellite observations, three-dimensional T-S fields, 4DVAR assimilation, mesoscale eddy

1 Introduction

By combining observations and numerical models, data assimilation not only can make up for the temporal discontinuity and spatial inhomogeneity of observations but also can improve the accuracy of numerical models. However, the effect of data assimilation depends on the quality of the observations and their spatiotemporal distribution. With the development of satellite technology, satellite data have experienced incredible growth (Ratheesh et al., 2012). Sea surface temperature (SST) and sea level anomaly (SLA) have become the indispensable data in operational systems, and the assimilation of sea surface salinity (SSS) can also enhance the forecasting effect on El Niño/Southern Oscillation (ENSO) (Tranchant et al., 2019). Compared with ship surveys or buoys, satellite can provide large-scale observations of the ocean surface with better time continuity. However, satellite cannot directly observe the subsurface, and the adjustment of the underwater structure of numerical models still relies on the observations of T/S profiles by instruments such as Argo buoys. However, the *in-situ* observations like Argo buoys are scarce. For example, only one Argo buoy was active in the South China Sea on May 18, 2018.

How to use satellite data to constrain the subsurface temperature and salinity structure is a complex problem in ocean data assimilation. One method is statistical, which establishes statistical relationships between surface and subsurface seawater states, including multiple linear regression and Empirical Orthogonal Functions (EOFs). Carnes et al. (1994) determined the EOF magnitude of the vertical temperature structure as a function of SST and sea surface height (SSH) by performing multivariate least squares regression on more than 33,000 historical T/S profiles in the Northwest Pacific and Northwest Atlantic Oceans. In this way, they projected the observation information of SST and SSH underwater. Based on the work of Carnes et al. (1994); Fox et al. (2002) developed the Modular Ocean Data Assimilation System (MODAS) and applied it to the US Navy Coupled Ocean Data Assimilation (NCODA) system. To reduce significant errors near the thermocline layer, the US Navy established an Improved Synthetic Ocean Profiles (ISOP) system. The system divided the ocean into mixed layers, thermocline, and quiescent layers. The vertical mapping model of the T/S profiles was established respectively by using multiple regression, one-dimensional variational data assimilation, and linear regression at three levels (Helber et al., 2013). Following the practice of the US Navy, China's National Marine Data Center successfully established MODAS and ISOP systems and used them to produce real-time analysis data. This approach proposed by Carnes et al. (1994) is not suitable for areas where observations are scarce. To address this problem, Ezer and Mellor (1994) utilized model output to establish the statistical relationship between the sea surface height anomaly and the subsurface temperature and salinity anomaly. Then they assimilated the data from satellite tracks to improve the temperature structure of the 500 m water layer. The simulation effect of this method is dependent on the performance of the model. In the ensemble method, the ensemble samples can also be employed to consider the relationship between the surface and the subsurface layers, and then the observation information can be passed down (Chen et al., 2018; Zhou et al., 2021).

Different from the statistical method, the dynamic method uses dynamic constraints to transfer the sea surface information downward. Common dynamic methods include the nudging approach (Holland and Malanotte-Rizzoli, 1989; Chen et al., 2020) and the dynamic conservation technique (Haines, 1991; Cooper and Haines, 1996; Weaver et al., 2005). In the Nudging approach, a nudging term is added to the right side of the dynamic equation, and the assimilated sea surface observation information is transferred to the deep layer only through the model dynamic framework. However, this approach may induce disturbances during the adjustment phase of the model, resulting in poor simulation of the subsurface structure (Cooper and Haines, 1996). In the dynamic conservation technique, the adjustment is performed based on the conservation properties or balance relationships of the ocean, such as potential vorticity conservation, geostrophic equilibrium, static equilibrium, etc. (Weaver et al., 2005; Liu et al., 2019; Liu et al., 2021). With the further development of the assimilation methods, the variational assimilation method was proposed. This method decomposes the assimilation increment into a balance term and a non-balance term. In the balance term, the balanced relationship is utilized to establish a multivariate balance operator to indirectly adjust the observed variables (Moore et al., 2011a; Moore et al., 2011b; Cummings and Smedstad, 2013). Dynamic conservation often adopts a simple dynamic framework, which is mainly suitable for adjusting the large-scale or small-scale components that satisfy the conservation relationship but is often not applicable to near shore.

In recent years, machine learning methods have been increasingly employed in meteorological and oceanographic applications, including bias correction of satellite observations (Vernieres et al., 2014; Le et al., 2020; Iqbal et al., 2022) and subsurface temperature and salinity reconstruction (Su et al., 2015; Chapman and Charantonis, 2017; Su et al., 2018; Bao et al., 2019). Compared with the multiple linear regression method used in MODAS, the machine learning method is more efficient in reconstruction and has a strong nonlinear regression capability and a high degree of fault tolerance and robustness. Bao et al. (2019) utilized the generalized regression neural network with fruit fly optimization (abbreviated as FOAGRNN) to reconstruct the salinity profiles based on sea surface data, and the reconstruction effect was better than that in multiple linear regression at the strong thermocline layer. To

address the shortcomings of the direct assimilation for insufficient constraints on the model underwater structure, a machine learning-based assimilation system was established in this study. We first reconstructed the satellite observations into underwater 3D T/S pseudo-profiles using the FOAGRNN algorithm. Then we assimilated the synthetic profiles into the ROMS using the I4D-Var data assimilation algorithm. To validate the accuracy and effectiveness of this system, the system results were compared with the experiment results of the direct assimilation and the real-time analysis data of the National Marine Data Center.

The article is organized in the following sections: Section 2 describes the data sources, machine learning algorithms, assimilation methods, and model setup. Section 3 evaluates the accuracy of the FOAGRNN reconstruction model. Section 4 conducts three groups of assimilation experiments to examine the simulation effect of the machine learning-based assimilation system on the model underwater structure. Finally, a summary and a discussion of our research findings are given in Section 5.

2 Details of the machine learning-based assimilation system

Compared with the conventional method of the direct assimilation of satellite data, we created a machine learning-based assimilation system (Figure 1), which was divided into three main steps: in the first step, historical satellite altimeter data and historical T/S profiles were collected, and then the reconstruction model was obtained by training with the FOAGRNN algorithm proposed by Bao et al. (2019); in the second step, the assimilation time period was selected, and the real-time satellite observations were employed as the input field of the reconstruction model to construct the T/S profiles for each day; finally, the synthetic T/S profiles were assimilated into the model using the I4D-Var method. The data source, reconstruction methods, and assimilation system configuration are described below.

2.1 Data

The ocean observations used in this study included SLA, SST, SSS, and *in situ* observations. The satellite SLA data were delayed time and gridded maps of sea level anomaly (MSLA) from Copernicus Marine Environment Monitoring Service (CMEMS) with a horizontal resolution of 0.25°; the satellite SST data were acquired from the gridded product released by United Kingdom Meteorological Office (UKMO) with a horizontal resolution of 1/20° and were interpolated to SLA gridded points to maintain a consistent horizontal resolution (Good et al., 2020); the satellite SSS data were obtained from the Soil Moisture Active Passion (SMAP) with the same spatial and temporal resolutions as the SST and SLA data (He et al., 2021). All of the above satellite products had a temporal resolution of one day. *In situ* observations included the EN4.2.1 T/S profile

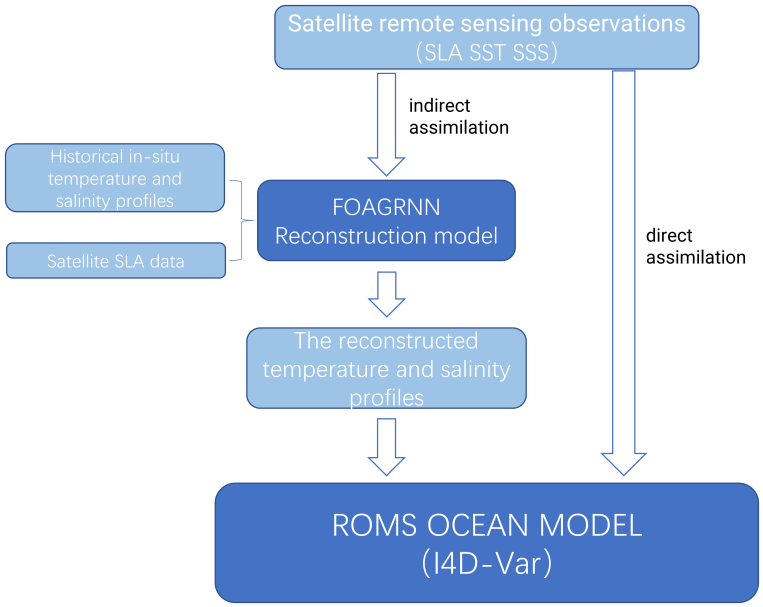


FIGURE 1
The flow chart of the machine learning-based assimilation system.

datasets from the Hadley Center (Good et al., 2013), and survey data from the Northwest Pacific. The EN4.2.1 T/S profile datasets were divided into 24 layers at different depths, namely 2, 5.01, 15.07, 25.28, 35.7, 46.61, 57.98, 70.02, 82.92, 96.92, 112.32, 129.49, 148.96, 171.40, 197.79, 229.48, 268.46, 317.65, 381.39, 465.91, 579.31, 729.35, 918.37, and 1139.15 m. To ensure the accuracy of the reconstruction model, the historical satellite altimeter data and EN4.2.1 historical T/S profiles from 2004 to 2018 were used as the training data of the reconstruction model, and the corresponding data in 2019 were utilized as the test data. Each set of assimilation experiments corresponds to a reconstructed model. The numbers of the *in-situ* EN4.2.1 T/S profiles for training data and the test data were shown in Table 1, respectively. Satellite data (SLA, SST, SSS) from the assimilation period were employed as input fields for the reconstruction model to construct real-time 3D T/S pseudo-profiles. We also utilized WOA13 climatology data, SODA3.4.2 reanalysis product, GREP (Global Reanalysis multi-model Ensemble Product) and real-time analysis data of the National Marine Data Center (hereinafter referred to as MODAS) to evaluate the reconstruction model accuracy and assimilation effects (Carton et al., 2018; Storto et al., 2019). Survey data from the Northwest Pacific were used as independent observations for the qualitative analysis of the assimilation effect.

2.2 Method

2.2.1 FOAGRNN

The Generalized Regression Neural Network (GRNN) is characterized by strong nonlinear mapping abilities, a flexible network structure, a high fault tolerance, and robustness. The theoretical basis of the GRNN is the nonlinear regression analysis. To maximize the effectiveness of GRNN, the key is the selection of smoothing parameters (Li et al., 2013). The fruit fly optimization algorithm is a new probabilistic method intended to find a global optimum based on the fruit fly's foraging behaviors. By setting the cost function for iterative optimization, the error between the output value and the actual value is gradually reduced, so as to determine the optimal parameters. Many researchers have used the FOA to optimize the parameters of artificial neural network models (Lin, 2013). Taking GRNN as the research framework and using the FOA algorithm to determine the optimal smoothing parameters, Bao et al. (2019) proposed the FOAGRNN, which was employed to create the reconstruction model in our study. Taking the reconstructed model of ExpBs as example, the specific flow chart is shown in Figure 2 and the steps are as follows:

Step 1: Sample data preprocessing. The sample data included satellite altimeter data and *in-situ* EN4.2.1 T/S profiles. The spatial range of the selected sample data was [19°N-27°N, 122°E-130°E], and the time span was from 2004 to 2019. The sample data from 2004 to 2018 was used as the training set of the reconstruction model, and the sample data in 2019 was utilized as the test set. The data input to the FOAGRNN model included sea surface data (EN4_SST, EN4_SSS, EN4_SLA) and location data (EN4_Lon, EN4_Lat). The uppermost temperature and salinity values of the *in-situ* EN4.2.1 T/S profiles were employed as the input value of EN4_SST and EN4_SSS, and EN4_SLA was obtained by interpolating satellite altimeter data to the location (EN4_Lon, EN4_Lat). Before training the reconstruction model, the input data was first normalized in the range of 0-1, and the output data were subsurface temperature or salinity values.

Step 2: Training of the FOAGRNN model. The initial smoothing parameter value of the FOAGRNN model was set in the range of [0.001, 1], and was dynamically adjusted by the FOA algorithm during the model training process. The smoothing parameter value was adjusted to the optimal value through the minimization of the cost function.

Step 3: Evaluation of the reconstruction model. The EN4_SST, EN4_SSS, EN4_SLA, and location data (EN4_Lon, EN4_Lat) of the 2019 sample data were used as the input field of the reconstruction model to construct three-dimensional pseudo-profiles. Then the EN4.2.1 T/S profiles in 2019 were utilized as the validation data to evaluate the accuracy of the reconstruction model.

Step 4: Real-time reconstruction based on satellite data. Based on the reconstruction model obtained from step 2, the satellite data (SLA, SST, SSS) and their location information (Lon, Lat) from October 2020 to November 2020 were employed as input data to reconstruct the T/S profiles in real-time. The vertical stratification was the same as the EN4.2.1 T/S profiles.

2.2.2 I4D-Var

Diagnostic variables in the ROMS model include potential temperature (T), salinity (S), horizontal velocity (u, v), and sea surface displacement (ζ). The state vector which is discretized onto the model grid at time t_i can be written as $\mathbf{x}(t_i) = (\mathbf{T,S,\zeta,u,v})$, which is integrated forward by the discretized nonlinear model

TABLE 1 The number of profiles being used for training and validation and the measurement error in the assimilation experiments.

Case name	Area	Training	Testing	T_MRMSE/°C	S_MRMSE/psu
ExpAs	[10°N-30°N, 124°E-160°E]	62723	4573	0.76	0.10
ExpBs	[19°N-27°N, 122°E-130°E]	6607	443	0.65	0.07
ExpCs	[23°N-28°N, 135°E-140°E]	4508	372	0.68	0.07

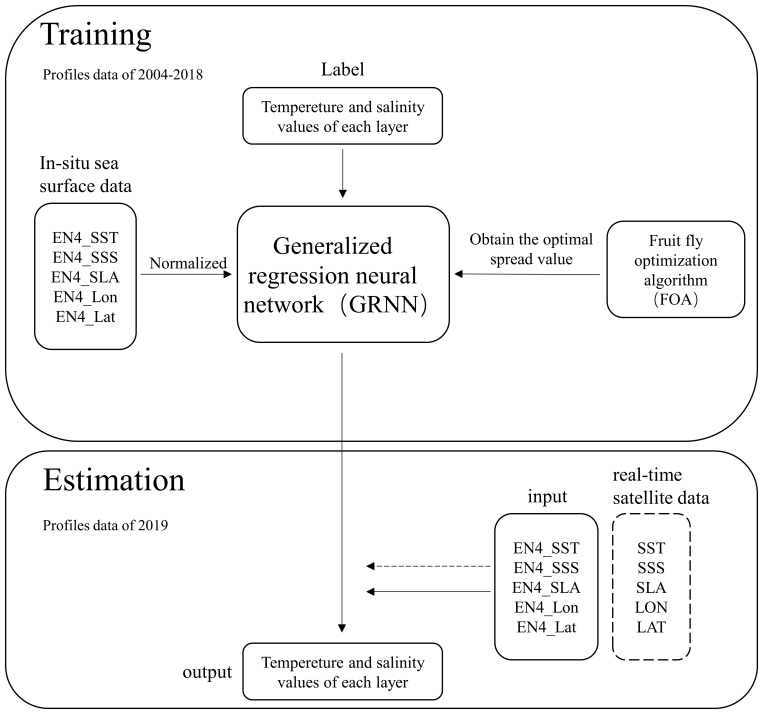


FIGURE 2
The flow chart of the reconstruction of 3D T/S profiles.

under the constraints of boundary conditions $\mathbf{b}(t_i)$ and forcing conditions $\mathbf{f}(t_i)$, and the integration process is expressed as:

$$\mathbf{x}(t_i) = \mathbf{M}(t_i, t_{i-1})(\mathbf{x}(t_{i-1}), \mathbf{f}(t_i), \mathbf{b}(t_i)) \tag{4}$$

The four-dimensional variational method used in this study is the I4D-Var of the original equation. I4D-Var aims to find the optimal estimate in the model space (Zhang et al., 2010; Moore et al., 2011b; Chen et al., 2014). The objective function of I4D-Var can be written as:

$$\begin{aligned} &J(\delta\mathbf{x}(t_0), \delta\mathbf{f}(t_1), \cdots, \delta\mathbf{f}(t_k), \delta\mathbf{b}(t_1), \cdots, \delta\mathbf{b}(t_k), \delta\boldsymbol{\eta}(t_1), \cdots, \delta\boldsymbol{\eta}(t_k)) \\ &= \tfrac{1}{2}\delta\mathbf{x}^T(t_0)\mathbf{B}_x^{-1}\delta\mathbf{x}(t_0) + \tfrac{1}{2}\sum_{k=1}^{N}\sum_{j=1}^{N}\{\delta\mathbf{f}^T(t_k)\mathbf{B}_f^{-1}(t_k, t_j)\delta\mathbf{f}(t_j) \\ &+ \delta\mathbf{b}^T(t_k)\mathbf{B}_b^{-1}(t_k, t_j)\delta\mathbf{b}(t_j) + \delta\boldsymbol{\eta}^T(t_k)\mathbf{Q}^{-1}(t_k, t_j)\delta\boldsymbol{\eta}(t_j)\} \\ &+ \frac{1}{2}\sum_{i=1}^{N}\sum_{j=0}^{N}\left((\mathbf{H}_i\delta\mathbf{x}(t_i) - \mathbf{d}_i)^T\mathbf{R}_{i,j}^{-1}\left(\mathbf{H}_i\delta\mathbf{x}(t_j) - \mathbf{d}_j\right)\right) \end{aligned} \tag{5}$$

where $\delta\boldsymbol{x}(t_k)$ represents the model increment, which is expressed as $\delta\mathbf{x}(t_k)$=$\mathbf{x}(t_k)$-$\mathbf{x}^b$ (t_k); $\mathbf{d}_i$ denotes the observation increment, which can be written as $\mathbf{d}_i = \mathbf{y}_i^0 - \mathrm{H}_i(\mathbf{x}^b(t_i))$; $\mathbf{H}_i$stands for the tangent linear operator of $\boldsymbol{H}_i$, and $\boldsymbol{H}_i$signifies the observation operator; $\mathbf{y}_i^0$symbolizes the observation at the moment t_i; $\mathbf{B}_x$, $\mathbf{B}_b$, $\mathbf{B}_f$, $\mathbf{Q}$, and $\mathbf{R}$ indicate the initial field, the boundary field, the forcing field, and the model and observation error covariance matrices, respectively. To simplify the objective function, the transformation is expressed as follows:

$$\mathbf{z}_a = \mathbf{z}_b + \delta\mathbf{z}^a \tag{6}$$

$$\delta\mathbf{z} = \left(\delta\mathbf{x}^T(t_0), \delta\mathbf{f}^T(t_1), \cdots, \delta\mathbf{f}^T(t_k), \delta\mathbf{b}^T(t_1), \cdots, \delta\mathbf{b}^T(t_k), \delta\boldsymbol{\eta}^T(t_1), \cdots, \delta\boldsymbol{\eta}^T(t_k)\right)^T \tag{7}$$

where $\mathbf{z}$ is the control variable, $\mathbf{z}_a$ is the value obtained from the analysis field, $\mathbf{z}_b$ is the value obtained from the background field, $\boldsymbol{\delta z}$ is expressed as the increment of the control variable, and the difference conversion increment $\mathbf{H}_i\boldsymbol{\delta}\mathbf{x}(t_i)$ can be written as $\mathbf{H}_i$ $\mathbf{M}(t_i,t_0)$=$\mathbf{G}_i\boldsymbol{\delta z}$. By introducing vectors $\mathbf{d} = (\cdots, \mathbf{d}_i^T, \cdots)$, matrices $\mathbf{G} = (\cdots, \mathbf{G}_i^T, \cdots)$, diagonal matrices $\mathbf{R}$ (with diagonal elements $\mathbf{R}_i$), and diagonal matrices $\mathbf{D}$ (with $\mathbf{B}_x$, $\mathbf{B}_b$, $\mathbf{B}_f$, $\mathbf{Q}$ as diagonal elements) into the cost function, the Eq. (5) can be simplified to the following equation:

$$J(\delta\mathbf{z}) = \frac{1}{2}\delta\mathbf{z}^T\mathbf{D}^{-1}\delta\mathbf{z} + \frac{1}{2}(\mathbf{G}\delta\mathbf{z} - \mathbf{d})^T\mathbf{R}^{-1}(\mathbf{G}\delta\mathbf{z} - \mathbf{d}) \tag{8}$$

The solution of the equation $\partial J/\partial\mathbf{z}$ =0 is the required solution of $\boldsymbol{\delta z}^{\mathrm{a}}$:

$$\partial J/\partial\mathbf{z} = \mathbf{D}^{-1}\delta\mathbf{z} + \mathbf{G}^T\mathbf{R}^{-1}(\mathbf{G}\delta\mathbf{z} - \mathbf{d}) \tag{9}$$

$$\delta\mathbf{z}^a = \left(\mathbf{D}^{-1} + \mathbf{G}^T\mathbf{R}^{-1}\mathbf{G}\right)^{-1}\mathbf{G}^T\mathbf{R}^{-1}\mathbf{d} \tag{10}$$

2.3 Experiment configuration

The simulated region was located in the Northwest Pacific [-10°S-45°N, 99°E-165°E] with a horizontal resolution of 1/6° × 1/6° and was divided into 48 layers in the vertical direction. The bathymetry field was generated using ETOPO2 data with a minimum depth of 10 m and a maximum depth of 5500 m. The model was integrated from January 5, 2014 to December 31, 2020 (without any data assimilation). The open boundary conditions were obtained from the SODA3.4.2 five-day averaged reanalysis product, and atmospheric forcing fields were obtained from ECMWF ERA-interim datasets (including wind stress, heat flux, and freshwater flux). We followed three purposes for this real simulation: (1) to provide a mean surface height for SLA assimilation; (2) to derive statistics regarding the climatic background error standard deviation; and (3) to provide a dynamically balanced initial condition for the subsequent assimilation experiments (Wang et al., 2021).

To evaluate the simulation effect of the machine learning-based assimilation system on the model underwater structure, we designed a set of assimilation experiments for quantitative analysis. The assimilation period was from October 1, 2020 to October 29, 2020. The experiment contained two cases (Table 2) in which the assimilated observations of ExpA1 were the satellite SST and SLA; the assimilated observations of ExpA2 were the synthetic T/S profiles in the satellite observational grid. The observation error was assumed to be spatially and temporally uncorrelated, which resulted in the fact that the observation error covariance matrix was specified as a combination of measurement error and representative error, which were additive. The measurement error was considered independent of the data source, and the standard deviations of the observations from the scatter assimilation experiments were as follows: 2 cm for Satellite_SLA, 0.48°C for Satellite_SST, respectively. Satellite_pseudo_TS standard deviations were given respectively based on the depth means of the RMSEs of the 2019 test datasets of 0.76°C and 0.10 psu (Dai et al., 2021). The representativeness error is the standard deviation of the observations that contribute to each super-observation.

In order to better visualize the effect of synthetic profiles assimilation on the improvement of the model underwater structure, we selected two area with mesoscale eddies for qualitative analysis. For regions with strong stratification, the selected area should not be too large, which could easily lead to inaccurate regression relationships between sea surface variables and underwater variables. Therefore, compared to the previous set of assimilation experiments, we reduced the selected area (Table 2). ExpB1 directly assimilated the satellite SST and SLA, and the adjustment of the underwater structure was carried out by model dynamical framework, while ExpB2 assimilated the synthetic T/S profiles based on SST and SLA. The period of this groups of experiments was from October 1, 2020 to November 19, 2020 with an assimilation window of 7 days. To more extensively verify the improvement effect of the machine learning-based assimilation system on the model underwater structure, another grid assimilation experiment was conducted. The period and the assimilation window are shown in Table 2. The assimilation data of ExpC1 and ExpC2 are the same as ExpB1 and ExpB2, respectively, but in different regions. To compare the assimilation effect, survey data for the Northwest Pacific were chosen as independent validation data for ExpBs and ExpCs. The measurement error of satellite observations for ExpB1 and ExpC1 is consistent with that of ExpA1 and the measurement error of synthetic profiles for ExpB2 and ExpC2 is shown in Table 1.

3 Evaluation of the accuracy of the synthetic profiles

Before assimilating the synthetic T/S profiles into the model, we evaluated the performance of the reconstruction model, including error analysis and characteristic analysis by comparing the synthetic profiles with the reanalysis product and the WOA13 climatology data.

3.1 Error analysis

To test the effectiveness of the synthetic model, the root mean square errors (RMSEs) of the 2019 synthetic T/S profiles, SODA3.4.2 reanalysis products, and WOA13 climatology data were calculated separately relative to the EN4.2.1 T/S profiles.

TABLE 2 Assimilation experiment setup.

Case name	assimilation period	window	assimilation data
ExpA1ExpA2	2020-10-01—2020-10-29	7day	Satellite_SST+Satellite_SLA Satellite_pseudo_TS
ExpB1ExpB2	2020-10-01—2020-11-19	7day	Satellite_SST+Satellite_SLA Satellite_pseudo_TS
ExpC1ExpC2	2019-11-01—2019-11-29	7day	Satellite_SST+Satellite_SLA Satellite_pseudo_TS

The following equation was utilized to calculate the skill score of the synthetic profiles and SODA3.4.2 reanalysis product relative to the WOA13 climatology data (Zhu et al., 2022).

$$Skill = 1 - \left(\frac{RMSE(m, o)}{RMSE(c, o)} \right) \quad (12)$$

where $RMSE(m,o)$ represents the root mean square error between the target data and the EN4.2.1 *in-situ* data, and $RMSE(c, o)$ stands for the root mean square error between the reference data and the EN4.2.1 *in-situ* data. We used the WOA13 climatology data as the reference data. When the RMSE of the target data was smaller than the RMSE of the reference data, the *Skill* was positive, and the closer it was to one, the greater the degree of improvement was.

The RMSE and the *Skill* of the synthetic T/S profiles are shown in Figure 3. In terms of temperature, the error distributions of the synthetic profiles and SODA3.4.2 reanalysis products are close and the depth averages of RMSE are basically equal (0.65 for SODA3.4.2 and 0.66 for the synthetic fields), indicating that the accuracy of the synthetic fields and the SODA3.4.2 product is comparable (Figure 3A). From the perspective of *Skill*, the synthetic temperature profiles show the advantage of accuracy in 150m-400m, while the accuracy of SODA3.4.2 reanalysis product is higher in the rest of the depths. The depth averages of *Skill* for the synthetic temperature profiles and the SODA3.4.2 reanalysis product are 0.34 and 0.36, respectively, which are almost equivalent (Figure 3C). From the perspective of salinity, the RMSE of the synthetic salinity profiles from 100 m to 300 m was similar to that of the WOA13 data but larger than that of the SODA3.4.2 product (Figure 3B). This is understandable since the 5 day-averaged SODA3.4.2 reanalysis product assimilates the satellite observations as well as the *in-situ* T/S profiles and the data used for comparison here belong to the *in-situ* T/S observations. In addition to the need to further enhance the accuracy near the thermocline layer, the accuracy of the synthetic salinity profiles was improved compared with the WOA13 data (Figure 3D).

Figure 4 exhibits the horizontal distribution of RMSEs for the synthetic data and SODA3.4.2 reanalysis product relative to the EN4.2.1 *in-situ* profiles in 2019. Overall, the RMSE of synthetic temperature profiles was smaller than 1°C with a mean value of about 0.67°C, and the RMSE of the synthetic salinity profiles was smaller than 0.15 psu with a mean value of about 0.07 psu. There were 437 EN4.2.1 T/S profiles in 2019 and the RMSE of about 37% for the synthetic temperature profiles was smaller than that for the SODA3.4.2 temperature product, and the RMSE of about 30% for the synthetic salinity profiles was smaller than that for the SODA3.4.2 salinity product.

3.2 Characteristic analysis

We quantitatively analyzed the accuracy of the reconstructed model in Section 3.1. In this section, combined with the three-dimensional structural features of mesoscale eddies, we qualitatively analyze the reconstruction ability of the FOAGRNN method for the mesoscale eddy. Figure 5 depicts

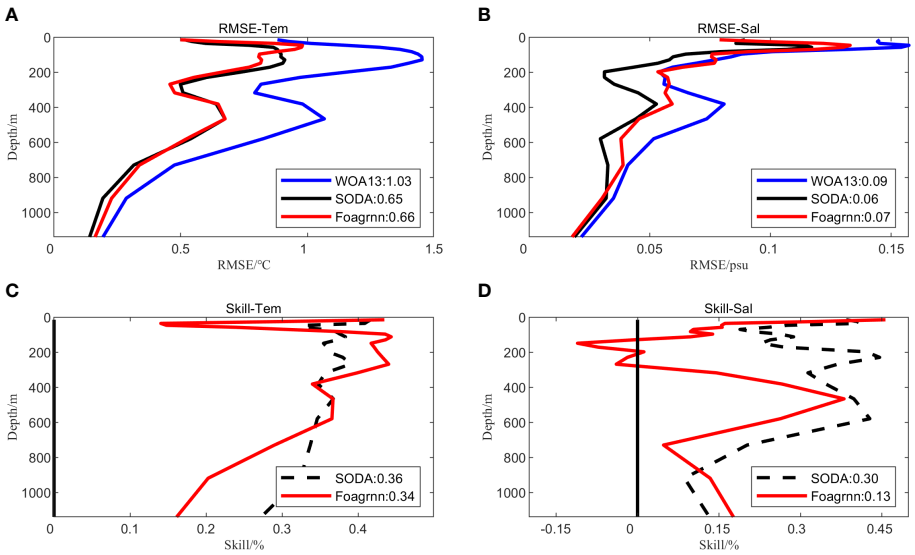


FIGURE 3
Temperature (left) and salinity (right) profiles root mean square errors (RMSEs) and skill scores as a function of depth. **(A)** temperature profiles RMSE, **(B)** salinity profiles RMSE, **(C)** skill score of temperature profiles relative to WOA13, **(D)** skill score of salinity profiles relative to WOA13. The black line represents SODA3.4.2 reanalysis data; the red line represents the synthetic profiles; the blue line represents WOA13 climatology data, and the numbers represent the depth mean of RMSE and skill score.

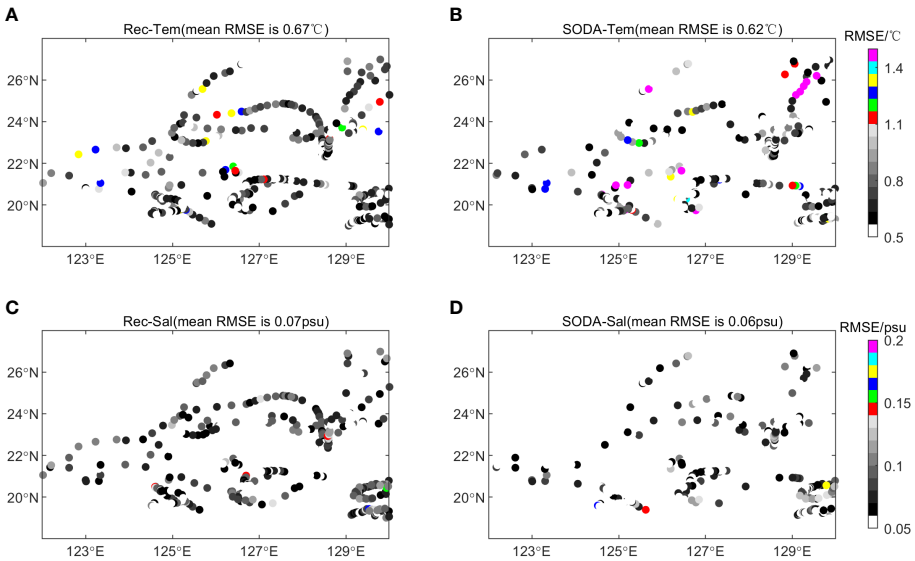


FIGURE 4
The distribution of RMSEs over space. Each row represents a different variable field. Each column represents a different data source: **(A, C)** the synthetic fields and **(B, D)** the SODA3.4.2 reanalysis data.

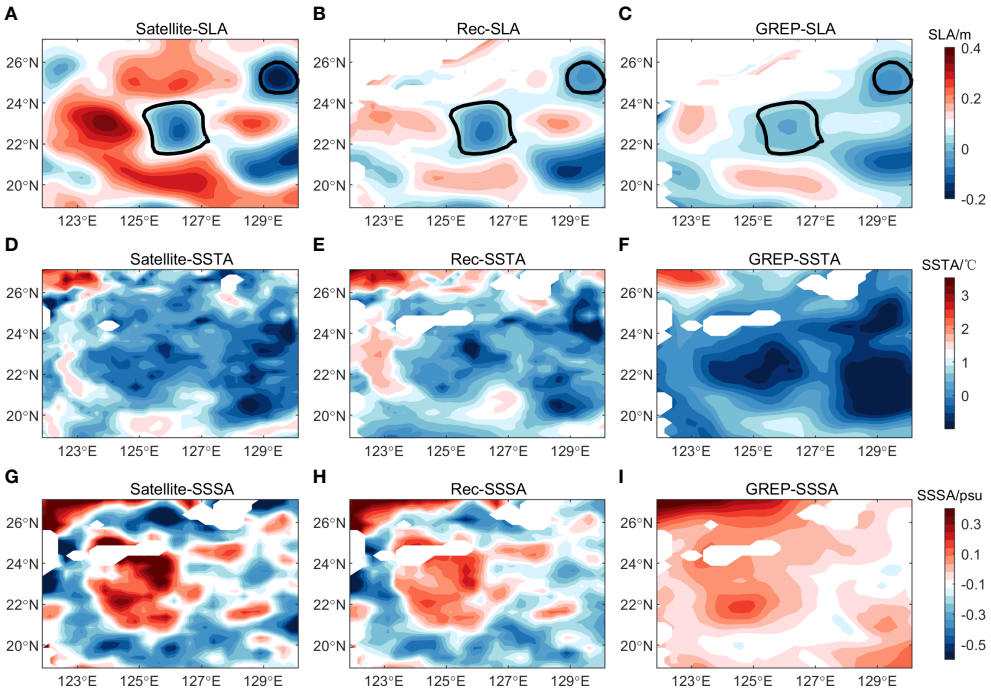


FIGURE 5
The distribution of the anomaly fields of sea surface level, temperature and salinity from satellite, synthetic fields and GREP reanalysis data on May 13, 2019. **(A, D, G)** the SLA, SSTA, and SSSA of the satellite observations, **(B, E, H)** the SLA, SSTA, and SSSA of the synthetic fields, **(C, F, I)** the SLA, SSTA, and SSSA of the GREP reanalysis product.

the distribution characteristics of SLA, sea surface temperature anomaly (SSTA), and sea surface salinity anomaly (SSSA). SSTA and SSSA are the anomalies of satellite data, synthetic fields, and GREP reanalysis product relative to WOA13 climatology data. The SLA for the synthetic data and GREP product is the anomaly of the vertically integrated dynamic height field relative to that calculated by WOA13 data. The synthetic profiles began at 10 m, thus, the SST and SSS of the synthetic data were approximately represented by the temperature and salinity values at this depth. It can be observed in Figure 5 that the characteristics (location and shape) of the mesoscale eddies calculated based on the synthetic temperature and salinity fields are closer to the observations than the results of the GREP product. One possible reason is that the reconstructed model incorporates much information from various surface observations. The changes with depth in the dynamic height anomaly field, temperature anomaly field, and salinity anomaly field are illustrated in Figure 6, respectively. The mesoscale eddies which were calculated by each element field of the synthetic data are quite consistent with the real sea surface eddies in both position and shape at these depth layers, indicating that the FOAGRNN algorithm has acceptable mesoscale eddy reconstruction capability.

4 Assimilation results of the synthetic T/S profiles

Through error analysis and characteristic analysis, section 3 suggested that the synthetic T/S profiles based on the machine learning method had sufficient accuracy. On this basis, we further conducted the following three groups of assimilation experiments.

4.1 Quantitative analysis

In order to make the RMSE statistically significant, we selected a relatively large area for the assimilation experiment. The RMSEs of the assimilation experiment results of each case relative to the EN4.2.1 *in-situ* T/S profiles are exhibited in Figure 7. In terms of temperature, ExpA1 has an error of more than 2°C at the thermocline, while ExpA2 has a reduced error at the 100-1100m depth layer. Compared to MODAS, ExpA2 shows higher accuracy at the thermocline and 300-800 m. The depth means of RMSE show that ExpA2 improves the simulation accuracy by 6.2% and 26.0% compared to MODAS and ExpA1, respectively (Figure 7A). From the perspective of salinity, compare with ExpA1, the experiment results of synthetic profiles not only reduce the error at the thermocline, but also have higher simulation accuracy at other depths with improvements in the simulation accuracy by 23.1%. The depth means of RMSE for ExpA2 and MODAS are basically equal (0.10 and 0.09, respectively), which shows that their simulation accuracy is comparable, and it may be related to both assimilating synthetic salinity profiles (Figure 7B).

4.2 Qualitative analysis

4.2.1 Region one

The results of the error analysis show that the assimilation of synthetic profiles can improve the simulation accuracy of the model underwater structure compared to the direct assimilation of the satellite observations. In section 4.2, the impact of the synthetic profiles on the model underwater structure is assessed from the qualitative analysis. Figure 8 shows the temperature anomaly and salinity sections along the observation route. The temperature anomalies were obtained by subtracting the WOA13 climatology data from the *in-situ* survey data, MODAS data, and the experimental results of ExpB1 and ExpB2. In terms of temperature, the shape characteristics of ExpB1 were inconsistent with the *in-situ* section structure and the distribution of the temperature anomaly in the deep layer was also relatively higher than any other data. Both ExpB2 and MODAS data simulated a cold eddy structure similar to the *in-situ* one in the corresponding position, but the shape feature of ExpB2 was more consistent with the *in-situ* one (Figures 9A–D). From the perspective of salinity, compared with MODAS data and the experiment results of ExpB1, the salinity section structure of ExpB2 was more consistent with the *in-situ* measurements in terms of depth distribution and microstructure (Figures 9E–H).

4.2.2 Region two

The section where the red dots in Figure 10 are located was taken as the research section. Figures 11A–D reveal the temperature anomaly sections of the *in-situ* survey data, ExpC2, MODAS data, and ExpC1, respectively. Taking the depth of 200 m as the dividing line, an obvious dual-core structure can be observed in the *in-situ* temperature anomaly section structure. Compared with the assimilation results of ExpC1, both ExpC2 and MODAS data simulated a dual-core structure, but the strength of the cold eddy structure is not in line with the actual measurement. From the perspective of salinity anomaly sections, we can see in Figures 11E–H that there is a low-salinity center near 400 m in the *in-situ* salinity anomaly section. The salinity anomaly section structure of ExpC1 did not simulate a single-core structure in the deep layer, and the salinity anomaly value of ExpC1 near the sea surface was much larger than that of the *in-situ* survey data. The cold eddy structure could not be simulated by MODAS data, and the salinity

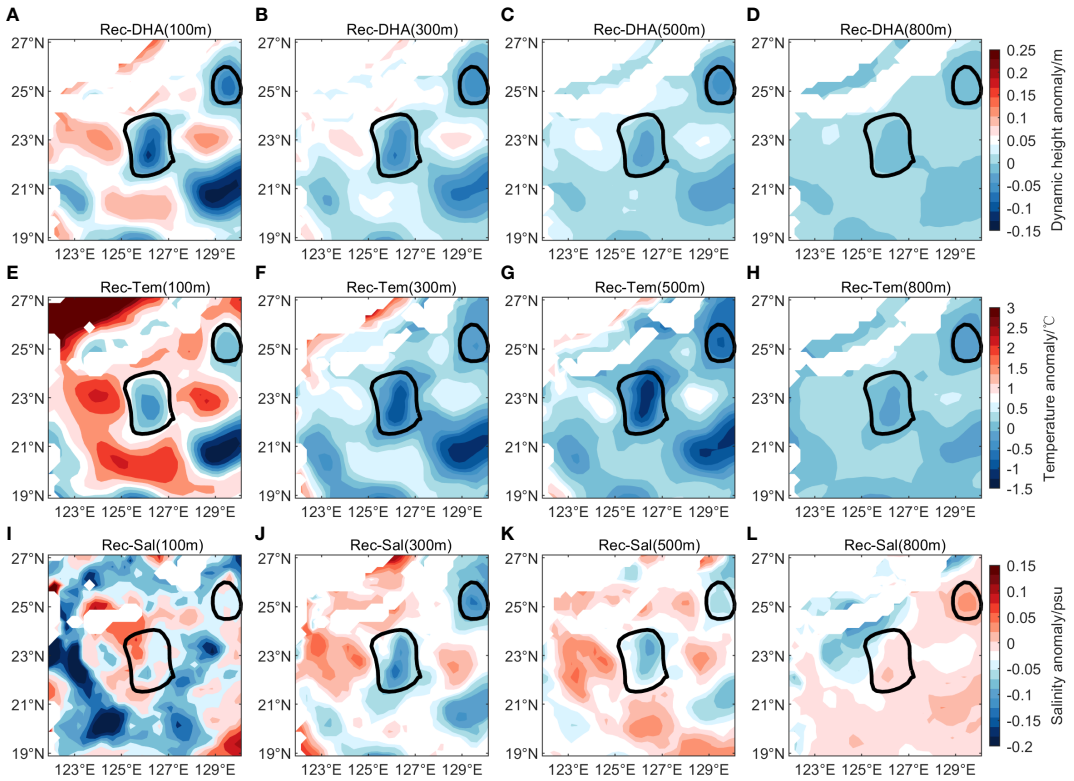


FIGURE 6
The distribution of the anomaly fields of dynamic height, temperature and salinity from the synthetic fields and on May 13, 2019. From left to right: 100m, 300m, 500m, and 800m. **(A–D)** the dynamic height anomaly fields of the synthetic fields, **(E–H)** the temperature anomaly fields of the synthetic fields, **(I–L)** the salinity anomaly fields of the synthetic fields. The circles in the figure are mesoscale eddies identified by satellite SLA.

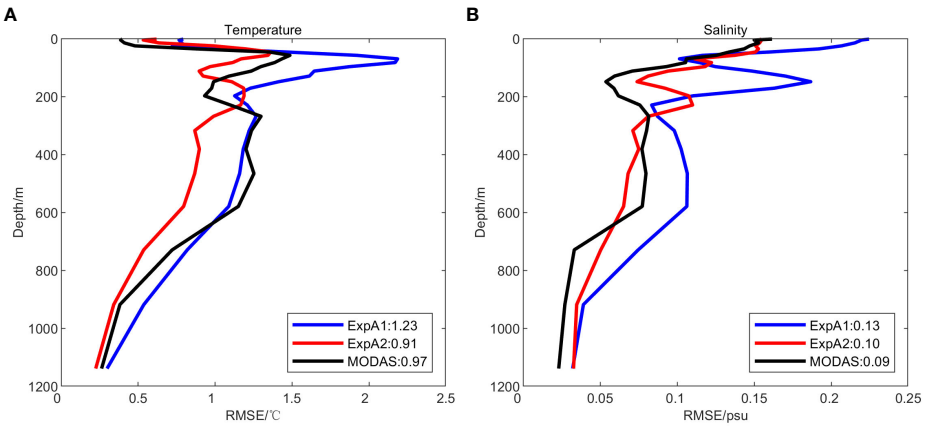


FIGURE 7
RMSEs of ROMS in nowcasting all the available EN4.2.1 T/S profiles. **(A)** temperature and **(B)** salinity (black line: MODAS data; red line: the experiment results of assimilating the synthetic T/S observations; blue line: the experiment results of assimilating sea surface observations). Numbers represent depth means of the RMSEs.

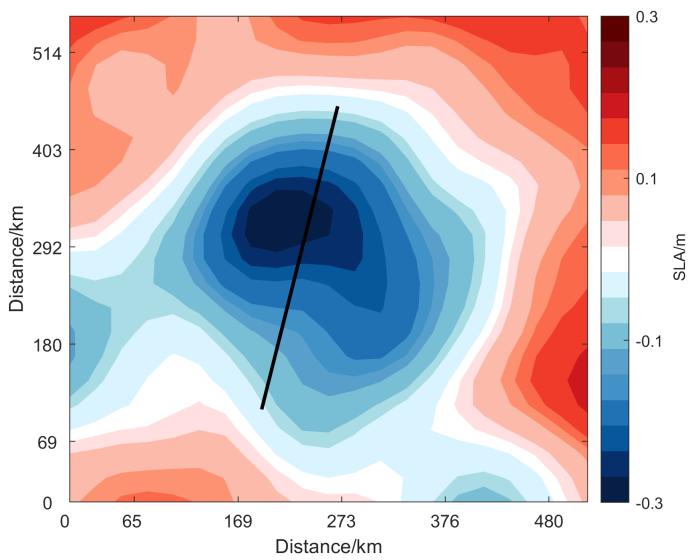


FIGURE 8
The distribution of the sea surface level anomalies on November 8, 2020 (unit: m). The black line represents the observation route of the Northwest Pacific survey data.

anomalies at almost all depths were near 0 psu. This indicated that the salinity field of MODAS data was not much different from the WOA13 salinity background field. Differently, ExpC2 simulated an obvious single-core structure, but the depth range covered by the structure was larger than that by the *in-situ* salinity anomaly section.

Figure 12 illustrates the horizontal distribution of temperature and salinity anomaly at depths of 10 m, 100 m,

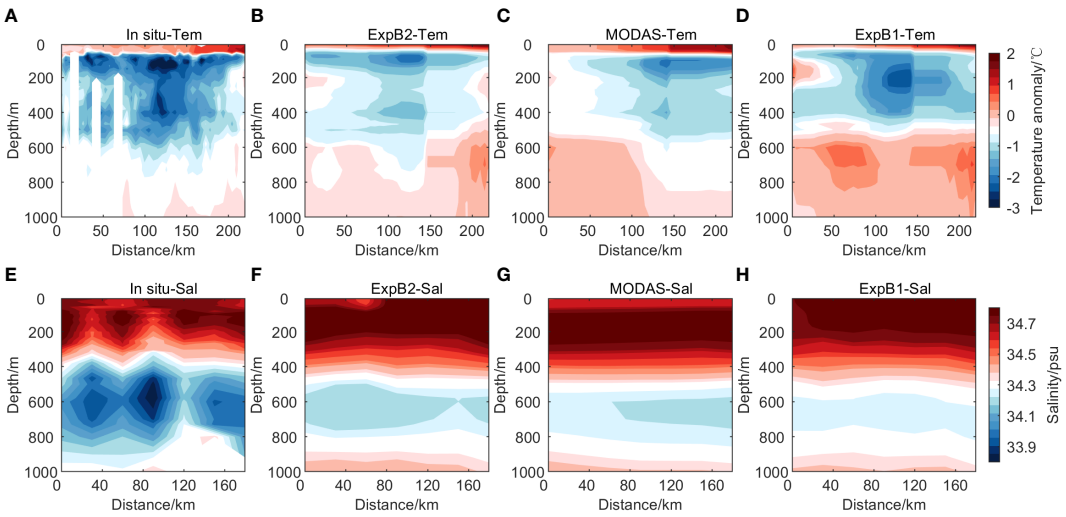


FIGURE 9
Temperature anomaly section and salinity section along the observation route. From top to bottom: temperature anomaly values and salinity values. **(A, E)** *in-situ* survey data, **(B, F)** the experiment results of assimilating the synthetic T/S observations, **(C, G)** the MODAS data of the National Marine Data Center, **(D, H)** the experiment results of assimilating the satellite observations. The abscissa represents the distance from the starting point of the survey data (unit: km).

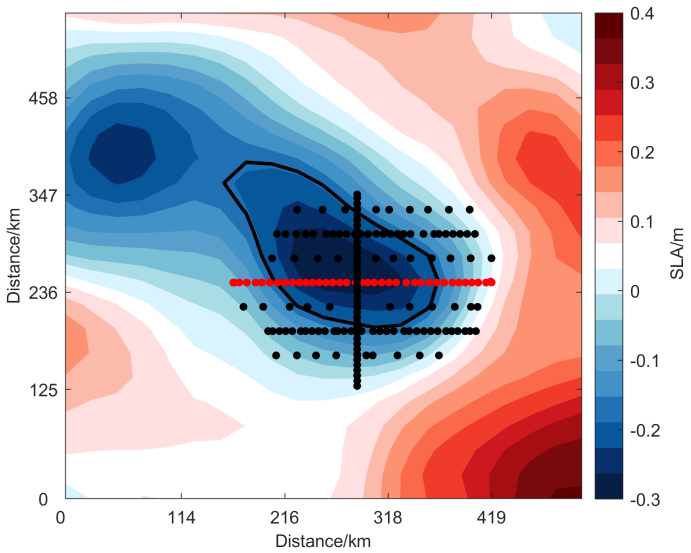


FIGURE 10
The distribution of survey stations of the Northwest Pacific cyclone eddy from November 13 to 15, 2019. The red dots represent the location of the study section, the color represents the sea surface level anomaly, and the circle in the figure is the mesoscale eddy identified by satellite SLA.

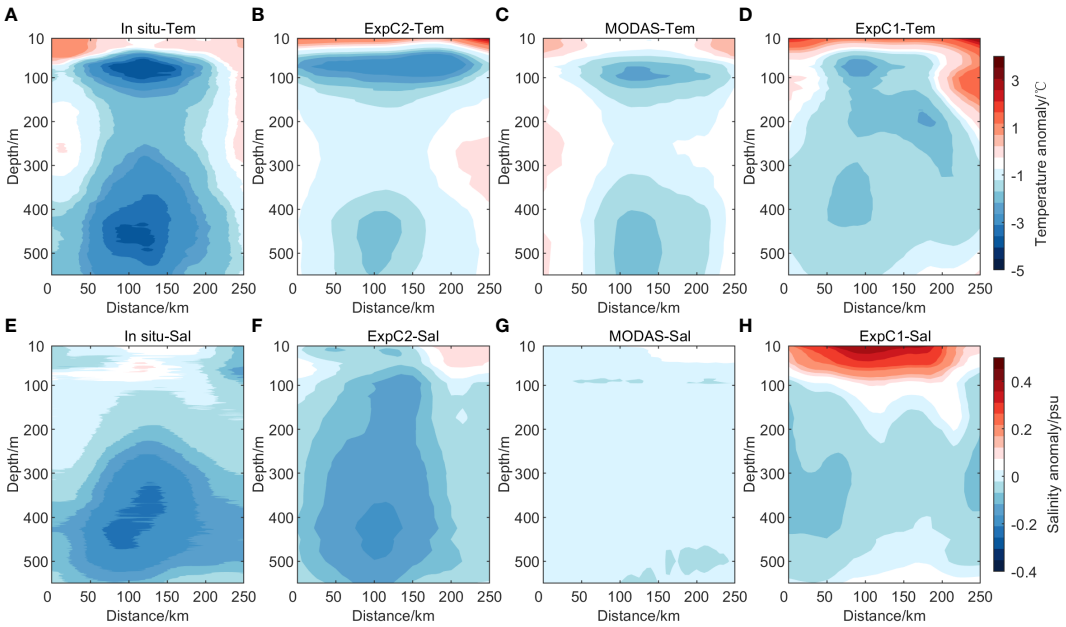


FIGURE 11
The section structure of the temperature and salinity anomaly at the position of the red dot in Figure 10. Top: temperature anomaly values; bottom: salinity anomaly values. **(A, E)** *in-situ* survey data, **(B, F)** the experiment results of assimilating the synthetic T/S observations, **(C, G)** the MODAS data of the National Marine Data Center, **(D, H)** the experiment results of assimilating the satellite observations.

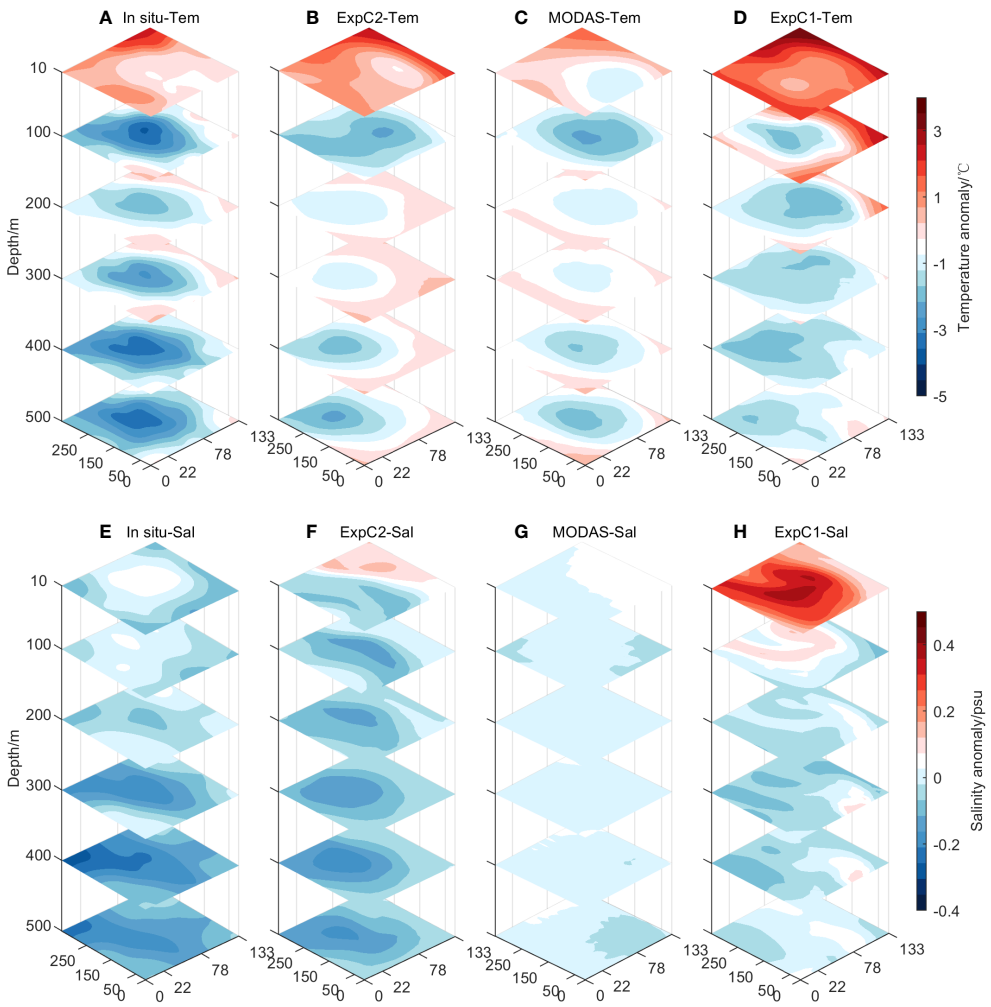


FIGURE 12
Temperature and salinity anomaly slices at 10m, 100m, 200m, 300m, 400m, and 500m. **(A, E)** the *in-situ* data, **(B, F)** the experiment results of assimilating the synthetic T/S profiles, **(C, G)** the MODAS data of the National Marine Data Center, and **(D, H)** the experiment results of the direct assimilation of satellite observations. The unit of abscissa and ordinate is km.

200 m, 300 m, 400 m, and 500 m. In terms of temperature anomaly, from 100m to 200m, the slice structure of ExpC1 was in good agreement with the *in-situ* data, suggesting that the direct assimilation of satellite data had a positive impact on the simulation of the model underwater structure at this depth layer. However, the cold eddy structure simulated by ExpC1 gradually disappeared in the subsequent depths. This indicated that the positive impact of the direct assimilation of satellite data was weakening with the change of depth, while MODAS data and ExpC2 could maintain the cold eddy structure with the change of depth (Figures 12A–D). From the perspective of salinity anomaly, ExpC1 could not simulate the cold eddy structure, and similar to Figure 11G, the salinity anomalies of MODAS data at all depths were around zero, indicating that MODAS salinity data were almost the same as WOA13 climatology salinity data. ExpC2 simulated both the cold eddy structure and the intensity variation trend close to the *in-situ* measurement (Figures 12E–H).

5 Conclusion and discussion

Satellite data are very important for marine operational forecasting systems, however, the traditional method of the direct assimilation of satellite observations cannot constrain the simulation of underwater structures well. To address this problem, we created a machine learning-based assimilation system. First, the historical EN4.2.1 *in-situ* T/S profiles and

historical satellite altimeter data were employed as the training data of the FOAGRNN to construct a reconstruction model. Then satellite observations (SST, SLA, SSS) were utilized as the input data of the reconstruction model to reconstruct three-dimensional T/S pseudo-profiles. Finally, the I4D-Var method was used to assimilate the synthetic data into the ROMS, and three groups of assimilation experiments were performed. A validation of the synthetic T/S profiles and the assimilation experiments results against the observations indicates that

1. In addition to the need to further enhance the accuracy near the thermocline, the accuracy of the synthetic profiles is comparable to the 5-day averaged SODA 3.4.2 reanalysis product and better than the WOA13 climatology data. The horizontal distribution of RMSE shows that the error of the synthetic temperature profiles is within 1°C, while the error for salinity is within 0.15 psu. The profiles with better quality than the SODA3.4.2 reanalysis product occupy an acceptable proportion in total. A validation against the GREP reanalysis products shows that the synthetic fields have better mesoscale eddy reconstruction ability. Moreover, the mesoscale eddies which were calculated by each element field of the synthetic data are quite consistent with the real sea surface eddies in both position and shape at the selected depths.
2. A validation against the EN4.2.1 *in-situ* T/S profiles (with 107 observed profiles) shows that compared with direct assimilation of satellite remote sensing observations, the simulation accuracy of assimilating synthetic profiles shows a significant improvement at the thermocline, with a 26.0% reduction for temperature and a 23.1% reduction for salinity in RMSE. Compared with MODAS, the simulation accuracy of assimilating synthetic profiles is improved by 6.2% in temperature and comparable in salinity.
3. Survey data from the Northwest Pacific were used as independent observations for the qualitative analysis of the assimilation effect, which demonstrates that compared with the direct assimilation of satellite remote sensing observations, the indirect assimilation based on machine learning can significantly improve the simulation effect of model underwater structure, and compared with MODAS, the machine learning-based assimilation system demonstrated a significant advantage in the simulation of underwater salinity structure.

Compared with the direct assimilation of satellite observations, the indirect assimilation based on the machine learning substantially improved the simulation effect of model underwater structures, which can provide a more accurate initial condition for ocean models to more accurately predict ocean phenomena, such as mesoscale eddies. Moreover, as an application example, our study can promote more scholars to explore the combination of machine learning and data assimilation in different ways, especially for applications of satellite data in the operational system. However, there is still a lot of optimization work to be done in the future.

1. There are mainly two error sources in the salinity profile estimation. First, the satellite SSS data still have various types of errors from the instrument's observations, brightness temperature Tb reconstruction, and salinity retrieval algorithm, especially at high latitudes. Second, the training data of the reconstruction model were based on years of *in-situ* T/S profiles, while the input data used to construct the real-time T/S pseudo-profiles originated from satellite observations. Different depth distributions of salinity values will inevitably cause errors in the synthetic profiles. Therefore, the following study can attempt to directly use the satellite observations as training data or to describe the characteristic of the relationship between *in-situ* SSS and satellite SSS.
2. when the machine learning algorithm reconstructs the T/S profiles in a complex stratified region, such as mesoscale eddies, the region of the selected sample data should be limited to a certain range. Otherwise, the empirical relationship between surface observations and underwater variable fields may not be representative, and the error of the synthetic profiles may be large. However, if the range is too small, the number of the *in-situ* T/S profiles in the study area will be scarce, and an effective reconstruction model may not be constructed. To compensate for this deficiency, it is possible to employ reanalysis products as training data to build a reconstruction model since the reanalysis data of each current platform have good simulation accuracy, which can provide favorable preconditions for the realization of this goal (Compo et al., 2011; Dee et al., 2011).
3. In terms of the assimilation, this study assumed that the observation errors were uncorrelated in time and space and prescribed a single scalar for all synthetic observations, which simplified the construction of the observation error covariance matrix, but in fact, the synthetic T/S profiles had a certain correlation, and a single scalar for all synthetic observations may cause under/overestimation of synthetic observations in the wrong places. Moreover, a large number of synthetic T/S profiles, which were not sparse enough, likely affected the assimilation efficiency.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

Data curation, ZC, PW, and SB. Visualization, ZC and PW. Software, ZC, PW, and SB. Writing – original draft, ZC. Review of the manuscript, PW, SB, and WZ. Funding, WZ. All authors have read and agreed to the published version of the manuscript.

Funding

This paper is supported by the National Key R&D Plan Project, Regional Multi-layer Coupling Data Assimilation (2021YFC3101500).

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Bao, S., Zhang, R., Wang, H., Yan, H., Yu, Y., and Chen, J. (2019). Salinity profile estimation in the pacific ocean from satellite surface salinity observations. *J. Atmospheric Oceanic Technol.* 36 (1), 53–68. doi: 10.1175/jtech-d-17-0226.1

Carnes, M. R., Teague, W. J., and Mitchell, J. L. (1994). Inference of subsurface thermohaline structure from fields measurable by satellite. *J. Atmospheric Oceanic Technol.* 11 (2), 551–566. doi: 10.1175/1520-0426(1994)011<0551:Iostsf>2.0.Co;2

Carton, J. A., Chepurin, G. A., and Chen, L. (2018). SODA3: A new ocean climate reanalysis. *J. Climate* 31 (17), 6967–6983. doi: 10.1175/jcli-d-18-0149.1

Chapman, C., and Charantonis, A. A. (2017). Reconstruction of subsurface velocities from satellite observations using iterative self-organizing maps. *IEEE Geosci. Remote Sens. Lett.* 14 (5), 617–620. doi: 10.1109/LGRS.2017.2665603

Chen, K., He, R., Powell, B. S., Gawarkiewicz, G. G., Moore, A. M., and Arango, H. G. (2014). Data assimilative modeling investigation of gulf stream warm core ring interaction with continental shelf and slope circulation. *J. Geophysical Research: Oceans* 119 (9), 5968–5991. doi: 10.1002/2014JC009898

Chen, X., Wang, H., Zheng, F., and Cai, Q. (2020). An ensemble-based SST nudging method proposed for correcting the subsurface temperature field in climate model. *Acta Oceanologica Sin.* 39 (3), 73–80. doi: 10.1007/s13131-020-1568-2

Chen, Y., Yan, C., and Zhu, J. (2018). Assimilation of Sea surface temperature in a global hybrid coordinate ocean model. *Adv. Atmospheric Sci.* 35 (10), 1291–1304. doi: 10.1007/s00376-018-7284-6

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., et al. (2011). The twentieth century reanalysis project. *Q. J. R. Meteorological Soc.* 137 (654), 1–28. doi: 10.1002/qj.776

Cooper, M., and Haines, K. (1996).). altimetric assimilation with water property conservation. *J. Geophysical Research: Oceans* 101 (C1), 1059–1077. doi: 10.1029/95JC02902

Cummings, J. A., and Smedstad, O. M. (2013). ""Variational data assimilation for the global ocean,"," in *Data assimilation for atmospheric, oceanic and hydrologic applications (Vol. II)*. Eds. S. K. Park and L. Xu (Berlin, Heidelberg: Springer Berlin Heidelberg), 303–343.

Dai, J., Wang, H., Zhang, W., Wang, P., and Luo, T. (2021). Three-dimensional structure of an observed cyclonic mesoscale eddy in the Northwest pacific and its assimilation experiment. *Acta Oceanologica Sin.* 40 (5), 1–19. doi: 10.1007/s13131-021-1810-6

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorological Soc.* 137 (656), 553–597. doi: 10.1002/qj.828

Ezer, T., and Mellor, G. L. (1994). Continuous assimilation of geosat altimeter data into a three-dimensional primitive equation gulf stream model. *J. Phys. Oceanography* 24 (4), 832–847. doi: 10.1175/1520-0485(1994)024<0832:Caogad>2.0.Co;2

Fox, D. N., Teague, W. J., Barron, C. N., Carnes, M. R., and Lee, C. M. (2002). The modular ocean data assimilation system (MODAS). *J. Atmospheric Oceanic Technol.* 19 (2), 240–252. doi: 10.1175/1520-0426(2002)019<0240:Tmodas>2.0.Co;2

Good, S., Fiedler, E., Mao, C., Martin, M. J., Maycock, A., Reid, R., et al. (2020). The current configuration of the OSTIA system for operational production of foundation Sea surface temperature and ice concentration analyses. *Remote Sens.* 12 (4), 720. doi: 10.3390/rs12040720

Good, S. A., Martin, M. J., and Rayner, N. A. (2013). EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. *J. Geophysical Research: Oceans* 118 (12), 6704–6716. doi: 10.1002/2013JC009067

Haines, K. (1991). A direct method for assimilating Sea surface height data into ocean models with adjustments to the deep circulation. *J. Phys. Oceanography* 21 (6), 843–868. doi: 10.1175/1520-0485(1991)021<0843:Admfas>2.0.Co;2

Helber, R. W., Townsend, T. L., Barron, C. N., Dastugue, J. M., and Carnes, M. R. (2013). *"Validation test report for the improved synthetic ocean profile (ISOP) system, part I: Synthetic profile methods and algorithm"* (Naval Research Lab Stennis Detachment Stennis Space Center Ms Oceanography Div).

He, Z., Wang, X., Wu, X., Chen, Z., and Chen, J. (2021). Projecting three-dimensional ocean thermohaline structure in the north Indian ocean from the satellite Sea surface data based on a variational method. *J. Geophysical Research: Oceans* 126 (1), e2020JC016759. doi: 10.1029/2020JC016759

Holland, W. R., and Malanotte-Rizzoli, P. (1989). Assimilation of altimeter data into an ocean circulation model: Space versus time resolution studies. *J. Phys. Oceanography* 19 (10), 1507–1534. doi: 10.1175/1520-0485(1989)019<1507:Aoadia>2.0.Co;2

Iqbal, Z., Shahid, S., Ahmed, K., Wang, X., Ismail, T., and Gabriel, H. F. (2022). Bias correction method of high-resolution satellite-based precipitation product for peninsular Malaysia. *Theor. Appl. Climatology* 148 (3), 1429–1446. doi: 10.1007/s00704-022-04007-6

Le, X.-H., Lee, G., Jung, K., An, H.-u., Lee, S., and Jung, Y. (2020). Application of convolutional neural network for spatiotemporal bias correction of daily satellite-based precipitation. *Remote Sens.* 12 (17), 2731. doi: 10.3390/rs12172731

Li, H.-z., Guo, S., Li, C.-j., and Sun, J.-q. (2013). A hybrid annual power load forecasting model based on generalized regression neural network with fruit fly optimization algorithm. *Knowledge-Based Syst.* 37, 378–387. doi: 10.1016/j.knosys.2012.08.015

Lin, S.-M. (2013). Analysis of service satisfaction in web auction logistics service using a combination of fruit fly optimization algorithm and general regression neural network. *Neural Computing Appl.* 22 (3), 783–791. doi: 10.1007/s00521-011-0769-1

Liu, L., Xue, H., and Sasaki, H. (2019). Reconstructing the ocean interior from high-resolution Sea surface information. *J. Phys. Oceanography* 49 (12), 3245–3262. doi: 10.1175/jpo-d-19-0118.1

Liu, L., Xue, H., and Sasaki, H. (2021). Diagnosing subsurface vertical velocities from high-resolution Sea surface fields. *J. Phys. Oceanography* 51 (5), 1353–1373. doi: 10.1175/jpo-d-20-0152.1

Moore, A. M., Arango, H. G., Broquet, G., Edwards, C., Veneziani, M., Powell, B., et al. (2011a). The regional ocean modeling system (ROMS) 4-dimensional variational data assimilation systems: Part II – performance and application to the

California current system. *Prog. Oceanography* 91 (1), 50–73. doi: 10.1016/j.pocean.2011.05.003

Moore, A. M., Arango, H. G., Broquet, G., Powell, B. S., Weaver, A. T., and Zavala-Garay, J. (2011b). The regional ocean modeling system (ROMS) 4-dimensional variational data assimilation systems: Part I – system overview and formulation. *Prog. Oceanography* 91 (1), 34–49. doi: 10.1016/j.pocean.2011.05.004

Ratheesh, S., Sharma, R., and Basu, S. (2012). Projection-based assimilation of satellite-derived surface data in an Indian ocean circulation model. *Mar. Geodesy* 35 (2), 175–187. doi: 10.1080/01490419.2011.637855

Storto, A., Masina, S., Simoncelli, S., Iovino, D., Cipollone, A., Drevillon, M., et al. (2019). The added value of the multi-system spread information for ocean heat content and steric sea level investigations in the CMEMS GREP ensemble reanalysis product. *Climate dynamics* 53 (1), 287–312. doi: 10.1007/s00382-018-4585-5

Su, H., Li, W., and Yan, X.-H. (2018). Retrieving temperature anomaly in the global subsurface and deeper ocean from satellite observations. *J. Geophysical Research: Oceans* 123 (1), 399–410. doi: 10.1002/2017JC013631

Su, H., Wu, X., Yan, X.-H., and Kidwell, A. (2015). Estimation of subsurface temperature anomaly in the Indian ocean during recent global surface warming hiatus from satellite measurements: A support vector machine approach. *Remote Sens. Environ.* 160, 63–71. doi: 10.1016/j.rse.2015.01.001

Tranchant, B., Remy, E., Greiner, E., and Legalloudec, O. (2019). Data assimilation of soil moisture and ocean salinity (SMOS) observations into the Mercator ocean operational system: focus on the El niño 2015 event. *Ocean Sci.* 15 (3), 543–563. doi: 10.5194/os-15-543-2019

Vernieres, G., Kovach, R., Keppenne, C., Akella, S., Brucker, L., and Dinnat, E. (2014). The impact of the assimilation of aquarius sea surface salinity data in the GEOS ocean data assimilation system. *J. Geophysical Research: Oceans* 119 (10), 6974–6987. doi: 10.1002/2014JC010006

Wang, P., Zhu, M., Chen, Y., Zhang, W., and Yu, Y. (2021). Ocean satellite data assimilation using the implicit equal-weights variational particle smoother. *Ocean Model.* 164, 101833. doi: 10.1016/j.ocemod.2021.101833

Weaver, A. T., Deltel, C., Machu, E., Ricci, S., and Daget, N. (2005). A multivariate balance operator for variational ocean data assimilation. *Q. J. R. Meteorological Soc.* 131 (613), 3605–3625. doi: 10.1256/qj.05.119

Zhang, W. G., Wilkin, J. L., and Arango, H. G. (2010). Towards an integrated observation and modeling system in the new York bight using variational methods. part I: 4DVAR data assimilation. *Ocean Model.* 35 (3), 119–133. doi: 10.1016/j.ocemod.2010.08.003

Zhou, W., Li, J., Xu, F., Shu, Y., and Feng, Y. (2021). The impact of ocean data assimilation on seasonal predictions based on the national climate center climate system model. *Acta Oceanologica Sin.* 40 (5), 58–70. doi: 10.1007/s13131-021-1732-3

Zhu, X., Zu, Z., Ren, S., Zhang, M., Zhang, Y., Wang, H., et al. (2022). Improvements in the regional south China Sea operational oceanography forecasting system (SCSOFSv2). *Geosci. Model. Dev.* 15 (3), 995–1015. doi: 10.5194/gmd-15-995-2022

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 10 October 2022
DOI 10.3389/fmars.2022.997537

Check for updates

OPEN ACCESS

EDITED BY
Shiqiu Peng,
State Key Laboratory of Tropical Oceanography, Chinese Academy of Sciences, China

REVIEWED BY
Xuefeng Zhang,
Tianjin University, China
Guijun Han,
Tianjin University, China

*CORRESPONDENCE
Yongzhi Liu
yzliu@fio.org.cn
Kai Fu
kfu@ouc.edu.cn

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

RECEIVED 19 July 2022
ACCEPTED 16 September 2022
PUBLISHED 10 October 2022

CITATION
Xu M, Liu Y, Zhao Z, Fu K and Lv X (2022) Advancing parameter estimation with Characteristic Finite Difference Method (CFDM) for a marine ecosystem model by assimilating satellite observations: Spatial distributions.
Front. Mar. Sci. 9:997537.
doi: 10.3389/fmars.2022.997537

COPYRIGHT
© 2022 Xu, Liu, Zhao, Fu and Lv. 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(s) 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.

Advancing parameter estimation with Characteristic Finite Difference Method (CFDM) for a marine ecosystem model by assimilating satellite observations: Spatial distributions

Minjie Xu[1], Yongzhi Liu[2]*, Zihan Zhao[3], Kai Fu[4]* and Xianqing Lv[3,5]

[1]School of Ocean, Yantai University, Yantai, China, [2]Center for Ocean and Climate Research, First Institute of Oceanography, Ministry of Natural Resources, and Laboratory for Regional Oceanography and Numerical Modeling, Pilot National Laboratory for Marine Science and Technology, Qingdao, China, [3]Frontier Science Center for Deep Ocean Multispheres and Earth System (FDOMES) and Physical Oceanography Laboratory, Ocean University of China, Qingdao, China, [4]School of Mathematical Sciences, Ocean University of China, Qingdao, China, [5]Qingdao National Laboratory for Marine Science and Technology, Qingdao, China

The ecosystem parameters are critical for precisely determining the marine ecological process and improving the simulations of the marine ecological model. In this study, based on the NPZD (nutrient, phytoplankton, zooplankton and detritus) model, the surface chlorophyll-a observations obtained from Sea-viewing Wide Field-of-view Sensor (SeaWiFS) data were assimilated to estimate spatially ecosystem parameters in the Bohai, Yellow, and East China Seas using an adjoint assimilation method with characteristic finite difference scheme. The experiments of the moving Gaussian hump indicated that the characteristic finite difference method (CFDM) can get rid of the limit of stability and permit using large time steps, which reduces long computation durations and large memory requirements. The model performance was significantly improved after data assimilation with CFDM using a large time step of 6 hours. Moreover, the distributions of parameters of the NPZD model in winter in the Bohai Sea, the Yellow Sea, and the East China Sea were simulated by our method. Overall, the developed method can efficiently optimize the ecosystem parameters and the results can be beneficial for determining reasonable parameters of the marine ecological model.

KEYWORDS

adjoint assimilation method, characteristic finite difference method, nutrient-phytoplankton-zooplankton-detritus (NPZD) model, parameter optimization, spatial distributions

Introduction

With the development of ocean exploitation, a series of ecological and environmental problems appeared, which caused serious harm to the coastal economic development. Therefore, it is very important to analyze the marine ecological environment data and rationally exploit marine resources. Marine biogeochemical models are useful tools that help to understand and predict marine environmental processes and are increasingly applied in ecological research, management advice, policy exploration, and environmental impact analysis (Link et al., 2011; Serpetti et al., 2017; Peck et al., 2018; Borja et al., 2020; Kytinou et al., 2020; Steenbeek et al., 2021). Riley et al. (1949) established the first marine ecosystem models and simulated the phytoplankton and zooplankton. Recently, marine ecological models have diversified from the simpler NPZ (nutrient, phytoplankton, zooplankton) model (Franks, 2002) or NPZD (nutrient, phytoplankton, zooplankton, and detritus) model to the more complicated North Pacific Ecosystem Model (Kishi et al., 2007). The model used in this paper is the NPZD model, which lies in the middle range of complexity (Heinle and Slawig, 2013).

Marine biogeochemical models usually comprise numerous parameters that describe biological and chemical rates of change such as growth, mortality, and degradation rates, and estimating their values is a non-linear problem (Matear, 1995; Athias et al., 2000; Jones et al., 2016). To tune and improve models, several studies have been carried out to estimate and optimize these parameters. Schartau and Oschlies (2003) used a micro-generic algorithm to estimate model parameters in the North Atlantic and found that there were different optimal parameter values at three locations. Rückelt et al. (2010) applied a hybrid quantum-evolutionary and deterministic optimization algorithm to a one-dimensional marine biogeochemical model of NPZD type and obtained the optimal parameter vectors lying in a wide range. Tashkova et al. (2012) estimated parameters in a nonlinear dynamic model of an aquatic ecosystem by four meta-heuristic optimization methods. Prieß et al. (2013) proposed one SBO approach to optimize parameters in a biogeochemical model of NPZD type in a single water column. Kuhn et al. (2015) estimated NPZD model parameters in different regions of the central North Atlantic by an evolutionary algorithm and the parameters varied in the space defined by the possible range of parameter values. Gharamti et al. (2017) developed an efficient data assimilation system and demonstrated that the estimated parameters varied spatially between different regions. Overall, the above *in situ* estimated parameters of the NPZD model vary by region in a range with apparent spatial variations.

Data assimilation, one approach of improving model fidelity for estimation, can determine the optimal parameter sets that minimize the difference between simulations and observations. Mattern et al. (2017) applied four-dimensional variational (4D-Var) data assimilation to improve the state of the NPZD model. The adjoint method is a typical four-dimensional variational data assimilation method and has been widely used to optimize uncertain parameters in numerical models (Qian et al., 2021; Wang et al., 2021; Wu et al., 2021). Pelc et al. (2012) provided a useful theoretical background for different 4D-Var approaches and showed how this adjoint method can be used to estimate ecosystem model parameters jointly with a large number of initial condition parameters. Lawson et al. (1996) applied the adjoint assimilation method to a five-component, time-dependent ecosystem model to get initial conditions and parameters. Gunson et al. (1999) applied the adjoint method to a 1-D marine biogeochemical model of NPZD type and adjusted parameter values *via* variational data assimilation. The variational adjoint technique was used to adjust six parameters of a five-component (phytoplankton, zooplankton, ammonium, nitrate, and detritus) ecosystem model in the study of Friedrichs (2001). Tjiputra et al. (2007) applied the adjoint method to a three-dimensional global ocean biogeochemical cycle model to optimize parameters based on Sea-viewing Wide Field-of-view Sensor (SeaWiFS) surface chlorophyll-a observation; The SeaWiFS chlorophyll-a data was assimilated into a simple NPZD model by the adjoint method in a climatological physical environment in the study of Fan and Lv (2009). Qi et al. (2011) estimated the spatially varying control parameters of a marine ecosystem dynamical model in the Bohai Sea, the Yellow Sea, and the East China Sea by using the adjoint method; Li et al. (2013) applied the adjoint variational method to a three-dimensional marine ecosystem dynamical model in North Pacific.

When using the adjoint assimilation method to treat problems of fluid flow, heat transfer, and pollutant transport, these are all governed by the convection-diffusion equation. Due to the large computational region and the long period of prediction, developing efficient and highly accurate numerical approaches to the problem is important and is a challenging task. Much effort has been made to solve convection-diffusion equations. The often-used methods include several implicit-explicit schemes, such as the first-order Lax-Friedrichs scheme (Lax, 1954), central difference scheme (Gao et al., 2015; Liu et al., 2017) and upwind difference scheme (Anderson et al., 1984; Wang et al., 2016), which are easy to implement. However, these kinds of schemes for calculating convection-diffusion problems were subject to severe Courant–Friedrichs–Lewy (CFL) restrictions (Ascher et al., 1995; Baba and Tabata, 1981; Kurganov and Tadmor, 2000; Celledoni and Kometa, 2009). Therefore, when the state variables of the NPZD model are simulated, small time step sizes have to be used, which causes a very high computational cost. The characteristic difference methods have been developed to overcome the CFL condition. Douglas and Russell (1982) first proposed a modified characteristic method to solve convection-diffusion equations. Shen et al. (2013) used the characteristic finite difference method to solve the variable-order fractional advection-diffusion

equation with a nonlinear source term. Fu et al. (2015) used an efficient time second-order characteristic finite element method for the nonlinear multicomponent aerosol dynamic equations. The characteristic difference methods incorporate the fixed Eulerian grids with Lagrangian tracking along the characteristics to treat the advective part of the equations, which allows one to use large time step sizes (Fu and Liang, 2019). The characteristic difference methods make use of the physical characteristics of the convection-diffusion equations and have no stability constraints required on the time step. Recently, several other methods were developed to achieve the numerical results of the advection-diffusion equations. Arbogast et al. (2020) developed a Runge–Kutta WENO scheme for advection–diffusion equations; Ebrahimijahan et al. (2020) proposed the compact local integrated radial basis functions (Integrated RBF) method for solving the system of non–linear advection-diffusion-reaction equations; Zhang and Ge (2021) used high-order compact difference method to solve the one-dimensional nonlinear advection diffusion reaction equation. However, for convection-diffusion problems in high dimensions, it is very difficult to achieve high order while maintaining a high order accuracy in both time and space. In this study, the adjoint method is used to estimate the spatially varying parameters of the NPZD model by combining a characteristic finite difference scheme, which permits the use of large time step sizes to get highly accurate solutions.

The paper is organized as follows. After the introduction, Section 2 describes the NPZD model and the adjoint assimilation method with the characteristic finite difference scheme. The numerical experiments are carried out and the results are analyzed in Section 3. Conclusions are given in Section 4.

Model and method

The marine ecological model used in this study is a four-compartment NPZD model. In this paper, the adjoint assimilation method includes the NPZD model, the adjoint model, and the assimilation processes. The NPZD model was used to simulate the distribution of phytoplankton with priori or adjusted parameters. Then, the optimal parameters were determined by comparing simulated values to observations. The adjoint model was used to compute the gradient of cost function on parameters. In the assimilation processes, the steepest descent method with the gradient was applied to adjust parameters.

The marine ecological model

Generally speaking, marine ecosystems are affected by physical, biological, chemical, and other processes. Based on nitrogen and dissolved inorganic nitrogen (N), phytoplankton (P), zooplankton (Z), and detritus (D), the governing equation of the ecosystem model is given as below (Gunson et al., 1999; Losa et al., 2006; Qi et al., 2011):

$$\frac{\partial C}{\partial t}+u\frac{\partial C}{\partial x}+v\frac{\partial C}{\partial y}+w\frac{\partial C}{\partial z}=\frac{\partial}{\partial x}\left(A_\rho\frac{\partial C}{\partial x}\right)+\frac{\partial}{\partial y}\left(A_\rho\frac{\partial C}{\partial y}\right)+\frac{\partial}{\partial z}\left(K_\rho\frac{\partial C}{\partial z}\right)+sms(C) \tag{1}$$

where C represents the state variables of the marine ecological model of nitrogen (N), phytoplankton (P), zooplankton (Z), and detritus (D); t is time, and x, y, z are components of the Cartesian coordinate system; u, v, w are the water velocity in the direction of x, y, z, respectively; A_ρ and K_ρ are the horizontal and vertical diffusivity coefficients, respectively. The last term on the right-hand side is the source-minus-sink term for each state variable (Franks and Chen, 2001; Fan and Lv, 2009; Qi et al., 2011) and is given by

$$sms(N)=-\frac{V_mN}{K_s+N}\cdot\frac{I}{I_o}\cdot\exp\left(1-\frac{I}{I_o}\right)\cdot AQ_{10}^{\frac{T-10}{10}}\cdot P+\theta\cdot G_m\cdot BQ_{10}^{\frac{T-10}{10}}\cdot(1-e^{-\phi P})\cdot Z+e_DD \tag{2}$$

$$sms(P)=\frac{V_mN}{K_s+N}\cdot\frac{I}{I_o}\cdot\exp\left(1-\frac{I}{I_o}\right)\cdot AQ_{10}^{\frac{T-10}{10}}\cdot P-G_m\cdot BQ_{10}^{\frac{T-10}{10}}\cdot(1-e^{-\phi P})\cdot Z-D_PP-w_p\frac{\partial P}{\partial z} \tag{3}$$

$$sms(Z)=\gamma G_m\cdot BQ_{10}^{\frac{T-10}{10}}\cdot(1-e^{-\phi P})\cdot Z-D_ZZ \tag{4}$$

$$sms(D)=(1-\gamma-\theta)G_m\cdot BQ_{10}^{\frac{T-10}{10}}\cdot(1-e^{-\phi P})\cdot Z+D_PP+D_ZZ-e_DD-w_d\frac{\partial D}{\partial z} \tag{5}$$

where T is the temperature of water; $I=I_{par}\exp(K_{ext}z)$ and I_{par} is the photosynthetically active radiation. The meaning of each term is listed in Table 1 and the values are organized according to previous studies (Franks and Chen, 2001; Franks, 2002; Fan and Lv, 2009; Qi et al., 2011).Constant boundary conditions are used at the inflow boundary Γ^{IN} , and non-gradient boundary conditions are used at the outflow boundary Γ^{OUT},

$$\begin{cases}\frac{\partial C}{\partial t}=0\ , & on\ \Gamma^{IN},\\ \frac{\partial C}{\partial n}=0\ , & on\ \Gamma^{OUT}.\end{cases} \tag{6}$$

The marine ecological equation (1) can be solved by several different numerical schemes. The central difference scheme is usually used in the adjoint assimilation method, but it is limited

TABLE 1 The meaning and initial values of parameters in the marine ecological model.

Parameter	Symbol	Value	Unit
maximum growth rate of phytoplankton (*P*)	V_m	1.0	day^{-1}
maximum growth rate of zooplankton (*Z*)	G_m	0.5	day^{-1}
mortality rate of phytoplankton (*P*)	D_p	0.1	day^{-1}
mortality rate of zooplankton (*Z*)	D_z	0.2	day^{-1}
remineralization rate of detritus (*D*)	e_D	0.05	day^{-1}
attenuation coefficient of light	K_{ext}	1.0	m^{-1}
half-saturation constant for *N* uptake	K_s	1.0	$mmol{\cdot}m^{-3}$
temperature coefficient for phytoplankton (*P*) growth at 10°C	AQ_{10}	2.08	\
temperature coefficient for zooplankton (*Z*) growth at 10°C	BQ_{10}	3.10	\
sinking velocity of phytoplankton (*P*)	w_p	0.73	$m{\cdot}day^{-1}$
sinking velocity of detritus (*D*)	w_d	1.00	$m{\cdot}day^{-1}$
Ivlev constant of zooplankton (*Z*)	ϕ	0.2	$m^3(mmol\ N)^{-1}$
assimilation ratio for zooplankton (*Z*)	γ	0.75	\
excretion ratio for zooplankton (*Z*)	θ	0.03	\

by the stability constraint and needs small time steps. To reduce the computation cost, we adopt a characteristic finite difference scheme that enables using large time steps. The variations of state variables (nitrogen (*N*), phytoplankton (*P*), zooplankton (*Z*), and detritus (*D*)) are small along the characteristic curve. Therefore, by computing along the characteristic direction, more accurate results of models can be obtained even using large time step sizes.

Let Δx , Δy and Δz be the spatial step size along *x*-, *y*-, and *z*-directions. The velocity $U=(u(x,y,z), v(x,y,z), w(x,y,z)$ is given at the center of the grid. Δt is the time step. As shown in Figure 1, let $C^n_{i,j,k}$ be the concentration of the state variables at $\vec{x} = (i\Delta x, j\Delta y, k\Delta z)$ and $t^n = n\Delta t$. Assuming that the concentration at each grid point at $t = t^n$ is known, we want to know the concentration at $t=t^{n+1}$. Let $X(\tau;\vec{x},t^{n+1})$ be the characteristic curve with the characteristic direction τ (Liang et al., 2016),

$$\frac{dX(\tau;\vec{x},t^{n+1})}{d\tau} = U\left(X(\tau;\vec{x},t^{n+1}),\tau\right), \tau \in [t^n,\ t^{n+1}], \quad (7)$$

$$X(t^{n+1};\vec{x},t^{n+1}) = \vec{x} \quad (8)$$

Denote the intersection point of $X(\tau;\vec{x},t^{n+1})$ with the time level t^n by $\vec{x}^*(\bar{x}_i,\bar{y}_j,\bar{z}_k)$ (point *D* in Figure 1). We solve $\vec{x}^*$ from the equations (7)-(8) by $\vec{x}^* = X(t^n;\vec{x},t^{n+1}) \approx \vec{x} - U(\vec{x},t^{n+1})\Delta t$

$$\vec{x}^* = X(t^n;\vec{x},t^{n+1}) \approx \vec{x} - U(\vec{x},t^{n+1})\Delta t \quad (9)$$

The concentration $\bar{C}^n_{i,j,k}$ at $\vec{x}^*$ is determined by the interpolation of the values of the points surrounding $\vec{x}^*$Then the characteristic finite difference scheme is given as:

$$\begin{cases} \frac{C1_{i,j,k}-\bar{C}^n_{i,j,k}}{\Delta t} = \frac{A_p}{\Delta x_j}\left(\left(\frac{C1_{i+1,j,k}-C1_{i,j,k}}{\Delta x_j}\right) - \left(\frac{C1_{i,j,k}-C1_{i-1,j,k}}{\Delta x_j}\right)\right) + sms\left(C^n_{i,j,k}\right), \\ \frac{C2_{i,j,k}-\overline{C1}_{i,j,k}}{\Delta t} = \frac{A_p}{\Delta y}\left(\left(\frac{C2_{i,j+1,k}-C2_{i,j,k}}{\Delta y}\right) - \left(\frac{C2_{i,j,k}-C2_{i,j-1,k}}{\Delta y}\right)\right), \\ \frac{C^{n+1}_{i,j,k}-\overline{C2}_{i,j,k}}{\Delta t} = \frac{K_p}{\Delta z_k}\left(\left(\frac{C^{n+1}_{i,j,k+1}-C^{n+1}_{i,j,k}}{\Delta z_{k+\frac{1}{2}}}\right) - \left(\frac{C^{n+1}_{i,j,k}-C^{n+1}_{i,j,k-1}}{\Delta z_{k-\frac{1}{2}}}\right)\right), \end{cases} \quad (10)$$

where

$$\begin{cases} \bar{C}^n_{i,j,k} = \frac{(\bar{x}_i - x_p)(\bar{x}_i - x_{p+1})}{2\Delta x_j^2} C^n_{p-1,j,k} - \frac{(\bar{x}_i - x_{p-1})(\bar{x}_i - x_{p+1})}{\Delta x_j^2} C^n_{p,j,k} \\ + \frac{(\bar{x}_i - x_{p-1})(\bar{x}_i - x_p)}{2\Delta x_j^2} C^n_{p+1,j,k}, \quad \left(x_p + \frac{\Delta x_j}{2} < \bar{x}_i < x_{p+1} - \frac{\Delta x_j}{2}\right) \\ \overline{C1}_{i,j,k} = \frac{(\bar{y}_j - y_q)(\bar{y}_j - y_{q+1})}{2\Delta y^2} C1_{i,q-1,k} - \frac{(\bar{y}_j - y_{q-1})(\bar{y}_j - y_{q+1})}{2\Delta y^2} C1_{i,q,k} \\ + \frac{(\bar{y}_j - y_{q-1})(\bar{y}_j - y_q)}{2\Delta y^2} C1_{i,q+1,k}, \quad \left(y_q + \frac{\Delta y}{2} < \bar{y}_j < y_{q+1} - \frac{\Delta y}{2}\right) \\ \overline{C2}_{i,j,k} = \frac{(\bar{z}_k - z_l)(\bar{z}_k - z_{l+1})}{2\Delta z_l \Delta z_{l+1}} C2_{i,j,l-1} - \frac{(\bar{z}_k - z_{l-1})(\bar{z}_k - z_{l+1})}{2\Delta z_{l-1}\Delta z_{l+1}} C2_{i,j,l} \\ + \frac{(\bar{z}_k - z_{l-1})(\bar{z}_k - z_l)}{2\Delta z_{l-1}\Delta z_l} C2_{i,j,l+1}, \quad (z_l + \frac{\Delta z_l}{2} < \bar{z}_k < z_{l+1} - \frac{\Delta z_{l+1}}{2}) \end{cases} \quad (11)$$

The adjoint model

According to the adjoint method, a cost function is defined to describe the difference between the simulated and observed surface phytoplankton:

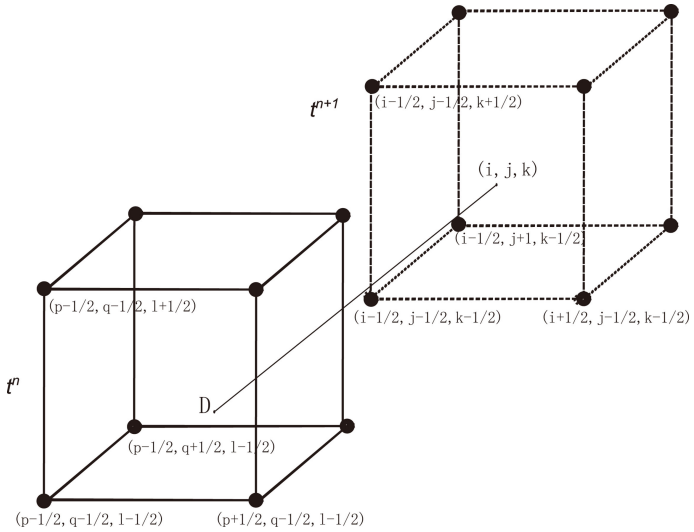


FIGURE 1
The process of constructing characteristic finite difference schemes.

$$J(\vec{S}) = \frac{1}{2}\int_{\Omega\times T} K(P - P_{obs})^2 d\Omega dt \tag{12}$$

where P is the simulated surface phytoplankton by the NPZD model and P_{obs} is the observed surface phytoplankton; Ω denotes the spatial domain and T is the time domain; K is the weighting matrix and the elements in K are 1 where observations are available and 0 otherwise. Rewrite equation (1),

$$F(C,\vec{S}) = \frac{\partial C}{\partial t} + u\frac{\partial C}{\partial x} + v\frac{\partial C}{\partial y} + w\frac{\partial C}{\partial z} - \frac{\partial}{\partial x}\left(A_\rho \frac{\partial C}{\partial x}\right) - \frac{\partial}{\partial y}\left(A_\rho \frac{\partial C}{\partial y}\right) - \frac{\partial}{\partial z}\left(K_\rho \frac{\partial C}{\partial z}\right) - sms(C) \tag{13}$$

where the parameters in the NPZD model are represented by $\vec{S}$. Based on the Lagrange multiplier method (Thacker and Long, 1988), the Lagrangian function is defined as:

$$L(\vec{\lambda}, N, P, Z, D, \vec{S}) = J(\vec{S}) + \langle \vec{\lambda}, F \rangle = J(\vec{S}) + \int_{\Omega\times T} (\lambda_P \cdot F(P,\vec{S}) + \lambda_N \cdot F(N,\vec{S}) + \lambda_Z \cdot F(Z,\vec{S}) + \lambda_D \cdot F(D,\vec{S}))\, d\Omega dt, \tag{14}$$

where $\vec{\lambda} = (\lambda_P, \lambda_N, \lambda_Z, \lambda_D)$ is the Lagrange multiplier of the state variables (nitrogen (N), phytoplankton (P), zooplankton (Z) and detritus (D)), respectively. Based on Lagrange multiplier theory, the first-order derivatives of the Lagrange function should be zero to minimize the cost function:

$$\frac{\partial L}{\partial \lambda_P} = 0, \quad \frac{\partial L}{\partial \lambda_N} = 0, \quad \frac{\partial L}{\partial \lambda_Z} = 0, \quad \frac{\partial L}{\partial \lambda_D} = 0, \tag{15}$$

$$\frac{\partial L}{\partial P} = 0, \quad \frac{\partial L}{\partial N} = 0, \quad \frac{\partial L}{\partial Z} = 0, \quad \frac{\partial L}{\partial D} = 0, \tag{16}$$

$$\frac{\partial L}{\partial \vec{S}} = 0 \tag{17}$$

Equation (15) is equation (1) of the NPZD model. The adjoint equations can be derived from (16),

$$-\frac{\partial \vec{\lambda}}{\partial t} - \frac{\partial}{\partial z}\left(K_\rho \frac{\partial \vec{\lambda}}{\partial z}\right) = \frac{\partial}{\partial x}(u\vec{\lambda}) + \frac{\partial}{\partial y}(v\vec{\lambda}) + \frac{\partial}{\partial z}(w\vec{\lambda}) + \frac{\partial}{\partial x}\left(A_\rho \frac{\partial \vec{\lambda}}{\partial x}\right) + \frac{\partial}{\partial y}\left(A_\rho \frac{\partial \vec{\lambda}}{\partial y}\right) + sms(\vec{\lambda}), \tag{18}$$

where the last term of each state variable's Lagrange multipliers (Fan and Lv, 2009; Qi et al., 2011) is

$$sms(\lambda_N) = \frac{V_m K_s}{(K_s + N)^2} \cdot \frac{I}{I_0} \cdot \exp\left(1 - \frac{I}{I_0}\right) \cdot AQ_{10}^{\frac{T-10}{10}} \cdot P \cdot (\lambda_P - \lambda_N) \tag{19}$$

$$sms(\lambda_P) = w_p \frac{\partial \lambda_P}{\partial z} + \frac{V_m N}{K_s + N} \cdot \frac{I}{I_0} \cdot \exp\left(1 - \frac{I}{I_0}\right) \cdot AQ_{10}^{\frac{T-10}{10}}(\lambda_P - \lambda_N) - D_P \cdot (\lambda_P - \lambda_D) - \phi \cdot e^{-\phi P} \cdot G_m \cdot BQ_{10}^{(T-10)/10} \cdot Z \cdot [\lambda_P - \gamma \cdot \lambda_Z - \theta \cdot \lambda_N - (1 - \gamma - \theta) \cdot \lambda_D] - K(P - P_{obs}) \quad (20)$$

$$sms(\lambda_Z) = -G_m \cdot BQ_{10}^{(T-10)/10}(1 - e^{-\phi P})[\lambda_P - \gamma \cdot \lambda_Z - \theta \cdot \lambda_N - (1 - \gamma - \theta) \cdot \lambda_D] - D_Z \cdot (\lambda_Z - \lambda_D), \quad (21)$$

$$sms(\lambda_D) = w_d \frac{\partial \lambda_D}{\partial z} + e_D \cdot (\lambda_N - \lambda_D) \quad (22)$$

We propose the characteristic finite difference schemes of (18)

$$\begin{cases} \frac{\vec{\lambda}1_{i,j,k} - \overline{\vec{\lambda}}^{n+1}_{i,j,k}}{\Delta t} = \frac{A_\rho}{\Delta x_j}\left(\left(\frac{\vec{\lambda}1_{i+1,j,k} - \vec{\lambda}1_{i,j,k}}{\Delta x_j}\right) - \left(\frac{\vec{\lambda}1_{i,j,k} - \vec{\lambda}1_{i-1,j,k}}{\Delta x_j}\right)\right) + sms\left(\vec{\lambda}^{n+1}_{i,j,k}\right), \\ \frac{\vec{\lambda}2_{i,j,k} - \overline{\vec{\lambda}}1_{i,j,k}}{\Delta t} = \frac{A_\rho}{\Delta y}\left(\left(\frac{\vec{\lambda}2_{i,j+1,k} - \vec{\lambda}2_{i,j,k}}{\Delta y}\right) - \left(\frac{\vec{\lambda}2_{i,j,k} - \vec{\lambda}2_{i,j-1,k}}{\Delta y}\right)\right), \\ \frac{\vec{\lambda}^n_{i,j,k} - \overline{\vec{\lambda}2}_{i,j,k}}{\Delta t} = \frac{K_\rho}{\Delta z_k}\left(\left(\frac{\vec{\lambda}^n_{i,j,k+1} - \vec{\lambda}^n_{i,j,k}}{\Delta z_{k+\frac{1}{2}}}\right) - \left(\frac{\vec{\lambda}^n_{i,j,k} - \vec{\lambda}^n_{i,j,k-1}}{\Delta z_{k-\frac{1}{2}}}\right)\right), \end{cases} \quad (23)$$

where,

$$\begin{cases} \overline{\vec{\lambda}}^{n+1}_{i,j,k} = \frac{(\bar{x}_i - x_p)(\bar{x}_i - x_{p+1})}{2\Delta x_j^2}\vec{\lambda}^{n+1}_{p-1,j,k} - \frac{(\bar{x}_i - x_{p-1})(\bar{x}_i - x_{p+1})}{\Delta x_j^2}\vec{\lambda}^{n+1}_{p,j,k} + \frac{(\bar{x}_i - x_{p-1})(\bar{x}_i - x_p)}{2\Delta x_j^2}\vec{\lambda}^{n+1}_{p+1,j,k}, \quad \left(x_p + \frac{\Delta x_j}{2} < \bar{x}_i < x_{p+1} - \frac{\Delta x_j}{2}\right) \\ \overline{\vec{\lambda}1}_{i,j,k} = \frac{(\bar{y}_j - y_q)(\bar{y}_j - y_{q+1})}{2\Delta y^2}\vec{\lambda}1_{i,q-1,k} - \frac{(\bar{y}_j - y_{q-1})(\bar{y}_j - y_{q+1})}{2\Delta y^2}\vec{\lambda}1_{i,q,k} + \frac{(\bar{y}_j - y_{q-1})(\bar{y}_j - y_q)}{2\Delta y^2}\vec{\lambda}1_{i,q+1,k}, \quad \left(y_q + \frac{\Delta y}{2} < \bar{y}_j < y_{q+1} - \frac{\Delta y}{2}\right) \\ \overline{\vec{\lambda}2}_{i,j,k} = \frac{(\bar{z}_k - z_l)(\bar{z}_k - z_{l+1})}{2\Delta z_l \Delta z_{l+1}}\vec{\lambda}2_{i,j,l-1} - \frac{(\bar{z}_k - z_{l-1})(\bar{z}_k - z_{l+1})}{2\Delta z_{l-1}\Delta z_{l+1}}\vec{\lambda}2_{i,j,l} + \frac{(\bar{z}_k - z_{l-1})(\bar{z}_k - z_l)}{2\Delta z_{l-1}\Delta z_l}\vec{\lambda}2_{i,j,l+1}, \quad \left(z_l + \frac{\Delta z_l}{2} < \bar{z}_k < z_{l+1} - \frac{\Delta z_{l+1}}{2}\right) \end{cases} \quad (24)$$

In the study of Fan and Lv (2009), the cost function is more sensitive to five constant parameters of V_m (maximum growth rate of phytoplankton), G_m (maximum grazing rate of zooplankton), D_p (mortality rate of phytoplankton), D_z (mortality rate of zooplankton) and e_D (remineralization rate of detritus) by sensitivity analysis. Therefore, these five parameters are selected to be optimized. From equation (17), the gradients of the cost function concerning the five constant parameters (Fan and Lv, 2009; Qi et al., 2011) are obtained as follows:

$$\frac{\partial J}{\partial V_m} = -\frac{N}{K_s + N} \cdot \frac{I}{I_0} \exp\left(1 - \frac{I}{I_0}\right) \cdot AQ_{10}^{\frac{T-10}{10}} \cdot P \cdot (\lambda_p - \lambda_N) \quad (25)$$

$$\frac{\partial J}{\partial G_m} = BQ_{10}^{\frac{T-10}{10}} \cdot (1 - \exp(-\phi P)) \cdot Z \cdot [\lambda_p - \gamma\lambda_Z - \theta\lambda_N - (1 - \gamma - \theta)\lambda_D] \quad (26)$$

$$\frac{\partial J}{\partial D_P} = P(\lambda_p - \lambda_D) \quad (27)$$

$$\frac{\partial J}{\partial D_Z} = Z(\lambda_Z - \lambda_D) \quad (28)$$

$$\frac{\partial J}{\partial e_D} = -D(\lambda_N - \lambda_D) \quad (29)$$

When the gradients are calculated, the spatially varying parameters are estimated using the steepest descent method. The details are shown in Wang et al. (2020).

Numerical experiments

Numerical experiment 1: The Gaussian pulse moving

To test the performance of the characteristic finite difference scheme, the transport of a Gaussian hump was simulated by the characteristic finite difference scheme and the results were compared to those simulated by the central difference scheme.

In this subsection, the transport of the Gaussian pulse of the problem (1) with $sms(C)$=0 was considered. The spatial domain is Ω=[0,2]×[0,2] and time domain is $t\in$(0,0.2]. The initial condition is given by

$$c(x, y, 0) = \exp\left[-\frac{(x - x_0)^2 + (y - y_0)^2}{2\sigma_0^2}\right], \quad (30)$$

where (x_0,y_0)=(0.3,0.3) , $\sigma_0 = \sqrt{2}/20$ and the velocity is u=v=6 . The exact solution was

$$c(x, y, t) = \frac{\sigma_0^2}{\sigma_0^2 + 2A_\rho t}\exp\left[-\frac{(x - x_0 - ut)^2 + (y - y_0 - vt)^2}{2(\sigma_0^2 + 2A_\rho t)}\right], \quad (31)$$

The diffusion coefficient is chosen as A_ρ=0.001. Let $C^n(x,y)$ denotes the approximate solution. The errors L_∞ and L_2 are defined as follows:

$$E_\infty^n = \max_{i,j}\{|c(x_i, y_i, t^n) - C^n(x_i, y_i)|\} \quad (32)$$

$$E_2^n = \sqrt{\sum_{i,j}\Delta x \Delta y (c(x_i, y_i, t^n) - C^n(x_i, y_i))^2} \quad (33)$$

To examine the convergence rates in time of the characteristic finite difference scheme and central difference scheme, the small spatial step sizes Δx=Δy=1/500 and the time step sizes Δt=T/5,T/10,T/15,T/20 were carried out. The

numerical results are presented in Table 2. When the large time steps were used, the results of the central difference scheme (CDS) were not stable. After adopting small time steps (Δt=T/20000,T/21000,T/22000,T/23000), the stable results of CDS are listed in Table 2. It is clearly shown that the characteristic finite difference method (CFDM) and CDS exhibited a first-order ratio of convergence in time. Moreover, the L_∞ and L_2 errors of CFDM with the large time steps were less than those of CDS using the small time steps, indicating the high accuracy of CFDM. For example, when using Δt=T/5, the CFDM produced 7.5769×10^{-4} of E_∞ and 6.2334×10^{-5} of E_2, while the E_∞ and E_2 of CDS using Δt=T/20000 were 1.3394×10^{-2} and 1.4925×10^{-3}, respectively.To further explore the effectiveness of CFDM, the transport of the Gaussian pulse of the problem (1) with $sms(C)$=0 in three-dimensional space was considered. The initial condition is given as:

$$c(x,y,z,0)=\exp\left[-\frac{(x-x_0)^2}{A_\rho}-\frac{(y-y_0)^2}{A_\rho}-\frac{(z-z_0)^2}{K_\rho}\right] \quad (34)$$

The exact solution of the problem is

$$c(x,y,z,t)=\left(\frac{1}{\sqrt{4t+1}}\right)^3 \quad (35)$$

$$\cdot\exp\left[-\frac{(x-x_0-ut)^2}{A_\rho(4t+1)}-\frac{(y-y_0-vt)^2}{A_\rho(4t+1)}-\frac{(z-z_0-wt)^2}{K_\rho(4t+1)}\right]$$

where, the spatial domain is Ω=[0,2]×[0,2]×[0,2], the initial center is (x_0,y_0,z_0)=(0.5,0.5,0.5), the velocity is (u,v,w)=(0.8,0.8,0.1), and the diffusivity coefficient is A_ρ=K_ρ=0.01. The spatial step sizes were taken as Δx=Δy=Δz=0.025. The time step size for CFDM was Δt=0.0625. With the limit of stability, a smaller time step size of Δt=0.00625 was set for CDS.

A group of surfaces obtained by CFDM and CDS in the section of y=x and contour plots at z = 0.6 at different times of t= 0.75, 1, and 1.25, which displayed the concentration distributions are shown in Figures 2, 3. The results calculated by CFDM almost coincided with the exact solutions and were more accurate than those of CDS with 10 times smaller time step, indicating that the Gaussian pulse could be simulated by CFDM very well even using a much large time step. Furthermore, the computed peaks and numerical errors in L_∞ and L_2 norms at different times shown in Table 3 demonstrated that the model performance was improved by CFDM. For example, at time t=1.25, where the maximum value of the exact solution was 0.0675, the computed peak of CFDM was 0.0707, while that of CDS was 0.0827. Besides, at time t=1.25, the errors of CFDM were 3.1538×10^{-3} in L_∞ -norm and 1.7879×10^{-3} in L_2-norms, which were one order of magnitude smaller than those of CDS (1.5155×10^{-2} in L_∞ -norm and 1.3649×10^{-2} in L_2 -norm). Overall, the simulation results of CFDM were much closer to the exact solutions than those of CDS, demonstrating that CFDM could significantly improve the model performance.

Numerical experiment 2: Parameter optimization of NPZD model

In this subsection, the adjoint assimilation method combining with CFDM was used to optimize the spatial parameters SP (V_m, G_m, D_p, D_z, e_D). The process of the numerical experiment can be described as:

Step 1. Give the guess of distribution of parameters SP^0 as the initialization of parameters.

Step 2. For n = 0, 1, ..., N (the number of iterations), do Step 3-Step 5.

Step 3. With the initial distribution of parameters SP^n, solve the NPZD model and get simulated results.

TABLE 2 Errors and ratios in time of the 2D Gaussian pulse for the experiment using large time steps by the CFDM and CDS method and using small time steps calculated by the CDS method.

	Δt	T/5	T/10	T/15	T/20
CFDM	E_∞	7.5769e-4	3.8919e-4	2.6460e-4	2.0196e-4
	Ratio	–	0.9611	0.9516	0.9389
	E_2	6.2334e-5	3.2124e-5	2.1990e-5	1.6945e-5
	Ratio	–	0.9564	0.9348	0.9060
CDS	E_∞	4.7570e+4	8.4835e+11	9.7634e+17	2.6198e+24
	Ratio	–	-2.4088	-3.4420	-5.1455
	E_2	2.5342e+3	2.6668e+10	3.7603e+16	1.9742e+22
	Ratio	–	-2.3327	-3.4921	-4.5784
	Δt	**T/20000**	**T/21000**	**T/22000**	**T/23000**
CDS	E_∞	1.3394e-2	1.2821e-2	1.2301e-2	1.1832e-2
	Ratio	–	0.8961	0.8899	0.8744
	E_2	1.4925e-3	1.4276e-3	1.3691e-3	1.3160e-3
	Ratio	–	0.9106	0.9004	0.8899

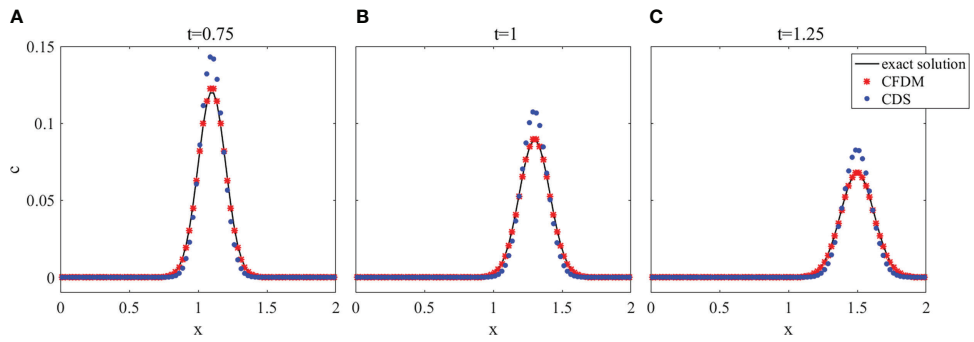


FIGURE 2
The surface plots in the section of $y = x$ of the 3D Gaussian hump at $z = 0.6$ with different times. **(A)** t=0.75, **(B)** t=1, **(C)** t=1.25.

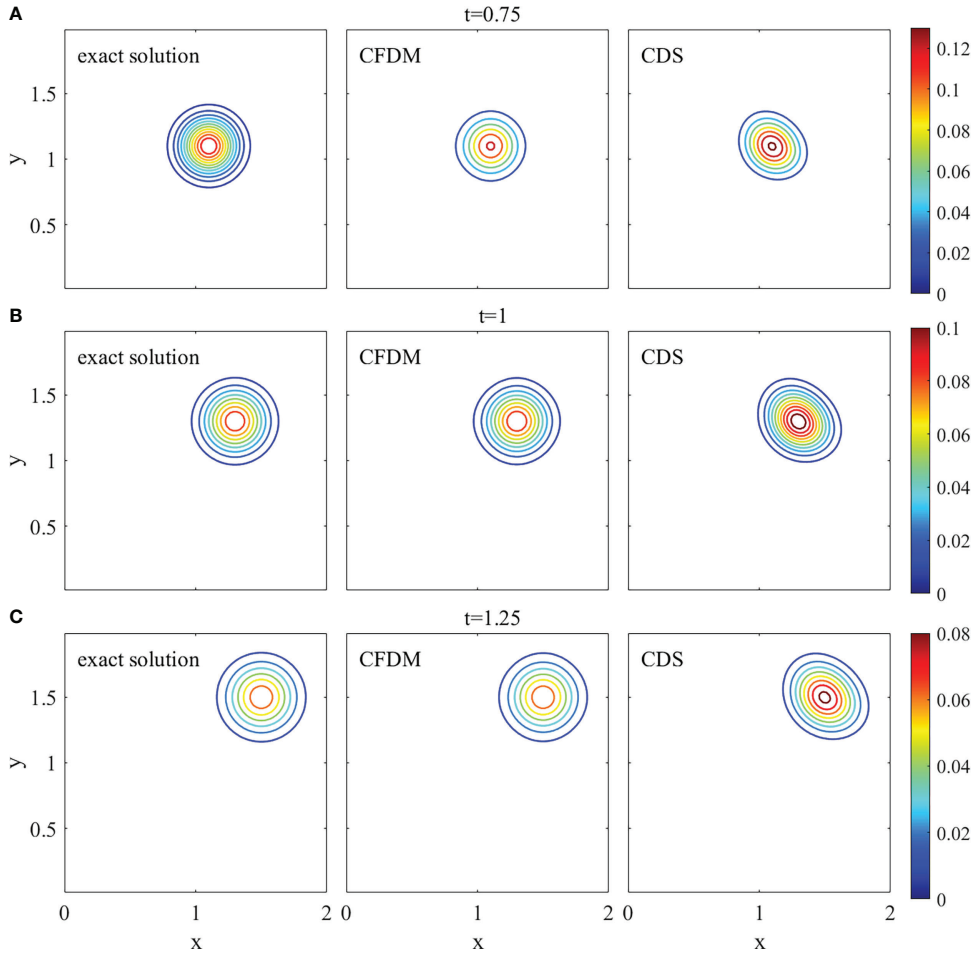


FIGURE 3
The contour plots of the 3D Gaussian hump at $z = 0.6$ with different times **(A)** t=0.75, **(B)** t=1, **(C)** t=1.25.

Step 4. Solve the cost function (12) by the simulations and observations. If the cost function decreases to 10^{-5}, exit the loop and run Step 6.

Step 5. Run the adjoint model backward in the time direction and calculate the gradient of the cost function on the parameters SP^{n}. Then get the optimized parameters SP^{n+1} by adjusting SP^{n} with the steepest descent method.

Step 6. Output the final optimized parameters.

Experiment design

The studied region (24°N–41°N,117.5°E–131°E) covers the Bohai Sea, the Yellow Sea, and the East China Sea. The horizontal resolution is 10′×10′ with grid numbers of 103 (south-north) ×82 (west-east). The vertical direction is 6 layers and the thickness of each layer from top to bottom is 10m, 10m, 10m, 20m, 25m, and 25m respectively. The data over the simulation period, such as the ambient physical velocities, the temperature, and eddy diffusivities, etc., were interpolated to the vertical grid using the results obtained by the three-dimensional Princeton Ocean Model (POM) (Blumberg and Mellor, 1987). In addition, river sources of Changjiang (http://xxfb.hydroinfo.gov.cn) were added to the model. The initial fields of nitrogen and phytoplankton were converted from monthly mean nitrate of World Ocean Atlas 2013 (WOA13) and monthly mean surface chlorophyll-a concentrations of SeaWiFS, respectively. The starting time was 1 January 2016. Besides, the NPZD model was run for 30 days for spin-up to obtain the initial fields of zooplankton and detritus. Then the adjoint model was run for 5 days backward in time from 1 February 2016 with the spin-up ocean state as the initial conditions.

During the experiment, four state variables of the marine ecological model of nitrogen (N), phytoplankton(P), zooplankton (Z), and detritus (D) were converted to nitrogen units($\mathrm{mmol\,N{\cdot}m^{-3}}$). For the NPZD model, satellite chlorophyll-a data ($\mathrm{mg{\cdot}m^{-3}}$) of SeaWiFS was converted to nitrogen units ($\mathrm{mmol\,N{\cdot}m^{-3}}$) using an equation (36) of the relation between chlorophyll-a and carbon proposed by Semovski and Wozniak (1995) and a constant phytoplankton carbon-to-nitrogen Redfield ratio (Redfield et al., 1963; Faugeras et al., 2004) $\frac{106\ \mathrm{mol\ C}}{16\ \mathrm{mol\ N}}$

$$C = \rho_{\max}\frac{Chl_a}{Chl_a + K_{1/2}}Chl_a \tag{36}$$

where $\rho_{\max}$=90 and the half-saturation coefficient is $K_{1/2}$=0.477 As indicated by Moisan et al. (2002), the parameters of the marine ecological model were related to temperature. Therefore, according to the initial values of parameters listed in Table 1 and the trend of the surface temperature field, two types of parameters were given to verify the accuracy of the adjoint data assimilation with CFDM.

Type 1:

$$\alpha(lon, lat) = [1.517 - 0.01083 \times lon(i) + 0.02508 \times lat(j)]\alpha_0 \tag{37}$$

Type 2:

$$\alpha(lon, lat) = [0.4827 + 0.01083 \times lon(i) - 0.02508 \times lat(j)]\alpha_0 \tag{38}$$

where α_0 was the initial value of the parameter shown in Table 1.

Single parameter inversion

To test the effectiveness of the adjoint data assimilation with CFDM, the experiments of single parameter inversion were carried out. In these experiments, one of the five parameters (V_m, G_m, D_p, D_z, e_D) was assumed to vary spatially by Type 1 or Type 2, and the other four parameters were constants in Table 1. The NPZD model was run for 5 days from 1 February 2016 and the time step was $\Delta t = 6h$. During the experiment, only the parameter with given spatial distributions was inversed, and initial guesses were set as the default values shown in Table 1. The observations used were the simulated concentrations of phytoplankton by the NPZD model with five assumed parameters. The calculated mean absolute errors (MAE) and mean relative errors (MRE) between the given parameter and the corresponding inversed parameter values are listed in Table 4. When the time step size was set to $6h$, the numerical oscillations occurred in CDS and the MAE and MRE were one order of magnitude larger than those of CFDM. Therefore, the experiment of CDS ($\Delta t = 30min$) was carried out.

As listed in Table 4, the MAEs between the inverted parameters (V_m, G_m, D_p, D_z, e_D) and the spatial values varying as Type 1 of CFDM ($\Delta t = 6h$) were 1.9907×10^{-3} , 1.1762×10^{-3} , 2.7284×10^{-4} , 1.2326×10^{-3} ,and 1.8694×10^{-4} $\mathrm{day^{-1}}$, respectively; the MAEs of CDS (Δt=30min) were much larger values of 7.9687×10^{-3}, 6.6683×10^{-3}, 1.1074×10^{-3}, 5.5154×10^{-3}, and

TABLE 3 The maximum and minimum values and errors of the 3D moving Gaussian pulse.

T	Exact solution		CFDM				CDS			
	Max.	Min.	Max.	Min.	L_∞-error	L_2-error	Max.	Min.	L_∞-error	L_2-error
0.75	0.1197	0.0	0.1265	0.0	7.7320e-3	3.2321e-3	0.1431	0.0	2.3984e-2	1.6208e-2
1.00	0.0886	0.0	0.0932	0.0	4.7603e-3	2.3515e-3	0.1074	0.0	1.8787e-2	1.4855e-2
1.25	0.0675	0.0	0.0707	0.0	3.1538e-3	1.7879e-3	0.0827	0.0	1.5155e-2	1.3649e-2

1.3193×10^{-3} day^{-1}, respectively. The mean MRE of the five parameters between the given Type1 and corresponding inverted results was reduced to 0.36 % in CFDM (Δt=6h) from 2.09 % in CDS (Δt=30min). Similar results could also be obtained in the Type2 experiments, indicating that the model performance was improved by CFDM even using large time steps.

Simultaneous inversion of five parameters

As indicated by Li et al. (2013), the distributions of V_m, D_z, and e_D were consistent and G_m and D_p had a similar distribution. Therefore, the five parameters were supposed to be spatially varying, which were estimated synchronously. In this subsection, the values of V_m, D_z, e_D were assumed to vary spatially by Type1, G_m, and D_p were distributed by Type 2. The observations used for data assimilation were the simulated concentrations of phytoplankton by the NPZD model with five assumed parameters. To verify the effect of data assimilation with CFDM, the data assimilation was implemented using CFDM with Δt=6h and CDS with Δt=30min, and the initial guess values of the five parameters were set to the default values shown in Table 1 in Case 1.

In addition, previous studies have indicated that reasonable initial guesses of target parameters can accelerate the optimization rate and improve the simulations. Therefore, the sensitivity analysis of the initial parameter values was designed in this subsection. The initial parameter values were set as 0.8 and 1.2 of the default values shown in Table 1 in Case 2 and Case 3. The distributions of inversion results and given parameters are shown in Figure 4. The estimated results of the three experiments were all in good agreement with the prescribed parameters. The cost function normalized by the value at the first iteration step was shown in Figure 5. The normalized cost functions of CFDM in Case 1 to Case 3 were both reduced by at least three orders of magnitude, which dropped faster than those of CDS. However, compared to Case 2 and Case 3, Case 1 had the highest efficiency of convergence, and only 42 iteration steps were used to reach the minimum. Consequently, the initial guess had a great impact on the efficiency of convergence and should be reasonably selected. In the following experiments, the initial guesses were all set as the default values shown in Table 1.

The errors in Case 1 are listed in Table 5. The inversed parameters after data assimilation with CFDM and CDS had similar spatial features with the given values, in which the high errors of CDS were near the south boundaries; conversely, the estimated parameters of CFDM were consistent with the given spatial variations. Besides, in CDS (Δt=30min), the MAEs between the inverted parameters (V_m, G_m, D_p, D_z, e_D) and the spatial values were increased to 8.0097×10^{-3}, 1.0031×10^{-2}, 1.6888×10^{-3}, 5.1425×10^{-3}, and 1.3063×10^{-3} day^{-1}, respectively, which were larger than 3.1432×10^{-3}, 4.6792×10^{-3}, 9.5625×10^{-4}, 1.4997×10^{-3}, and 6.1339×10^{-4} day^{-1} of CFDM (Δt=6h). The mean MRE of the five parameters was reduced to 0.88 % by CFDM (Δt=6h) from 2.08 % by CDS (Δt=30min).

In addition, The MAE between simulated chlorophyll-a and observations and running time of CFDM and CDS in Case 1 are listed in Table 6. For CFDM, the normalized cost function was less than 0.0016 after running 1351.55 s. For CDS, the normalized cost function was less than 0.0826 after running 26697.83 s. The MAE between the simulated phytoplankton and chlorophyll-a was 0.2434 mmol $N{\cdot}m^{-3}$ (0.2823 mmol $N{\cdot}m^{-3}$) before data assimilation in CFDM (CDS); after assimilation, the MAE was decreased to 0.0036 mmol $N{\cdot}m^{-3}$ (0.0197 mmol $N{\cdot}m^{-3}$) in CFDM (CDS), indicating that the model performance was improved with a reduction of 98.52 % (93.02 %) in overall simulation error. Using about 1/30 of the time of the CDS method, more accurate results were obtained by CFDM, which further shows that CFDM can optimize five parameters efficiently.

Influence of errors of observations

In all the above experiments, the observations of chlorophyll-a were perfect without errors. However, in practice, there might be

TABLE 4 The errors between the inversion results and the given parameters in the experiment of optimizing parameters separately.

	Parameter	CFDM (Δt = 6h)		CDS (Δt = 6h)		CDS (Δt = 30min)	
		MAE (day^{-1})	MRE (%)	MAE (day^{-1})	MRE (%)	MAE (day^{-1})	MRE (%)
Type1	V_m	1.9907e-3	0.2077	5.2134e-2	6.8101	7.9687e-3	0.9491
	G_m	1.1762e-3	0.2481	2.9389e-2	6.8627	6.6683e-3	1.6249
	D_p	2.7284e-4	0.2967	5.3237e-3	6.2515	1.1074e-3	1.3328
	D_z	1.2326e-3	0.6673	1.2844e-2	7.4986	5.5154e-3	3.3307
	e_D	1.8694e-4	0.4009	3.4199e-3	8.0113	1.3193e-3	3.2013
Type2	Vm	5.5599e-3	0.5035	6.0603e-2	5.3297	9.3864e-3	0.8142
	Gm	1.2637e-3	0.2444	2.8248e-2	5.0008	8.4338e-3	1.4545
	Dp	2.2292e-4	0.2172	6.5089e-3	5.7599	1.5961e-3	1.3772
	Dz	1.4915e-3	0.6978	1.2206e-2	5.3740	4.4056e-3	1.8926
	eD	2.0576e-4	0.3896	2.9709e-3	5.1449	1.0034e-3	1.7255

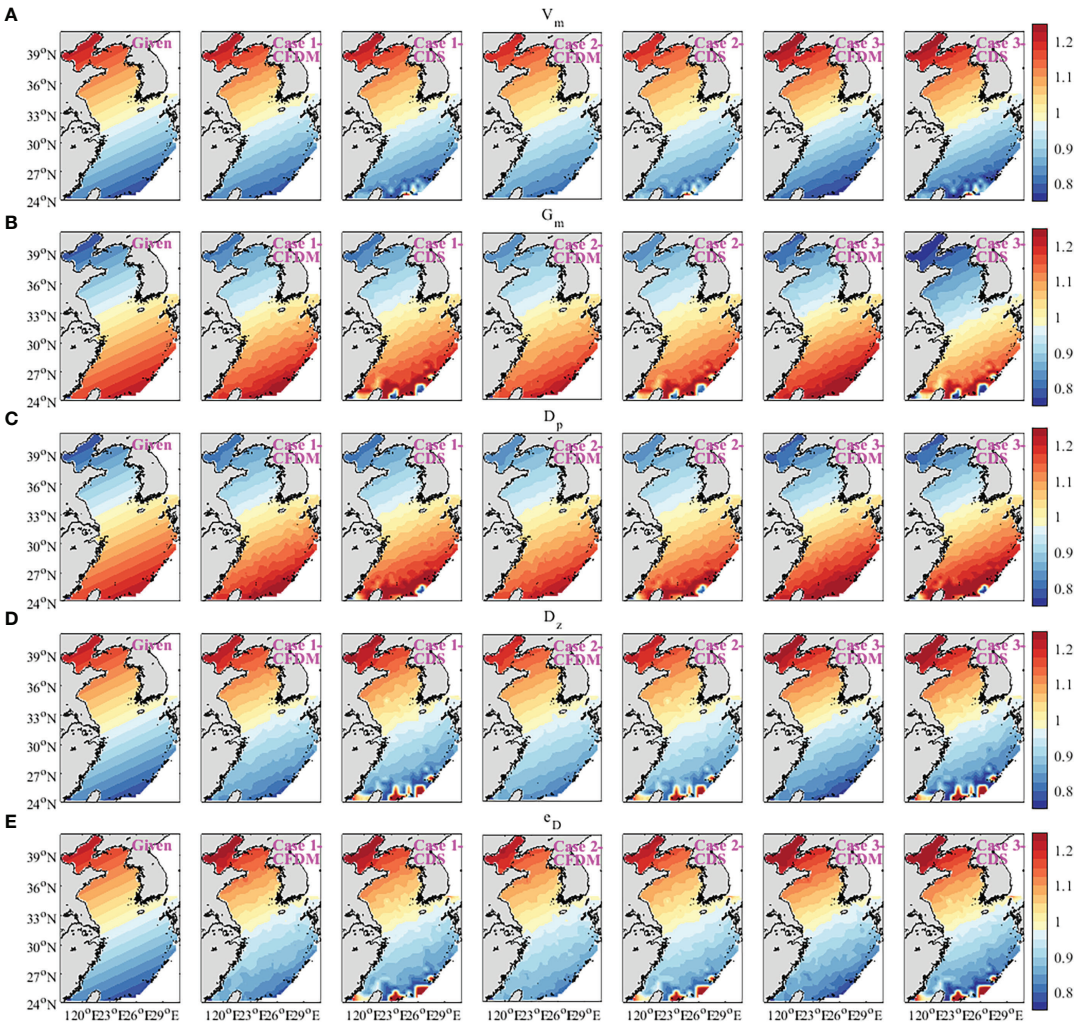


FIGURE 4
Spatial variation of five inversion parameters in the experiment of optimizing parameters simultaneously (unit: day^{-1}). **(A)** V_m, **(B)** G_m, **(C)** D_p, **(D)** D_z, and **(E)** e_D.

20%-30% errors in the chlorophyll-a concentrations obtained by SeaWiFS (Gregg and Casey, 2004; Cui et al., 2014), due to its digitization round-off and noise errors (Hu et al., 2001). If the available observations are not accurate, the estimation of parameters might not conform to the actual distribution. Thus, it is necessary to discuss the influence of observation errors on the estimated parameters. To partly reflect reality, the observations over random error within ±20%, ±30%, ±35%, and ±40% were considered in Case 4 to Case 7. In these experiments, the values of V_m, D_z and e_D were assumed to vary spatially by Type1, G_m and D_p were distributed by Type 2. During the experiment, five parameters were inversed simultaneously and initial guesses was set as the default values shown in Table 1. The observations used for data assimilation were the simulations of phytoplankton by the NPZD model with five assumed parameters with ±20%, ±30%, ±35%, and ±40% random errors.For experiments of Case 4 to Case 7, when the percentage error became larger, the estimated results worsened (Figure 6). With the 30% errors, the estimated results in CFDM (Δt=6h) were close to the prescribed parameters, and the high values of errors only occurred near the boundary. There were lots of misfits between the estimated results in CDS (Δt=30min) and the prescribed parameters. When the maximum percentage of errors was up to 35%, the estimations of e_D differed greatly from prescribed e_D. When the maximum percentage of errors was up to 40%, the estimations of G_m differed greatly from the prescribed G_m.

As listed in Table 7, with the 30% errors, the estimated parameters contained more errors than in experiments without observation errors, where the maximum MRE of CFDM (Δt=6h) between the estimated results and the prescribed parameters decreased from 1.34% to 2.92%, indicating that the estimated

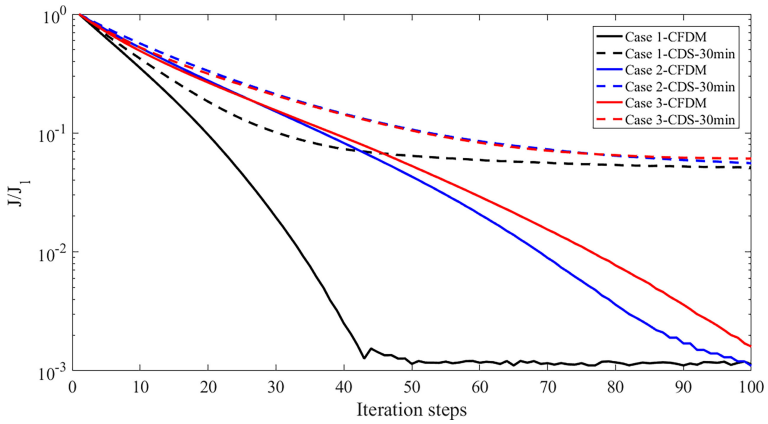


FIGURE 5
The values of the normalized cost function (J/J_1) versus the iteration steps.

parameters were still acceptable. The errors of CFDM were smaller than those of the CDS method. For example, the MAE and MRE of V_m obtained by CFDM were 9.6420×10^{-3} day^{-1} and 1.0184% respectively, while those of the CDS method were $3.0106\times10^{-2}day^{-1}$ and 3.2083%. Based on this analysis, we concluded that the degree of errors of chlorophyll-a obtained by SeaWiFS used in this paper was acceptable and the present model and method could partly bear the influence of errors of observations.

Practical experiments and results analysis

In practical experiments, the monthly mean climatological SeaWiFS data of the period 1997-2016 were interpolated into daily data and the daily SeaWiFS data (Sathyendranath et al., 2020) was used to correct the interpolated data. These observations were assimilated to optimize the parameters (V_m, G_m, D_p, D_z, e_D) and to improve simulation results. The CFDM was selected in the optimization algorithm. The initial conditions of nitrogen and phytoplankton were obtained through spatial interpolation of the

TABLE 5 Same as Table 4, but for the experiment of Case 1 of optimizing parameters simultaneously.

Method	Parameter	MAE of parameters (day^{-1})	MRE of parameters (%)
CFDM	V_m	3.1432e-3	0.3388
	G_m	4.6792e-3	0.9405
	D_p	9.5625e-4	0.9356
	D_z	1.4997e-3	0.8363
	e_D	6.1339e-4	1.3363
CDS	*Vm*	8.0097e-3	0.9454
	Gm	1.0031e-2	1.8279
	Dp	1.6888e-3	1.5532
	Dz	5.1425e-3	3.0414
	eD	1.3063e-3	3.0513

TABLE 6 The J/J_1, MAE of chlorophyll-a and running time assimilated by different methods.

Method	J/J_1	MAE of chlorophyll-a(mmol $N\cdot m^{-3}$)		Running time (s)
		Before assimilation	After assimilation	
Case 1-CFDM	1.5952e-3	0.2434	0.0036	1351.55
Case 1-CDS	8.2648e-2	0.2823	0.0197	26697.83

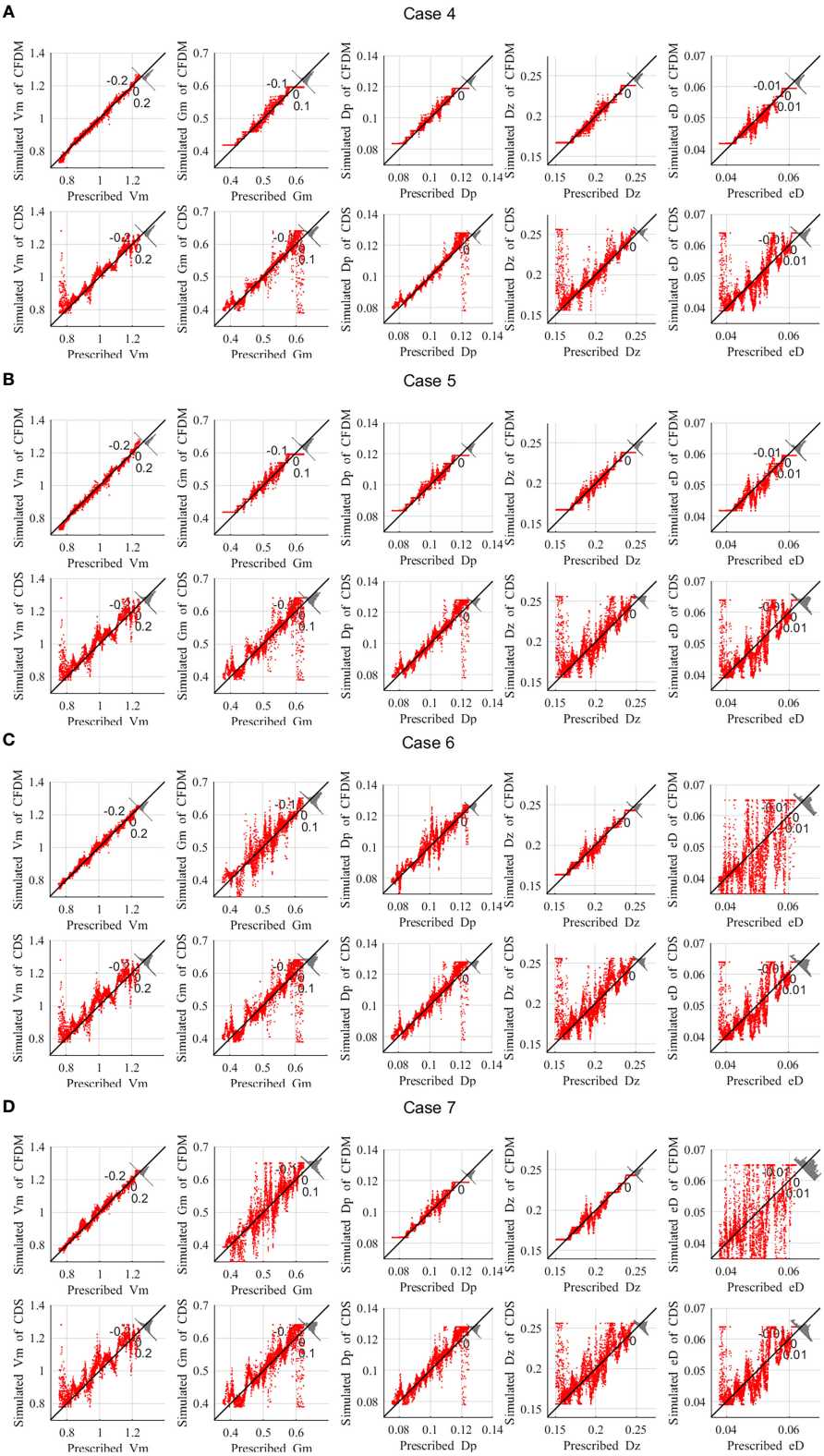


FIGURE 6
Comparison of the five parameters obtained by the CFDM and CDS method and the prescribed values in **(A)** Case 4, **(B)** Case 5, **(C)** Case6, **(D)** Case 7. This image shows the misfits between the estimations and their real values. (unit: day−1).

TABLE 7 Same as Table 5, but for the experiment of Case 5 with 30% observation errors.

Method	Parameter	MAE of parameters (day^{-1})	MRE of parameters (%)
CFDM	V_m	9.6420e-3	1.0184
	G_m	9.2129e-3	1.8401
	D_p	1.8402e-3	1.8414
	D_z	3.6162e-3	1.9518
	e_D	1.3783e-3	2.9177
CDS	*Vm*	3.0106e-2	3.2083
	Gm	1.7848e-2	3.4028
	Dp	2.8342e-3	2.6650
	Dz	1.0120e-2	5.5695
	eD	3.4486e-3	7.3006

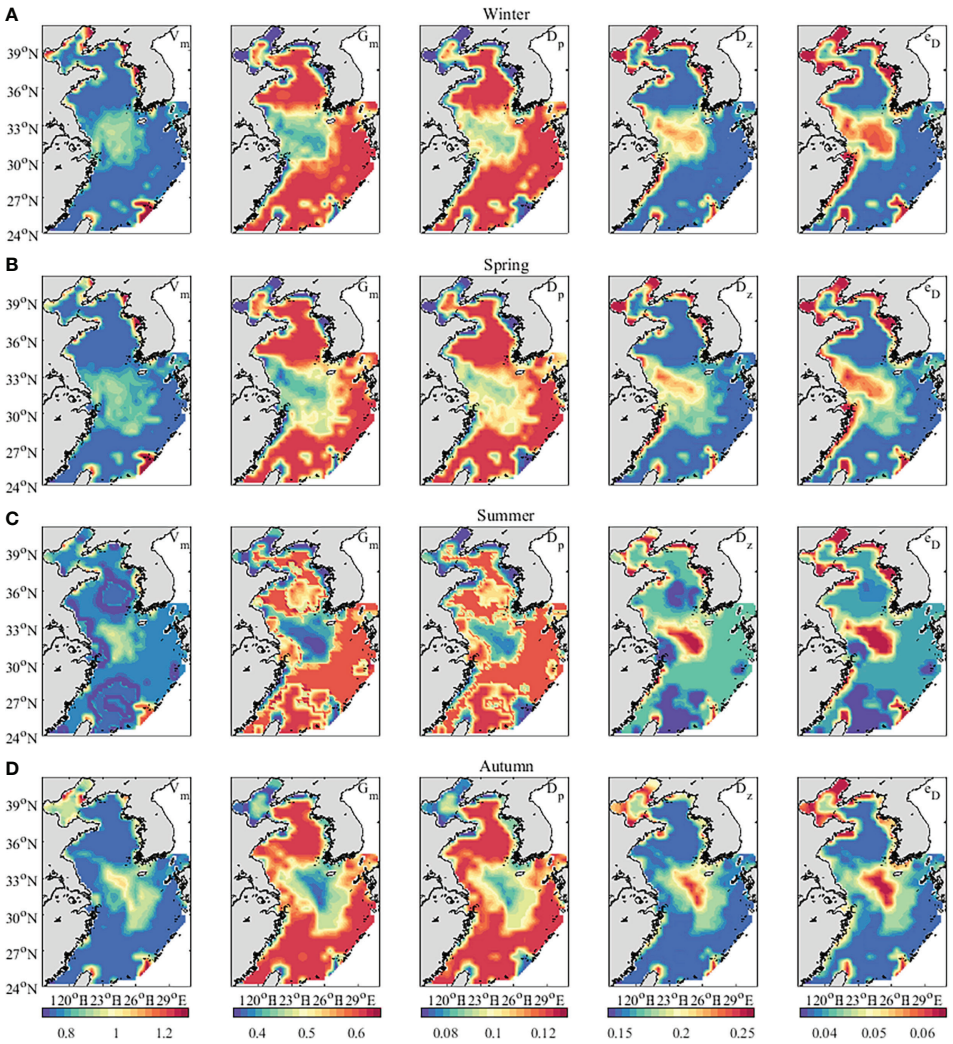


FIGURE 7
The distribution of five parameters inverted by the CFDM (unit: day^{-1}). **(A)** Winter (Jan-Mar); **(B)** Spring (Apr-Jun); **(C)** Summer (Jul-Sep); **(D)** Autumn (Oct-Dec).

WOA13 and SeaWiFS data measured at the initial time. Because of the lack of the distributions of zooplankton and detritus, the NPZD model was run for one months from December 1st, 2015 to January 1st, 2016, in which the initial guesses of zooplankton and detritus in the surface layer were 0.2 and 0.1 mmol N·m^{-3} and the concentration decreased exponentially with the increase of depth $P(k)=P(1)e^{-(z(k)-z(1))/z_{ch}}$ (k=1,2,…,6), with z_{ch} = 100m (Losa et al., 2006). The remaining model parameters were set as default empirical values (Table 1). Then the simulated results were taken as the initial conditions to optimize five parameters in 2016.

Figure 7 depicts the seasonal means of the five parameters (V_m, G_m, D_p, D_z, e_D), estimated in the practical experiments. The estimated distributions of the five parameters showed a seasonal cycle. The relatively high values of V_m, G_m and D_p appeared in winter, then decreased through the spring and summer, and increased again during autumn. Conversely, The D_z and e_D increased in spring and summer and decreased in autumn. In winter, high values of the maximum growth rate of phytoplankton (V_m) appeared in the near seas (~12 nautical miles territorial seas) north of 34°N. The nutrient concentrations varied seasonally and the highest concentrations of dissolved inorganic nitrogen, dissolved inorganic phosphorus, and dissolved silicate in the Bohai Sea was in winter (Zheng et al., 2020; Ding et al., 2021). Therefore, the values of V_m were large in the Bohai Sea, especially in Liaodong Bay. Previous studies showed that temperature had a great effect on zooplankton in the Yellow Sea during winter (Chen et al., 2011; Shi et al., 2018). Affected by the Yellow Sea Warm Current, Taiwan Warm Current, and Kuroshio, the values of G_m (maximum growth rate of zooplankton) were large in the East China Sea and the middle of the Yellow Sea.

Besides, based on the inversion results, the correlation coefficients (R) between the five parameters were calculated. There were strong correlations between V_m, D_z, and e_D, in which

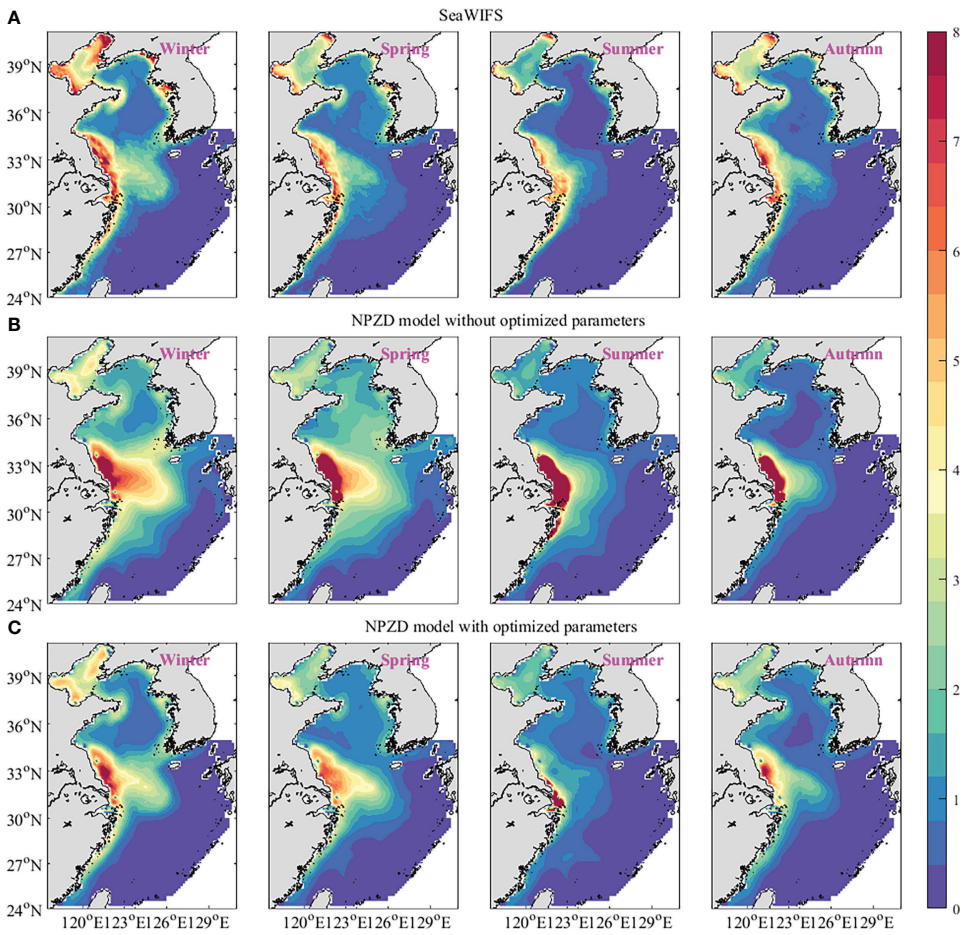


FIGURE 8
The sea surface phytoplankton obtained from SeaWiFS data **(A)**, simulations of the NPZD model without optimized parameters **(B)** and simulations of the NPZD model without optimized parameters **(C)**. (unit: mmol N·m^{-3}).

R of V_m and D_z were 0.73, 0.67, 0.81 and 0.77, respectively, R of V_m and e_D were 0.69, 0.63, 0.75, and 0.76, respectively, and R of D_z and e_D were 0.92, 0.92, 0.93, and 0.96, respectively, in different seasons. G_m and D_p were strongly related. R of G_m and D_p were 0.90, 0.89, 0.94, and 0.89 respectively, in different seasons. The result is consistent with the actual situation. When zooplankton increased, the concentrations of phytoplankton would decrease. Therefore, G_m and D_p (mortality rate of phytoplankton) change consistently. The concentrations of phytoplankton and detritus increased as zooplankton mortality increased. The V_m and e_D (remineralization rate of detritus) would vary with D_z (mortality rate of zooplankton).

To further verify the accuracy of the result, the distributions of phytoplankton simulated by the NPZD model with optimized parameters were compared with those of the NPZD model without optimized parameters. In the data assimilation experiment, one-tenth of the observation data randomly selected were not assimilated and used as independent observation to test the results, which was called 'test observation'. The seasonal sea surface phytoplankton obtained from SeaWiFS data, simulated by the NPZD model without optimized parameters and the NPZD model without optimized parameters are shown in Figure 8. Compared with the observations of SeaWiFS data, the results simulated by the NPZD model with optimized parameters matched the real data. The MAEs between the simulation of the NPZD model with optimized parameters and the test observation were 0.45, 0.50, 0.49 and 0.40 mmol N·m^{-3} in winter, spring, summer and autumn, respectively, which were smaller than 1.01, 1.24, 2.18 and 0.97 mmol N·m^{-3} of the NPZD model without optimized parameters.

The seasonal observations and simulations of the average value of phytoplankton on the sea surface are shown in Figure 8. The simulations reproduced the concentrations of the sea surface phytoplankton and the seasonal cycle where the high values appeared in winter and then decreased through the spring and summer, and increased again during autumn. This variety tendency was the same as that of Vm. The microbiology points out that the maximum growth rate occurs when the population density is optimum (Weaver and Grime, 1980). Once the population density is larger than the optimum density, the growth rate will reduce because of the limited nutrient and environmental conditions. The estimations of Vm implied that there was a connection between the growth rate and the phytoplankton concentration. The experiments show that the adjoint assimilation method with CFDM can invert the parameter values in the ecosystem model very well using large time steps.

Conclusions

Based on the adjoint data assimilation method with CFDM, the estimation of the parameters of the NPZD model is studied in this paper. The CFDM reduced the calculation time by large time steps and generated accurate numerical solutions. The experiment of the Gaussian pule moving shows the effectiveness of CFDM. Further, a series of experiments are carried out to evaluate the adjoint data assimilation method with CFDM and CDS and to examine the influential factors on the inversions of the five parameters (V_m, G_m, D_p, D_z, e_D) by assimilating the chlorophyll-a concentrations obtained by the SeaWiFS data. Considering the inversion errors and convergence rates synthetically, the CFDM performs better than CDS. Whether the five parameters were optimized separately or simultaneously, the parameters obtained by CFDM were more consistent with the distribution of the given parameters. The errors in observations have little influence on the estimated results, and the estimated results are all satisfactory, indicating that this model has a strong parameter estimation ability. According to the results of the twin experiments, the adjoint data assimilation method with CFDM is applied in a practical experiment to estimate the five parameters in February 2016. The results indicate that the improved data assimilation method can optimize the parameter values in the ecosystem model efficiently and the CFDM which gets rid of the limitation of stability provides an efficient choice for the study of high resolution model. Future work will focus on the optimization of parameters of the NPZD model in larger areas, longer time scales, and extreme weather conditions using the adjoint method, thus further improving the numerical simulations of the marine ecosystem.

Data availability statement

The raw data supporting the conclusions of this article will be made available upon reasonable request and with authors' permission. Open source datasets, includes World Ocean Atlas 2013 (WOA13; available online at https://www.nodc.noaa.gov/OC5/woa13) and Sea-Viewing Wide Field-of-View Sensor (SeaWIFS; available online at https://seawifs.gsfc.nasa.gov/ and https://esa-oceancolour-cci.org/).

Author contributions

MX conducted the data analysis, software, model validation, writing, and visualization. YL, KF, and XL conducted the formal analysis, software, and supervision. ZZ prepared the formal analysis and model validation. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the National Natural Science Foundation of China 42106146, the Natural Science Foundation

of Shandong Province of China ZR2021MA005 and ZR2020MD056, and the National Natural Science Foundation of China (Grant No. 11601497, 42076011 and U1806214).

Acknowledgments

We acknowledge all the reviewers for their valuable suggestions.

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2022.997537/full#supplementary-material

References

Anderson, D. A., Tannehill, J. C., and Pletcher, R. H. (1984) Computational fluid mechanics and heat transfer *Comput. fluid Mech. heat Transf.* (Hemisphere Pub. Corp.) 1(4),101–234. doi: 10.2307/2008017

Arbogast, T., Huang, C., Zhao, X., and King, D. N. (2020). A third order, implicit, finite volume, adaptive runge–kutta WENO scheme for advection–diffusion equations. *Comput. Methods Appl. Mech. Eng.* 368. doi: 10.1016/j.cma.2020.113155

Ascher, U. M., Ruuth, s. J., and Wetton, B. T. R. (1995). Methods for time dependent partial differential equations. *SIAM J. Numer. Anal.* 32, 797–823. doi: 10.1090/psapm/022/533052

Athias, V., Mazzega, P., and Jeandel, C. (2000). Selecting a global optimization method to estimate the oceanic particle cycling rate constants. *J. Mar. Res.* 58, 675–707. doi: 10.1357/002224000321358855

Baba, K., and Tabata, M. (1981). On a conservation upwind finite element scheme for convective diffusion equations. *RAIRO. Anal. numérique* 15, 3–25. doi: 10.1051/m2an/1981150100031

Blumberg, A. F., and Mellor, G. L. (1987). A description of a three-dimensional coastal ocean circulation model. *Three-Dimens. Coast. Model. Coast. Estuar. Ser.*, 1–16. doi: 10.1029/co004p0001

Borja, A., Andersen, J. H., Arvanitidis, C. D., Basset, A., Buhl-Mortensen, L., Carvalho, S., et al. (2020). Past and future grand challenges in marine ecosystem ecology. *Front. Mar. Sci.* 7. doi: 10.3389/fmars.2020.00362

Chen, H., Qi, Y., and Liu, G. (2011). Spatial and temporal variations of macro- and mesozooplankton community in the huanghai Sea (Yellow Sea) and East China Sea in summer and winter. *Acta Oceanol. Sin.* 30, 84–95. doi: 10.1007/s13131-011-0108-5

Cui, T., Zhang, J., Tang, J., Sathyendranath, S., Groom, S., Ma, Y., et al. (2014). Assessment of satellite ocean color products of MERIS, MODIS and SeaWiFS along the East China coast (in the yellow Sea and East China Sea). *ISPRS J. Photogramm. Remote Sens.* 87, 137–151. doi: 10.1016/j.isprsjprs.2013.10.013

Ding, X., Guo, X., Gao, H., Gao, J., Shi, J., Yu, X., et al. (2021). Seasonal variations of nutrient concentrations and their ratios in the central bohai Sea. *Sci. Total Environ.* 799, 149416. doi: 10.1016/j.scitotenv.2021.149416

Douglas, J.Jr., and Russell, T. F. (1982). Numerical methods for convection-dominated diffusion problems based on combining the method of characteristics with finite element or finite difference procedures. *SIAM J. Numer. Anal.* 19, 871–885. doi: 10.1137/0719063

Ebrahimijahan, A., Dehghan, M., and Abbaszadeh, M. (2020). Compact local integrated radial basis functions (Integrated RBF) method for solving system of non–linear advection-diffusion-reaction equations to prevent the groundwater contamination. *Eng. Anal. Bound. Elem.* 121, 50–64. doi: 10.1016/j.enganabound.2020.09.003

Fan, W., and Lv, X. (2009). Data assimilation in a simple marine ecosystem model based on spatial biological parameterizations. *Ecol. Modell.* 220, 1997–2008. doi: 10.1016/j.ecolmodel.2009.04.050

Faugeras, B., Bernard, O., Sciandra, A., and Lévy, M. (2004). A mechanistic modelling and data assimilation approach to estimate the carbon/chlorophyll and carbon/nitrogen ratios in a coupled hydrodynamical-biological model. *Nonlinear Process. Geophys.* 11, 515–533. doi: 10.5194/npg-11-515-2004

Franks, P. J. S. (2002). NPZ models of plankton dynamics: Their construction, coupling to physics, and application. *J. Oceanogr.* 58, 379–387. doi: 10.1023/A:1015874028196

Franks, P. J. S., and Chen, C. (2001). A 3-d prognostic numerical model study of the georges bank ecosystem. part II: Biological-physical model. *Deep. Res. Part II Top. Stud. Oceanogr.* 48, 457–482. doi: 10.1016/S0967-0645(00)00125-9

Friedrichs, M. A. M. (2001). Assimilation of JGOFS EqPac and SeaWiFS data into a marine ecosystem model of the central equatorial pacific ocean. *Deep. Res. Part II Top. Stud. Oceanogr.* 49, 289–319. doi: 10.1016/S0967-0645(01)00104-7

Fu, K. A. I., and Liang, D. (2019). A mass-conservative temporal second order and spatial fourth order characteristic finite volume method for atmospheric pollution advection diffusion problems. *SIAM J. Sci. Comput*, 41, B1178–B1210. doi: 10.1137/18M121914X.

Fu, K., Liang, D., Wang, W., and Cui, M. (2015). The time second-order characteristic fem for nonlinear multicomponent aerosol dynamic equations in environment. *Int. J. Numer. Anal. Model.* 12, 211–229.

Gao, X., Wei, Z., Lv, X., Wang, Y., and Fang, G. (2015). Numerical study of tidal dynamics in the south China Sea with adjoint method. *Ocean Model.* 92, 101–114. doi: 10.1016/j.ocemod.2015.05.010

Gharamti, M. E., Tjiputra, J., Bethke, I., Samuelsen, A., Skjelvan, I., Bentsen, M., et al. (2017). Ensemble data assimilation for ocean biogeochemical state and parameter estimation at different sites. *Ocean Model.* 112, 65–89. doi: 10.1016/j.ocemod.2017.02.006

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

Gunson, J., Oschlies, A., and Garçon, V. (1999). Sensitivity of ecosystem parameters to simulated satellite ocean color data using a coupled physical-biological model of the north Atlantic. *J. Mar. Res.* 57, 613–639. doi: 10.1357/002224099321549611

Heinle, A., and Slawig, T. (2013). Internal dynamics of NPZD type ecosystem models. *Ecol. Modell.* 254, 33–42. doi: 10.1016/j.ecolmodel.2013.01.012

Hu, C., Carder, K. L., and Muller-Karger, F. E. (2001). How precise are SeaWiFS ocean color estimates? implications of digitization-noise errors. *Remote Sens. Environ.* 76, 239–249. doi: 10.1016/S0034-4257(00)00206-6

Jones, E. M., Baird, M. E., Mongin, M., Parslow, J., Skerratt, J., Lovell, J., et al. (2016). Use of remote-sensing reflectance to constrain a data assimilating marine biogeochemical model of the great barrier reef. *Biogeosciences* 13, 6441–6469. doi: 10.5194/bg-13-6441-2016

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. Modell.* 202, 12–25. doi: 10.1016/j.ecolmodel.2006.08.021

Kuhn, A. M., Fennel, K., and Mattern, J. P. (2015). Model investigations of the north Atlantic spring bloom initiation. *Prog. Oceanogr.* 138, 176–193. doi: 10.1016/j.pocean.2015.07.004

Kurganov, A., and Tadmor, E. (2000). New high-resolution central schemes for nonlinear conservation laws and convection-diffusion equations. *J. Comput. Phys.* 160, 241–282. doi: 10.1006/jcph.2000.6459

Kytinou, E., Sini, M., Issaris, Y., and Katsanevakis, S. (2020). Global systematic review of methodological approaches to analyze coastal shelf food webs. *Front. Mar. Sci.* 7. doi: 10.3389/fmars.2020.00636

Lawson, L. M., Hofmann, E. E., and Spitz, Y. H. (1996). Time series sampling and data assimilation in a simple marine ecosystem model. *Deep. Res. Part II Top. Stud. Oceanogr.* 43, 625–651. doi: 10.1016/0967-0645(95)00096-8

Lax, P. D. (1954). Weak solutions of nonlinear hyperbolic equations and their numerical computation. *Commun. Pure Appl. Math.* 7, 159–193. doi: 10.1002/cpa.3160070112

Liang, D., Fu, K., and Wang, W. (2016). Modelling multi-component aerosol transport problems by the efficient splitting characteristic method. *Atmos. Environ.* 144, 297–314. doi: 10.1016/j.atmosenv.2016.08.043

Link, J. S., Bundy, A., Overholtz, W. J., Shackell, N., Manderson, J., Duplisea, D., et al. (2011). Ecosystem-based fisheries management in the Northwest Atlantic. *Fish Fish.* 12, 152–170. doi: 10.1111/j.1467-2979.2011.00411.x

Li, Y., Peng, S., Yan, J., and Xie, L. (2013). On improving storm surge forecasting using an adjoint optimal technique. *Ocean Model.* 72, 185–197. doi: 10.1016/j.ocemod.2013.08.009

Liu, Y. Z., Shen, Y. L., Lv, X. Q., and Liu, Q. (2017). Numeric modelling and risk assessment of pollutions in the Chinese bohai Sea. *Sci. China Earth Sci.* 60, 1546–1557. doi: 10.1007/s11430-016-9062-y

Losa, S. N., Vézina, A., Wright, D., Lu, Y., Thompson, K., and Dowd, M. (2006). 3D ecosystem modelling in the north Atlantic: Relative impacts of physical and biological parameterizations. *J. Mar. Syst.* 61, 230–245. doi: 10.1016/j.jmarsys.2005.09.011

Matear, R. J. (1995). Parameter optimization and analysis of ecosystem models using simulated annealing: a case study at station p. *J. Mar. Res.* 53, 571–607. doi: 10.1357/0022240953213098

Mattern, J. P., Song, H., Edwards, C. A., Moore, A. M., and Fiechter, J. (2017). Data assimilation of physical and chlorophyll a observations in the California current system using two biogeochemical models. *Ocean Model.* 109, 55–71. doi: 10.1016/j.ocemod.2016.12.002

Moisan, J. R., Moisan, T. A., and Abbott, M. R. (2002). Modelling the effect of temperature on the maximum growth rates of phytoplankton populations. *Ecol. Modell.* 153, 197–215. doi: 10.1016/S0304-3800(02)00008-X

Peck, M. A., Arvanitidis, C., Butenschön, M., Canu, D. M., Chatzinikolaou, E., Cucco, A., et al. (2018). Projecting changes in the distribution and productivity of living marine resources: A critical review of the suite of modelling approaches used in the large European project VECTORS. *Estuar. Coast. Shelf Sci.* 201, 40–55. doi: 10.1016/j.ecss.2016.05.019

Pelc, J. S., Simon, E., Bertino, L., El Serafy, G., and Heemink, A. W. (2012). Application of model reduced 4D-var to a 1D ecosystem model. *Ocean Model.* 57–58, 43–58. doi: 10.1016/j.ocemod.2012.09.003

Prieß, M., Piwonski, J., Koziel, S., Oschlies, A., and Slawig, T. (2013). Accelerated parameter identification in a 3D marine biogeochemical model using surrogate-based optimization. *Ocean Model.* 68, 22–36. doi: 10.1016/j.ocemod.2013.04.003

Qian, S., Wang, D., Zhang, J., and Li, C. (2021). Adjoint estimation and interpretation of spatially varying bottom friction coefficients of the M2 tide for a tidal model in the bohai, yellow and East China seas with multi-mission satellite observations. *Ocean Model.* 161, 101783. doi: 10.1016/j.ocemod.2021.101783

Qi, P., Wang, C., Li, X., and Lv, X. (2011). Numerical study on spatially varying control parameters of a marine ecosystem dynamical model with adjoint method. *Acta Oceanol. Sin.* 30, 7–14. doi: 10.1007/s13131-011-0085-8

Redfield, A. C., Ketchum, B. H., and Richards, F. A. (1963). The influence of biomineralisation on the composition of seawater. *M.N. Hill (Ed.) Sea. Intersci.*, 26–77. doi: 10.1007/978-94-009-7944-4_5

Riley, G. A., Stommel, H., and Bumpus, D. F. (1949). Quantitative ecology of the plankton of the western north Atlantic. *Bull. Bingham Oceanogr. Collect.* 12, 1–169.

Rückelt, J., Sauerland, V., Slawig, T., Srivastav, A., Ward, B., and Patvardhan, C. (2010). Parameter optimization and uncertainty analysis in a model of oceanic CO2 uptake using a hybrid algorithm and algorithmic differentiation. *Nonlinear Anal. Real World Appl.* 11, 3993–4009. doi: 10.1016/j.nonrwa.2010.03.006

Sathyendranath, S., Jackson, T., Brockmann, C., Brotas, V., Calton, B., Chuprin, A., et al. (2020). "ESA Ocean colour climate change initiative (Ocean_Colour_cci): Global chlorophyll-a data products gridded on a sinusoidal projection, version 4.2," in *Cent. environ. data anal.* Available at: https://catalogue.ceda.ac.uk/uuid/99348189bd33459cbd597a58c30d8d10.

Schartau, M., and Oschlies, A. (2003). Simultaneous data-based optimization of a 1D-ecosystem model at three locations in the north Atlantic: Part I - method and parameter estimates. *J. Mar. Res.* 61, 765–793.

Semovski, S. V., and Wozniak, B. (1995). Model of the annual phytoplankton cycle in the marine ecosystem-assimilation of monthly satellite chlorophyll data for the north Atlantic and Baltic. *Oceanologia* 37, 3–31.

Serpetti, N., Baudron, A. R., Burrows, M. T., Payne, B. L., Helaouët, P., Fernandes, P. G., et al. (2017). Impact of ocean warming on sustainable fisheries management informs the ecosystem approach to fisheries. *Sci. Rep.* 7(1), 13438. doi: 10.1038/s41598-017-13220-7

Shen, S., Liu, F., Anh, V., Turner, I., and Chen, J. (2013). A characteristic difference method for the variable-order fractional advection-diffusion equation. *J. Appl. Math. Comput.* 42, 371–386. doi: 10.1007/s12190-012-0642-0

Shi, Y. Q., Zuo, T., Yuan, W., Sun, J. Q., and Wang, J. (2018). Spatial variation in zooplankton communities in relation to key environmental factors in the yellow Sea and East China Sea during winter. *Cont. Shelf Res.* 170, 33–41. doi: 10.1016/j.csr.2018.10.004

Steenbeek, J., Buszowski, J., Chagaris, D., Christensen, V., Coll, M., Fulton, E. A., et al. (2021). Making spatial-temporal marine ecosystem modelling better – a perspective. *Environ. Model. Software* 145, 105209. doi: 10.1016/j.envsoft.2021.105209

Tashkova, K., Šilc, J., Atanasova, N., and Džeroski, S. (2012). Parameter estimation in a nonlinear dynamic model of an aquatic ecosystem with meta-heuristic optimization. *Ecol. Modell.* 226, 36–61. doi: 10.1016/j.ecolmodel.2011.11.029

Thacker, W. C., and Long, R. B. (1988). Fitting dynamics to data. *J. Geophys. Res.* 93, 1227. doi: 10.1029/jc093ic02p01227

Tjiputra, J. F., Polzin, D., and Winguth, A. M. E. (2007). Assimilation of seasonal chlorophyll and nutrient data into an adjoint three-dimensional ocean carbon cycle model: Sensitivity analysis and ecosystem parameter optimization. *Global Biogeochem. Cycles* 21, GB1001. doi: 10.1029/2006GB002745

Wang, D., Li, N., Shen, Y., and Lv, X. (2016). The parameters estimation for a PM2.5 transport model with the adjoint method. *Adv. Meteorol.* 2016, 1–13. doi: 10.1155/2016/9873815

Wang, D., Zhang, J., Mao, X., Bian, C., and Zhou, Z. (2020). Simultaneously assimilating multi-source observations into a three-dimensional suspended cohesive sediment transport model by the adjoint method in the bohai Sea. *Estuar. Coast. Shelf Sci.* 241, 106809. doi: 10.1016/j.ecss.2020.106809

Wang, D., Zhang, J., and Wang, Y. P. (2021). Estimation of bottom friction coefficient in multi-constituent tidal models using the adjoint method: Temporal variations and spatial distributions. *J. Geophys. Res. Ocean.* 126(5), e2020JC016949. doi: 10.1029/2020JC016949

Weaver, T., and Grime, J. P. (1980). Plant strategies and vegetation processes. *J. Range Manage.* 33, 159. doi: 10.2307/3898436

Wu, X., Xu, M., Gao, Y., and Lv, X. (2021). A scheme for estimating time-varying wind stress drag coefficient in the ekman model with adjoint assimilation*J. Mar. Sci. Eng,* 9, 1220. doi: 10.3390/jmse9111220

Zhang, L., and Ge, Y. (2021). Numerical solution of nonlinear advection diffusion reaction equation using high-order compact difference method. *Appl. Numer. Math.* 166, 127–145. doi: 10.1016/j.apnum.2021.04.004

Zheng, L.w., Zhai, W.d., Wang, L.f., and Huang, T. (2020). Improving the understanding of central bohai Sea eutrophication based on wintertime dissolved inorganic nutrient budgets: Roles of north yellow Sea water intrusion and atmospheric nitrogen deposition. *Environ. pollut.* 267, 115626. doi: 10.1016/j.envpol.2020.115626

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 24 November 2022
DOI 10.3389/fmars.2022.1050028

Check for updates

OPEN ACCESS

EDITED BY
Yasumasa Miyazawa,
Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Japan

REVIEWED BY
Tianran Liu,
Kyushu University, Japan
Hideyuki Kawamura,
Japan Atomic Energy Agency, Japan

*CORRESPONDENCE
Kaijun Ren
renkaijun@nudt.edu.cn
Wen Zhang
zhangwen06@nudt.edu.cn

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

RECEIVED 21 September 2022
ACCEPTED 31 October 2022
PUBLISHED 24 November 2022

CITATION
Wang H, Ren K, Lina M, Yu C, Chen W and Zhang W (2022) Long-term mean circulation in the Japan Sea as reproduced by multiple eddy-resolving ocean circulation models.
Front. Mar. Sci. 9:1050028.
doi: 10.3389/fmars.2022.1050028

COPYRIGHT
© 2022 Wang, Ren, Lina, Yu, Chen and Zhang. 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(s) 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.

Long-term mean circulation in the Japan Sea as reproduced by multiple eddy-resolving ocean circulation models

Haodi Wang[1], Kaijun Ren[1,2]*, Ma Lina[1], Chen Yu[1], Wen Chen[1] and Wen Zhang[1]*

[1]College of Meteorology and Oceanography, National University of Defense Technology, Changsha, China, [2]College of Computer Science and Technology, National University of Defense Technology, Changsha, China

The capacity of four eddy-resolving ocean circulation models—HYCOM (HYbrid Coordinate Ocean Model), MRI.COM (Meteorological Research Institute Community Ocean Model), OFES (Ocean General Circulation Model for the Earth Simulator), and NEMO (Nucleus for European Modeling of the Ocean)—to simulate the long-term mean hydrographic conditions and circulation patterns in the Japan Sea is investigated in this study. The assessment of this study includes the evaluation of mean vertical profiles and time series of temperature and salinity at the representative monitoring stations. Different model products from 1993 to 2015 are compared with *in situ* measurements provided by historical cruises and monitoring stations. After that, we compared the observed and simulated surface current velocities over the basin and volume transports through the key straits in the Japan Sea. Simulated current velocities are validated against 15 years of Acoustic Doppler Current Profiler (ADCP) measurements near the longshore and offshore branches of the East Korea Warm Current (EKWC). Furthermore, the atmospheric forcing data of the four ocean circulation models are validated against the satellite wind product. We found that the vertical profiles and long-term variations of temperature and salinity reproduced by MRI.COM and HYCOM are closer to *in situ* measurements. All models simulate temperature well in upper ocean, but salinity simulations are of lower quality from OFES and NEMO at several stations. Simulated current velocities predominantly lie within the standard deviation of ADCP measurements at two locations. However, the sea surface currents are underestimated by four models compared with Drifter data. Although simulated hydrographic profiles agree well with *in situ* observations, the mean circulation patterns greatly differ between the models, which highlight the need for additional evaluation and corrections based on the long-term current measurements. Because of the lack of ocean

current measurements, only the baroclinic velocities simulated by each model are reliable. The substantial part of the differences in barotropic velocities among the simulate result of four models is explained by the differing wind velocities from the corresponding atmospheric forcing datasets.

KEYWORDS

Japan Sea, long-term mean circulation, ocean circulation models, current velocity measurements, hydrographic conditions

1 Introduction

Knowledge of long-term variations in heat, matter, and salt transport by highly variable ocean currents is essential to understanding the ocean climate of a semi-closed marginal sea. The Japan Sea is one of the largest marginal seas in the Northwest Pacific with a mean depth of 1,750 m, which is divided by the seamounts into three large basins: the Japan Basin, the Yamato Basin, and the Ulleung Basin (Figure 1). Recent studies have revealed a series of anomalous ocean environment changes occurring in the Japan Sea that are potentially related to the global warming trend. The sea surface temperature (SST) averaged over the Japan Sea has risen by 1.3°C–1.7°C during the last century (Kida et al., 2020). A long-term ocean acidification trend has been detected along the Japan coast (Ishida et al., 2021). Shipboard measurements also indicate a rapid freshening trend of the Japan Sea Intermediate Water and a decreasing trend of dissolved oxygen in bottom water of Japan Sea. All the above phenomena are significantly influenced by the ocean current, which exists in the form of a northeastward Japan Sea Throughflow (JSTF) due to the semi-closed topography in the Japan Sea and exhibits an interannual intensification trend similar to that of SST and sea surface height (SSH)

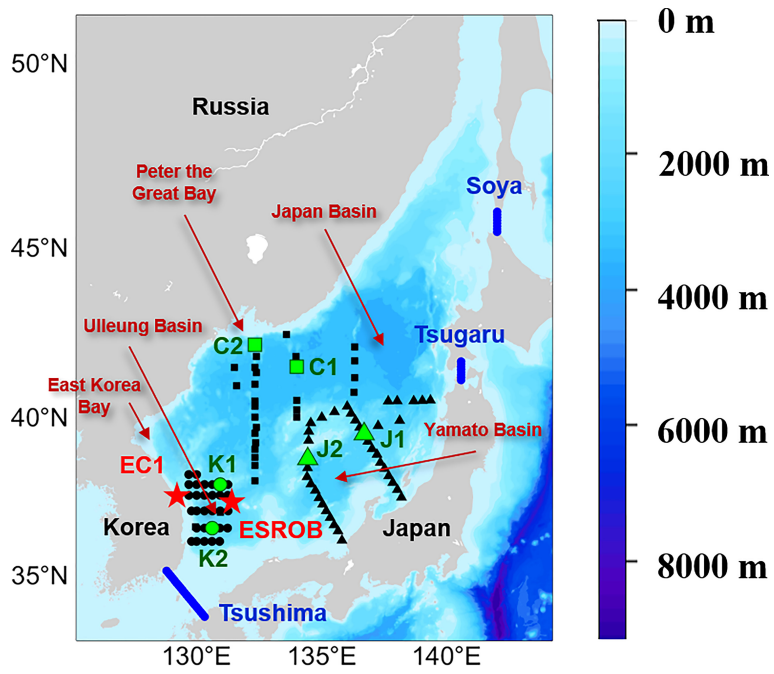


FIGURE 1
Locations of the used monitoring stations for the assessment of temperature and salinity (green marks) and current velocities (red pentagons) as well as the analyzed transects in three major straits (blue lines). Black marks denote hydrographic stations with more than 36 months of observation data. Round, square, and triangle circles denote KODC stations, CREAMS/EAST-I cruise stations, and JMA cruise stations, respectively. The background blue shading shows the bathymetry (m) from ETOPO1. See text for further information.

(Kida et al., 2020). Therefore, it is of great importance to study the long-term variability of Japan Sea circulation patterns and predict its future, developing trend under the global warming background, which is mainly achieved by using numerical tools and high-quality reanalysis data.

Because *in situ* measurements of JSTF are scarce in space and time, numerical modeling has become a necessary tool for studying the dynamics of ocean circulation in the Japan Sea. The first diagnostic model of horizontal circulation in the Japan Sea is based on a simple geostrophic balance applied by Yi, (1966), which is the first attempt to estimate the mean values and seasonal variability of volume transport into the Japan Sea by using dynamic calculation. With the help of three-dimensional (3D) models on the basis of linear barotropic equations, Minato and Kimura (1980) and Ohshima (1987) investigated the driving mechanism of inflow through the Tsushima Strait and outflow through the Soya Strait. By establishing a barotropic model with shallow water equations, Ohshima (1994) forced the diagnostic model by using observed sea level difference across the strait and found that the geopotential anomaly between subtropical and subpolar gyres serves as the primary driver of the mean horizontal circulation in the Japan Sea. Despite the absence of atmospheric forcing in the models, their simple modeling experiments still obtained reasonable mean volume transport and vertical current profiles that are highly consistent with *in situ* measurements. In addition, with diagnostic models, Sekine (1986) evaluated the influence of wind-driven circulation on the branching of the Tsushima Warm Current and found that the winter monsoon critically determines the intensity of boundary currents in the Japan Sea. A similar conclusion was drawn from Spall (2002), where sensitivity experiments indicate that wind stress critically determines the formation of eastern boundary current in the Japan Sea, whereas the buoyancy forcing is primarily responsible for its maintenance.

Further modeling studies of the JSTF variability and dynamics are primarily based numerical models rather than analytical models. Kim and Yoon (1999) simulated the separation point of the East Korea Warm Current (EKWC) using Modular Ocean Model (MOM) with an isopycnal mixing scheme forced by the observed heat flux and wind stress data. Using Research Institute for Applied Mechanics Ocean Model (RIAMOM) with the boundary conditions from ADCP, Kawamura et al. (2009) well reproduced the branch structure of JSTF and the variations of sea level along the Japan coast without any data assimilation. Numerical experiments of Park et al. (2013) and Kim et al. (2020) revealed that the surface heat flux has significant influence on the mean upper ocean circulation patterns in the Japan Sea by affecting the thickness of mixing layer.

Lagrangian passive tracer is a useful tool to trace the origin and paths of large ocean current systems, which has been involved into the 3D regional models by Stepanov et al. (2020) in the middle Japan Sea and by Prants et al. (2022) and Fayman et al. (2019) in the Peter the Great Bay. Stepanov et al. (2020) explained the clustering phenomena of floating tracer in the middle Japan Sea by identifying a mesoscale ocean current field with various eddy structures. Prants et al. (2022) tracked the formation and deep convection of dense shelf water in winter by using Lagrangian maps in the Regional Ocean Modeling System (ROMS). They found that the modeling results may be affected by the atmospheric forcing, the water exchange in the strait, and the sea ice formation near the northern coast. In addition, by the Lagrangian experiments of ROMS, Fayman et al. (2019) confirmed the existence of mesoscale eddies near the Ussuri Bay and its specific function to carry dense water masses in Japan Sea.

Because of the narrow width of strait throughflow and boundary currents, dynamic studies on JSTF increasingly rely on eddy-resolving models. With the advances of supercomputing power, a large number of high-resolution numerical experiments on JSTF have appeared in recent studies. Examples include the topography experiment in Finite Volume Coastal Ocean Model by Han et al. (2018), the heat sensitivity experiment in RIAMOM by Kim et al. (2020) and Hirose et al. (2021), and the beta effect experiment in MOM by Kim et al. (2018). At the same time, many institutions are focusing on the development and update of ocean reanalysis and forecast products, most of which cover a quasi-global domain and have reached eddy resolution in the horizontal direction. However, a comprehensive assessment on these models and products is still absent in Japan Sea to find which models could better simulate and reproduce the dynamical and hydrographic environment in this semi-closed "Miniature Ocean". Hence, validating the model simulations against the real ocean environment is of great necessary, especially in the current stage when the data assimilation effect is severely limited by the sparce *in situ* observation in Japan Sea.

This study focuses on the assessment of the long-term mean state of ocean circulation in the Japan Sea as reproduced by different eddy-resolving ocean circulation models, utilizing different mixing parameterization schemes, vertical coordinates, atmospheric forcing, and boundary conditions. Four eddy-resolving ocean circulation models are involved in the assessment, including the HYbrid Coordinate Ocean Model (HYCOM), the western North Pacific version of the MRI Community Ocean Model (MRI.COM-WNP, referred to as MRI.COM below), the Ocean General Circulation Model for the Earth Simulator (OFES), and the Nucleus for European Modeling of the Ocean (NEMO). These models were established for different purposes and have been widely used for dynamic studies in the not only the Japan Sea but also other marginal seas (Cheng et al., 2015; Fujii et al., 2016; Usui et al., 2016; Cheon, 2020; Menezes, 2021). The performance of the four eddy-resolving Ocean General Circulation Models (OGCMs) in simulating the variability of salinity and temperature profiles as well as the mean state of volume transport and circulation patterns is investigated, using the long-term gridded data from 1993 to 2014.

For this purpose, we compared the model outputs against multi-source *in situ* measurements and took a closer sight into the quality of wind forcing data driving each model. Because there are no long-term wind observations at fixed locations, we used the QuikSCAT satellite wind product as the ground truth. It should be noted that we are not conducting an intercomparison of different models due to the distinct configuration of model initializations, grid coordinates, sub-grid parameterizations, as well as hydrographic and atmospheric forcing. In this study, we can only speculate the possible reasons for the differences in model outputs, which provide some reference for the development of next-generation high-resolution ocean circulation models with higher scientific value in Japan Sea. Therefore, this study aims to compare the state estimates and uncertainties of ocean circulation as well as the hydrographic conditions, rather than comparing models on shorter time scales, like what Pätsch et al. (2017) and Myrberg and Andrejev (2006) have conducted in the Baltic Sea.

This article is organized as follows. Ocean circulation models and observation databases involved in this study, as well as the overall strategy of evaluation, are briefly introduced in Section 2. The evaluation results are presented in Section 3. The possible causes for the differences in the mean circulation simulated by different models are demonstrated in Section 4. Finally, the article is ended with a summarization of major conclusions in Section 5 and a prospect for the development of eddy-resolving ocean circulation models in Japan Sea in Section 6.

2 Data and methods

2.1 Ocean circulation models

For the present assessment, four eddy-resolving ocean circulation models are taken into account; all data were interpolated to a consistent regular coordinate with 2-m vertical resolution and 1/12° horizontal resolution. The overall configurations of each model are summarized in Table 1.

The HYCOM is a primitive equation OGCM that develops from the Miami Isopycnic-Coordinate Ocean Model, which is widely used for ocean climate studies (Halliwell et al., 1998; Halliwell et al., 2000; Bleck, 2002). HYCOM employs hybrid vertical coordinates, reverting smoothly from isopycnals in stratified open seas to z-level coordinates in mixed layers and further to terrain-following coordinate in weak stratification areas. A tri-pole latitudinal grid is employed as horizontal coordinate in which all 1D submodels are embedded. The used vertical mixing parameterization is K-Profile Parameterization (KPP) (Large et al., 1994). Lateral advection and diffusion of salinity, temperature, and momentum were presented by a combination of biharmonic and Laplacian diffusivity. The bulk formula of Kara et al. (2000) is used for heat flux parameterizations. The sea ice thermodynamics are simulated by energy loan ice model. Atmospheric forcing data are the 1-hourly National Centers for Environmental Prediction (NCEP) Climate Forecast System Reanalysis (CFSR) data, which has a spatial resolution of ~38 km. In Global Ocean Forecast System (GOFS) 3.0, the temperature and salinity profiles, altimeter SSH, and satellite SST were assimilated by the Navy Coupled Ocean Data Assimilation (NCODA) using 3D variational (3DVar) assimilation scheme. Outputs of the GOFS3.0 forecast system include daily 3D temperature, salinity, currents, and 2D SSH. To ensure a consistent time period with other models, we selected GOFS 3.0 Glbu Reana lysis data from 1993 to 2013 and extend the time span to 2015 by supplying the GOFS 3.0 GLBu Analysis data. The GOFS 3.0 Glbu data has 40 vertical layers with a horizontal resolution of 1/12°. The vertical resolution reduces from 2 m near the sea surface at 1,000 m to the bottom at 5,500 m.

The Meteorological Research Institute Community Ocean Model (MRI.COM) is a multilevel primitive equation model

TABLE 1 Fundamental configurations of four models assessed in this study, including the information on assessed product, the wind forcing dataset, the horizontal and vertical grid schemes, the parameterization schemes, the data assimilation method, and the model domain.

	HYCOM	MRI.COM	OFES	NEMO
Product	GOFS 3.0 GLBu Reanalysis/Analysis	FOR A-WNP30	OFES ver. 1	GLORYS12 version 1
Domain	Global	117°E to 160°W, 15°N to 65°N	0°E to 180°W 75°S to 75°N	0°E to 180°W 85°S to 85°N
Wind Forcing	NCEP CFSR	JRA-55	NCEP Reanalysis	ERA-Interim
Horizontal Resolution in Japan Sea	1/12°	1/10°	1/10°	1/12°
Vertical Grid	Hybrid coordinate scheme	Terrain-following vertical coordinates	Z-coordinate	Z-coordinate
Horizontal Mixing Parameterization	Biharmonic and Laplacian diffusivity	Biharmonic Smagorinsky viscosity	Scale-selective damping of biharmonic operator	Rotated Laplacian operator
Vertical Mixing Parameterization	KPP	Turbulent closure scheme of Noh and Kim H.J. (1999)	KPP	TKE turbulent closure scheme
Data Assimilation	NCODA (3DVar)	MOVE-4DVar	None	Multivariate reduced-order Kalman filter

with hydrostatic and Boussinesq assumptions. It was used for large-scale simulations of oceanic phenomena as an ocean component of coupled climate models (Ishikawa, 2005; Tsujino et al., 2006). The 4D variational (4DVar) Ocean Re-Analysis for the Western North Pacific over 30 years (FORA-WNP30) is used in this study, which is the output of the western North Pacific version of the MRI Community Ocean Model (MRI.COM-WNP) (Usui et al., 2017). The configurations of MRI.COM-WNP are fundamentally consistent with MRI.COM (Tsujino et al., 2006) except for denser bottom layers and incorporation of sea-ice model. The horizontal mixing of tracers and momentum were parameterized by a biharmonic operator and biharmonic Smagorinsky viscosity, respectively. The vertical viscosity and diffusivity were provided by the turbulent closure scheme from Noh and Kim (1999) after eliminating the dependency of vertical mixing coefficients and bottom friction on the background state. Heat fluxes were obtained from bulk formula (Kondo, 1975). The model domain covers the Northwest Pacific zonally from 117°E to 160°W and meridionally from 15°N to 65°N. MRI.COM uses terrain-following vertical coordinates. There are 54 vertical layers with the interval increasing from the surface at 1 m to the bottom at 600 m. Horizontally, MRI.COM-WNP applied variable grid scheme for a cost-effective simulation of the Northwest Pacific circulation at high resolution, setting zonal resolution of 1/10° from 117°E to 160°E and 1/6° from 160°E to 160°W and meridional resolution of 1/10° from 15°N to 50°N and 1/6° from 50°N to 65°N. The *in situ* temperature and salinity profiles above 1,500 m, gridded SST, altimeter SSH, and sea ice concentration data were assimilated into MRI.COM-WNP by using 4DVar analysis scheme version of the MOVE system (MOVE-4DVar) with the first Gauss from the analysis fields of MOVE-3DVar. Previous evaluation studies shows that FORA-WNP30 incorporated with MOVE-4DVar has higher accuracy than the 3DVar product (Usui et al., 2017). Atmospheric forcing was taken from daily JRA-55 atmospheric reanalysis product with a regular horizotal grid of ~0.56° resolution (Kobayashi et al., 2015). Boundary conditions were created by a North Pacific model with a horizontal resolution of 1/2° using a one-way nesting method (Hirose et al., 2016).

The OFES is based on the MOM version 3 for a long-term, eddy-resolving hindcast of the global ocean circulation with the available supercomputing resources. OFES utilizes the vertical z-level coordinate and solves 3D primitive equations in the horizontal spherical coordinates under hydrostatic and Boussinesq approximations. The domain of OFES model is a quasi-global region from 75°S to 75°N excluding the polar regions. The horizontal resolution is 0.1°. The number of vertical layers is 54 from 2.5 to 6,065 m with the intervals varying from the surface at 5 m to the bottom at 330 m. The horizontal mixing of momentum and tracers was parameterized by scale-selective damping of biharmonic operator to suppress computational noise. The vertical mixing parameterization used KPP (Large et al., 1994). The bulk formula of Rosati and Miyakoda (1988) is used to calculate surface heat flux. The OFES-based hindcast experiments were initialized by the 50-year climatological spin-up integration (Masumoto et al., 2004). Atmospheric forcing is obtained from daily NCEP/NCAR reanalysis products (Kalnay et al., 1996). The hindcast products of OFES1 was assessed in this study.

The NEMO is a state-of-the-art modeling framework in ocean and climate sciences (Madec, 2008). The ocean engine of NEMO (NEMO-OCE) is a primitive equation model intended to be used for modeling studies on the ocean and its interactions with the earth climate system in various spatiotemporal scales (Madec et al., 2017). As a state-of-the-art representative of NEMO-based products, the GLORYS12 version 1 reanalysis product generated by the PSY4V3 forecast system based on NEMO 3.1 (Lellouche et al., 2018; Jean-Michel et al., 2021) is used in this study. NEMO 3.1 for GLORYS12 adopts rotated Laplacian operator for the horizontal parameterization of both tracers and momentum advection, and the TKE turbulent closure scheme is used for the vertical mixing parameterization (Lellouche et al., 2018). A quasi-isotropic grid was utilized in the horizontal direction, and a z-coordinate is applied in the vertical. The horizontal resolution of GLORYS12 is 1/12°. There are 50 vertical layers from the surface at 0 m to the bottom at 5,700 m. The model is driven by the European Center for Medium-Range Weather Forecasts ERA-Interim atmospheric reanalysis (Dee et al., 2011). A multivariate reduced-order Kalman filter is applied to assimilate temperature and salinity profiles provided by the Coriolis Ocean Dataset for Reanalysis (CORA), altimeter SSH and AVHRR satellite SST in a 7-day assimilation cycle (Lellouche et al., 2018). In the present study, we use an overlapping period of 1993–2015 for all the four models.

2.2 Observations

2.2.1 Temperature and salinity profiles

To evaluate the physical conditions simulated by the selected eddy-resolving ocean circulation models, we took the temperature and salinity profiles collected by the Conductivity Temperature Depth (CTD) from 1993 to 2014 *via* 19 research cruises conducted through the CREAMS (Circulation Research of the East Asian Marginal Seas) and EAST-I (East Asian Seas Timeseries I) and 59 cruises conducted by the Japan Meteorological Agency. The CTD stations cover most areas of the northern (Figure 1, square dots) and southeastern Japan Sea (Figure 1, triangle dots). All CTD data were carefully calibrated based on raw data except for those collected by cruises from 1999 June to 2000 February, for which only preliminarily processed data are available. They were processed by using standard procedures, such as those of the Sea-Bird Electronics (SBE).

Most CTD data collected after 2002 were processed by pre-cruise and post-cruise calibrations following the typical SBE data processing sequences (Morison et al., 1994). Ancillary temperature and salinity along five geostationary observation lines from 1968 to 2021 were taken from the Korea Oceanographic Data Center (KODC). The KODC data were sampled bimonthly or more frequently, whereas Japan Meteorological Agency (JMA) and CREAMS/EAST-I data had more irregular sampling intervals (Kang et al., 2008). For the present study, we selected hydrographic data from the monitoring stations C1 (134.0°E, 41.5°N) and C2 (132.3°E, 42.2°N) in the Japan Basin, J1 (136.7°E, 39.5°N) and J2 (134.4° E, 38.7°N) in the Yamato Basin, as well as K1 (130.9°E, 37.9°N) and K2 (130.6°E, 36.5°N) in the Ulleung Basin. These locations are illustrated by green dots in Figure 1. The selected stations cover conditions in the northern, the southeastern, and the southwestern Japan Sea and differ primarily in the influence of ocean current, thermal conditions, and atmospheric conditions. All profiles were preprocessed into standard depths with a vertical resolution of 2 m from 0 m to 5,000 m. For each *in situ* measurement data, the collocated model data were selected by searching and averaging in a 3-day time window and a spatial radius of 1/6°. A 30-day moving average has been applied on both the time series of collocated observation and model data to eliminate small-scale noise and focus on the mean state and obtain monthly time series.

2.2.2 Ocean current

For the assessment of surface circulation, we used the gridded Drifter ocean current product as the ground truth. After the launch of Global Drifter Program in 1979, a total number of 907 buoys were deployed by various institutions in the Japan Sea, covering at least 80% of the total area (Wang et al., 2020). Each Drifter buoy carries a floating anchor at 15 m to record its real-time location and transmit raw data through the Global Telecommunication System. The buoy trajectories reflect the near surface current velocities because the impact of wind-driven current and Stokes flow has been minimized, and the actual Drifter velocity data are composed of geostrophic velocity, Ekman velocity, and other non-geostrophic velocity.

For the evaluation of simulated ocean current, we used the velocity measurements from Acoustic Doppler Current Profiler (ADCP) at two long-term monitoring stations: the EC1 located near the offshore branch of the EKWC and the East Sea Real-time Ocean Buoy (ESROB) located near the longshore branch of the EKWC (Figure 1, red pentagons). The EC1 is a deep mooring buoy deployed to the north between Ulleung and Dok Islands in 1996 during a cooperative experiment conducted by the Korea Ocean Research and Development Institute, the Woods Hole Oceanographic Institution, and the Seoul National University (SNU). It works autonomous and is equipped with various instruments at 400, 1,400, and 2,000 m (Chang et al., 2002), which records hourly velocity data from 1993 to present. The ESROB is a long-term, continuous, and real-time ocean monitoring buoy maintained by the SNU Ocean Observation Laboratory to observe multiple oceanographic phenomena and processes in the Japan Sea (Kim et al., 2005). It is deployed 10 km off the coast at a depth of ~100 m from 1999 to present. ESROB is equipped with 300-kHz ADCP and several SBE37 CTDs to measure 3D subsurface currents and hydrographic properties at 26 vertical layers for every 1 to 10 min.

For the present study, we used ADCP hourly current velocity data from 2000 through 2014. The 15-year period is well overlapped by each ocean current models. We selected the EC1 data at 400 m and ESROB data at 5, 20, 40, 60, and 100 m, respectively, where the largest amount of velocity data was collected. To compare with the ADCP data, the model data were taken from 3D fields at approximately the same location and depth as the EC1 and ESROB stations.

2.2.3 Surface wind velocity

To estimate the deviations of four atmospheric wind forcing datasets—the Climate Forecast System Reanalysis (CFSR) for HYCOM, the Japanese 55-year Reanalysis (JRA-55) for MRI.COM, the NCEP/NCAR Reanalysis 1 (NCEP R1) for OFES, and the ERA-Interim for NEMO, we compared these reanalysis data with the QuikSCAT sea surface wind product. For the comparison purpose, both reanalysis and satellite products were interpolated into daily time series with a consistent spatial resolution of 0.25°.

2.2.4 Taylor diagram

The Taylor diagram, which presents three statistical evaluation parameters in one semicircle or quarter chart, including the Pearson correlation coefficient, the standard deviation, and the root mean square deviations (RMSDs), is employed for this assessment. The Pearson correlation coefficient quantifies the similarity between the observed and simulated time series. It is represented by straight lines related to the azimuthal angle. Values from 0 to 1 indicate from no correlation to 100% agreement. The normalized standard deviation of each model time series relative to the observations was calculated to obtain a unified representation of deviations in one Taylor diagram. This value is proportional to the radial distance from the origin of the Taylor diagram, which indicates the similarity between the amplitude of simulated and observed time series. The centered RMSD in the Taylor diagram is proportional to the distance from the corresponding arc with the x-axis. See the study by Taylor (2001) for a detailed description on Taylor diagram. The closer distance represents the smaller RMSD of simulated time series relative to the observation result.

2.2.5 Extended cost function

As the offsets between the time series of simulated and observed data are not included in the Taylor diagram, we defined an extended cost function as an additional measure for the quality assessment of simulated time series. According to Eilola et al. (2011), the original cost function (C) was computed for each model (i) by

$$C(i) = \left| \frac{M_i - O}{STD} \right| \quad (1)$$

where the M_i indicates the monthly simulated time series and O indicates the monthly observed time series. STD is the standard deviation of observational data. The cost function values were calculated at each depth to evaluate the similarity between observations and simulations. Values between 0 and 1 indicate good agreement, values between 1 and 2 indicate reasonable quality, and values larger than 2 reveal poor quality. Therefore, a good accordance is defined by the fact that the deviation between observed and simulated time series is smaller than twice the STD values of observation time series. On the basis of the normalized STD, the Pearson correlation coefficient and the RMSD given by the Taylor diagram, an extended cost function value is defined as follows

$$CE(i) = \frac{1}{3}\left(\left| \frac{M_i - O}{STD} \right| + \left| \frac{STD_i - STD}{STD} \right| + \left| \frac{RMSD_i}{STD} \right| \right) \quad (2)$$

where STD_i and $RMSD_i$ represent the standard deviation and RMSD of simulated time series, respectively.

2.3 Evaluation strategies

To assess the capacity of four eddy-resolving ocean circulation models to reproduce long-term mean state of physical and hydrographic environment in Japan Sea, a 22-year time period from 1993 through 2014 was selected. First, the quality of simulated temperature and salinity data is evaluated by comparing the corresponding model reanalysis products with *in situ* measurements provided by JMA cruises in the southeast, CREAMS/EAST-I cruises in the north, and KODC monitoring stations in the southwest. The evaluation process includes the validation of vertical profiles and the statistical comparison of time series at different depths of the selected monitoring stations.

Next, the simulated sea surface current velocities throughout the Japan Sea and the volume transport into three major straits were evaluated. Note the assessment of middle and deep circulation patterns is not in the scope of this study as model simulations in deep layers tend to contain large errors in Japan Sea with inadequate data assimilation and sparce vertical model layers in deep ocean. In addition, we also analyzed the vertical structure of simulated ocean currents across the major straits. Locations of three transactions used for the calculation of volume transport are marked as blue lines in Figure 1. The Tartar Strait, located at the northern end of the Japan Sea with more than 10 km width, is excluded from this study due to the negligible volume transport.

Finally, the current velocities simulated by four eddy-resolving ocean models were compared against ADCP measurements at two monitoring stations, and the atmospheric wind forcing data driving the ocean models were analyzed and compared with satellite wind product from QuikSCAT.

3 Results

3.1 Temperature and salinity at monitoring stations

3.1.1 Mean vertical profiles

Figure 2 shows the 22-year mean vertical temperature and salinity profiles along with their standard deviations at six typical monitoring stations. Both the mean values and standard deviations were calculated from the monthly time series between January 1993 and December 2014. The mean temperatures at these stations vary from approximately 1°C to 18°C with the maximum standard deviations of ~5°C near sea surface. The near surface temperature from 0 to 100 m is well reproduced by each model at almost each station with the maximum deviations no larger than 1°C. However, the exception occurs at the station C2 where 0–40 m is a positive error and 40–80 m is a significant negative error. The temperature is slightly overestimated at station C1 by all models except HYCOM and is slightly underestimated at station K2 by HYCOM. In addition, OFES has worse performance at almost each station except the K2 station with both significant positive and negative biases.

At all stations, the temperature monotonously decreases with depth, but the thermocline stretches at different vertical locations depending on each station. For most stations, the thermocline simulated by NEMO lies 10–30 m lower than that observations from six stations. Below the thermocline, each model almost perfectly fits to the observations at all stations with the deviations no larger than 1°C. The standard deviations of temperature at very large depths are smaller than 1°C for both simulations and observations.

The 22-year mean salinity values range from approximately 33.6 to 34.3 g kg^{-1} at the sea surface, whereas, at the bottom, the salinity values are highly concentrated near ~34.05 g kg^{-1}. At most stations, the standard deviations of salinity are much smaller than 0.3 g kg^{-1}. Exceptions occur at the sea surface of C1 station, at 0–30 m depth of C2 station and the sea surface at K2 station, where the maximum standard deviations all exceed 0.5 g kg^{-1}. The mean salinity simulated by four models rarely

10.3389/fmars.2022.1050028

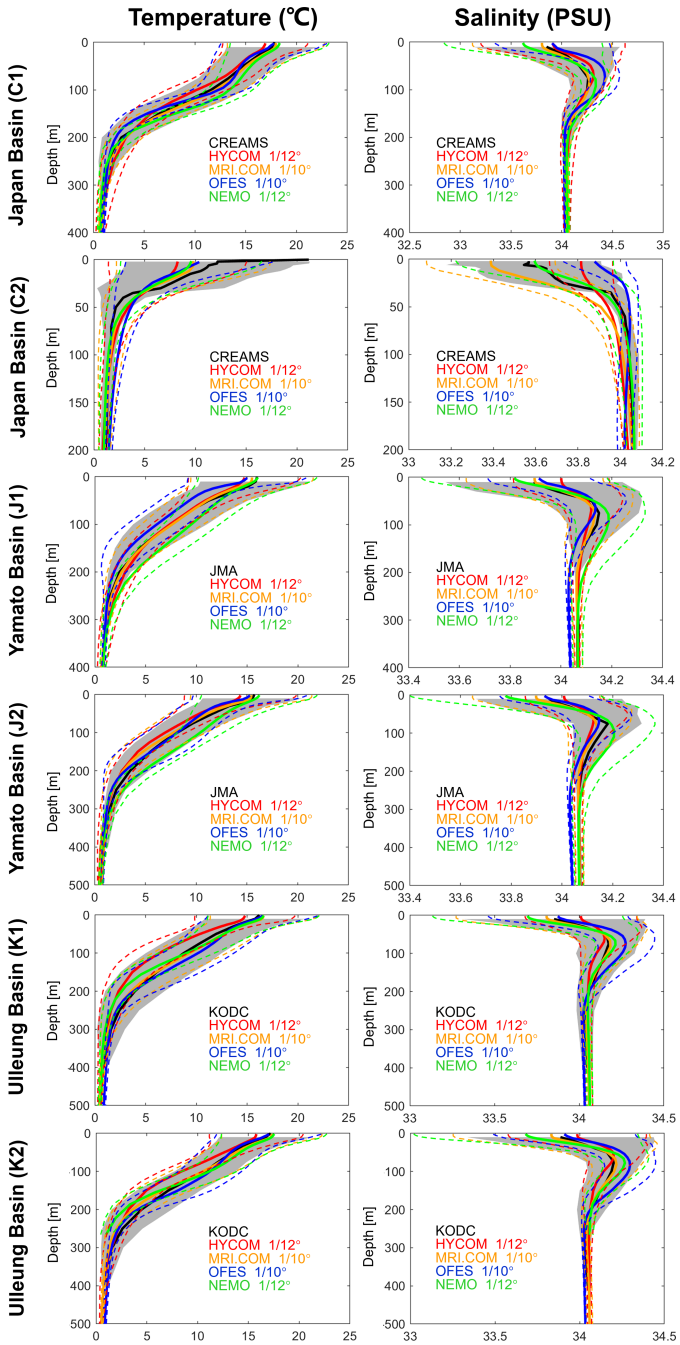


FIGURE 2
Mean vertical profiles of temperature (left) and salinity (right) observed by six representative stations (black) and simulated by four eddy-resolving ocean models HYCOM (red) and OFES (blue) with 1/12° resolution, MRI.COM (orange) and NEMO (green) with 1/12° resolution with 1/2° resolution for the time period from 1993 through 2014. Results for the monitoring stations C1 and C2 in the Japan Basin, J1 and J2 in the Yamato Basin, and K1 and K2 in the Ulleung Basin are displayed from bottom to top. Mean values are illustrated by solid lines. Standard deviations are indicated by the gray-shaded area for the observation data and by the colored dashed lines for the models.

exceeds the range of observed standard deviations. However, almost all models fail to reproduce the upper salinity structure at C2 station, which significantly bends at 40 m depth instead of smoothly stretches from sea surface to 70 m depth. Among them, the negative biases of HYCOM and MRI.COM at 30–80 m and the positive biases of OFES at 30 through 70 m slightly exceed the range of observed standard deviation. Moreover, NEMO presents the location of salinity thermocline at

relatively low depths, especially at the southeastern stations J1 and J2. At very large depths, an overestimation can be noticed from NEMO at C2 station, whereas OFES undersetimates the deep salinity of all stations by the magnitude of 0.05 g kg^{-1}.

In summary, the mean temperature and salinity simulated by MRI.COM are partly beyond the range covered by the standard deviation of observation, whereas the means simulated by HYCOM and NEMO exceed the observed standard deviations more often than MRI.COM at various stations and depths. Moreover, HYCOM has an overall better performance than NEMO in simulating both temperature and salinity profiles. By contrast, the OFES data are in relatively poor quality especially in terms of salinity at very large depths. It is possibly due to the absence of data assimilation in OFES model to correct the simulated profiles at those monitoring stations.

3.1.2 Statistical assessment

To further assess the capacity of four models to simulate the temporal variations of temperature and salinity at six stations, the Taylor diagram and the extended cost function are applied here. Figure 3 shows the Taylor diagram of upper ocean temperature and salinity relative to the observation data as the black reference point located on the x-axis (normalized standard deviation = 1, Pearson correlation coefficient = 100%). The Taylor diagram of temperature shows high consistency between simulated and observed profiles at each station. The Pearson correlation coefficients of SST are generally higher than 0.7. Among them, the MRI.COM values at surface to subsurface reach as high as 0.95, and the normalized standard deviations for temperature are closest to 1 at J1, K1, and K2 stations.

The cost function in Figure 4 further verifies that eddy-resolving ocean circulation models can well reproduce the temperature profiles at each station. At almost every depth, the cost function values of temperature lie between 0 and 1, indicating good data quality. Below 100 m, different models begin to show diversified capacities to simulate the temperature at different depths. MRI.COM presents the overall smallest cost function values, which are mainly distributed between 0 and 0.5 with a few large deviations below 50 m at the C2 station. This proves the very good quality of SST data simulated by MRI.COM. The temperature cost function values are generally smaller than 1 and are maximized for OFES at 200 m depth of C2 station, for HYCOM at the surface of C2 station, and for NEMO at 200 m depth of the J2 station. At the above locations, the cost function values are larger than 1.5, which represents reasonable data quality. It is noted that all the four models have the lowest cost function values at C2 station, which corresponds to the obvious deviation of simulated profiles in Figure 2.

The Taylor diagrams and cost function values reveal stronger differences of salinity among four eddy-resolving ocean models. MRI.COM has the overall best performance at almost each station with the highest correlations, the smallest spread of the normalized standard deviation around the observation value of 1, as well as the smallest RMSD values appearing in almost all depths. Most cost function values are significantly smaller than 0.5, indicating a very good accordance of salinity reanalyzed by MRI.COM with the observation data.

The statistics of salinity reproduced by HYCOM and NEMO are comparable to each other with the overall data quality of HYCOM slightly higher than that of NEMO. Their correlation coefficients are mainly between 0.5 and 0.95. Most deviations from the normalized standard deviation of 1 are no more than 0.5, and the RMS errors lie predominantly between 0.5 and 1.3. The cost functions are predominantly no larger than 1.5 but still reveal the shortcomings of two models in reproducing salinity at certain depths and stations.

Comparatively, OFES has the overall worst performance of salinity simulations, which is reflected by the generally lower correlation coefficients (most below 0.8), the higher deviations from the normalized standard deviation of 1 (up to about 0.7 in deep layers), and the highest RMS errors (up to 1.4) in the Taylor diagrams. Similarly, the cost function values are higher in OFES than in other models, indicating a poor agreement between OFES simulations and observations. It is interesting to note that MRI.COM and NEMO are complementary to each other in the salinity data quality at most stations, especially the C2, J1, and J2 stations. Moreover, it can be concluded from the hydrographic simulation results of four models that the spread of cost function values for temperature is smaller than that of salinity.

The mean cost function values for temperature and salinity also support the above results. As shown in Table 2, MRI.COM presents the lowest values, indicating the best data quality. The cost function values of HYCOM and NEMO are mid-table, whereas OFES presents the highest cost function values that indicate the lowest data quality.

3.1.3 Long-term variability

We further analyzed temperature and salinity at near-surface depths of the ESROB station, including the 5, 20, 40, 60, and 100 m depths for both models and observations. Figure 5 shows the time series of monthly mean temperature from 2000 to 2014 and that for salinity is shown in Figure 6. During this period, the observed temperature varies from 5°C to 25°C with the minimum occurring in winter and maximum appearing in summer, and the range of observed salinity extends from 33 g kg^{-1} in summer to 34.5 g kg^{-1} in winter. In upper layers, the seasonal variations of both temperature and salinity have much stronger magnitudes than that of interannual variations. As the depth increases, the interannual variations become more pronounced in both ESROB observations and model simulations, which even overtake the magnitude of seasonal variations at 100 m depth. This conforms to the fact that

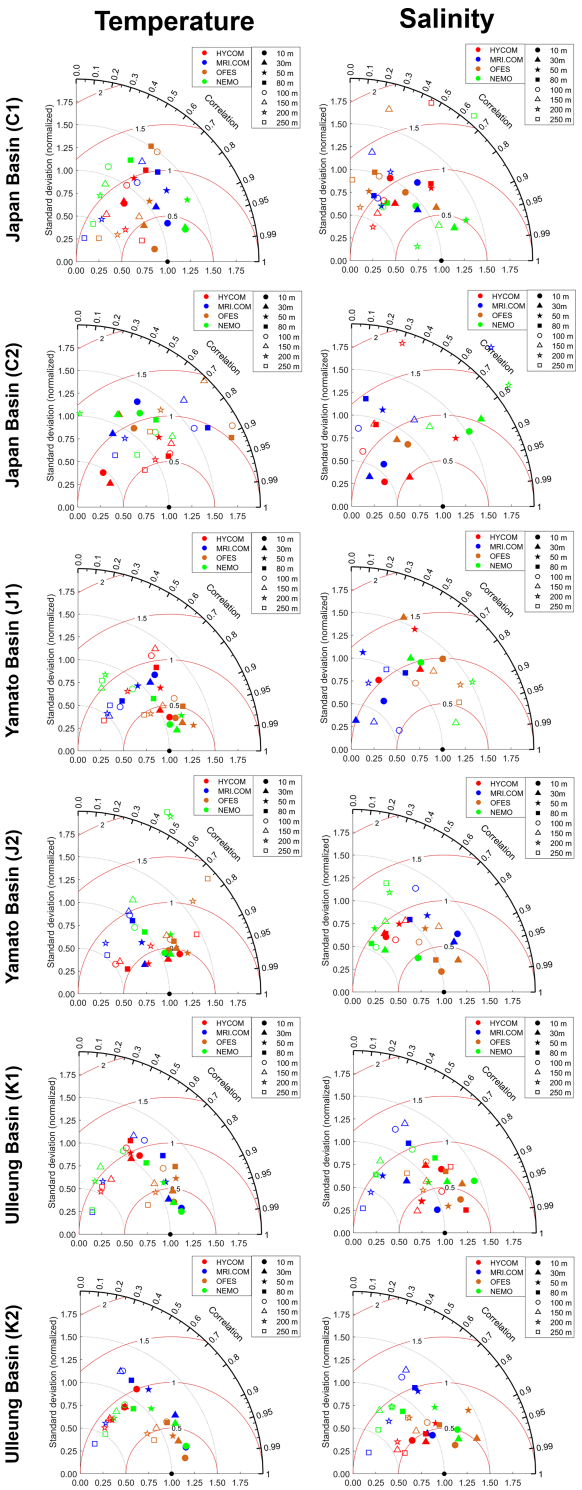


FIGURE 3
Taylor diagrams plotted from monthly time series of temperature (left) and salinity (right) from 1993 to 2014 at standard depths at the monitoring station C1 and C2 in the Japan Basin, J1 and J2 in the Yamato Basin, and K1 and K2 in the Ulleung Basin. Correlation, normalized standard deviation, and centered RMS difference of the simulated time series from the ocean models HYCOM (red) and NEMO (green), MRI.COM (orange), and OFES (blue) compared with the observational data are shown. Different markers refer to different standard depths. See text for more details.

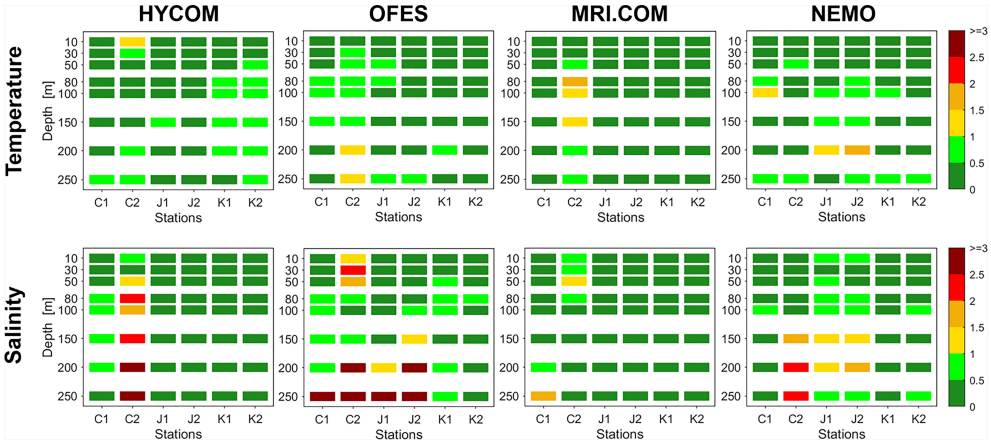


FIGURE 4
Values of Cost function derived from monthly time series of temperature (top) and salinity (bottom) from 1993 to 2014 at the six representative monitoring stations. Results from the ocean models HYCOM, OFES, MRI.COM, and NEMO (from left to right) compared with the observation data are shown at the same depths as illustrated in the Taylor diagrams in Figure 3. See text for more details.

subsurface water is less affected by the surface wind and heat flux forcing with strong seasonal variations (Byju et al., 2018). For both temperature and salinity, it is interesting to note that OFES agrees very well with the ESROB data both in magnitude and variability, whereas MRI.COM has an unexpectedly worse performance. Sudden changes in temperature and salinity, for instance, at the beginning of 1994 or 1998 are reproduced very realistically. At deeper layers, the differences between observation and simulation data become significantly larger.

The HYCOM and MRI.COM models partly reproduce the observed magnitudes very well but do not show the reasonable variations that are visible in the ESROB and OFES data. They overestimate or underestimate temperature at each depth from time to time by up to 5°C. NEMO generally overestimates the temperature by a maximum from 1°C at sea surface to 9°C at 100 m. Salinity at each depth is well reproduced by HYCOM, OFES, and MRI.COM with maximum deviations from the observations of no more than 0.3 g kg^{-1}. The best accordance of models and observations occurs at sea surface where OFES and MRI.COM simulate the observed seasonal variations precisely. Again, NEMO still presents the largest deviations compared with ESROB data, especially in deep layers where the maximum exceeds 0.5 g kg^{-1} from time to time.

3.2 Mean circulation

3.2.1 Horizontal surface circulation patterns

A comparison with Drifter ocean current at 15 m is conducted to evaluate the surface circulation patterns simulated by the four eddy-resolving models. The ocean current structure observed by Drifter buoys at 15 m is shown in Figure 7. Statistics show that most sea surface areas of the Japan Sea are covered by current velocities varying from 0 to 0.3 m s^{-1}. Maxima velocity of up to 1 m s^{-1} appears around the Tsugaru Strait, up to 0.4 m s^{-1} in the Tsushima Strait, and up to 0.4 m s^{-1} in Peter the Great Bay. The majority of large flows that were previously observed in the Japan Sea has been well depicted by Drifter data, including the Liman Cold Current (LCC) along the Russia coast, the Nearshore Branch (NB), and Offshore Branch (OB) near the Japan coast, the EKWC and North Korea Cold Current (NKCC) along the Korea coast, as well as the SubPolar Front Current (SPFC) in the middle and wind-driven current near Vladivostok.

The left column of Figure 8 shows the difference in surface current velocities between model simulations and Drifter data. The overall strength of the surface circulation throughout the Japan Sea has been underestimated by all models as indicated by

TABLE 2 Average cost function values for temperature and salinity based on Equation 2 for the ocean models HYCOM, OFES, MRI.COM, and NEMO.

	HYCOM	OFES	MRI.COM	NEMO
Temperature	0.48	0.97	0.23	0.44
Salinity	0.37	0.38	0.24	0.46

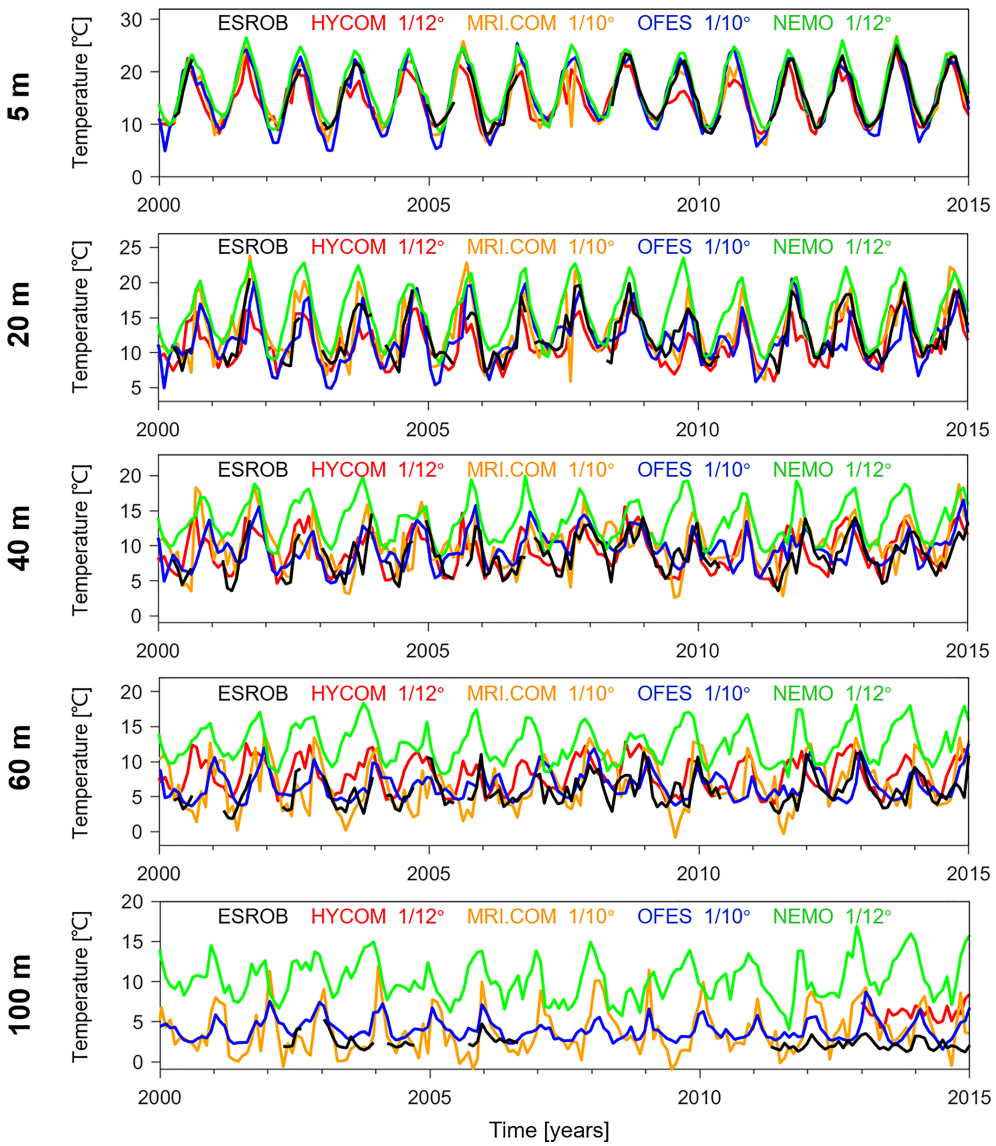


FIGURE 5
Monthly mean temperature from ESROB observation data (black) and the ocean models HYCOM (red), MRI.COM (orange), NEMO (green), and OFES (blue) at 5, 20, 40, 60, and 100 m depths for the time period from 2000 to 2014.

the large blue areas. Generally, most models have underestimated the strength of EKWC in the East Korea Bay and the SPFC at the south boundary of Japan Basin. Compared with other models, the difference between the 15-m current velocities simulated by MRI.COM and that observed by Drifter buoys shows relatively smaller deviations of about −0.30 m s^{-1} in maximum and −0.05 m s^{-1} on average. HYCOM and NEMO differ stronger with Drifter surface currents throughout the whole Japan Sea with the maximum deviation of −0.76 and −0.69 m s^{-1}, respectively. Both models overestimate the strength of NB along the whole Japan coast, but HYCOM underestimates the strength of the EKWC near the Korea coast, whereas NEMO tend to overestimate it. OFES still has the largest difference with Drifter. OFES simulates mainly smaller current speeds than other models. The average deviation is −0.07 m s^{-1}, and the maximum negative difference is approximately −0.94 m s^{-1}. The middle and right columns of Figure 8 show the differences in warm season (June through August) and cold season (December through February). The statistics are generally consistent between cold and warm seasons with similar distribution of bias values. The largest discrepancy appears south of Vladivostok where strong winter monsoon flows across the mountain gap, brings strong wind-driven current, and causes larger differences between simulated and observed current

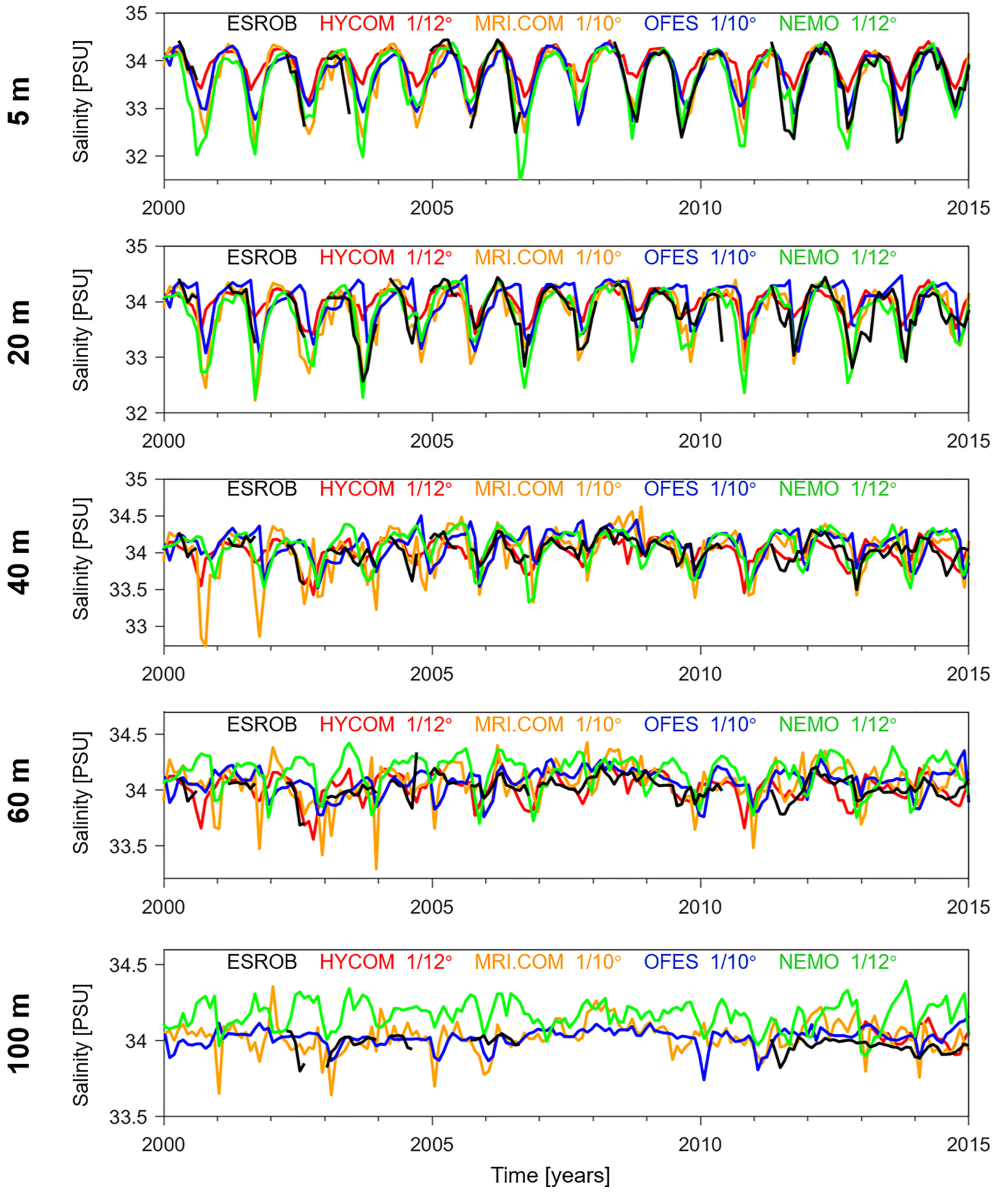


FIGURE 6
The same as Figure 5 but for the monthly mean salinity.

velocities. This is possibly related to the sparce resolution of wind forcing data because the monsoon and wind-driven current occupy a small area near the northwest coast of Japan Sea.

3.2.2 Current velocity through major straits

Figure 9 shows the average current velocities through the three major straits as simulated by four eddy-resolving ocean circulation models. Positive values indicate eastward currents for the Tsugaru and Soya Strait and northeastward currents for the Tsushima Strait, respectively.

The 22-year mean zonal velocity through the Tsushima Strait as simulated by all the four models shows a bifurcation structure in the downstream region and a weak countercurrent near the Tsushima Island. Generally, the current is northeast directed (positive) in the middle part of both channels maximizing approximately between 0 and 50 m depth at about 0.5 m s^{-1} for MRI.COM and OFES and 0.4 m s^{-1} for HYCOM. In the west channel, the current velocity simulated by HYCOM, MRI.COM, and OFES is generally consistent, with the maximum speed of 0.4 m s^{-1} appearing in the middle part of the sea surface. However, NEMO describes a very distinct flow

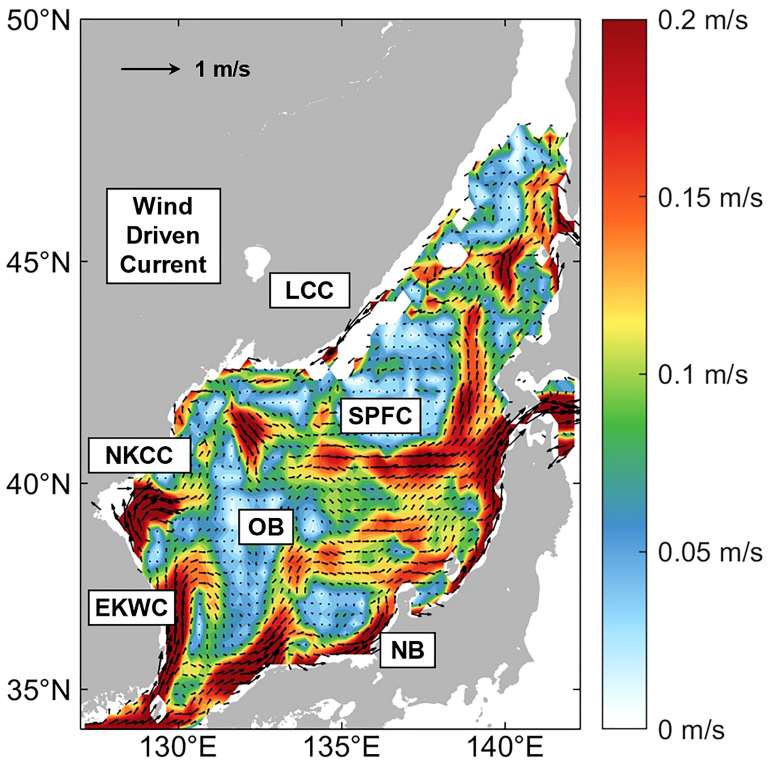


FIGURE 7
The climatological ocean circulation patterns at 15 m depth as represented by the Drifter interpolation dataset with the major flows labeled by white boxes. EKWC, East Korea Warm Current; NKCC, North Korea Cold Current; NB, Nearshore Branch; OB, Offshore Branch; SPFC, SubPolar Front Current; LCC, Liman Cold Current.

pattern with a maximum of 0.9 m s^{-1} and a strong flow core located near 25 m depth very close to the coastline. Near the Tsushima Island in the east channel, the current is southward directed (negative) and the negative velocities range from 0 m at the surface to 90 m at the bottom, which is also overestimated by NEMO. The middle part of the east channel is characterized by a northeastward current near the sea surface with a maximum of 0.4 m s^{-1}. Near the Japan coast is the southwestward-directed (negative) current extending to the bottom. However, for NEMO, the velocities are negligible whereas HYCOM, MRI.COM, and OFES simulate a maximum velocity about 0.1 m s^{-1}. Overall, the currents simulated by NEMO are stronger than HYCOM, MRI.COM, and OFES. The vertical profile simulated by four models is comparable to that observed by ADCP (Ostrovskii et al., 2009) except for the common overestimations of velocity in the east channel.

The zonal current through the Tsugaru Strait is predominantly eastward (positive) over the whole transect with a maximum of 0.5–1.3 m s^{-1} depending on each model. MRI.COM simulates the weakest current throughout the transect with a maximum of 0.5 m s^{-1}. HYCOM and NEMO simulate comparatively stronger magnitudes than MRI.COM with the maximums appearing at the same location. OFES simulated the strongest current, but large velocities are mainly distributed near the southern coast with a narrow range both horizontally and vertically. This is not necessarily due to the 1/10° horizontal resolution of OFES because the MRI.COM model with the same horizontal resolution also captures the narrowest transect across the Tsugaru Strait, which shows similar distributions of velocity compared with HYCOM and NEMO. Nor it is likely due to the z-coordinates adopted by OFES in shallow water because reasonable simulation results are found in the Tsushima Strait, where the lateral boundaries are steep for z-coordinates. It is likely because OFES assimilates no observational data, or, possibly, it is related to the fact that the bathymetries near each strait are individually adapted to the numerical solvers of each numerical model. Similar to the Tsugaru Strait, the zonal currents through the Soya Strait reveal a unidirectional throughflow structure with a strong flow core in the middle of the sea surface. The velocity direction is predominantly eastward with the maximum of 0.5 m s^{-1} for MRI.COM and NEMO, 0.6 m s^{-1} for HYCOM, and 1.0 m s^{-1} for OFES. In principle, the patterns simulated by MRI.COM and NEMO are similar with HYCOM having slightly stronger magnitudes. OFES simulates the throughflow with a significant shift to the west and with stronger current velocities but smaller meridional and vertical ranges than the other models.

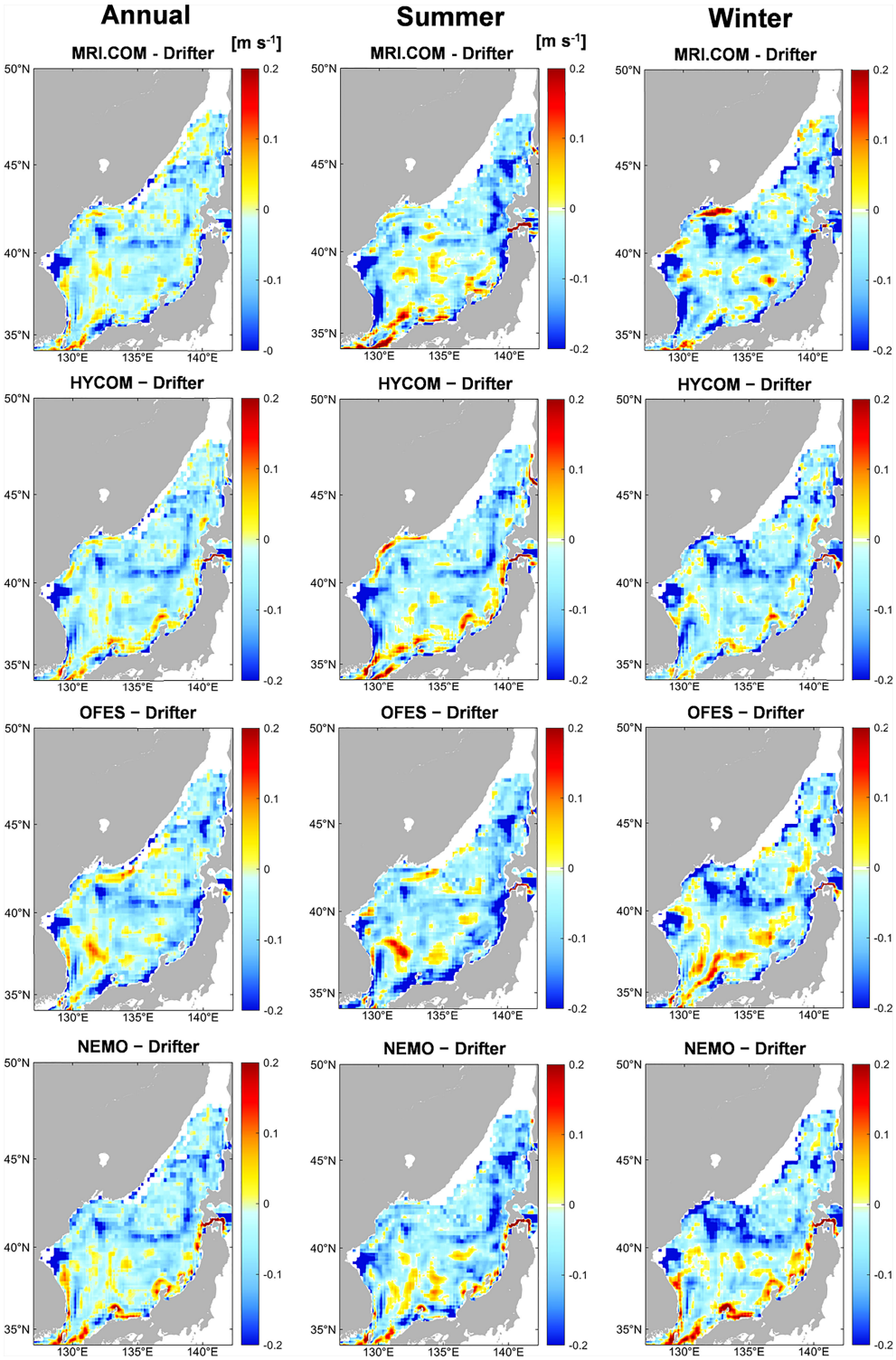


FIGURE 8
Twenty-two–year mean surface velocity at 15 m for the years 1993 through 2014, including the annual mean (left), the summer mean (middle), and the winter mean (right). The panels from top to bottom show the differences of the individual models MRI.COM, HYCOM, OFES, and NEMO from the Drifter reference data, respectively.

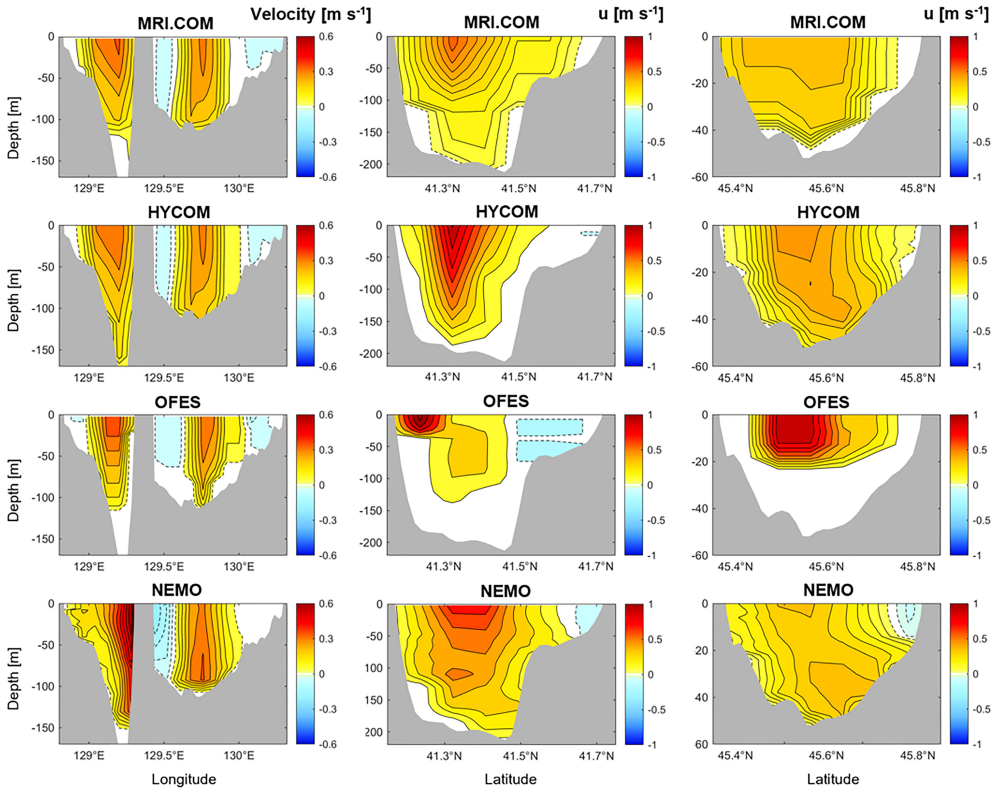


FIGURE 9
Mean current velocity profiles across the Tsushima Strait (left), Tsugaru Strait (middle), and Soya Strait (right) for the time period from 1993 to 2014 for the models MRI.COM, HYCOM, OFES, and NEMO. Positive values denote northeastward velocity for Tsugaru Strait and eastward velocity for Tsugaru Strait and Soya Strait. Isolines are solid for positive values and dashed for negative values. Note that the shown profiles cover different ranges of depth. The contour interval is 0.1 m/s for each graph.

3.2.3 Volume transport through main straits

Figure 10 shows the simulated 22-year average vertical profiles of volume transport per unit depth that are integrated horizontally along the three straits each. For the Tsushima Strait, the vertical profiles of MRI.COM, HYCOM, and NEMO are highly similar with maximum inflows of about 30,000 m^2 s^{-1} at sea surface, whereas OFES has the weakest magnitudes (the maximum is around 20,000 m^2 s^{-1}) but similar distributions. In the upper layers, all models simulate inflows into the Japan Sea with OFES having the narrowest horizontal and vertical ranges.

When only the east channel of the Tsushima Strait is considered, the vertical profiles greatly differ from the profiles for the whole strait. Because of the exclusion of northeastward inflow in the east channel, the Tsushima Strait east transect is characterized by a weaker inflow into the Japan Sea over the whole depth and the vertical range decreases from 170 m for the whole transect to no more than 150 m for the east channel. The almost unchanged and even increased volume transport per unit depth between the surface and 90 m depth simulated by HYCOM, MRI.COM, and NEMO is partly caused by the weakened countercurrent at the eastern coast near the Tsushima Island, which is strongest pronounced in NEMO near the sea surface. Therefore, NEMO maximizes in about 12,000 m^2 s^{-1} at 80 m depth. At a shallower depth of 40 m, HYCOM and MRI.COM reveal maximum inflows of about 11,000 and 10,000 m^2 s^{-1}, respectively, whereas OFES simulates maximum inflow of only 9,500 m^2 s^{-1} at the sea surface.

At the Tsugaru Strait, the vertical structure of volume transport per unit depth is, in principle, comparable to that at the Tsushima Strait except for the weaker magnitude. Because of the deeper topography, the flow range extends further down to the 200 m depth, but the maximum outflow still occurs near the sea surface for all models. The outflow magnitudes reach from almost 20,000 m^2 s^{-1} for NEMO to about 15,000 m^2 s^{-1} for HYCOM and MRI.COM and only 13,000 m^2 s^{-1} for OFES. The flow range extends down to 110 m for OFES, 210 m for NEMO, and 190 m for HYCOM and MRI.COM.

The mean vertical profiles of volume transport per unit depth at the Soya Strait have a similar shape with that of Tsushima and Tsugaru Strait. Maximum outflow from the Japan Sea to the Northwest Pacific takes place near the sea surface and extends to the depth of about 50 m for MRI.COM,

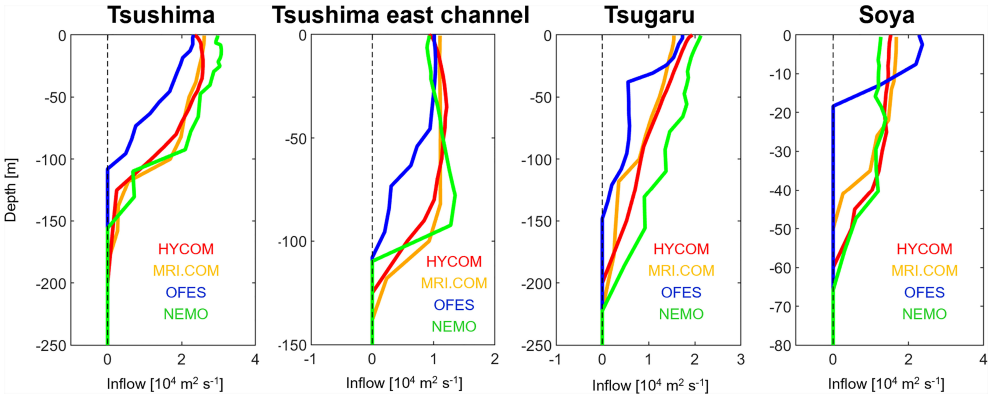


FIGURE 10
Vertical profiles of horizontally integrated velocities per unit depth along the Tsushima Strait, Tsugaru Strait, Soya Strait, as well as along the east channel of the Tsushima Strait only (between Tsushima Island and Japan) as simulated by the ocean models HYCOM (red), MRI.COM (orange), NEMO (green), and OFES (blue) for the time period from 1993 to 2014.

60 m for HYCOM, and 65 m for NEMO, but only 20 m for OFES. These are consistent with the current velocities illustrated by Figure 9. The maximum outflow ranges from almost 22,000 m^2 s^{-1} for OFES to about 17,000 m^2 s^{-1} for HYCOM and MRI.COM and 16,000 m^2 s^{-1} for NEMO.

In addition, the 22-year mean depth-integrated volume transport per unit length in three main straits is shown in Figure 11. For the Tsushima Strait, the volume transport simulated by each model is predominantly northeastward directed (positive) between Korea and the western coast of the Tsushima Island. The magnitudes for MRI.COM and HYCOM are generally the same, whereas OFES simulates a smaller transport and NEMO simulates a larger transport. Between the eastern coast of the Tsushima Island and Japan, a countercurrent structure is visible with southwestward transport along the eastern coast of the Tsushima Island, whereas the rest transport is dominantly northeastward directed. Here, the transports differ very slightly among the four models with the maximum close to 35 m^2 s^{-1}. For the Tsugaru Strait and Soya Strait, the whole transects are dominated by eastward (positive) outflow with the volume transport differing greatly in both magnitude and meridional distributions. The maximum outflow ranges from almost 40 m^2 s^{-1} for OFES to about 130 m^2 s^{-1} for HYCOM in the Tsugaru Strait, whereas, in the Soya Strait, there is a narrower range from 15 m^2 s^{-1} for OFES to about 27 m^2 s^{-1} for MRI.COM. Note that, compared with other models, MRI.COM simulates a significantly larger proportion of outflow through the Soya Strait, and a rather smaller proportion is allocated to the Tsugaru Strait.

We further analyzed the water balance through Japan Sea straits as simulated by four models between 1997 and 2008. The observed volume transport data are taken from ADCP measurements (Fukudome et al., 2010) in the Tsushima Strait and sea level difference across the Tsugaru Strait (Han et al., 2016). The observed volume transport is unavailable in the Soya Strait due to the lack of a tide gauge on the north side and the insufficient ADCP measurements. Figure 12 shows the differences between the transport entering the Japan Sea and the transport flowing out through the Tsugaru Strait and the Soya Strait from the observation data and four model simulations. All models reach a fine balance between the inflow transport into the Tsushima Strait and outflow transport into the Tsugaru and Soya Strait with the in-out differences smaller than 0.2×10^6 m^3 s^{-1}. Among them, OFES shows an underestimation of total transport by ~0.8×10^6 m^3 s^{-1}, NEMO slightly overestimates by ~0.2×10^6 m^3 s^{-1}, whereas both HYCOM and MRI.COM transport are comparable with the observations. The proportion of outflow transport through the Tsugaru Strait has been overestimated by each model, which is close to 55% from observation data but has exceeded 70% from the simulations of four models. The above suggests that simulations of numerical models satisfy the physical conservation of volume transport, but the allocation of outflow transport between the Tsugaru and Soya Strait cannot reach an agreement with the observations.

3.3 Evaluation of current velocities

The current velocities simulated by each model were validated against the mooring ADCP measurements at ESROB and EC1 stations, which locate at the axis of the longshore EKWC branch and the offshore EKWC branch, respectively.

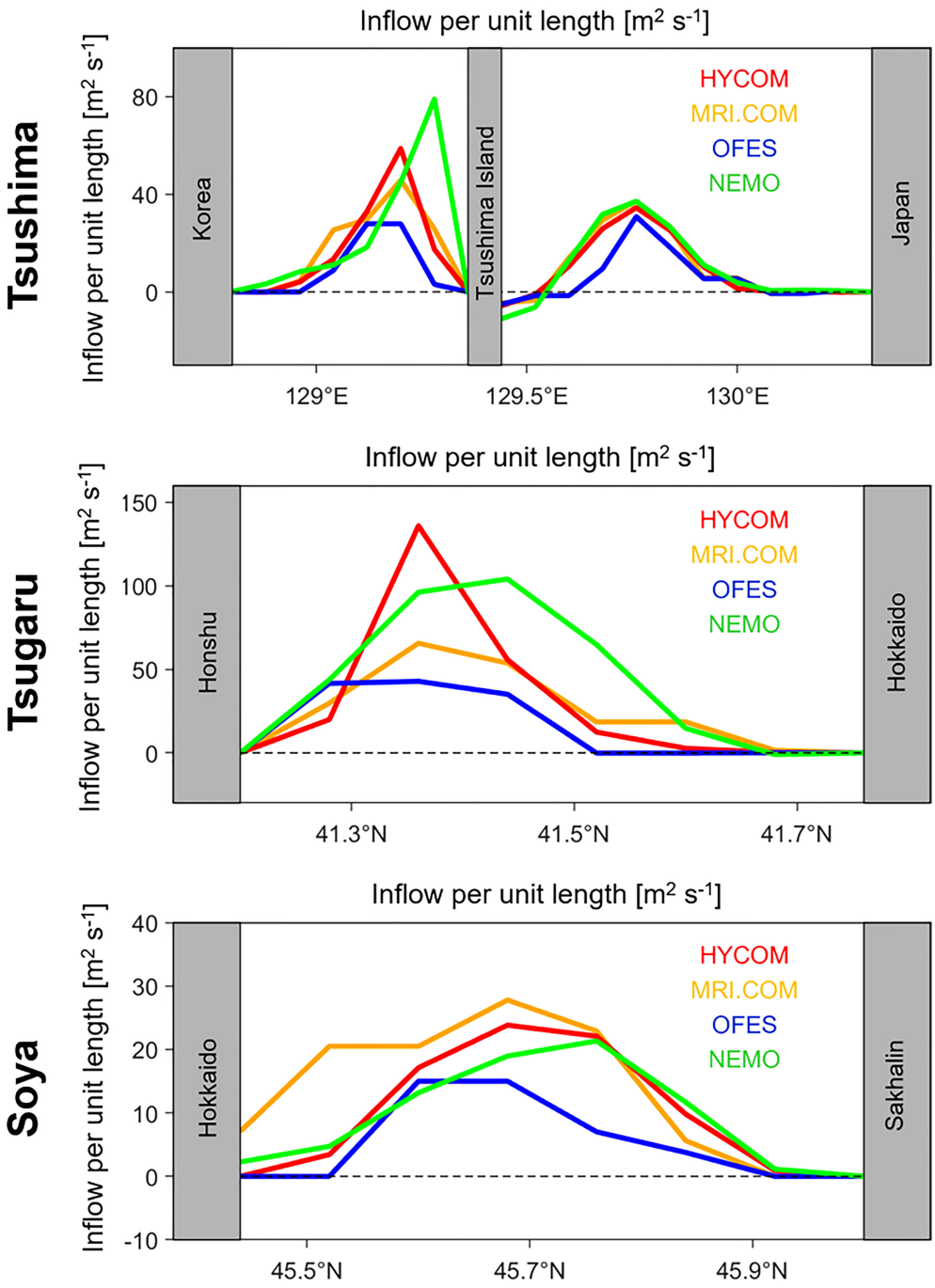


FIGURE 11
Depth-integrated velocities per unit length orthogonal across three major straits as simulated by the ocean models HYCOM (red), MRI.COM (orange), OFES (blue), and NEMO (green) for the time period from 1993 to 2014.

3.3.1 ESROB

The statistics of ADCP data are shown on the left-hand side of Figure 13 labeled by black markers, which are calculated from monthly time series at 20 m depth below the sea surface. The 10-year mean observed zonal current velocity is close to 1.5 cm s^{-1} with a standard deviation of about 6 cm s^{-1}. Each model simulates a similar magnitude of zonal current, but the standard deviations vary from 3 cm s^{-1} for HYCOM to 12 cm s^{-1} for NEMO. The relative frequency of the observed and simulated zonal currents shows a very good agreement with

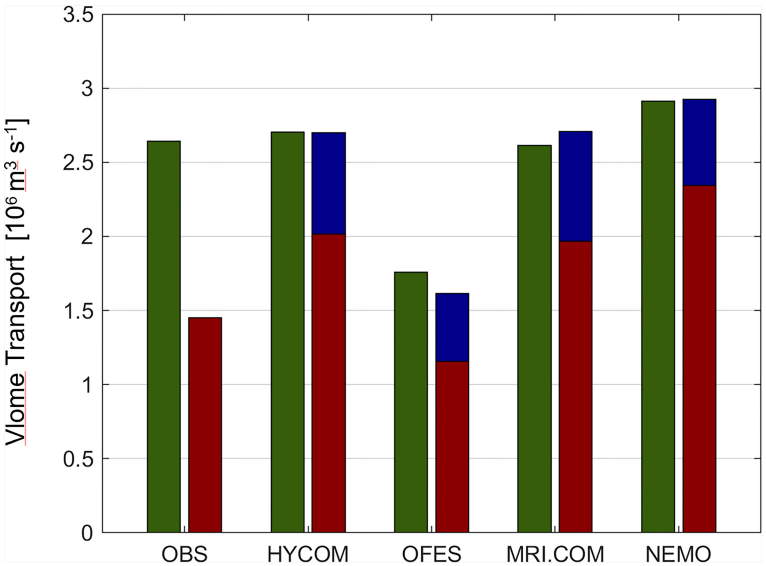


FIGURE 12
Differences between the volume transport entering the Japan Sea through the Tsushima Strait (green) and the outgoing transport through the Tsugaru (red) and Soya/La Perouse (blue) Straits for observed data, reanalyses from the four models, and consolidated estimations.

similar mean values and consistent spread. Note that a few outlier values of ADCP between −0.3 and −0.4 cm s^{-1} are not included in this histogram.

The mean meridional current velocities observed and simulated at 20 m depth are less in accordance with each other. The ADCP measurements are distributed around the mean value of 2 cm s^{-1} with a standard deviation of about 10 cm s^{-1}. MRI.COM and OFES lie very close to these values. HYCOM shows a similar variation with the observations but varies around the mean value of 1 cm s^{-1}. NEMO has a distinct mean value of about −11 cm s^{-1}, and the standard deviation is much higher, which is almost twice as large as that of ADCP data. The relative frequency values reflect good agreement of each model with ADCP measurements, except for the

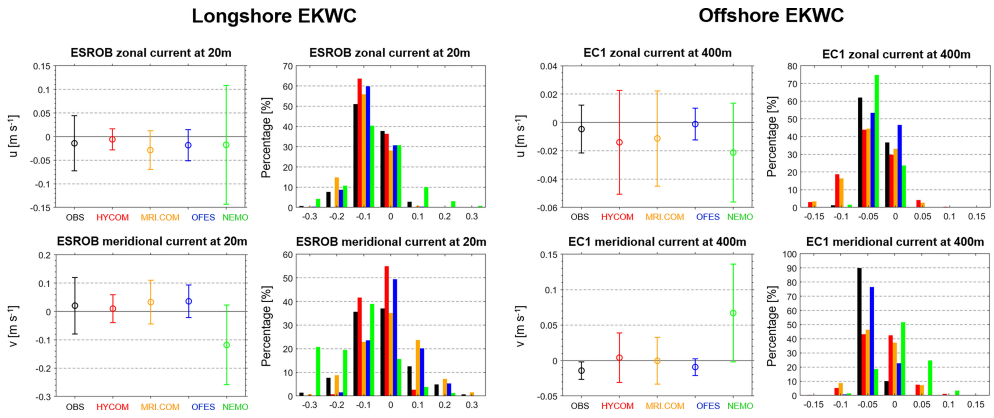


FIGURE 13
Mean current velocities with their standard deviations as well as relative frequency of these velocities from geostationary observations (black) from and model simulations (colors) at 20 m depth of the ESROB station located near the longshore branch of EKWC (left) and at 400 m depth of the EC1 located near the offshore branch of EKWC (right). Results for zonal components are shown in the top panels and for meridional components in the bottom panels. See text for further information.

NEMO model that presents significantly larger relative frequency values.

3.3.2 EC1

The statistics of ADCP observations at the EC1 station near the offshore branch of EKWC are shown on the right-hand side of Figure 13, which are similarly based on monthly averaged data, but at the depth of 400 m. NEMO still has a poor performance in simulating both the zonal and meridional current velocities with very different mean values and standard deviations. For the rest three models, the 15-year mean meridional current velocities and the standard deviations are very close to zero for both the ADCP observations and model simulations. The small range of deviation is also reflected by the large relative frequency from −5 to 0 cm s^{-1}. Larger discrepancies between ADCP measurements and model simulations occur for the zonal current velocities. From the ADCP data, it varies around the mean value of −0.5 cm s^{-1} with a standard deviation of about 1.5 cm s^{-1}. OFES has most similar mean values and variation range with ADCP. In contrast, HYCOM and MRI.COM models simulate larger mean values (~−1.5 cm s^{-1}) and standard deviations larger than 6 cm s^{-1}.

3.4 Evaluation of atmospheric forcing data

Finally, the wind velocities of ERA-Interim, CFSR, NCEP R1, and JRA-55, which drive the ocean circulation models, are evaluated by a comparison with QuikSCAT satellite data. The mean statistics over the whole Japan Sea, which are considered for this assessment, have been summarized in Table 3. They reveal partly substantial differences between QuikSCAT data and reanalysis forcing data. The average zonal wind from both observations and reanalysis is eastward directed with generally the same magnitude. The mean meridional wind is southward directed and reanalysis data reproduce similar magnitudes with observations. However, the wind speed averaged over the whole Japan Sea is significantly underestimated by each reanalysis dataset with the largest differences and RMSDs that occur for CFSR (HYCOM) and ERA-Interim (NEMO). In addition, the standard deviations, both for the component speed and total speed, are larger in observations than in reanalysis. Note that the smaller differences for JRA-55 (MRI.COM) and NECP R1 (OFES) correspond to the smaller deviations of MRI.COM and OFES current velocities with Drifter data.

The annual cycle and standard deviations of wind speed averaged throughout the whole Japan Sea are illustrated in Figure 14. Generally, all the four reanalysis datasets agree well with the QuikSCAT observations, but reanalysis datasets systematically underestimate the magnitude of mean wind speed, especially in autumn and winter months when the northwest monsoon prevails over the Japan Sea. NCEP R1 and JRA-55 are closer to the satellite observations but still underestimate the mean wind speed in autumn and winter and slightly overestimate it in May and July.

4 Discussions

In this assessment, the overall best simulation performance of temperature and salinity profiles is found for MRI.COM-WNP, a North Pacific regional ocean circulation model using the MOVE-4DVar data assimilation scheme to ingest a large amount of *in situ* measurements in Japan Sea. The temperature and salinity profiles utilized for this assessment belong to the World Ocean Database 13 (WOD13), which is a subset of the assimilation database for MRI.COM-WNP. Compared with the other three models with vertically adaptive coordinates, OFES and NEMO have the disadvantage of applying z-coordinates in the vertical direction, which may generate spurious mixing across the isopycnals (diapycnical mixing) in the interior of the ocean (Gräwe et al., 2015) especially in Japan Sea where steep sea floors are widely distributed. This unreal numerical mixing might result in the anomalous thermocline structures and mixed layer depths

TABLE 3 Mean zonal wind speed (u), meridional wind speed (v), and total wind speed (ws) together with their standard deviations for QuikSCAT, CFSR, ERA-Interim, JRA-55, and NCEP R1 dataset, as well as the mean and root mean square deviation (RMSD) between the reanalysis datasets (mod) and the satellite observations (quik) for the zonal, meridional, and total wind speed.

	QuikSCAT		CFSR		ERA-Interim		JRA-55		NCEP R1	
	Mean	**STD**	**Mean**	**STD**	**Mean**	**STD**	**Mean**	**STD**	**Mean**	**STD**
u	1.58	2.63	1.23	1.48	1.45	1.45	1.65	1.65	1.75	1.75
v	−0.52	1.73	−0.57	1.39	−0.71	1.43	−0.45	1.58	−0.38	1.81
ws	3.11	1.56	2.29	1.31	2.28	1.32	2.54	1.43	2.76	1.41
			Mean	**RMSD**	**Mean**	**RMSD**	**Mean**	**RMSD**	**Mean**	**RMSD**
u_{mod}-u_{quik}			−0.33	0.81	−0.13	0.82	0.07	0.57	0.17	0.35
v_{mod}-v_{quik}			0.05	0.36	0.19	0.37	−0.07	0.23	−0.14	0.05
ws_{mod}-ws_{quik}			−0.82	0.46	−0.83	0.62	−0.57	0.21	−0.36	0.24

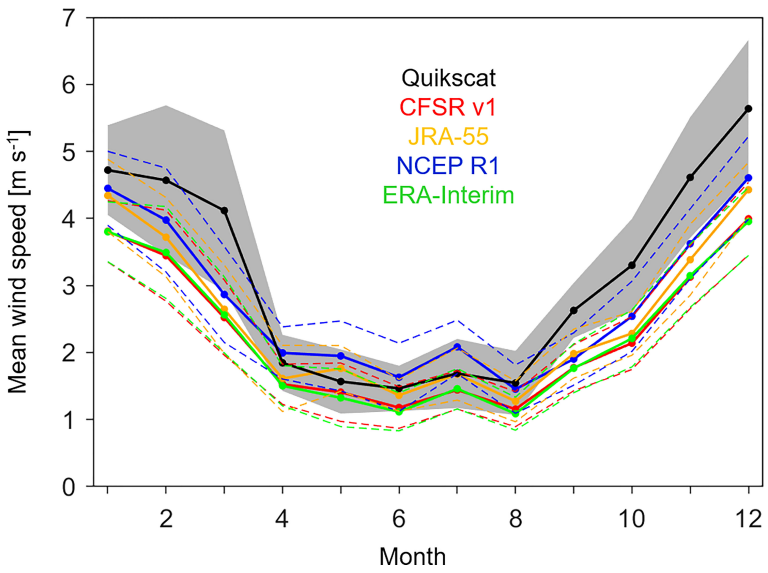


FIGURE 14
Annual cycle of monthly mean wind speed over the Japan Sea for the observations (black), CFSR v1 (red), JRA-55 (yellow), NECP R1 (blue), and ERA-Interim (green). The solid lines indicate the monthly mean and the shading indicates ±1 standard deviation.

(Figure 2). However, the thermocline reproduced by OFES is less biased than NEMO, which is possibly because the OFES model adopts a scale-selective damping of biharmonic operator to suppress the spurious computational noise. This opens the possibility to explore the effects of applying the biharmonic operator in NEMO to improve the parameterizations for surface waves and Langmuir circulation. Although numerical mixing is also present in HYCOM and MRI.COM, this mixing might mimic the effects of physical mixing and, therefore, corresponds better to the *in situ* observations than the deep water mixing parameterizations used in z-coordinate models.

Because of the assimilation of a very large volume of *in situ* temperature and salinity measurements in the Japan Sea, we assume that the baroclinic current velocity fields simulated by MRI.COM are closest to observations. However, the barotropic currents, which are driven by the wind stress curl minus the bottom friction curl, are not influenced by data assimilation and, consequently, not constrained by temperature and salinity measurements.

In the following, several hypotheses related to the diversities in both hydrographic and dynamical conditions among four ocean circulation models are proposed:

1. Because all models tend to reproduce temperature much better than salinity, it is induced that the calibration of heat balance is easier than that of water balance in the Japan Sea. However, there are detailed differences among different eddy-resolving models. Largest deviations in temperature are found from OFES and NEMO data possibly because the z-coordinates are not well simulating the vertical mixing processes. In contrast, smaller biases are found for MRI.COM and HYCOM, which applies terrain-following coordinates in the shallow water.
2. Mean surface velocities in the southern part of the Japan Sea, where most large surface currents are distributed, are significantly stronger in NEMO than in the other models. A possible explanation is the zooming of vertical layers toward the sea surface in NEMO. The minimum layer thickness can reach 1 m, and there are more than 20 layers above 100 m, which enables a better representation of the wind shear effect on the upper layers. Whether the stronger barotropic velocities simulated by NEMO are more realistic than that simulated by other models cannot be determined as mentioned above. Because OFES also has dense layers near the sea surface but simulates a weaker barotropic circulation in the south part than MRI.COM, either the bottom fraction or the horizontal viscosity in OFES is significantly larger than in MRI.COM. Moreover, all the four models have underestimated the ocean current at 15 m compared with Drifter observations, especially near the Peter the Great Bay where wind-driven current is dominant. The possible reason behind might be the weaker wind speeds of the atmospheric forcing in each model compared with the reality.
3. Compared with the other five stations, both the temperature and salinity at the C2 station located in

the Peter the Great Bay (PGB) are not well reproduced by each of the four model. This problem might be caused by a different type of water mass formation process in PGB under the combined influence of strong monsoon, land contaminations, submesoscale eddies, and the dynamics of the Primorye Current. These processes have not been fully considered or ideally simulated by the present version of eddy-resolving ocean circulation models. Therefore, a regional ocean circulation model with higher resolution is expected to be established in the Japan Sea to better reproduce the detailed local ocean environment characteristics.

4. We identified several feature differences of near-surface ocean currents among the models. For example, the southwestward-directed LCC along the coast of Russia is less centered and stronger in OFES than in the other three models. Another detail is the stronger loop flow of the EKWC in both the Ulleung Basin and the East Korea Bay in OFES and NEMO than in MRI.COM. This might suggest the potential influence of dense vertical layers near the sea surface on the simulation quality of EKWC. However, similarly, using the terrain-followed coordinates in upper ocean, in HYCOM, the number of vertical layer near above 100 m is locally larger than that of MRI.COM, but the strength of EKWC does not differ a lot between two models, and thus, the change in upper layer density has no larger effect. Furthermore, the barotropic circulation in the Tsugaru Strait shows very large differences among the four models, which are possibly due to the differing bottom topography across the strait transect and how the used bathymetries adapt to the numerical solvers of each model.
5. The OFES and NEMO model has worse performed in the Japan Sea compared with HYCOM and MRI.COM. The most possible reason might be that the application of z-coordinates in OFES and NEMO is less optimized for an accurate simulation of the surface ocean circulation patterns and hydrographic properties in the Japan Sea. The absence or the inadequacy of data assimilation in the Japan Sea might only be responsible for a part of the total deviations. Although NEMO has assimilated CORA v4.1 database into the dynamic framework similar to that of OFES with similar model setup, subgrid scale parameterizations, and parameter settings, it still has worse performance compared with HYCOM and MRI.COM. Hence, there are no reasons to assume that OFES with the same data assimilation will perform as good as HYCOM or MRI.COM.

5 Summary

In this study, we conducted a comprehensive assessment of long-term mean circulation in the Japan Sea as a typical semi-closed marginal sea as reproduced by multiple eddy-resolving ocean circulation models: the HYCOM, MRI.COM, OFES, and NEMO. The capability of these models to simulate hydrographic and dynamic conditions of the Japan Sea is evaluated over the 22-year period from 1993 to 2015. Simulation results of temperature and salinity at representative ocean monitoring stations are evaluated by comparing the vertical profiles and statistical time series at various depths with *in situ* measurements by introducing the Taylor diagrams and cost functions. The *in situ* measurements are provided by the post-processed data observed by CREAMS, JMA cruises, and KODC stations.

Results show that the observed temperature and salinity data are most realistically simulated by the northwest pacific version of MRI.COM model, which holds for temporal variations, magnitudes, and vertical profiles of these parameters. HYCOM and NEMO models present larger deviations from *in situ* measurements than that of MRI.COM at certain depths of representative ocean monitoring stations, but they still agree with the observation data better than OFES model. Generally, each of the four models well reproduces the temperature profiles from sea surface to 200 m depth. Salinity simulations are of predominantly good to reasonable quality for MRI.COM and HYCOM independently of the depth except for stations located in the north part of the Japan Basin. The largest deviations occur for salinity simulated by OFES.

The investigation of the surface circulation below 30 m and of the depth-integrated volume transport from 0 to 300 m throughout the Japan Sea shows partly different circulation patterns among the four models with OFES and NEMO, revealing stronger differences to both MRI.COM simulations and Drifter observations. The best agreement of currents and transport is found between HYCOM and MRI.COM. Large deviations are mainly distributed along the Japan coast, in the East Korea Bay, and along the south coast of Russia, where the large current velocities are distributed.

Furthermore, the mean current velocities and flow structures across three major straits around the Japan Sea, including the Tsushima Strait, the Tsugaru Strait, and the Soya Strait, are examined. Generally, all of the four models show similar velocity patterns across the Tsushima Strait although the magnitude and location of some circulation patterns are diversified. Smallest differences to MRI.COM are found for HYCOM, and strongest deviations occur for NEMO. In the Tsugaru and Soya Strait, OFES reproduced very different current structure and much stronger magnitude of velocity compared with the other three

models. The horizontal and vertical distributions of volume transport generally show consistent shapes among four models except for the significant underestimation from OFES at each of the three straits. It is also found that all the four models overestimate the proportion of outflow transport into the Tsugaru Strait by around 10 percent compared with that of observed results.

The variations of simulated meridional and zonal current velocities from ocean models with mooring ADCP measurements at two monitoring stations present an overall acceptable accordance. Except for the too large standard deviations and inconsistent mean values from NEMO, the mean current velocities simulated by the other three models lie predominantly within the standard deviation of ADCP measurements. More observations at other depths are expected to allow a more comprehensive evaluation of current velocities at key locations in the future.

The assessment of atmospheric products that drive each ocean circulation model shows that the average wind speed over the entire Japan Sea is slightly underestimated by each forcing dataset. Compared with QuikSCAT satellite observations, the CFSR and ERA-Interim products, which drives the HYCOM and NEMO models, respectively, present larger deviations at both zonal and meridional directions. Seasonal comparison shows that the underestimations are more significant in winter time when the southeastward monsoon prevails throughout the Japan Sea and is less reproduced by atmospheric reanalysis products. This indicates that satellite wind products might be more suitable to drive regional modeling in the Japan Sea than atmosphere reanalysis products.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding authors.

Author contributions

HW performed the data analysis, wrote the initial draft, and modified the manuscript. WZ proposed the main ideas and modified the manuscript. ML and CY provided high-performance data processing and computing platforms. KR supervised all processes. All authors contributed to the article and approved the submitted version.

Funding

The present study is supported by the National Key Research Program of China (2017YFA0604100) and is partly supported by the National Key Research and Development Program of China (No. 2018YFC1406206). This research is partially supported by the National Natural Science Foundation of China (Grant No. 61802424). This study is also supported by The science and technology innovation Program of Hunan Province (ID: 2022RC3070).

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Bleck, R. (2002). An oceanic general circulation model framed in hybrid isopycnic-Cartesian coordinates. *Ocean Model.* 4 (1), 55–88. doi: 10.1016/S1463-5003(01)00012-9

Byju, P., Dommenget, D., and Alexander, M. (2018). Widespread reemergence of sea surface temperature anomalies in the global oceans, including tropical regions forced by reemerging winds. *Geophysical Res. Lett.* 45 (15), 7683–7691. doi: 10.1029/2018GL079137

Chang, K. I., Hogg, N. G., Suk, M. S., Byun, S. K., Kim, Y. G., and Kim, K. (2002). Mean flow and variability in the southwestern East Sea. *Deep Sea Res. Part I: Oceanographic Res. Papers* 49 (12), 2261–2279. doi: 10.1016/S0967-0637(02)00120-6

Cheng, Y., Plag, H. P., Hamlington, B. D., Xu, Q., and He, Y. (2015). Regional sea level variability in the bohai Sea, yellow Sea, and East China Sea. *Continental Shelf Res.* 111, 95–107. doi: 10.1016/j.csr.2015.11.005

Cheon, W. G. (2020). The Summer/Fall variability of the southern East/Japan Sea in the ENSO period. *Ocean Sci. J.* 55 (3), 341–352. doi: 10.1007/s12601-020-0027-5

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-interim reanalysis: Configuration and performance of the data assimilation system. *Q. J. R. Meteorological Soc.* 137 (656), 553–597. doi: 10.1002/qj.828

Eilola, K., Gustafsson, B. G., Kuznetsov, I., Meier, H., Neumann, T., and Savchuk, O. (2011). Evaluation of biogeochemical cycles in an ensemble of three state-of-the-art numerical models of the Baltic Sea. *J. Mar. Syst.* 88 (2), 267–284. doi: 10.1016/j.jmarsys.2011.05.004

Fayman, P., Ostrovskii, A., Lobanov, V., Park, J. H., Park, Y. G., and Sergeev, A. (2019). Submesoscale eddies in Peter the great bay of the Japan/East Sea in winter. *Ocean Dynamics* 69 (4), 443–462. doi: 10.1007/s10236-019-01252-8

Fujii, Y., Mitsudera, H., Nakamura, T., Nishigaki, H., Miyama, T., Wagawa, T., et al. (2016). *Pathway of the kuroshio water traveling to the Bering Sea in a western north pacific eddy-resolving model analyzed with the tangent linear and adjoint models*.

Fukudome, K. I., Yoon, J. H., Ostrovskii, A., Takikawa, T., and Han, I. S. (2010). Seasonal volume transport variation in the tsushima warm current through the tsushima straits from 10 years of ADCP observations. *J. oceanography* 66 (4), 539–551.

Gräwe, U., Holtermann, P., Klingbeil, K., and Burchard, H. (2015). Advantages of vertically adaptive coordinates in numerical models of stratified shelf seas. *Ocean Model.* 92, 56–68. doi: 10.1016/j.ocemod.2015.05.008

Halliwell, G., Bleck, R., and Chassignet, E. (1998). "Atlantic Ocean simulations performed using a new hybrid-coordinate ocean model", in *EOS, Fall 1998 AGU Meeting*.

Halliwell, G., Bleck, R., Chassignet, E., and Smith, L. (2000). Mixed layer model validation in Atlantic ocean simulations using the hybrid coordinate ocean model (HYCOM). *Eos* 80, OS304. doi: 10.1007/s10872-010-0045-5

Han, S., Hirose, N., and Kida, S. (2018). The role of topographically induced form drag on the channel flows through the East/Japan Sea. *J. Geophysical Research: Oceans* 123 (9), 6091–6105. doi: 10.1029/2018jc013903

Han, S., Hirose, N., Usui, N., and Miyazawa, Y. (2016). Multi-model ensemble estimation of volume transport through the straits of the East/Japan Sea. *Ocean Dynamics* 66 (1), 59–76. doi: 10.1007/s10236-015-0896-9

Hirose, N., Liu, T., Takayama, K., Uehara, K., Taneda, T., and Kim, Y. H. (2021). Vertical viscosity coefficient increased for high-resolution modeling of the Tsushima/Korea strait. *J. Atmospheric Oceanic Technol.* 38 (6), 1205–1215. doi: 10.1175/JTECH-D-20-0156.1

Hirose, N., Takatsuki, Y., Usui, N., Wakamatsu, T., Tanaka, Y., Toyoda, T., et al. (2016). Four-dimensional variational ocean ReAnalysis for the Western north pacific over 30 years (FORA-WNP30). *AGU Fall Meeting Abstracts* 2016, NG41B-1737.

Ishida, H., Isono, R. S., Kita, J., and Watanabe, Y. W. (2021). Long-term ocean acidification trends in coastal waters around Japan. *Sci. Rep.* 11 (1), 5052. doi: 10.1038/s41598-021-84657-0

Ishikawa, I. (2005). Meteorological research institute community ocean model (MRI.COM) manual. *Tech Rep. Meteorol. Res. Inst.* 47, 1–189.

Jean-Michel, L., Eric, G., Romain, B. B., Gilles, G., Angélique, M., Marie, D., et al. (2021). The Copernicus global 1/12° oceanic and sea ice GLORYS12 reanalysis. *Front. Earth Sci.* 9. doi: 10.3389/feart.2021.698876

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., et al. (1996). The NCEP/NCAR 40-year reanalysis project. *Bull. Am. Meteorological Soc.* 77 (3), 437–472. doi: 10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2

Kang, S. K., Cherniawsky, J. Y., Foreman, M. G. G., So, J. K., and Lee, S. R. (2008). Spatial variability in annual sea level variations around the Korean peninsula. *Geophysical Res. Lett.* 35 (3). doi: 10.1029/2007GL032527

Kara, A. B., Rochford, P. A., and Hurlburt, H. E. (2000). Efficient and accurate bulk parameterizations of air–Sea fluxes for use in general circulation models. *J. Atmospheric Oceanic Technol.* 17 (10), 1421–1438. doi: 10.1175/1520-0426(2000)017<1421:EAABPO>2.0.CO;2

Kawamura, H., Ito, T., Hirose, N., Takikawa, T., and Yoon, J.-H. (2009). Modeling of the branches of the tsushima warm current in the Eastern Japan Sea. *J. Oceanography* 65 (4), 439–454. doi: 10.1007/s10872-009-0039-3

Kida, S., Takayama, K., Sasaki, Y., Matsuura, H., and Hirose, N. (2020). Increasing trend in Japan Sea throughflow transport. *J. Oceanography* 2), 1–9. doi: 10.1007/s10872-020-00563-5

Kim, Y. Y., Cho, Y. K., and Kim, Y. H. (2018). Role of cold water and beta-effect in the formation of the East Korean warm current in the East/Japan Sea: a numerical experiment. *Ocean Dynamics* 68 (8), 1013–1023. doi: 10.1007/s10236-018-1175-3

Kim, K., Kim, Y. B., Park, J. J., Nam, S., Park, K. A., and Chang, K. I. (2005). Long-term and real-time monitoring system of the East/Japan sea. *Ocean Sci. J.* 40 (1), 25–44. doi: 10.1007/BF03023463

Kim, D., Shin, H. R., Kim, C. H., and Hirose, N. (2020). Characteristics of the East Sea (Japan Sea) circulation depending on surface heat flux and its effect on branching of the tsushima warm current. *Continental Shelf Res.* 192, 104025. doi: 10.1016/j.csr.2019.104025

Kim, C. H., and Yoon, J. H. (1999). A numerical modeling of the upper and the intermediate layer circulation in the East Sea. *J. Oceanography* 55 (2), 327–345. doi: 10.1023/A:1007837212219

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., et al. (2015). The JRA-55 reanalysis: General specifications and basic characteristics. 気象集誌. 第*2*輯 93 (1), 5–48. doi: 10.2151/jmsj.2015-001

Kondo, J. (1975). Air-sea bulk transfer coefficients in diabatic conditions. *Boundary-Layer Meteorology* 9 (1), 91–112. doi: 10.1007/BF00232256

Large, W. G., McWilliams, J. C., and Doney, S. C. (1994). Oceanic vertical mixing: A review and a model with a nonlocal boundary layer parameterization. *Rev. Geophysics* 32 (4), 363–403. doi: 10.1029/94RG01872

Lellouche, J. M., Greiner, E., Le Galloudec, O., Garric, G., Regnier, C., Drevillon, M., et al. (2018). Recent updates to the Copernicus marine service global ocean monitoring and forecasting real-time 1⁄ 12° high-resolution system. *Ocean Sci.* 14 (5), 1093–1126. doi: 10.5194/os-14-1093-2018

Madec, G. (2008). *NEMO reference manual, ocean dynamics component: NEMO-OPA. preliminary version* Vol. 27 (France: Note du Pole de modélisation, Institut Pierre-Simon Laplace (IPSL), 1288–1161.

Madec, G., Bourdallé-Badie, R., Bouttier, P.-A., Bricaud, C., Bruciaferri, D., Calvert, D., et al. (2017). NEMO ocean engine.

Masumoto, Y., Sasaki, H., Kagimoto, T., Komori, N., Ishida, A., Sasai, Y., et al. (2004). A fifty-year eddy-resolving simulation of the world ocean: Preliminary outcomes of OFES (OGCM for the earth simulator). *J. Earth Simulator* 1, 35–56. doi: 10.32131/jes.1.35

Menezes, V. V. (2021). Advective pathways and transit times of the red Sea overflow water in the Arabian Sea from Lagrangian simulations. *Prog. Oceanography* 199, 102697. doi: 10.1016/j.pocean.2021.102697

Minato, S., and Kimura, R. (1980). Volume transport of the Western boundary current penetrating into a marginal Sea. *J. Oceanogr. Soc. Japan.* 36 (4), 185-195. doi: 10.1007/BF02070331

Morison, J., Andersen, R., Larson, N., D'Asaro, E., and Boyd, T. (1994). The correction for thermal-lag effects in Sea-bird CTD data. *J. Atmospheric Oceanic Technol.* 11 (4), 1151–1164. doi: 10.1175/1520-0426(1994)011<1151:TCFTLE>2.0.CO;2

Myrberg, K., and Andrejev, O. (2006). Modelling of the circulation, water exchange and water age properties of the gulf of bothnia. *Oceanologia* 48 (S). doi: 10.3390/fluids7040134

Noh, Y., and Kim, H. J. (1999). Simulations of temperature and turbulence structure of the oceanic boundary layer with the improved near-surface process. *J. Geophysical Research: Oceans* 104 (C7), 15621–15634. doi: 10.1029/1999JC900068

Ohshima, K. I. (1987). On the stability of the soya warm current. *J. Oceanographical Soc. Japan* 43 (1), 61–67. doi: 10.1007/BF02110634

Ohshima, K. I. (1994). The flow system in the Japan Sea caused by a sea level difference through shallow straits. *J. Geophysical Research: Oceans* 99 (C5), 9925–9940. doi: 10.1029/94JC00170

Ostrovskii, A. G., Fukudome, K., Yoon, J. H., and Takikawa, T. (2009). Variability of the volume transport through the Korea/Tsushima strait as inferred from the shipborne acoustic Doppler current profiler observations in 1997–2007. *Oceanology* 49 (3), 338–349. doi: 10.1134/S0001437009030060

Park, Y. G., Park, J. H., Lee, H. J., Min, H. S., and Kim, S. D. (2013). The effects of geothermal heating on the East/Japan Sea circulation. *J. Geophysical Research: Oceans* 118 (4), 1893–1905. doi: 10.1002/jgrc.20161

Pätsch, J., Burchard, H., Dieterich, C., Gräwe, U., Gröger, M., Mathis, M., et al. (2017). An evaluation of the north Sea circulation in global and regional models relevant for ecosystem simulations. *Ocean Model.* 116, 70–95. doi: 10.1016/j.ocemod.2017.06.005

Prants, S. V., Fayman, P. A., Budyansky, M. V., and Uleysky, M. Y. (2022). Simulation of winter deep slope convection in Peter the great bay (Japan Sea). *Fluids* 7 (4), 134. doi: 10.1175/1520- 0485(1988)018<1601:AGCMFU>2.0.CO;2

Rosati, A., and Miyakoda, K. (1988). A general circulation model for upper ocean simulation. *J. Phys. Oceanography* 18 (11), 1601–1626. doi: 10.1029/2000JD900719

Sekine, Y. (1986). Wind-driven circulation in the Japan Sea and its influence on the branching of the tsushima current. *Prog. Oceanography* 17 (3), 297–312. doi: 10.1016/0079-6611(86)90051-0

Spall, M. A. (2002). Wind- and buoyancy-forced upper ocean circulation in two-strait marginal seas with application to the Japan/East Sea. *J. Geophysical Research: Oceans* 107 (C1), 6–1-6-12. doi: 10.1029/2001JC000966

Stepanov, D. V., Ryzhov, E. A., Berloff, P., and Koshel, K. V. (2020). Floating tracer clustering in divergent random flows modulated by an unsteady mesoscale ocean field. *Geophysical Astrophysical Fluid Dynamics* 114 (4-5), 690–714. doi: 10.1080/03091929.2020.1786551

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *J. Geophysical Research: Atmospheres* 106 (D7), 7183–7192.

Tsujino, H., Usui, N., and Nakano, H. (2006). Dynamics of kuroshio path variations in a high-resolution general circulation model. *J. Geophysical Research: Oceans* 111 (C11). doi: 10.1029/2005JC003118

Usui, N., Ogawa, K., Yasuda, T., Hirose, N., and Kuragano, T. (2016). *Mechanism of interannual to decadal sea level variability along the Japanese coast* American Geophysical Union 2016, PO43B-06.

Usui, N., Wakamatsu, T., Tanaka, Y., Hirose, N., Toyoda, T., Nishikawa, S., et al. (2017). Four-dimensional variational ocean reanalysis: a 30-year high-resolution

dataset in the western north pacific (FORA-WNP30). *J. Oceanography* 73 (2), 205–233. doi: 10.1007/s10872-016-0398-5

Wang, H., Chen, S., Wang, N., Yu, P., Yang, X., Wang, Y., et al. (2020). "Evaluation of multi-model current data in the East/Japan sea," in *2020 IEEE 3rd International Conference on Information Communication and Signal Processing (ICICSP)* (IEEE). 486–491.

Yi, S. (1966). Seasonal and secular variations of the water volume transport through the Korea strait. *J. Oceanological Soc. Korea* 1 (1).

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 01 December 2022
DOI 10.3389/fmars.2022.1023694

Check for updates

OPEN ACCESS

EDITED BY
Yasumasa Miyazawa,
Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Japan

REVIEWED BY
Teruhisa Okada,
Central Research Institute of Electric Power Industry (CRIEPI), Japan
Shiqiu Peng,
South China Sea Institute of Oceanology, Chinese Academy of Sciences, China

*CORRESPONDENCE
Xuefeng Zhang
xuefeng.zhang@tju.edu.cn
Wei Li
liwei1978@tju.edu.cn

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

RECEIVED 20 August 2022
ACCEPTED 16 November 2022
PUBLISHED 01 December 2022

CITATION
Hu Y, Zhang X and Li W (2022) Variational parameter estimation in a two-equation turbulence model: A case study with a 3D primitive-equation ocean model. *Front. Mar. Sci.* 9:1023694. doi: 10.3389/fmars.2022.1023694

COPYRIGHT
© 2022 Hu, Zhang and Li. 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(s) 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.

Variational parameter estimation in a two-equation turbulence model: A case study with a 3D primitive-equation ocean model

Yan Hu, Xuefeng Zhang* and Wei Li*

School of Marine Science and Technology, Tianjin University, Tianjin, China

A three-dimensional and complete adjoint model of the Princeton Ocean Model with a generalized coordinate system (POMgcs) is developed to construct the 4D-Variational data assimilation (4D-Var) algorithm in this study. Uncertain parameters in the Mellor-Yamada 2.5 turbulence submodel (MY-2.5) which is enclosed in POMgcs, are tentatively estimated *via* the 4D-VAR algorithm within a biased model framework. Here, the control variables in the biased model are set to two uncertain wave-affected parameters (wave energy factor α and Charnock coefficient β) in the MY-2.5 turbulence model, which play a crucial role in modulating the heat content distribution in the upper coastal sea. First of all, the ocean temperature and salinity in a typical coastal sea, Bohai Sea, are simulated by the model to validate the rationally of the MY-2.5 parameterization scheme for both constructing the "truth model" and generating the "pseudo-observations" in the data assimilation studies. Then, after thoroughly testing the ability of the 4D-Var to optimize the initial state fields of the POMgcs model, a series of parameter estimation experiments are carried out to investigate whether and to what degree the parameters embedded in high-order turbulence models can be significantly optimized. Results of parameter estimation with perfect initial fields show the two estimated parameters in the MY-2.5 submodel can successfully converge to the "truth" value. The local minimum of the cost function can be effectively and efficiently jumped out once two kinds of optimization algorithms, LBFGS and LMBM, are jointly used. In addition, the estimated parameter will converge to the optimal value rather than the truth one to compensate for the initial field error when the state-parameter are estimated simultaneously. Further, the performance of the parameter estimation is also deeply discussed when the observation samples are noised. Finally, prescribing the initial field and parameter as error source, a forecasting experiment for sea temperature is performed. The experiment results indicate that assimilating "pseudo-observations" to the model based on 4D-Var can significantly improve the sea temperature simulation. Moreover, adjusting the initial field and parameter leads to a better result than the only initial field, and this conclusion is more evident at the surface than in the deeper ocean.

KEYWORDS

numerical model, adjoint method, 4D-Var, data assimilation, parameters estimated

Introduction

The purpose of data assimilation is to provide a better estimation of the ocean (atmosphere) states by combining the numerical model and observations. With the deployment of ocean observing systems and remote sensing techniques, an increasing number of oceanic data is becoming available. These data provide a promising prospect for initial field optimization and parameters estimated through data assimilation. In the last few decades, enormous progress has been made in improving the data assimilation method. Until now, variational methods and the Kalman filter have been widely used in numerical prediction and reanalysis.

The four-dimensional variational (4D-Var) method is one of the most influential and robust schemes among all the data assimilation methods. It has the advantage of assimilating various observations distributed in time and space into the model and maintaining the dynamical and physical consistency with the model. More critically, 4D-Var applies the adjoint technique to get the complicated gradient of the cost function. The adjoint method was proposed with the prognostic equation as the strong constraint at the earliest (Sasaki, 1970), and then LeDimet and Talagrand (1986) first applied this approach for analysis and assimilation of the meteorological observations. In the following years, the application of the adjoint method became more and more common for improving weather forecasts (Talagrand and Courtier, 1987; Thepaut and Courtier, 1991; Navon et al., 1992; Rabier and Courtier, 1992; Courtier et al., 1994; Andersson et al., 1994; Zou and Xiao, 1999; Peng and Zou, 2002; Peng and Zou, 2004). Although the adjoint method is implemented in oceanographic studies later than in atmospheric, its development is also remarkable in the former as in the latter (Bennett and McIntosh, 1982; Yu and O'Brien, 1991; Das and Lardner, 1991; Yu and O'Brien, 1992; Seiler, 1993; Lardner and Song, 1995; Lu and Hsieh, 1997; Lu and Hsieh, 1998a; Lu and Hsieh, 1998b; Heemink et al., 2002 and Zhang et al., 2002; Zhang et al., 2003). It is worth noting that if the numerical model is complex, developing the adjoint code needs a great deal of effort, and the portability of the adjoint model is poor. Therefore, most of the research mainly focused on investigating the feasibility of the adjoint approach under a simplified model based on one-dimensional or two-dimensional assumptions (Bennett and McIntosh, 1982; Yu and O'Brien, 1991; Das and Lardner, 1991; Yu and O'Brien, 1992). Zhang et al. (2002); Zhang et al. (2003) assimilate predicted coastal tidal elevation and coastal subtidal water level data into a linear two-dimensional Princeton Ocean Model (POM) to estimate the lateral tidal open boundary conditions and wind drag coefficient using the adjoint data assimilation method. With the advent of more powerful supercomputing capabilities, Peng and Xie, (2006) developed the adjoint model of the three-dimensional, time-dependent, nonlinear POM to build the 4D-Var method for storm surge forecast. In the subsequent studies, Peng et al. (2007), Peng et al. (2013) corrected the error of the initial conditions and estimated the parameters of wind stress and drag coefficient in the storm surge forecast using the adjoint technique based on the three-dimension POM again. It is worth noting that Mellor Yamada 2.5 order (MY-2.5) turbulent closed scheme is enclosed in POM. Due to the high nonlinear and discontinuity of the vertical turbulence, the nonphysical noise might be produced, and thus result in numerical instability during linearizing the MY-2.5. To avoid the problem, Peng and Xie (2006) neglects the variation of the vertical diffusion coefficients in the linearization of the vertical turbulence scheme replaced by a pre-run of POM with MY-2.5 to determine the value of the diffusion coefficient. In that study, the noise generated by the linear approximation of the turbulence closure scheme has a negligible impact on the storm surge. However, research about using the whole adjoint model of the three-dimensional POM to construct 4D-Var for investigating the feasibility of the adjoint model of MY-2.5 is scarce.

In this study, we developed the three-dimensional and complete adjoint model of the Princeton Ocean Model with the generalized coordinate system (POMgcs). On this basis, constructing the 4D-Var method to estimate the uncertain parameter used in the MY-2.5 turbulence enclosed scheme. The enhanced-turbulent kinetic energy is an important factor in controlling the profile pattern of surface layer circulation and temperature field. Several vertical mixing parameterization schemes can be used to model the coastal circulation and thermohaline structure (Qiao et al., 2004). The appropriate parameterization scheme contributes to simulating the surface layer structure of the ocean temperature. Craig and Banner (1994) and Craig (1996) proposed a scheme to model wave-enhanced turbulence, which imposed a surface diffusion boundary condition (CB boundary condition) into a two-equation turbulence model. It is worth noting that the CB boundary condition introduced two uncertain parameters, the wave energy factor α and the Charnock coefficient β. It is essential to estimate the two parameters accurately for modulating the heat content distribution in the upper coastal sea.

The paper is organized as follows: the following section describes the 4D-Var, POMgcs, and its adjoint model. In section 3, a series of correctness test is performed to evaluate the adjoint

model of POMgcs. Secondly, the truth simulation and biased simulation are conducted, and then a biased assimilation experiment is performed to identify the capability of 4D-Var to optimize the initial field. Moreover, the sensitivity of simulated temperature to parameters is investigated. Then, a series of parameter estimation experiments are performed, and the corresponding results are discussed. Finally, the forecast experiments of sea temperature are evaluated by prescribing the initial field and parameter as error sources. Discussion and summary are presented in section 4.

4D-Var, the nonlinear POMgcs, and its adjoint model

In general, the 4D-Var can be attributed to the minimization of the cost function as follow (Bouttier and Courtier, 1999):

$$J(x) = \frac{1}{2}(x - x_b)^T B^{-1}(x - x_b) + \frac{1}{2}\sum_{i=0}^{n} (y_i - H_i[x_i])^T R_i^{-1}(y_i - H_i[x_i]). \quad (1)$$

It can be found that the 4D-Var is a simple generalization of 3D-Var for observations that are distributed in time. Where x is the analysis variable. x_b and B represent the background value and background error covariance matrix, respectively. In the given assimilation window, the observations are distributed over n intervals, and the subscript i denotes the series of time levels. For $\forall i$, x_i=$M_{0\to i}(x)$, and $M_{0\to i}$ is the model forecast operator from the initial time to i^{th} time level, the 4D-Var is a nonlinearly constrained optimization problem that is difficult to solve in the general case. y_i, H_i, and R_i^{-1} represent the observation, linear interpolation operator, and the inversion of the observation error covariance matrix at the time i, respectively. The implementation of the optimization algorithm requires the participation of the gradient of the cost function. The gradient of the cost function can be deduced:

$$\nabla J(x) = B^{-1}(x - x_b) + \sum_{i=0}^{N} M_{i\to 0}^T H_i^T R_i^{-1}(H_i M_{0\to i}(x) - y_i), \quad (2)$$

where M is the tangent linear model (TLM), i.e. the differential of M, M^T is the adjoint model (ADM) of M. Similarly, H^T is the adjoint of H . The development of the ADM is difficult, especially for the complicated forward model.

In this study, the forward model used in the 4D-Var algorithm is POMgcs, which incorporates the MY-2.5 turbulence closure scheme for vertical mixing. More detail has been discussed in Mellor and Yamada (1982); Galperin et al. (1988), and Mellor (1989). The generalized coordinate system in which sigma-and/or z-level coordinates can be chosen (Ezer and Mellor, 2004) is employed on the vertical level. The governing equation of the POMgcs can be written as follows:

$$\begin{cases} \mathcal{J}(1) = 0 \\ \mathcal{J}(U) - fVs_k + gs_k\frac{\partial \eta}{\partial x} + g\frac{s_k}{\rho_o}\int_k^0 [s_k\frac{\partial \rho'}{\partial x} - (s_x+\eta_x)\frac{\partial \rho'}{\partial k'}]dk' = \frac{\partial}{\partial k}[\frac{K_M}{s_k}\frac{\partial U}{\partial k}] + F_x \\ \mathcal{J}(V) - fUs_k + gs_k\frac{\partial \eta}{\partial y} + g\frac{s_k}{\rho_o}\int_k^0 [s_k\frac{\partial \rho'}{\partial y} - (s_y+\eta_y)\frac{\partial \rho'}{\partial k'}]dk' = \frac{\partial}{\partial k}[\frac{K_M}{s_k}\frac{\partial V}{\partial k}] + F_y \\ \mathcal{J}(T) = \frac{\partial}{\partial k}[\frac{K_H}{s_k}\frac{\partial T}{\partial k}] + F_T - \frac{\partial R}{\partial k} \\ \mathcal{J}(S) = \frac{\partial}{\partial k}[\frac{K_H}{s_k}\frac{\partial S}{\partial k}] + F_S \\ \mathcal{J}(q^2) = \frac{\partial}{\partial k}[\frac{K_q}{s_k}\frac{\partial q^2}{\partial k}] + \frac{2K_M}{s_k}[(\frac{\partial U}{\partial k})^2 + (\frac{\partial V}{\partial k})^2] + \frac{2g}{\rho_0}K_H\frac{\partial \overline{\rho}}{\partial k} - \frac{2s_k q^3}{B_1 l} + F_q \\ \mathcal{J}(q^2 l) = \frac{\partial}{\partial k}[\frac{K_q}{s_k}\frac{\partial q^2 l}{\partial k}] + E_1 l(\frac{K_M}{s_k}[(\frac{\partial U}{\partial k})^2 + (\frac{\partial V}{\partial k})^2] + E_3\frac{g}{\rho_0}K_H\frac{\partial \rho^2}{\partial k})\overline{W} - \frac{q^3 s_k}{B_1} + F_l . \end{cases} \quad (3)$$

These equations are called continuity equation, momentum equation, temperature equation, salinity equation, and MY-2.5 turbulent closure equation (from top to bottom) respectively. Where

$$\mathcal{J}(\varphi) \equiv \frac{\partial s_k\varphi}{\partial t} + \frac{\partial Us_k\varphi}{\partial x} + \frac{\partial Vs_k\varphi}{\partial y} + \frac{\partial \omega\varphi}{\partial k}, \quad (4)$$

$$\omega \equiv W - (s_x+\eta_x)U - (s_y+\eta_y)V - (s_t+\eta_t) = s_k\frac{dk}{dt}. \quad (5)$$

W is the wall proximity function, which can be prescribed according to W=1+$E_2(l/\kappa L)^2$, L^{-1}=$(\eta-z)^{-1}$+$(H+z)^{-1}$, where κ =0.41 is the von Kármán constant. E_1, E_2, and E_3 are empirical constants. $\partial\overline{\rho}/\partial k \equiv \partial\rho/\partial k - c_s^{-2}\partial p/\partial k$, where c_s is the speed of sound, p is pressure. ρ and ρ_0 are the density and reference density, respectively. s represents the generalized coordinate system. x, y, and k are the horizontal and vertical coordinates, respectively. f is the Coriolis parameter, and g is the gravitational acceleration. η, u, v, T, S, q^2, and q^2l are surface elevation, velocity, temperature, salinity, turbulent kinetic energy, and macroscale, respectively, and F_x, F_y, F_T, F_S, F_q and F_l represent the horizontal diffusion of them except for surface elevation. They are defined according to:

$$F_x \equiv \frac{\partial}{\partial x}(s_k\tau_{xx}) + \frac{\partial}{\partial y}(s_k\tau_{xy}), \quad (6)$$

$$F_y \equiv \frac{\partial}{\partial x}(s_k\tau_{xy}) + \frac{\partial}{\partial y}(s_k\tau_{yy}), \quad (7)$$

where

$$t_{xx} = 2A_M\frac{\partial U}{\partial x}\, t_{xy}, t_{yy} = 2A_M\frac{\partial V}{\partial x}, = A_M(\frac{\partial U}{\partial y} + \frac{\partial V}{\partial x}). \quad (8)$$

Also,

$$F_\varphi \equiv \frac{\partial}{\partial x}(s_k A_H\frac{\partial \varphi}{\partial x}) + \frac{\partial}{\partial y}(s_k A_H\frac{\partial \varphi}{\partial y}). \quad (9)$$

φ represents T, S, q^2, q^2l. A_M, A_H are the horizontal kinematic viscosity and horizontal heat diffusivity coefficient, respectively.

K_M, K_H, and K_q are the vertical kinematic viscosity, the vertical diffusivity, and the vertical mixing coefficient for turbulence, respectively, and they can be defined by

$$K_M = lqS_M, K_H = lqS_H, K_q = lqS_q\,. \tag{10}$$

S_M and S_H are functions of a Richardson number, given by

$$S_M(6A_1A_2G_M) + S_H(1 - 2A_2B_2G_H - 12A_1A_2G_H) = A_2, \tag{11}$$

$$S_M(1 + 6A_1^2G_M - 9A_1A_2G_H) - S_H(12A_1^2G_H + 9A_1A_2G_H) = A_1(1 - 3C_1), \tag{12}$$

G_M and G_H can be defined as

$$G_M = \frac{l^2}{q^2}((\frac{\partial u}{\partial z})^2 + (\frac{\partial v}{\partial z})^2)^{1/2,} \tag{13}$$

$$G_H = \frac{l^2}{q^2}\frac{g}{\rho_o}\frac{\partial \rho}{\partial z}. \tag{14}$$

The five empirical constants are assigned (A_1,A_2,B_1,B_2,C_1)= (0.92,0.74.16.6,10.1,0.08) (Blumberg and Mellor 1987). A complete description of POM can be found in Blumberg and Mellor, 1987 and Mellor, 2002.

To model the wave-breaking-enhanced turbulence, the input of turbulence kinetic energy and surface roughness length to the surface boundary condition (CB boundary condition) of the MY-2.5 turbulent closure equation should be introduced respectively. They reflect breaking waves' impact on the magnitude of turbulent kinetic energy and the influence depth, respectively. The CB boundary condition for q^2 is (Craig and Banner, 1994):

$$K_q\frac{\partial q^2}{\partial z} = 2\alpha u_\tau^3,\ z = 0, \tag{15}$$

where u_τ is the friction velocity; α is the wave energy factor, which has $O(10^2)$ magnitude. The second one is for l (Terray et al., 1996, Terray et al., 1999):

$$l = \max(\kappa z_w, l_z), \tag{16}$$

where l_z is the conventional empirical length scale, which is calculated prognostically by the MY-2.5 turbulence closure scheme, and z_w is the surface roughness length, and denotes it as:

$$z_w = (\beta \times 10^5)\frac{u_\tau^2}{g}, \tag{17}$$

where β is Charnock coefficient, g is gravitation acceleration. Both α and z_w are set as 0 in the absence of a surface wave (Blumberg and Mellor, 1987). In contrast, when the effect of the surface wave is considered, both α and z_w are defined as a constant and vary with the state of the surface wave. The surface boundary conditions for q^2 and l are given by Eq. (15) and Eq. (16), and the bottom boundary conditions are $q^2 = B_1^{2/3}u_{\tau b}^2$ and $l=\kappa z_0$ respectively. Where B_1=16.6 (Blumberg and Mellor, 1987) and $u_{\tau b}$ is the friction velocity associated with the bottom frictional stress; z_0 is taken as 0.1, representing the bottom roughness parameter.

The TLM of the POMgcs can be obtained by linearizing the POMgcs forecast model Eq. (3) about the state variable and the boundary condition:

$$\frac{\partial x'}{\partial t} = \frac{\partial F(x)}{\partial x}x',\tag{18}$$

$$x'|_{t=0} = x_0',\tag{19}$$

$$x'(t)|_\Gamma = y'(t),\tag{20}$$

where x represents the state variable of the model, x_0 defines the initial condition at the initial time t_0, $y(t)$ represents the boundary condition on Γ, and the prime represents the perturbations of the variables.

For the variables w and z in a linear space, the linear operator M and its adjoint operator M^* can be defined as:

$$\langle z, Mw\rangle = \langle M^*z, w\rangle\,. \tag{21}$$

The adjoint operator M^* is equivalent to the transpose of M, i. e. $M^*=M^T$. Thus, the ADM of the POMgcs can be written as:

$$\frac{\partial \bar{x}}{\partial t} = -(\frac{\partial F(x)}{\partial x})^{\mathrm{T}}\bar{x},\tag{22}$$

$$\bar{x}|_{t=S} = 0,\tag{23}$$

$$\bar{x}(t)|_\Gamma = 0,\tag{24}$$

The $\bar{x}$ represents an adjoint variable, S is the terminal time of the forward integration of the POMgcs model. The negative sign on the right side of Eq. (22) indicates that the ADM integrates backward in time.

The ADM can be constructed by discretizing the continuous adjoint equation. However, this method to derive the ADM is feasible for simple models rather than a complex three-dimensional POM (Zou, 1997). On the one hand, POM is tedious, and the various physics options with more than one expression are included. On the other hand, the accuracy of the gradient will be limited by the accuracy of the difference scheme used in the discretize procedure. In practical application, the Tangent and Adjoint Model Compiler (TAMC) (Giering and Kaminski, 1998) combined with a hand-coding correction was used to construct the ADM. TAMC can simplify the construction procedure and avoid human errors, which always occur during direct coding. To avoid some errors induced by some irregular expressions of the forward numerical model hand-coding correction is essential, for example, the iterative use of intermediate arrays and the partial array assignment.

In addition, the run time of the ADM can be shortened due to recording the values of the intermediate results into memory in place of recomputed, and transferring the local variables and arrays into global can improve the computational efficiency of the ADM (Zhang et al., 2014).

In this study, the cost function is defined as:

$$J(T^1, T^2, \alpha, \beta) = \frac{1}{2}\sum_{i=1}^{2}(T^i - T_b^i)^T B_i^{-1}(T^i - T_b^i) + \frac{1}{2}\sum_{i=3}^{N}[T^i(T^1, T^2, \alpha, \beta) - T_{obs}^i]R_i^{-1}[T^i(T^1, T^2, \alpha, \beta) - T_{obs}^i], \quad (25)$$

where T^i is the initial temperature field at the i th time step, respectively. T_b^i is the background temperature filed at i th time step. Due to the variety of sources of temperature observation, it is more sufficient and easier to obtain than other state variables. In addition, the temperature is sensitive to α and β, so sea temperature is used as the initial field in this study. B_i represent the background error covariance matrix of the T_b^i. It determines to what extent the background fields will be corrected to match the observations. In an ideal experiment, the perfect observation distribution can make up for the role of the background field, so B_i is set to an identity matrix, rather than a sophisticated matrix. $T^i(T^1,T^2,\alpha,\beta)$ denotes T^i as a function of the control variables, and $T^i(T^1,T^2,\alpha,\beta)=M_{0\to i}(T^1(\alpha,\beta),T^2(\alpha,\beta))$. The background values will derive from the model run, and the initial values at the two consecutive time steps are considered as the control variables to be estimated optimally. Otherwise, the inconsistency of the initial value at the two-time steps may induce initial shocks of the model states during the variational estimation. (Robert, 1966). The third term of Eq. (25) measures the misfit between the control variables and the observation at certain time intervals within the assimilation window, where the subscript i and N are the time level of observation and the total of them, respectively. R_i is the observation error covariance matrix, which is set to a diagonal matrix with diagonal elements 10^{-4} if only one parameter is optimized, otherwise, it is also set to an identity matrix. The wave-affect parameters α and β are implicitly expressed in the above equation. In theory, the cost function has the following form if the wave-affect parameters have the background value:

$$J(T^1, T^2, \alpha, \beta) = \frac{1}{2}\sum_{i=1}^{2}(T^i - T_b^i)^T B_i^{-1}(T^i - T_b^i) + \frac{1}{2}\sum_{i=3}^{N}[T^i(T^1, T^2, \alpha, \beta) - T_{obs}^i]R_i^{-1}[T^i(T^1, T^2, \alpha, \beta) - T_{obs}^i] + \frac{1}{2}K_\alpha(\alpha - \alpha_b)^2 + \frac{1}{2}K_\beta(\beta - \beta_b)^2, \quad (26)$$

where α_b and β_b are the background values of α and β, respectively. K_α and K_β are the coefficients controlling the best fits for the parameter. For simplicity, Eq. (25) is used as the cost function of this study.

Synthetic experiment

In this section, the feasibility of the 4D-Var based on the complete ADM of POMgcs is evaluated thoroughly by estimating the wave-affected parameters. Meanwhile, its ability to optimize the initial field is simply investigated. Before doing that, the ocean temperature and salinity in a typical coastal sea, Bohai Sea, are simulated by the model to validate the rationally of the MY-2.5 parameterization scheme for both constructing the "truth model" and generating the "pseudo-observations" in the data assimilation studies. Among all the experiments involved in this study, the truth model is POMgcs with a "truth" initial field and parameters, where the spin-up output is regarded as the "truth" initial field, and the "truth" wave-affected parameters in high-order turbulence closure are set as α=200 and β=2. The "pseudo-observation" is generated by the "truth model", which is perfect and exists at every geographic location. The model domain covers the Bohai Sea from 117.52°E to 122.47°E and from 37.083°N to 41.033°N. The horizontal resolution is $1/20^o \times 1/20^o$. The maximum depth is set to 65m, with 6 vertical levels. The vertical levels are 0.0, 5.0, 15.0, 25.0, 35.0, and 65.0 m. The model starts with a "cold start" (i.e., without initialization) at 1Z January 1, 2005, and then the model is integrated 5/9 -h, and the output was used as the initial condition of assimilation experiments. The control variables include temperature variables and parameters, the total number of them is 38990 and 2, respectively. When the control variable is temperature, the initial temperature field is generated by adding 1°C perturbation to the "truth" initial field. When the control variable is the wave-affected parameter, the initial value of α and β are set to 100 and 0.5, respectively. The assimilation windows and the sampling frequency of temperature observation both are 1h. Table 1 lists all the assimilation experiments. The process of the assimilation experiment can be outlined as follows:

(1) Integrating POMgcs 5/9 -h with cold start and perfect parameter (α=200 and β=2), the output of temperature was used as the "truth" initial field.

(2) The biased initial field is generated by adding 1°C perturbation to the "truth" initial field, and biased parameters are set to α=100 and β=0.5.

(3) Integrating the forward model in a fixed time window to calculate the cost function Eq. (25).

(4) Integrating the ADM of the forward model backward in time to obtain the gradient of the cost function with respect to the control variables ($\nabla J(T^1,T^2,\alpha,\beta)$).

(5) Inputting the values of the cost function and the gradient of it to Limited Memory Bundle Method (LMBM) (Haarala et al., 2004) to update the control variables.

TABLE 1 Design of identical twin experiment, where "Truth" indicate the truth experiment, and "B_IN", "B_ β ", "B_ α " "B_ $\alpha\beta$ ", and "B_IN β " indicate the bias coming from initial field, β, α, both α and β, and initial field together with β, respectively.

Name	Initial field	α	β
Truth	–	200	2
B_IN	biased	200	2
B_ β	perfect	200	0.5
B_ α	perfect	100	2
B_ $\alpha\beta$	perfect	100	0.5
B_IN β	biased	200	0.5

(6) Repeating processes (3)-(5) until a quasi-equilibrium state is reached, the cost function and the norm of the gradient tend to be stable.

(7) Same as (5) but using Limited memory Broyden-Fletcher-Goldfarb-Shanno (LBFGS) quasi-Newton minimization algorithm (Liu and Nocedal, 1989).

(8) Repeating processes (3), (4), and (7) until the convergence condition is met.

In this study, when only one optimization algorithm is used, steps (5) and (6) or (7) and (8) are directly omitted.

Correctness test of the gradient of the cost function, TLM, and ADM

The correctness of the gradient of the cost function and TLM can be checked through the following formula respectively:

$$\varphi(\epsilon) = \frac{J\left(x_0 + \frac{\epsilon \nabla J}{\| \nabla J \|}\right) - J(x_0)}{\epsilon \| \nabla J \|} \approx 1, \tag{27}$$

$$\varphi(\epsilon) = \frac{\| M(x_o + \epsilon h) - M(x_o) \|}{\| \epsilon \mathbf{M} h \|} \approx 1 . \tag{28}$$

Where x_0 is the control variable, h is the perturbation of x_o, the value of φ shall converge to 1 as ϵ tend to zero, $\|\cdot\|$ represents the two norm. Table 2 and Table 3 show the correctness of the gradient of the cost function and TLM with respect to α, β, and temperature initial field, respectively. For both of them, the $\varphi(\epsilon)$ convergence to around 1 as ϵ decrease from 10^{-2}, and with the further reduction of ϵ, $\varphi(\epsilon)$ is far away from 1 due to the influence of calculation accuracy. From Table 2 and Table 3, one can see that the gradient and TLM of the initial field at least have a 7 digits accuracy, but two parameters in MY-2.5 turbulence enclosed scheme only 3 or 4 digits accuracy. The decrease in accuracy is mainly due to the direct linearization of a highly nonlinear and discontinuous turbulent kinetic energy scheme, which may produce nonphysical noise.

The correctness check for ADM should satisfy the criteria:

$$< \mathbf{M}h, \mathbf{M}h >=< \mathbf{M}^{\mathrm{T}}\mathbf{M}h, h > . \tag{29}$$

Where $\langle\cdot\rangle$ represents the inner product. Here, for the whole POM concerning the initial field, the left-hand side is 0.144525709394330 and the right-hand side is 0.144525709394331. Equation (29) should be held with at least 13 digits of the left-hand side being the same as those of the right-hand side, so the accuracy satisfies the criteria. However, due to the high nonlinear and discontinuous of MY-2.5, α and β only 4 digits accuracy.

TABLE 2 The correctness test of the gradient of the cost function with respect to α, β, and initial field.

ϵ	α	β	Initial field
10^{-2}	0.999687645800464	–	1.00003593381916
10^{-3}	0.999952870019137	–	1.000000359409132
10^{-4}	0.999979701401808	–	1.00000036152066
10^{-5}	0.999985011370854	–	1.00000008556950
10^{-6}	0.999983107245816	–	0.999999936467773
10^{-7}	1.00023174457549	–	0.999999315210570
10^{-8}	1.00388936764555	–	1.00003433152567
10^{-9}	–	1.02381871045560	1.00067592078292
10^{-10}	–	0.999972446877359	1.00180547933443
10^{-11}	–	0.999907036093396	–
10^{-12}	–	0.999900507173298	–

TABLE 3 The correctness test of the TLM with respect to α, β, and initial field.

ϵ	α	β	Initial field
10^{-2}	0.999255539669947	–	1.04806018243665
10^{-3}	0.999592379110171	–	0.999999802687119
10^{-4}	0.999631784976628	–	0.999999799428105
10^{-5}	0.9996610171244403	–	0.999999799284839
10^{-6}	1.00006899468010	1.01109567384988	0.999999796497366
10^{-7}	1.00181297747990	1.00890917645884	0.999999794830625
10^{-8}	1.12911093018209	1.00866255613313	1.00000000548787
10^{-9}	–	1.00841520418398	1.00000039858141
10^{-10}	–	1.00541927348315	1.00002627308542
10^{-11}	–	0.999713066328589	–
10^{-12}	–	1.03163510253935	–

Truth simulation and biased simulation

POMgcs simulated the horizontal and profile structure of temperature and salt in winter and summer, respectively, based on the MY-2.5 parameterization scheme with the CB boundary. Figures 1, 2 show the results of the truth model simulation, where Figures 1A, B are the horizontal distribution of temperature and salt in January, respectively. Figures 1C, D are the same as Figures 1A, B, but for the profile distribution in 38.5°N . Figures 2A–D are the same as Figures 1A–D, but in August, respectively. One can see that the temperature and salt are uniformly mixed in the vertical in January. In August, the temperature formed a distinct thermocline in the area away from the coast, and the thickness of the upper mixing layer is about 10m. The water temperature at the bottom of the ocean remains cold as winter. However, the thermocline of salt is indistinctive. In winter, the horizontal construction of temperature shows that the closer it gets to the coast, the colder it gets, and the conclusion in summer is the opposite of that in winter. The above conclusions are consistent with those obtained from actual measured temperatures for many years (Su and Yuan, 2005). Therefore, this model is rational.

In this section, the biased simulation uses the model with biased wave-affected parameters, and the other configurations of the model are the same as the truth model. Comparing the biased simulation with perfect simulation aims to test the effect of incorrect parameters setting. Figure 3 shows the vertical

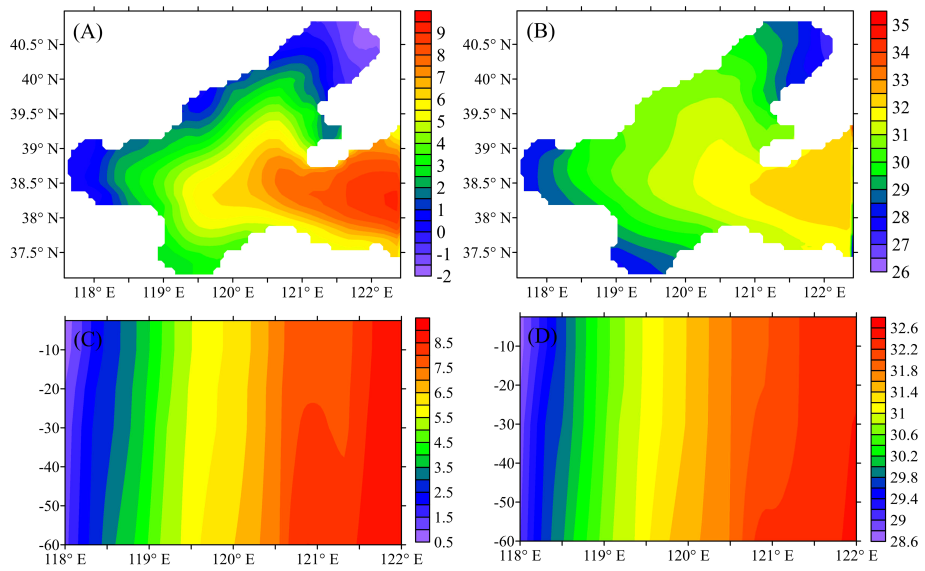


FIGURE 1
Simulated temperature (°C) **(A–C)** and salt **(B–D)** horizontal **(A, B)** and 38.5°N vertical **(C, D)** profile in January.

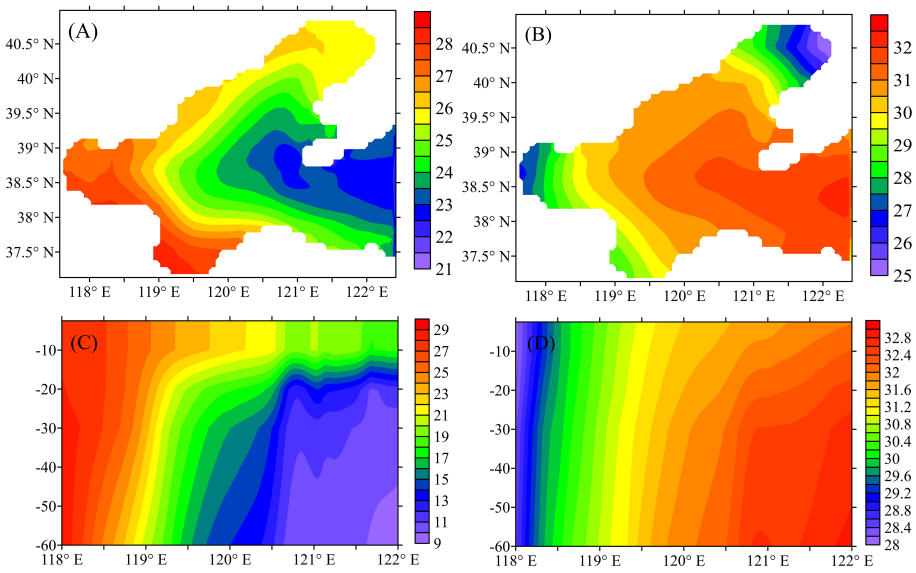


FIGURE 2
Simulated temperature (°C) **(A, C)** and salt **(B, D)** horizontal **(A, B)**) and 38.5oN vertical **(C, D)** profile in August.

structure of the difference in root-mean-square-error (RMSE) of the simulated temperature between the perfect model and the biased model with α=100 , β=0.5 in January. It can be seen that the difference in the temperature between the two simulations is pronounced at the sea surface and 10-m depth. This phenomenon is because the values of wave-affected parameters used in the biased model are smaller than that in the truth model, which indicates that the turbulent kinetic energy is too weak to mix the surface and subsurface water well in the biased simulation. Below 20m, the difference is tiny, suggesting that the turbulent kinetic energy generated by the breaking wave only affects near the sea surface and cannot penetrate the deeper sea. However, the wave-affected parameters play a vital role in the simulated upper layer structure of temperature. It is necessary to estimate the two parameters accurately.

Initial field optimization

Figure 4 shows the variation of the cost function and norm of the gradient with several iterations for B_IN experiment. The value of the cost function decreases rapidly from 2810 to 0.045 within two iterations, and it keeps the low value (0.045) steadily after the second iteration (Figure 4A). Moreover, the norm of the

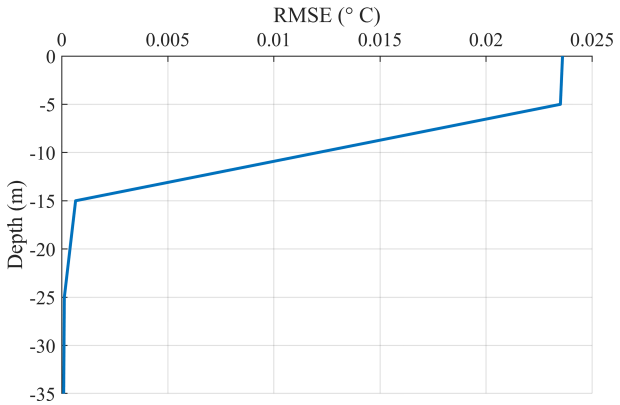


FIGURE 3
The vertical structure of the difference in RMSE of the temperature between the B_$\alpha\beta$ and "truth" model simulation in January.

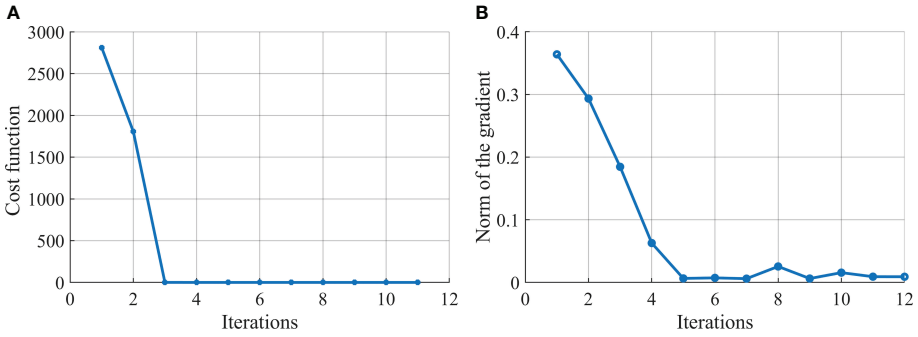


FIGURE 4
The variation of **(A)** the cost function and **(B)** the norm of the gradient with the number of iterations for B_IN experiment.

gradient declines rapidly and then slightly oscillates to search for the optimal declining direction. The norm of the gradient nearly becomes stable after the 4th iteration (Figure 4B), and the minimization process stops after 11 iterations, indicating the local minima of the initial field for that day.

Figure 5 depicts the difference in temperature between the "truth" model simulation and biased initial field simulation without assimilation, and optimal initial field simulation with 4D-Var assimilation, respectively. It is evident that the temperature from the optimal initial field is closer to the "truth" simulation than that from a biased simulation. The above results show the usefulness of 4D-Var based on POMgcs for optimizing the initial field.

Sensitivity check

It is essential to investigate temperature sensitivities with respect to parameters being optimized before parameter estimation. Otherwise, if the insensitive physical variable to the parameter were used to perform parameter estimation, it would be hard to obtain an optimal solution. Figure 6 shows the distribution of the cost function with different α and β. One can see that it is increased with increasing parameters in general. However, the local minimum of the cost function is located in the region where α and β are closed to "truth" values (200 and 2). The existence of a local minimum indicates that it will probably estimate the optimal value of α and β if the gradient of the cost function can be calculated correctly through the ADM. However, α and β are not independent variables, the former represents the turbulence strength induced by the breaking-wave, and the latter is the influence range of turbulence, so the two parameters cannot be determined independently.

The gradient of the cost function with α and β can be used to investigate the model sensitivities with respect to the two parameters. Table 4 shows the value of the gradient with different parameters. The closer the value is to zero, the more sensitive it is to the corresponding parameter value. When the wave-affected parameters are exactly set to 200 and 2, respectively, both gradient values are the closest to zero. It can be found that the sensitivities value of β is several orders of magnitude greater than α. It indicates that α is more vulnerable

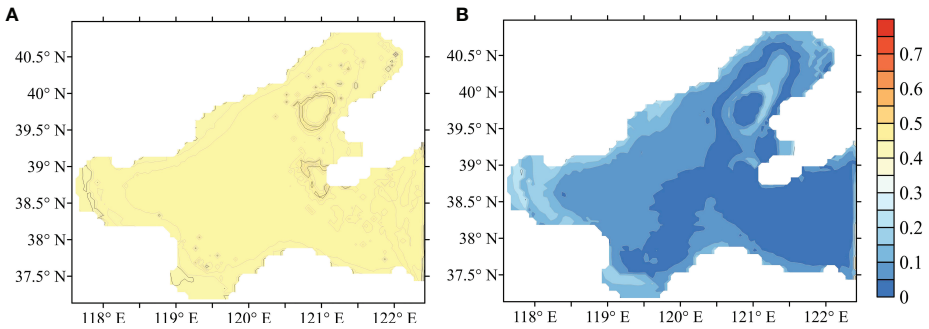


FIGURE 5
Temperature fields from the differences between **(A)** Truth and B_ IN, **(B)** Truth and optimal initial field simulation.

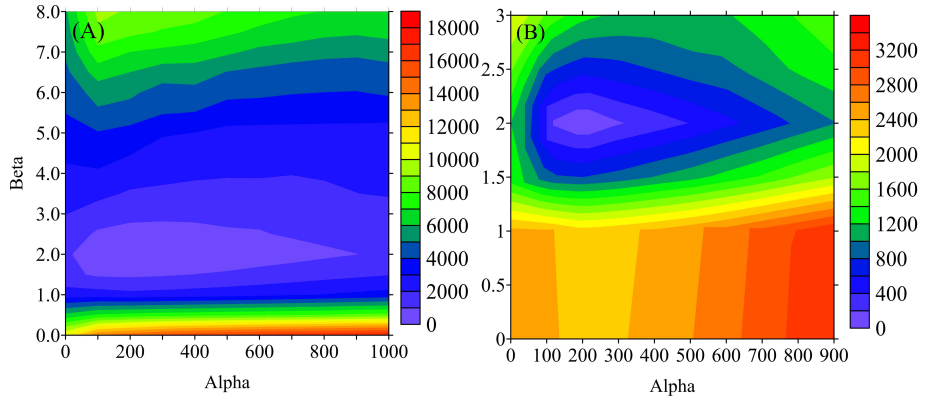


FIGURE 6
The value of the cost function on α and β for **(A)** $10 \geq \beta \geq 0$ and **(B)** $3 \geq \beta \geq 0$.

to being disturbed by the error that arises from observation and initial field during parameter estimation.

Parameters estimation

In this section, B_α and B_β experiments are conducted to determine whether the wave-affect parameters can be estimated correctly or not, where the initial field is perfect. Figure 7 depicts the iteration variations of the cost function, norm of the gradient, and the value of the parameter when the control variable is α or β. Where the left side of Figure 7 shows B_α experiment result, and the other side is B_β . When the control variable is only α, both the cost function and the norm of the gradient decrease dramatically in the first four iterations and keep stable after the 5^{th} iteration, the wave-affected parameter α from the initial value of 0.5 converges to the "truth" value within 6 times iterations with the 4D-Var method, while β from the initial value of 100 converges to its "truth" value after about 4 times iterations.

Figure 8 shows the variation of the cost function, the norm of the gradient, and two wave-affected parameters with the number of the iterations for $B_\alpha\beta$ experiment, where only LBFGS algorithms is used. Due to the nonlinearity of the model, the cost function is not strictly convex, and a nonconvex cost function may have many local minimums. The parameter estimation strongly relies on the initial value of the parameter. Under this condition, when both wave-affected parameters are selected as the control variables, neither α nor β reaches their "truth" value. In this study, the two optimization algorithms LBFGS and LMBM, are applied to enhance the estimation of the double parameters. LBFGS is suitable for solving large-scale optimization problems but has not been proved to be globally convergent for nonconvex or nonsmooth cost functions (Haarala et al., 2007). LMBM combines the variable metric bundle method and the limited memory variable metric method. It is not only suitable for solving large-scale nonsmooth but also globally convergent for nonconvex unconstrained optimization problems. The utilization of the two optimization algorithms speeds up the convergence as well as probably jumps out the local extremum near the initial value of the parameters, thus the parameters are more likely to converge to the "truth" value. Figures 9A, B shows the iteration series of the cost function with

TABLE 4 Dependence of the sensitivity on the initial values of the parameters α and β.

Initial values of (α, β)	α	β
(100,1)	-0.39	1642.11
(100,2)	1.39	-117,22
(100,3)	-2.60	-594.09
(200,1)	-0.17	-1717.57
(200,2)	2.25×10^{-2}	3.71
(200,3)	1.05	-392.61
(300,1)	-0.61	1630.66
(300,2)	-0.45	133.41
(300,3)	0.30	-290.85

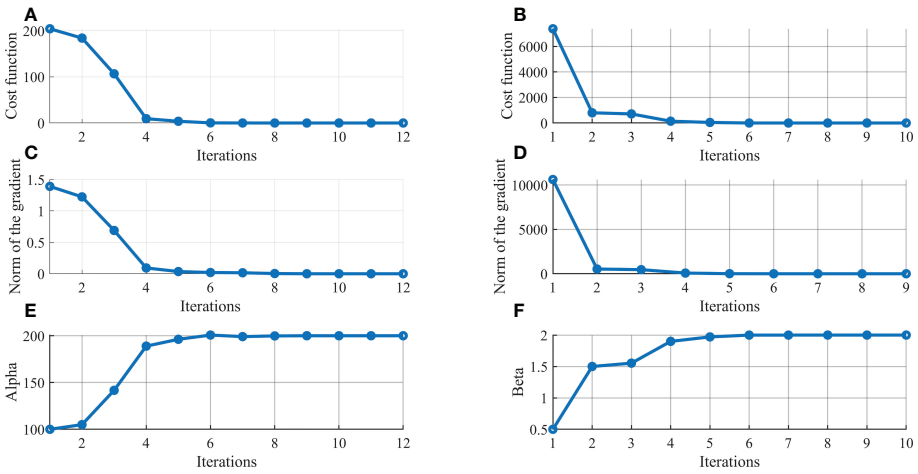


FIGURE 7
The variation of the cost function, the norm of the gradient, and the value of the wave-affected parameter with the number of iterations for B_ α **(A, C, E)** and B_ β **(B, D, F)** experiment, respectively.

the LMBM and LBFGS, respectively. It can be found that the cost function dramatically oscillates in the beginning and becomes stable from the ninth cycle with LMBM method, and then, with LBFGS, it continues decreasing until converging to 0. Figures 9C–F depicts the variation of the wave-affected parameters with LMBM and LBFGS algorithms. One can see that α remarkably boost from the 5^{th} to 9^{th} iteration, then decrease and tend to be stable. However, the parameter failed to converge to the "truth" value with a single optimization algorithm. As shown in Figure 9D, optimizing the cost function with the LBFGS algorithm, α gradually reaches the truth value. The evolution of the value of β with the number of iterations is similar to that of α.

When the initial condition and parameter are regarded as control variables simultaneously, the accuracy of the parameter estimation is restricted by that of the state estimated. In this case, the parameter hardly reaches the "truth" value and merely attains the optimal value of the parameter to compensate for the error derived from the state variable. Figure 10 shows the B_INβ experiment result. It can be found that β does not converge to the "truth" value even though two optimal algorithms are used. However, β has almost reached the

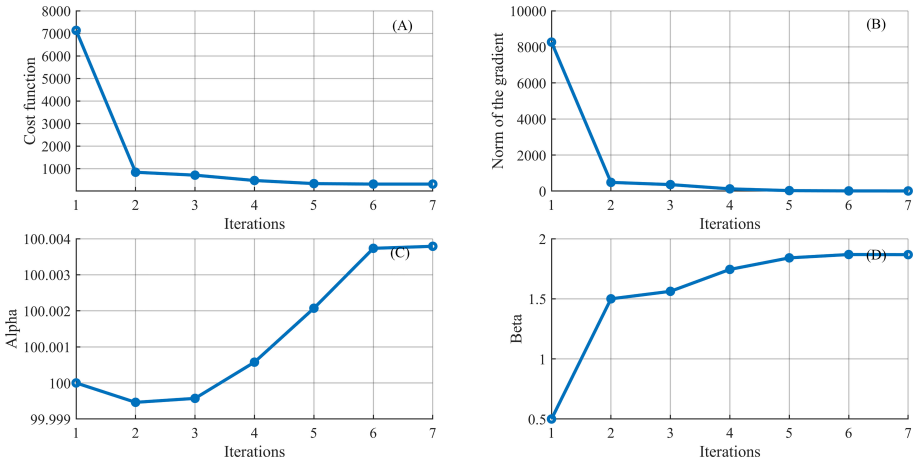


FIGURE 8
The variation of **(A)** the cost function, **(B)** the norm of the gradient, and wave-affected parameters **(C)** α and **(D)** β with LBFGS for *B_αβ* experiment.

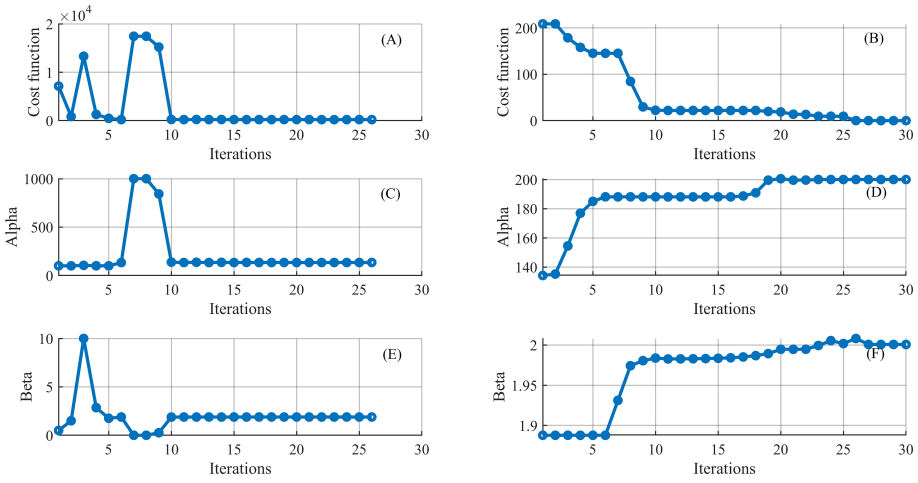


FIGURE 9
The evolution of the cost function, α, and β with LMBM **(A, C, E)** and LBFGS **(B, D, F)** for *B_αβ* experiment, respectively.

optimal value close to the truth value from the eighth iteration with LMBM.

Figure 11 depicts the evolution of parameters α and β based on B_α and B_β with the number of iterations when their initial values are set to (50, 150, 250, 300) and (1, 1.5, 2.5, 3), respectively. It clearly shows that the parameters converge to the "truth" value no matter what the initial values are. Therefore, the 4D-Var based on 3D-POMgcs is feasible for one wave-affected parameter estimated with different initial parameter values and the perfect initial field.

To investigate the impact of the temperature observation noise on wave-affected parameters, the next experiment is designed based on B_β and B_α, and adds different standard deviation noises into temperature observation. Table 5 shows the dependence of parameter estimation on the standard deviation of temperature observation. The relative error is obtained by the absolute error divided by the true value (i.e. the greater the relative error, the lower the reliability). One can see from the table that the relative error increases with increasing the standard deviation. When the standard deviation exceeds 0.5, both reliabilities of α and β are quite low, which indicates the noise level is not acceptable for assimilation purposes. However, the variation of the relative error of β is more slowly than that of α as the standard deviation goes up. In other words, the effect of observational noise

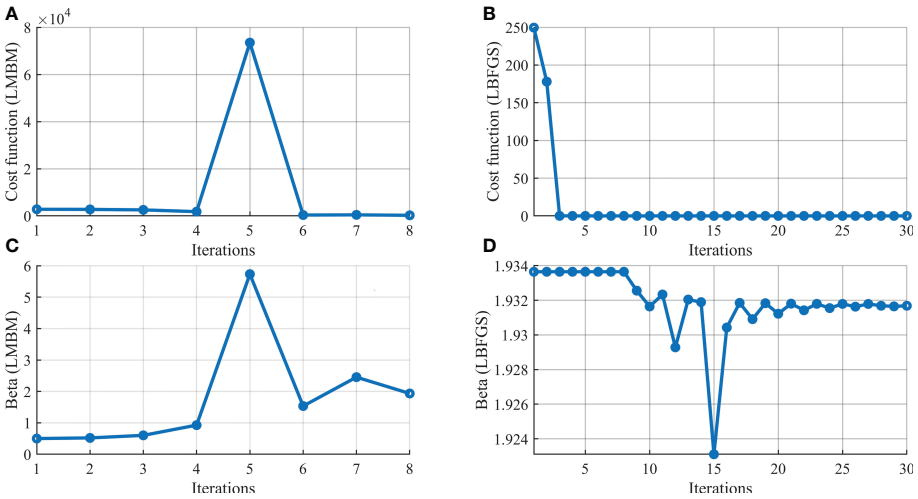


FIGURE 10
The evolution of **(A, B)** the cost function and **(C, D)** β value with LMBM and LBFGS algorithms for B_IN β experiment, respectively.

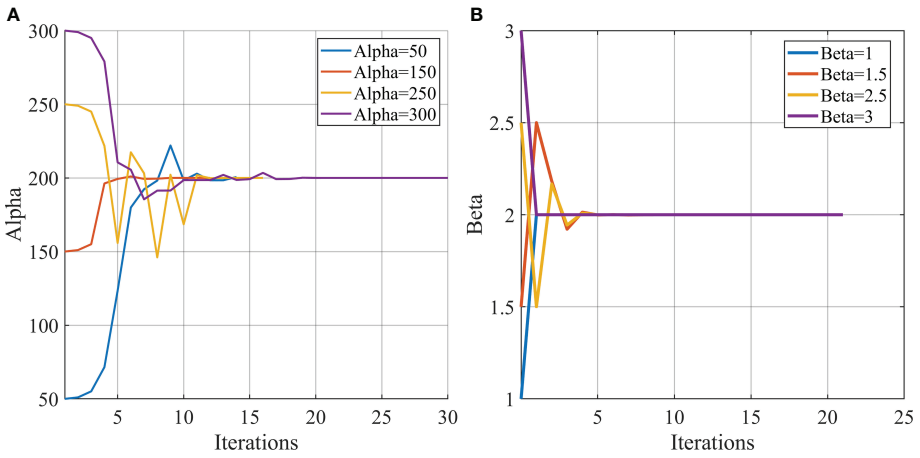


FIGURE 11
The evolution of the estimated wave-affected parameters **(A)** α and **(B)** β for different initial values with the number of iterations.

on the estimation is more severe on α than on β, which means the positive signal of α is difficult to capture during the optimization process, when the noise dominates the model and observation.

Forecast experiment

A 72-h forecast of the upper 35m level temperature is performed with the optimal initial field and parameter from the 4D-Var algorithms. In the forecast experiment, the error source is the initial field and β parameter, where the initial field is obtained by adding 0.35 perturbation to "truth" value, and β parameter is set as 0.3. Since α is more susceptible than β to error from observation or initial field, when both α and β are taken as control variables, there is almost no improvement in α, so only parameter β is involved in this experiment. The control variable is the initial temperature field for the first forecast experiment (EINP-IN). For another forecast experiment (EINP-INP), the initial temperature field and β parameter are optimized simultaneously. A control run with a biased initial field and β parameter without assimilation is called control (CTRL). In these forecast experiments, the assimilation window, assimilation period, and observation source are the same as the assimilation experiments mentioned in the above section. The analysis result of assimilation is used as the initial field and parameter for the 72-h sea temperature forecast.

The RMSE of the 72-h forecast of the sea temperature with respect to the corresponding "truth" value for CTRL, EINP-IN, and EINP-INP is shown in Figures 12, 13. Figure 12 depicts the time series of RMSE of sea surface temperature (SST) starting from 01Z January 01 to 02Z January 04 for three experiments. Compared to CTRL, significant improvements can be seen in EINP-IN and EINP-INP, with EINP-INP outperforming EINP-IN because β reached 2.09 in EINP-INP, which is more close to the truth value. Figure 13 is the same as Figure 12, but for the 35m level. Similar results are found at the 35m level, but in the first few hours, the advantage of the EINP-INP over EINP-IN is not obvious. In the 35m depth, the RMSE of the initial field for the 72-h forecast is 0.1488 and 0.1486 for EINP-IN and EINP-INP, respectively. However, that is 0.2233 and 0.1531 on the sea surface, respectively. This is due to the fact that the breaking wave has a more significant impact on the sea surface than on the deep sea, and the biased parameter leads to larger RMSE of analysis results on the sea surface than at the 35m levels. With the accumulation of initial field and parameter error, the advantage of EINP-INP is shown after three hours at the 35m

TABLE 5 Dependence of the optimally estimated on the standard deviation of the temperature observation.

Std dev of temperature observation	Estimated value of α	Estimated value of β	Relative error of α	Relative error of β
10^{-3}	199.976	1.991	0.02%	0.45%
10^{-2}	201.291	1.981	0.65%	0.95%
0.05	175.779	1.819	12.12%	9.05%
0.1	101.214	2.292	49.39%	14.6%
0.5	99.809	6.089	50.10%	–

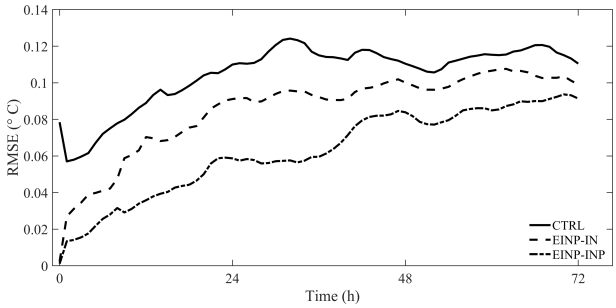


FIGURE 12
The RMSE of 72-h forecast of surface temperature with respect to "truth" value for CTRL, EINP-IN, and EINP-INP.

level. Figure 14 shows the vertical structure of the vertical mixed coefficient KH for EINP-IN and EINP-INP. The simulated averaged vertical mixing coefficient for temperature are much larger for the upper-15-m layer in EINP-INP than in EINP-IN. Below 15m, KH obtained by EINP-INP is still slightly larger than that obtained by EINP-IN, due to too low turbulent mixing strength, the improvement is not shown in Figure 14. The increased KH indicate the enhanced of turbulent mixing strength, which make the seawater of upper layer more vertically homogeneous, so the model performance can be effectively improved by EINP-INP.

Summary and discussion

In this study, the three-dimensional and complete ADM of POMgcs is developed to build 4D-Var. Due to the high nonlinear and discontinuity of the vertical of the MY-2.5 turbulence enclosed scheme, nonphysical noise might be produced and then lead to the numerical oscillation during linearizing it. Although developing the adjoint code is complex, the effective and robust ADM has been obtained by TAMC and hand-coding correction. To evaluate the feasibility of the 4D-Var based on POMgcs with the MY-2.5, the two uncertain parameters in the MY-2.5 parameterization scheme, wave energy factor α and Charnock coefficient β, are tentatively estimated *via* the 4D-VAR algorithm. The two wave-affected parameters determine the magnitude and effective depth of turbulent kinetic energy, respectively. The turbulence kinetic energy can modulate the vertical structure distinctly in the upper ocean, the dissipation of which is adjusted by surface gravity waves under breaking waves. In order to investigate the upper ocean mixed layer, it is essential to obtain the optimal value of the wave-affect parameters in the turbulence-closed scheme.

First of all, the distribution of the temperature and salinity in Bohai is simulated by POMgcs for evaluating the rationally of the MY-2.5 turbulence enclose scheme. Based on that, the "truth model" and "pseudo-observations" are constructed for the following identical twin experiment. After thoroughly testing the capability of the 4D-Var for optimizing the initial field, a suite of parameter estimation experiments is performed. Within an identical twin experiment framework, when the single parameter is being optimized, the "truth" value of the wave-affected mixing parameters can be estimated successfully, no matter what the value of the initial parameter is. However, when the two parameters are used as the control variable simultaneously, the parameters fail to

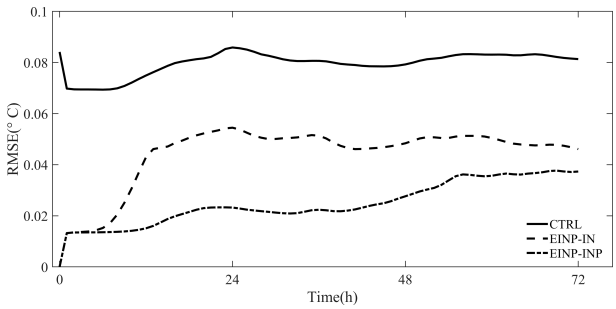


FIGURE 13
As in Figure 12, but for 35m levels.

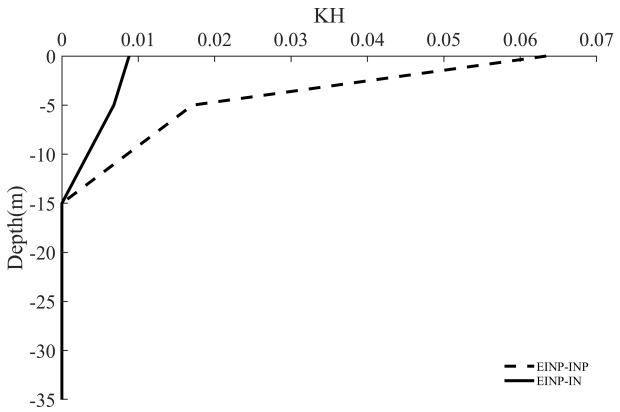


FIGURE 14
Vertical profiles of the simulated averaged the vertical mixing coefficient for temperature KH (m^2 s^{-2}) from EINP-IN and EINP-INP.

reach the "truth" value with the LBFGS optimization algorithm. In this study, both LMBM and LBFGS optimization algorithms are used within one assimilation period, which can speed up the convergence and jump out the local minimum of the cost function. When both the initial field and wave-affected are set to the control variable, parameter estimation is limited by the accuracy of the initial field. In that case, the parameter cannot converge to the "truth" value. However, 4D-Var can fit the model results to the observations, and the optimal value of the parameter can be estimated to compensate for part of the error arising from the initial field of the numerical model. Furthermore, the wave-affected parameter can also reach the optimal value when the observation error is acceptable. It is worth noting that α is more susceptible than β to error from observation or initial field. When α and the initial field are optimized, the parameter cannot converge to an optimal value, even if LMBM and LBFGS algorithms are used simultaneously. The 4D-Var algorithm aims to obtain an 'optimal' initial field or parameter for a better forecast. Therefore, the forecast experiment is performed to further demonstrate 4D-Var, where the forecast errors are attributed to incorrect initial fields and parameters. The results indicate that optimizing both of them or the initial field by setting them as control variables is effective for 4D-Var to improve sea temperature simulation, whether sea surface or deep sea and adjusting the initial field and parameter outperforms only the initial field.

The above results imply that the complete ADM of POMgcs developed in this study is feasible. However, since the precision of the TLM, ADM, and gradient of the cost function for the two parameters is lower due to the high nonlinearity of the MY-2.5, the optimization of α and β is more suitable for short integration time. The process of parameter optimization is not only model-dependent but also observation-dependent. It is well known that high-frequency observations are essential for the study of turbulence. In the real observation experiment, the ability of parameter optimization can be improved by high-frequency observations from surface drifting buoys, *in situ* subsurface buoy, etc. Thus more observations can be assimilated into the model in less integration time. In addition, reducing the initial field error is essential for obtaining the "optimal" α, a more comprehensively designed background error covariance based on flow-dependent or multiscale may enhance the effectiveness of the initial field. These will be explored in our further studies.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

Conceptualization: XZ, YH. Methodology: YH, XZ. Software: XZ, YH. Formal analysis: YH. Writing-original draft: YH. Writing – review & editing: WL, XZ. All authors contributed to the article and approved the submitted version.

Funding

This research was funded by grants from the National Key Research and Development Program (2021YFC3101501) and the National Natural Science Foundation (41876014) of China.

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Andersson, E., Pailleux, J., The´paut, J.-N., R Eyre, J., McNally, A. P., Kelly, G. A., et al. (1994). Use of cloud-cleared radiances in three/four-dimensional variational data assimilation. *Quart. J. R. Meteor. Soc* 120, 627–653. doi: 10.1002/qj.49712051707

Bennett, A. F., and McIntosh, P. C. (1982). Open ocean modeling as an inverse problem: tidal theory. *J. Phys. Oceanogr.* 12, 1004–1018. doi: 10.1175/1520-0485(1982)012<1004:OOMAAI>2.0.CO;2

Blumberg, A., and Mellor, G. (1987). A description of a three dimensional coastal ocean circulation model. *Three-dimensional coastal ocean models. Coast. Estuar. Sciences Amer. Geophys. Union* 4, 1–16. doi: 10.1029/CO004-p0001

Bouttier, F., and Courtier, P. (1999). Data assimilation concepts and methods. meteorological training course lecture series. *Printed January* 9 2001, 33–35.

Courtier, P., The´paut, J.-N., and Hollingsworth, A. (1994). A strategy for operational implementation of 4D-var, using an incremental approach. *Quart. J. R. Meteor. Soc* 120, 1367–1388. doi: 10.1002/qj.49712051912

Craig, P. D. (1996). Velocity profiles and surface roughness under wave breaking. *J. Geophys. Res.* 101, 1265–1277. doi: 10.1029/95JC03220

Craig, P. D., and Banner, M. L. (1994). Modeling wave-enhanced turbulence in the ocean surface layer. *J. Phys. Oceanogr.* 24, 2546–2559. doi: 10.1175/1520-0485(1994)024,2546:MWETIT.2.0.CO;2

Das, S. K., and Lardner, R. W. (1991). On the estimation of parameters of hydraulic models by assimilation of periodic tidal data. *J. Geophys. Res.* 96, 15187–15196. doi: 10.1029/91JC01318

Ezer, T., and Mellor, G. L. (2004). A generalized coordinate ocean model and a comparison of the bottom boundary layer dynamics in terrain-following and in z-level grids. *Ocean Modell.* 6, 379–403. doi: 10.1016/S1463-5003(03)00026-X

Galperin, B., Kantha, L. H., Hassid, S., and Rosati, A. (1988). A quasi-equilibrium turbulent energy model for geophysical flows. *J. Atmos. Sci.* 45, 55–62. doi: 10.1175/1520-0469(1988)0452.0.CO;2

Giering, R., and Kaminski, T. (1998). Recipes for adjoint code construction. ACMTrans. *Math. Software* 24, 437–474. doi: 10.1145/293686.293695

Haarala, N., Miettinen, K., and Mäkelä, M. M. (2004). New limited memory bundle method for large-scale nonsmooth optimization. *Optim. Methods Softw* 19, 673–692. doi: 10.1080/10556780410001689225

Haarala, N., Miettinen, K., and Mäkelä, M. M. (2007). Globally convergent limited memory bundle method for large-scale nonsmooth optimization. *Math. Program* 109, 181–205. doi: 10.1007/s10107-006-0728-2

Heemink, A. W., Mouthaan, E. E. A., Roest, M. R. T., Vollebregt, E. A. H., Robaczewska, K. B., and Verlaam, M. (2002). Inverse 3D shallow water flow modeling of the continental shelf. *Continental Shelf Res.* 22, 465–484. doi: 10.1016/S0278-4343(01)00071-1

Lardner, R. W., and Song, Y. (1995). Optimal estimation of eddy viscosity and friction coefficients for a quasi-three-dimensional numerical tidal model. *Atmos.-Ocean* 33, 581–611. doi: 10.1080/07055900.1995.9649546

LeDimet, F. X., and Talagrand, O. (1986). Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus* 38A, 97–110. doi: 10.1111/j.1600-0870.1986.tb00459.x

Liu, D. C., and Nocedal, J. (1989). On the limited memory BFGS method for large scale optimization. *Math. Program* 45, 503–528. doi: 10.1007/BF01589116

Lu, J., and Hsieh, W. E. (1997). Adjoint data assimilation in coupled atmosphere–ocean models: determining initial model parameters in a simple equatorial model. *Quart. J. R. Meteor. Soc* 123, 2115–2139. doi: 10.1002/qj.49712354316

Lu, J., and Hsieh, W. E. (1998a). Adjoint data assimilation in coupled atmosphere–ocean models: determining initial conditions in a simple equatorial model. *J. Meteor. Soc Jpn.* 76, 737–748. doi: 10.2151/jmsj1965.76.5_737

Lu, J., and Hsieh, W. E. (1998b). On determining initial conditions and parameters in a simple couples atmosphere–ocean model by adjoint data assimilation. *Tellus* 50A, 534–544. doi: 10.1034/j.1600-0870.1998.00011.x

Mellor, G. L. (1989). "Retrospect on oceanic boundary layer modeling an second moment closure," in *Hawaiian Winter workshop on "Parameterization of small-scale processes"* (Honolulu, Hawaii: University of Hawaii).

Mellor, G. L. (2002). Users guide for a three-dimensional, primitive equation, numerical ocean model. *Prog. Atmos. Ocean. Sci.. Princeton Univ.* 17 (1), 4–42.

Mellor, G. L., and Yamada, T. (1982). Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys. Space Phys.* 20, 851–875. doi: 10.1029/RG020i004p00851

Navon, I. M., Zou, X., Derber, J., and Sela, J. (1992). Variational data assimilation with an adiabatic version of the NMC spectral model. *Mon.Weather Rev.* 120, 1433–1446. doi: 10.1175/1520-0493(1992)1202.0.CO;2

Peng, S. Q., and Xie, L. (2006). Effect of determining initial conditions by four-dimensional variational data assimilation on storm surge forecasting. *Ocean Modell.* 14, 1–18. doi: 10.1016/j.ocemod.2006.03.005

Peng, S.-Q., and Zou, X. (2002). Assimilation of NCEP multi-sensor hourly rainfall data using 4D-var approach: a case study of the squall line on April 5 1999. *Meteor. Atmos. Phys.* 81, 237–255. doi: 10.1007/s00703-002-0545-y

Peng, S.-Q., and Zou, X. (2004). Assimilation of ground-based GPS zenith total delay and rain gauge precipitation observations using 4D-var and their impact on short-range QPF. *J. Meteor. Soc Jpn.* 82, 491–506. doi: 10.2151/jmsj.2004.491

Peng, S.-Q., Xie, L., and Pietrafesa, L. J. (2007). Correcting the errors in the initial conditions and wind stress in storm surge simulation using an adjoint optimal technique. *Ocean Modell* 18, 175–193. doi: 10.1016/j.ocemod.2007.04.002

Peng, S.-Q., Li, Y., and Xie, L. (2013). Adjusting the wind stress drag coefficientin storm surge forecasting using an adjoint technique. *J. Atmos. Oceanic Technol.* 30, 590–608. doi: 10.1175/JTECH-D-12-00034.1

Qiao, F., Yuan, Y., Yang, Y., Zheng, Q., Xia, C., and Ma, J. (2004). Wave-induced mixing in the upper ocean: distribution and application to a global ocean circulation model. *Geophysical Res. Lett.* 31, L11303. doi: 10.1029/2004GL019824

Rabier, F., and Courtier, P. (1992). Four-dimensional assimilation in the presence of baroclinic instability. *Quart. J. R. Meteor. Soc* 118, 649–672. doi: 10.1002/qj.49711850604

Robert, A. J. (1966). The integration of a low-order spectral form of the primitive meteorological equations. *J. Meteor. Soc Japan* 44, 237–245. doi: 10.2151/jmsj1965.44.5_237

Sasaki, Y. (1970). Some basic formalisms in numerical variational analysis. *Mon. Weather Rev.* 98, 875–883. doi: 10.1175/1520-0493(1970)098<0875:sbfinv>2.3.co;2

Seiler, U. (1993). Estimation of open boundary conditions with the adjoint method. *J. Geophys. Res.* 98, 22855–22870. doi: 10.1029/93jc02376

Su, J. L., and Yuan, L. Y. (2005). *China Offshore hydrology* (Beijing: Ocean Press).

Talagrand, O., and Courtier, P. (1987). Variational assimilation of meteorological observations with the adjoint vorticity equation. part I: Theory. *Quart. J. R. Meteor. Soc* 113, 1311–1328. doi: 10.1002/qj.49711347812

Terray, E. A., Donelan, M. A., Agarwal, Y., Drennan, W. M., Kahma, K., Williams, A. J.III, et al. (1996). Estimates of kinetic energy dissipation under breaking waves. *J. Phys. Oceanogr.* 26, 792–807. doi: 10.1175/15200485(1996)026,0792:EOKEDU.2.0.CO;2

Terray, E. A., Drennan, W. M., and Donelan, M. A.. (1999). The vertical structure of shear and dissipation in the ocean surface layer. In M. L. Banner Ed. *The Wind-Driven Air-Sea Interface: Electromagnetic and Acoustic Sensing, Wave Dynamics and Turbulent Fluxes* University of New South Wales, 239–245.

Thepaut, J.-N., and Courtier, P. (1991). Four-dimensional variational data assimilation using the adjoint of a multilevel primitive-equation model. *Quart. J. R. Meteor. Soc* 117, 1225–1254. doi: 10.1002/qj.49711750206

Yu, L., and O' Brien, J. J. (1991). Variational estimation of the wind stress drag coefficient and the oceanic eddy viscosity profile. *J. Phys. Oceanogr.* 21, 709–719. doi: 10.1175/1520-0485(1991)0212.0.CO;2

Yu, L., and O' Brien, J. J. (1992). On the initial condition in parameter estimation. *J. Phys. Oceanogr.* 22, 1361–1364. doi: 10.1175/1520-0485(1992)022<1361:oticip>2.0.co;2

Zhang, X. F., Han, G. J., and Li, D. (2014). Variational estimation of wave-affected parameters in a two-equation turbulence model. *J. Atmos. Oceanic Technol.* 32, 528–546. doi: 10.1175/JTECH-D-14-00087.1

Zhang, A., Parker, B. B., and Wei, E. (2002). Assimilation of water level data into a coastal hydrodynamic model by an adjoint optimal technique. *Continent. Shelf Res.* 22, 1909–1934. doi: 10.1016/S0278-4343(02)00067-5

Zhang, A., Wei, E., and Parker, B. B. (2003). Optimal estimation of tidal open boundary conditions using predicted tides and adjoint data assimilation technique. *Cont. Shelf Res.* 23, 1055–1070. doi: 10.1016/S0278-4343(03)00105-5

Zou, X. (1997). Tangent-linear and adjoint of on-off process and their feasibility for use in 4-dimensional variational data assimilation. *Tellus* 49A, 3–31. doi: 10.1034/j.1600-0870.1997.00002.x

Zou, X., and Xiao, Q. (1999). Studies on the initialization and simulation of a mature hurricane using a variational bogus data assimilation scheme. *J. Atmos. Sci.* 57, 836–860. doi: 10.1175/1520-0469(2000)057<0836:sotias>2.0.co;2

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 02 December 2022
DOI 10.3389/fmars.2022.1009852

Check for updates

OPEN ACCESS

EDITED BY
Dongxiao Zhang,
Cooperative Institute for Climate, Ocean and Ecosystem Studies/ University of Washington and NOAA/ Pacific Marine Environmental Laboratory, United States

REVIEWED BY
Peter R. Oke,
Oceans and Atmosphere (CSIRO), Australia
Xiaobiao Xu,
Florida State University, United States

*CORRESPONDENCE
Shiqiu Peng
speng@scsio.ac.cn

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

RECEIVED 02 August 2022
ACCEPTED 17 November 2022
PUBLISHED 02 December 2022

CITATION
Zhu Y, Peng S and Li Y (2022) A vortex-implanted initialization scheme for the mesoscale eddy prediction: Idealized experiments. *Front. Mar. Sci.* 9:1009852. doi: 10.3389/fmars.2022.1009852

COPYRIGHT
© 2022 Zhu, Peng and Li. 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(s) 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.

A vortex-implanted initialization scheme for the mesoscale eddy prediction: Idealized experiments

Yuhang Zhu[1,2,3], Shiqiu Peng[1,2,3,4]* and Yineng Li[1,4]

[1]State Key Laboratory of Tropical Oceanography, South China Sea Institute of Oceanology, Chinese Academy of Sciences, Guangzhou, China, [2]Laboratory for Regional Oceanography and Numerical Modeling, Qingdao National Laboratory for Marine Science and Technology, Qingdao, China, [3]Guangxi Key Laboratory of Marine Disaster in the Beibu Gulf, Bubei Gulf University, Qinzhou, China, [4]Key Laboratory of Science and Technology on Operational Oceanography, Chinese Academy of Sciences, Guangzhou, China

Mesoscale eddy prediction has been a big challenge to oceanographers and marine environment forecasters. Although the traditional initialization for the prediction, i.e., through assimilating the satellite-derived sea level anomalies (SLA) into a model, has some improvement, it is yet unable to predict well the main characteristics of a mesoscale eddy, including its three-dimensional (3D) structure, moving track, size, and intensity. In this study, a vortex-implanted initialization scheme for the mesoscale eddy prediction (VISTMEP) is developed. With the VISTMEP, a bogus vortex is first constructed in terms of 3D SLA-derived currents, and then it is implanted into the model initial field to obtain a more accurate 3D current field of a mesoscale eddy for prediction. The results from idealized experiments show that the VISTMEP can significantly improve prediction of the mesoscale eddy with a longer valid prediction length up to 30 days compared to the experiment with the traditional initialization. Detailed analysis indicates that, as the model is integrated forward, a more "realistic" 3D structure of the eddy in terms of both current and temperature fields is formed when the VISTMEP is employed, leading to the improvement of the eddy prediction regarding to the moving track, size, and intensity of the eddy, which is largely influenced by the accuracy of the initial current field of the eddy obtained by the VISTMEP. This study provides an innovative method for the mesoscale eddy prediction, which could have great potential application in operational services of the marine environments.

KEYWORDS

mesoscale eddy prediction, initialization, bogus vortex, 3D structure, data assimilation

Introduction

As an important part of the marine environment, oceanic mesoscale eddies not only have direct impact on the distributions of ocean temperature, salinity and current, but also play a key role in the transports of mass, momentum, heat and other tracers in the ocean (Wang et al., 2005). In addition, marine activities, such as marine engineering, sailing and fishing, are also influenced greatly by mesoscale eddies. Improving prediction of mesoscale eddies by numerical models is of great importance and practical value from both perspectives of scientific research and operational services of the marine environments.

Oceanic mesoscale eddies have always been an important research object for the oceanographers. The development of oceanic satellite remote sensing technology since the 1990s makes it possible for wide-range, quasi-synchronous and long-term marine observations, from which the Sea Level Anomaly (SLA) observed by the satellite is widely used in the research of mesoscale eddies (Morrow et al., 2004; Chaigneau and Pizarro, 2005; Chelton et al., 2007; Halo et al., 2014). The satellite-derived SLA can be used to estimate the sea surface characteristics of mesoscale eddies, such as the location, amplitude, radius, lifecycle and movement speed. However, it cannot reflect the vertical structures of mesoscale eddies. The accumulation of Argo buoy/CTD/XBT/observed temperature/salinity (T/S) profiles in the last several decades provides valuable information for understanding the interior of the ocean. With the satellite-derived SLA and T/S profiles, a number of studies on mesoscale eddies are carried out, including their generation, development, movement, and dissipation (e.g. 2007; 2011; Chelton et al., 2000; Liang et al., 2012; Xu et al., 2016; Zhang et al., 2016). However, although a few studies tried to describe the 3D structures of mesoscale eddies (e.g. Isern-Fontanet et al., 2004; Chaigneau et al., 2011; Dong et al., 2012; Yang et al., 2013; Zhang et al., 2013), the observations beneath the sea surface are still far enough to construct their 3D structures exactly. Moreover, mesoscale eddies are usually in states of continuous movement with life cycles of several months, a specific mesoscale eddy is hardly being fully observed by the current observation network.

With the development of computer technique, marine numerical models are able to simulate or predict most of the marine environment elements or features in an acceptable accuracy, such as temperature, salinity, currents, waves, tropical cyclones (TC), and so on (e.g. Lee et al., 2018; Sandhya et al., 2018; Peng et al., 2019; Zhu et al., 2020). However, it is still difficult for models to predict the moving track and intensity of a mesoscale eddy like a TC. The reasons could be: 1) the generation, movement and dissipation of oceanic mesoscale eddies, which involve the cascade and inverse cascade of energy between different scales, are one of the most complicated dynamic processes in the ocean, of which the related physical mechanisms are still elusive; 2) an accurate 3D structures of oceanic mesoscale eddies cannot be obtained for the model in the initial time due to the insufficient observations beneath the sea surface, which seriously affects the prediction for oceanic mesoscale eddies.

Improving the accuracy of initial conditions is one of the most effective ways to reduce marine forecasting errors of numerical models (Shriver et al., 2007; Wang et al., 2014). Data assimilation, which incorporates available observations into numerical models, is a common way to generate initial conditions of higher accuracy for marine forecast (Usui et al., 2006; Martin et al., 2007; Chassignet et al., 2009). The previous studies on improving the initial conditions for the simulation of mesoscale eddies usually rely on the assimilation of satellite-derived SLA. However, the assimilation of SLA can only improve the T/S or current structure of mesoscale eddies in the near-surface layer. It is crucial to assimilating observations like T/S profiles or current beneath the sea surface for improving the structure of mesoscale eddies in deeper layers. Unfortunately, these observations beneath the sea surface are still scarce for describing the structure of a specific mesoscale eddy currently. Therefore, the 3D vortex structure of an eddy could hardly be represented accurately in the initial field through regular methods such as data assimilation.

Mesoscale eddies are somewhat similar to TCs in terms of vortex dynamics. Similarly, the 3D structure of a TC is also crucial to the initialization of atmospheric models for TC forecast, but it could not be well represented in the initial fields of atmospheric models 30 years ago due to the coarse resolution of data for generating initial fields and insufficient observations within a 3D TC. To overcome this, Kurihara et al. (1993) proposed an initialization scheme, called the "Bogus" scheme, for the TC forecast, which implants a false but more accurate 3D tropical cyclone into the initial field of the atmospheric numerical model to replace the original one. This "Bogus" scheme was proved to improve the tropical cyclone track forecast significantly. Inspired by the "Bogus" scheme, this study develops a vortex-implanted initialization scheme for the mesoscale eddy prediction (VISTEMP) in the Northwestern Pacific Ocean. Through the VISTEMP, the 3D structure of an eddy retrieved from the satellite-derived SLA is implanted into the initial field of ocean model to replace the original one, generating a new initial field that gives a more accurate description of the 3D structure of the eddy. The impact of VISTEMP on eddy prediction is then evaluated through the idealized eddy prediction experiments.

The rest of this paper is organized as follows. The next section gives a brief introduction of the model and data assimilation schemes used in this study, followed by a description of the VISTEMP scheme in Section 3. The experimental setup is described in section 4, and the experimental results as well as associated discussion are presented in section 5. A summary is given in the final section.

Model and data assimilation

Model configurations

The model used in the study is the Regional Oceanic Modeling System (ROMS, 2005; Shchepetkin and McWilliams, 2003), with version of ROMS_ARGIF 3.1.1 (http://www.romsagrif.org/. Penven et al., 2006; Debreu et al., 2012). The radiational open boundary scheme (Orlanski, 1976; Raymond and Kuo, 1984) and K-Profile Parameterization (KPP) vertical mixing scheme (Large et al., 1994) are chosen for the model. The model domain covers the Northwestern Pacific Ocean (128°-143°E, 16°-24.8°N) with a horizontal resolution of 1/12°×1/12° and 32 sigma vertical levels. The topography of the domain is set as flat base with a constant depth of 4800 m, which is approximately equal to the mean depth of the domain. For the initial filed, the zonal- and meridional- components of the current (u and v) and the sea surface height (SSH) are set to 0, while the salinity is set to 34 psu and the temperature is set to be horizontally homogeneous but vertically varying, of which the temperature at each vertical level is taken from region-mean temperature of the CSIRO Atlas of Regional Seas 2009 (CARS2009, www.cmar.csiro.au/cars) dataset. There is no heat/momentum/mass exchange between the atmosphere and the ocean in the model. In the lateral boundary, the temperature, salinity, SSH and v are set to be equal to those of the initial field, while the u is set to -0.1 m/s to generate westward current. No sponge layer is defined.

Data assimilation scheme

The data assimilation scheme used is the 3D variational assimilation (3DVAR) scheme developed by Li et al. (2008a; 2008b). The cost function of 3DVAR is represented in the form of increment:

$$J(\delta x) = \frac{1}{2}\delta x^{\mathrm{T}}\mathbf{B}^{-1}\delta x + \frac{1}{2}(\mathrm{H}\delta x - \delta y)^{\mathrm{T}}\mathbf{R}^{-1}(\mathrm{H}\delta x - \delta y) \quad (1)$$

In which $\delta x = x - x^b$ is the increment of the model state vector x relative to the background state vector x^b , $\delta y = y - \mathrm{H}x^{\mathrm{b}}$ is difference between the observational vector y and the corresponding model state vector $\mathrm{H}x^{\mathrm{b}}$, in which $\mathrm{H}x^{\mathrm{b}}$ represents the model state vector at the same location of y, and H is the Jacobian matrix of the nonlinear observational operator. The superscript T represents the matrix transposition. B and R are background error covariance matrix and observational error covariance matrix, respectively.

The control variables in 3DVAR include increments of temperature (T), salinity (S), SSH (ζ), stream function (ψ) and potential function (χ):

$$\delta x_c^{\mathrm{T}} = \left(\delta x_{\zeta'}^{\mathrm{T}}, \delta x_{\psi''\chi''}^{\mathrm{T}}, \delta x_{TS}^{\mathrm{T}}\right) \quad (2)$$

The ψ and χ are calculated from the ageostrophic terms of the whole current field. The number of vertical levels in 3DVAR is set to 36 from sea surface to 2000m depth unequally. In addition, two dynamical constraints, i.e., the hydrostatic balance and geostrophic balance, are considered in the 3DVAR.

The construction of B matrix

The **B** matrix in 3DVAR is decomposed into the correlation and standard deviation matrixes due to its extremely vast size:

$$\mathbf{B}_{\zeta_n} = \sum_{\zeta_n} C_{\zeta_n}^{xy} \sum_{\zeta_n} \quad (3)$$

$$\mathbf{B}_V = \sum_V C_V^{xyz} \sum_V \quad (4)$$

in which V represents the variables of ψ, χ , T or S; Σ is the standard deviation matrixes; C is the two- or three- dimensional self- and co- correlation matrixes; x, y and z represent zonal, meridional and vertical directions, respectively. The C has a huge order of magnitude in storage. Thus, the Kronecker method (Graham, 1981) is applied to decompose C into components of C^x, C^y and C^z corresponding to x, y and z directions, respectively. Therefore, the computation of **B** is simplified as the computations of Σ and C, and could also be represented theoretically as:

$$\mathbf{B} = \left\langle \left(x^{\mathrm{f}} - x^{\mathrm{t}}\right)\left(x^{\mathrm{f}} - x^{\mathrm{t}}\right)^{\mathrm{T}} \right\rangle \quad (5)$$

In which x^{f} and x^{t} represent the forecast and the "true" of the ocean, respectively;<> is the variance and superscript T is the transpose of matrix. Since the forecast errors could not be calculated directly due to lack of an adequate number of observations, the NMC method (National Meteorological Center, Parrish and Derber, 1992) is used to obtain a proxy based on the differences of model states between different forecast times. In this study, the computation of **B** is further simplified. The one-dimensional (1D) correlation matrixes of C^x and C^y are assumed as Gaussian-distribution and isotropic, in which the horizontal correlation coefficient C_s of each variable between two model grids (r_1, r_2) is defined as (Daley, 1993):

$$C_s(r_1, r_2) = \exp\left[-(r_2 - r_1)^2/(2L^2)\right] \quad (6)$$

In which L is the horizontal decorrelation length which represents the length when C_s decreases from 1 to $e^{-1/2}$. The value of L is set as 80km in our study. The definition of C^z is similar to C^x and C^y which decreases with the increase of distance between two vertical levels:

$$C^z = e^{-\left(\frac{(dep(i)-dep(i-1))^2}{2L_z^2}\right)}, \; i = 2, 3, 4, \ldots, N \quad (7)$$

in which *dep* is the depth of the vertical level in 3DVAR, L_z is the vertical decorrelation length (L_z=200m), i is the vertical level and N (N=32) is the number of vertical levels. The root-mean-square deviation (*RMSD*) is isotropic horizontally but varying vertically. The vertical profile of *RMSD* is defined as:

$$RMSD = RMSD0 * e^{-\frac{dep(i)^2}{dep(N)}}, i = 1, 2, 3, \ldots, N \tag{8}$$

In which *RMSD*0 is the *RMSD* of each variable at sea surface. The values of *RMSD*0 are listed in Table 1.

Design of the VISTEMP

The VISTEMP scheme consists of two steps. The first step is to retrieve the current fields of 3D eddy from the SLA of mesoscale eddies using the method proposed by Zhang et al. (2013).The second step is to implant the retrieved current fields of 3D eddy into the model initial field to generate a new initial field with more accurate 3D eddy.

The retrieval of 3D eddy

According to the universal structure of oceanic mesoscale eddies (Zhang et al., 2013), the pressure field of a 3D eddy (p_n) can be decomposed into horizontal and vertical structure functions (R and H) under the assumption that the mesoscale eddies are upright in the ocean. The decomposition are represented as:

$$p_n(r_n, z) = R(r_n) \cdot H(z) \tag{9}$$

$$r_n = r / R_0 \tag{10}$$

in which z is the vertical coordinate, r the radial coordinate and R_0 the radius of the mesoscale eddy. $p_n(r_n,z)$ is the standard pressure anomaly obtained by standardizing the pressure anomaly $p'(r,z)$ with R_0 and the eddy center pressure P_0:

$$p_n(r_n, z) = p'(r/R_0, z)/P_0 \tag{11}$$

$$P_0 = \rho_0 g \eta_0 \tag{12}$$

Here ρ_0 is sea surface density, which is set to 1025kg/m^3, g is the gravitational acceleration and η_0 is the SLA at eddy center.

R could be expressed by the theoretical model considering large-time asymptotics of potential vorticity under effects of horizontal diffusion (Kloosterziel, 1990):

$$R(r_n) = \left(1 - r_n^2/2\right) \cdot exp\left(-r_n^2/2\right) \tag{13}$$

H is defined in a stretched coordinate (formula 14) as suggested by a theoretical study of eddies within the quasi-geostrophic framework (Flierl, 1987):

$$z_s = \int_0^z (N/f) dz \tag{14}$$

$$H(z_s) = H_0 \cdot \sin(k z_s + \theta_0) + H_{ave} \tag{15}$$

In which H_0 、 k、 θ_0 and H_{ave} are unknown parameters.

In the practice of the 3D eddy retrieval, the formula (9) should be further deduced. The hydrostatic balance under the sea water density anomaly ρ' is:

$$\rho' = -\frac{1}{g}\frac{\partial p'}{\partial z} \tag{16}$$

Combining formula (9), (10), (11), (12) and (16), ρ' could be expressed as:

$$\rho'(r, z) = -\rho_0 \eta_0 \cdot R(r/R_0) \frac{dH(z)}{dz} = \eta(r) \cdot J(z) \tag{17}$$

in which $\eta(r)$ and $J(z)$ are the SLA of mesoscale eddy and the vertical structure function of ρ':

$$\eta(r) = \eta_0 \cdot R(r/R_0) \tag{18}$$

$$J(z) = -\rho_0 \frac{dH(z)}{dz} \tag{19}$$

in which η_0 represents the SLA in the eddy center. Once the η_0 and R_0 are given, $\eta(r)$ could be calculated using formula (18) to generate a horizontally isotropous mesoscale eddies. However, in the retrieval of the real mesoscale eddies, $\eta(r)$ is equal to satellite-derived gridded SLA of the mesoscale eddies and could be obtained easily. Therefore, the key point of obtaining ρ' is the calculation of $J(z)$. Combining ρ' and the mean density $\bar{\rho}$, we can get the density field of 3D eddy ρ , and then get the current field of 3D eddy through the geostrophic balance relationship:

$$u = -\frac{1}{f\rho}\frac{\partial P}{\partial y} \tag{20}$$

$$v = \frac{1}{f\rho}\frac{\partial P}{\partial x} \tag{21}$$

TABLE 1 The surface RMSD of each variable defined in the 3DVAR.

variable	Temperature	Salinity	u&v	SSH
RMSD0	0.5°C	0.2psu	0.1m/s	0.05m

The implantation of 3D eddy

Before the implantation procedure, the SLA in the whole model domain is first assimilated into the initial field by 3DVAR to achieve a good representation of the near-sea surface structure of mesoscale eddy, and then the retrieved current field of 3D eddy is implanted. To reduce the model shock caused by the imbalance of the initial field, the current of the initial field is decomposed into the large-scale (seasonal climate state) and small-scale components (perturbation field). Only the small-scale component of initial current field of the original mesoscale eddy is replaced by the retrieved current field of 3D eddy, and the large-scale component remains unchanged.

The detailed process of vortex-implantation is shown as Figure 1. First, the SLA in the whole domain is assimilated model initial field (denoted as MINI) by 3DVAR to generate a high-quality analysis field (denoted as ANA) for the model initial conditions; Second, the current field of ANA is decomposed into the seasonal climate state (denoted as CLIM_ANA) and the perturbation field (denoted PERT_ANA); Third, the retrieved current field of 3D eddy is implanted into PERT_ANA to generate a new perturbation field (denoted as PERT1_ANA) by adjusting the currents of PERT_ANA near the location of the eddy center through the following formula:

$$V = x \star EP + (1 - EP) \star y, \tag{22}$$

where V, x and y represent the currents of PERT1_ANA, the constructed 3D eddy and PERT_ANA, and EP is a ratio coefficient defined as:

$$EP = 1 - e^{-\left(\frac{r}{R_0} - 2.5\right)^8} \tag{23}$$

The values of EP regarding to the radius are shown in Figure 2. Based on Eq. (23) and the values of EP, the current field in the PERT_ANA is replaced by the retrieved current field of 3D eddy within 1.25 times eddy radius, and keeps unchanged beyond the 2 times eddy radius. It is a transitional region between the 1.25 and 2 times eddy radius to avoid an abrupt change of the current field. Finally, the PERT1_ANA is recombined with the CLIM_ANA to generate a new model initial field (denoted as NMINI) for eddy forecast.

Design of observational system simulation experiment (OSSE)

The purpose of OSSE is to assess the influence of an observational system or a forecast technique on the forecast skill, in which the "true" state is obtained from the model simulation or other mathematic ways (Arnold and Dey, 1986; Masutani et al., 2010). Before performing the OSSE, the 3D eddy should be constructed first because the mesoscale eddies could not be generated automatically by the free run under idealized model setting of ROMS. Here, two idealized 3D eddies including the warm and cold eddies were constructed separately for the OSSEs through the theoretical formulas (17)-(21). The $\eta(r)$ and $J(z)$ of the idealized 3D eddies are calculated using (18) and (19), in which the related parameters are set as follows: R_0 =50km, η_0 = ± 0.3m, N=0.008, θ_0 =π, k=(3π/2- θ_0)/1800, H_0 =10 and H_{ave} =1, and the integral range in (19) is from sea surface to 1800m depth. The parameters for the constructions of the idealized warm and cold eddies are the same except for the sign of η_0 , which "+" represents the warm eddy and "-" represents the cold eddy. The profile of $\bar{\rho}$ is calculated using the T/S profile from the initial field of the experiment. The potential temperature of the idealized 3D eddies is calculated using a simplified state equation of sea water $\rho=\rho_0(1-\alpha T)$, in which ρ_0 is 1025kg/m^3 and α is 0.0003°C^{-1}, The salinity of the idealized 3D eddies is set to a constant of 34 psu. Figure 3 shows the vertical structure of $J(z)$ of the idealized 3D eddies, and Figure 4 shows the 3D ρ' , potential temperature anomaly, u and v of the idealized 3D warm eddy.

The first step of the OSSEs is to generate the "true" state. A 30-day free run is conducted first to get a background field with steady westward current of u=-0.1m/s. Then, the SLA,

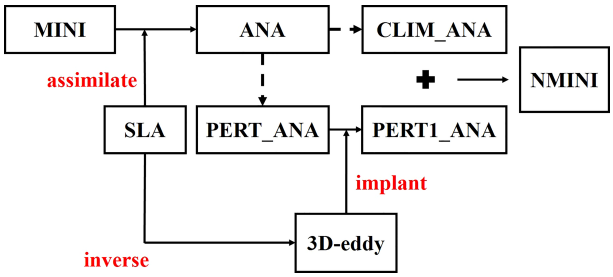


FIGURE 1
The diagrams of the vortex-implantation.

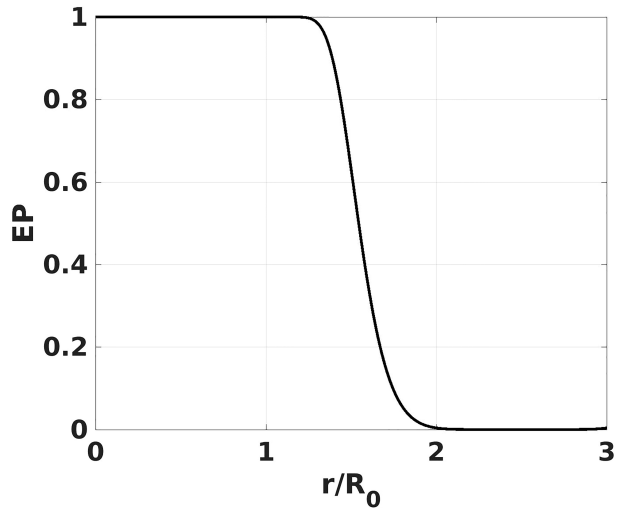


FIGURE 2
The function distribution of the implantation proportion coefficient EP of the 3D eddy.

temperature anomaly and current of the idealized 3D warm or cold eddy constructed before are added into the background field at the location 138°E, 21°N, followed by another 52-day simulation, which denotes as the "true" state run. The results between the 20th to 52th (33 days) day from the "true" state run are chosen as the "true" state for the OSSEs (Figure 5). Based on the background field, five experiments are designed for the OSSEs (Table 2). In the first pair of experiments (DASLA_W and DASLA_C), only the SLAs in the model domain obtained from the two "true" states are assimilated. In the second pair of experiments (ED_BOG_W and ED_BOG_C), the VISTEMP scheme is applied, of which the implanted 3D eddies is retrieved as follows: 1) the $\eta(r)$ of the implanted 3D eddies is obtained from the SLA of the "true" state, 2) the parameters N and f of $J(z)$ are calculated based on the mean T/S and location of the "true" 3D eddies at the implanting moment, and 3) the other parameters are set as follows to fit the vertical structure of "true" 3D eddies at the implanting moment: $k=\pi/18000$, $\theta_0=-1800k$, $H_0=1/\sin(\theta_0)$, $H_{ave}=0$. The values of $J(z)$ for the

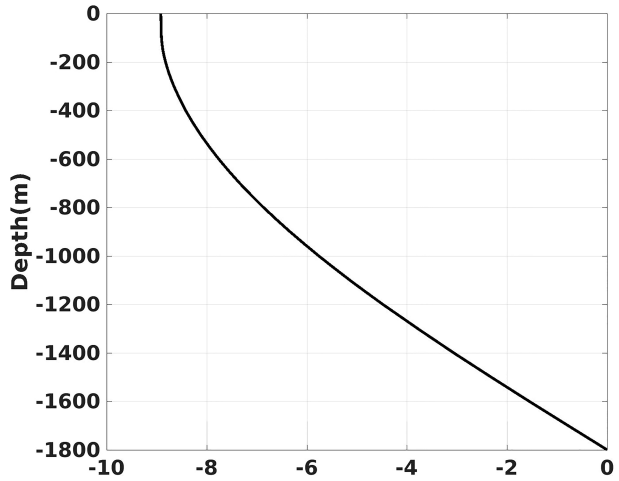


FIGURE 3
The vertical structure function J(z) of potential density of the idealized 3D eddy.

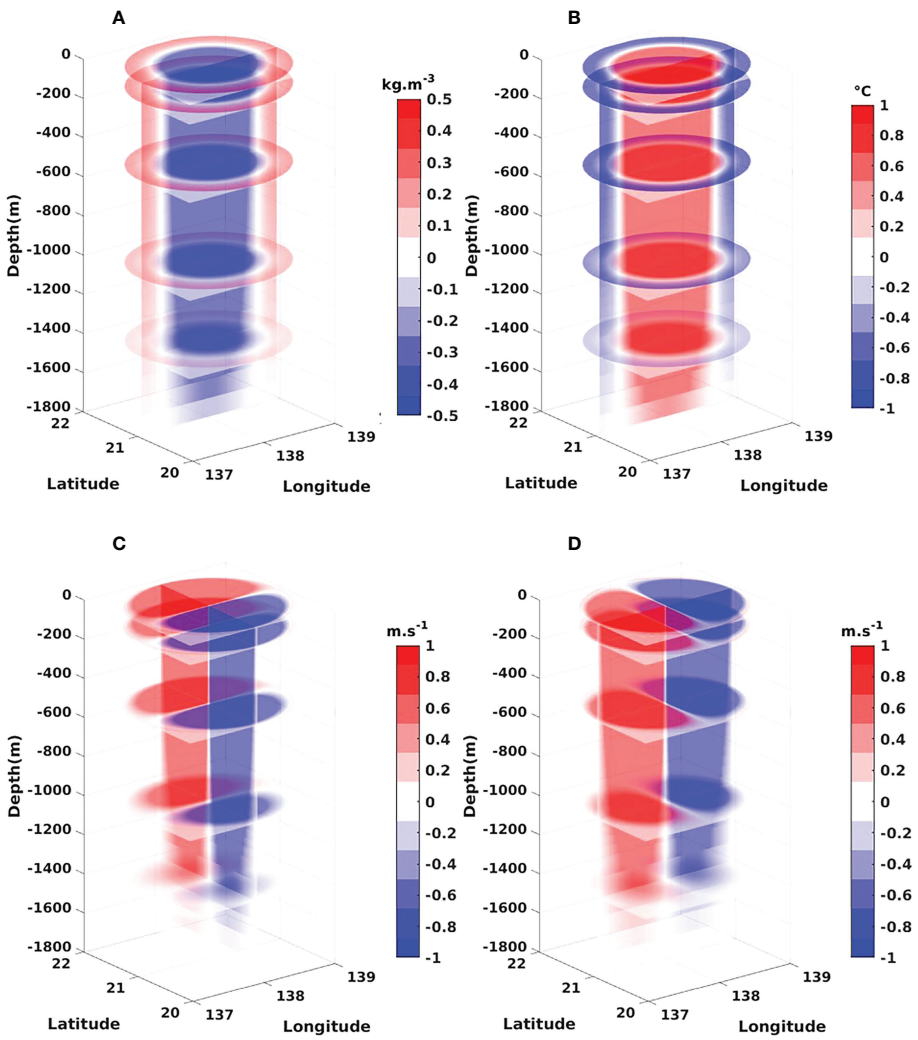


FIGURE 4
The 3D fields of potential density anomaly **(A)**, potential temperature anomaly **(B)**, current u-component **(C)** and current v-component **(D)** of the idealized 3D warm eddy.

implanted 3D eddies are plotted in Figure 6, from which we can see that the implanted 3D eddies dissipates greatly below the depth of 200m comparing to the idealized 3D eddies. As indicated in Section 3, $J(z)$ is the key to the accuracy of the implanted 3D eddy, and in practice the estimate of $J(z)$ could exist considerable biases. Therefore, the fifth experiment (ED_BOG_W_2J) is designed to investigate the impact of the $J(z)$ estimate biases on the eddy prediction by erroneously doubling the absolute values of $J(z)$ below the depth of 200m in the ED_BOG_W. All the five experiments are initialized with the results from 30-day free run. The assimilation of SLA or the application of the VISTEMP scheme is carried out at 0000 UTC of each day for 3 days, followed by a 30-day eddy prediction, as schematically illustrated in Figure 5.

Results and discussion

The evaluation of eddy prediction

For the evaluation of eddy prediction, the Okubo-Weiss (OW) method is applied to the detection of eddies in the predicted fields. The feature variables, including the track, eddy center SLA (ECSLA), 0-500m averaged relative vorticity (RV), sea surface eddy kinetic energy (SSEKE), amplitude, rotation speed (ROS), size, and zonal/meridional moving speeds (mu and mv) are selected for the evaluation. Figure 7 shows the sea surface current and the vertical cross sections of u, v and temperature from "true" 3D warm eddy, DASLA_W and ED_BOG_W at the 10th day. For the "true" 3D warm eddy,

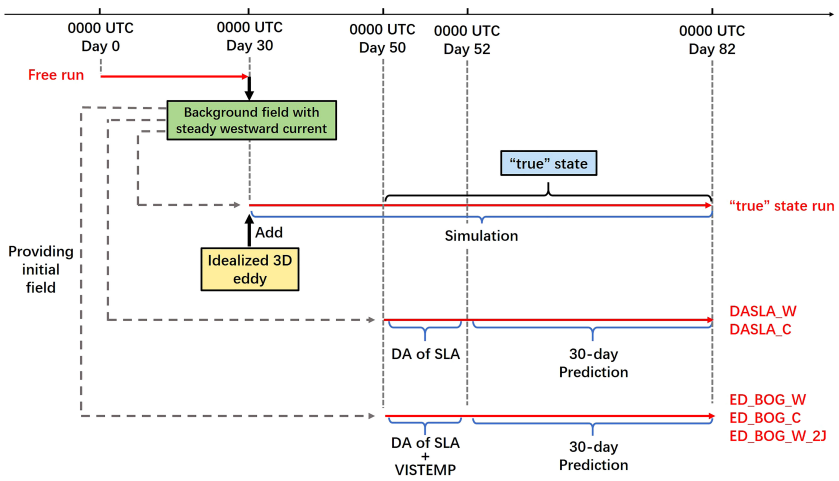


FIGURE 5
The flowchart of generation of "true" state and idealized prediction experiments.

the sea surface current presents an asymmetrical distribution that is stronger in the south and weaker in the north, which is attributed to the westward background current (Figure 7A); the vertical sections of u and v show that the depth of the "true" 3D warm eddy reaches down to 1000m depth (Figures 7D, G). For DASLA_W, the sea surface current is much weaker than that in the "true" 3D warm eddy (Figure 7B), and the assimilation of SLA can only adjust the current down to a depth of about 300m beneath the sea surface through the **B**-matrix and dynamical constrains in the 3DVAR system (Figures 7E, H). For ED_BOG_W, the vertical structure of predicted eddy in terms of current field ((Figures 7C, F, I) and temperature field is very close to the "true" 3D warm eddy (Figures 7J, K, L). The results from the experiments for 3D cold eddy are similar (not shown).

The 30-day averaged prediction biases of eddy feature variables for the warm and cold eddies are listed in Tables 3 and 4, respectively. The implement of the VISTEMP significantly reduces the biases of the track, ECSLA, RV, SSEKE, amplitude and ROS of the predicted warm eddy, but has little effect on biases of size, mu and mv. Similar results are obtained for the experiment with a cold eddy (ED_BOG_C); in addition, the bias of size is also reduced. Figures 8 and 9 show the eddy tracks from "true" eddies and different experiments as well as the time series of the prediction biases of eddy feature variables for different experiments. The "true" warm (cold) eddy moves southwestward (northwestward) under the influences of background current and β-effect (Chelton et al., 2011; Liu et al., 2012). The track predicted by DASLA_W (DASLA_C) presents a significant southern (northern) deviation, which is corrected by the VISTEMP scheme in ED_BOG_W (ED_BOG_C) (Figures 8 and 9A, B). The eddy size is underestimated (overestimated) in DASLA_W (DASLA_C), while it is overestimated in ED_BOG_W (Figure 8I) and well predicted in ED_BOG_C (Figure 9I). Generally, the larger the eddy scale, the further poleward the eddy moves due to the stronger β-effect (Chelton et al., 2011). It is obvious that, while DASLA_W and DASLA_C significantly underestimate the ECSLA, RV, SSEKE, amplitude and ROS of the eddy, ED_BOG_W and ED_BOG_C provide a much better prediction of the eddy regarding to these feature variables (Figures 8/9E, F, G, H, J).

It should be aware that the VISTEMP is developed under the assumption that the mesoscale eddies are upright in the

TABLE 2 The design of idealized prediction experiment.

SettingExperiment	Assimilation of SLA	Eddy implantation	$J(z)$
DASLA_W	√	×	×
DASLA_C	√	×	×
ED_BOG_W	√	√	Calculated from "true" warm eddy
ED_BOG_C	√	√	Calculated from "true" cold eddy
ED_BOG_W_2J	√	√	Twice the $J(z)$ in ED_BOG_W beneath the depth of 200m

The symbols "√" and "×" represent the application and absence of the corresponding items in the first line of the table, respectively.

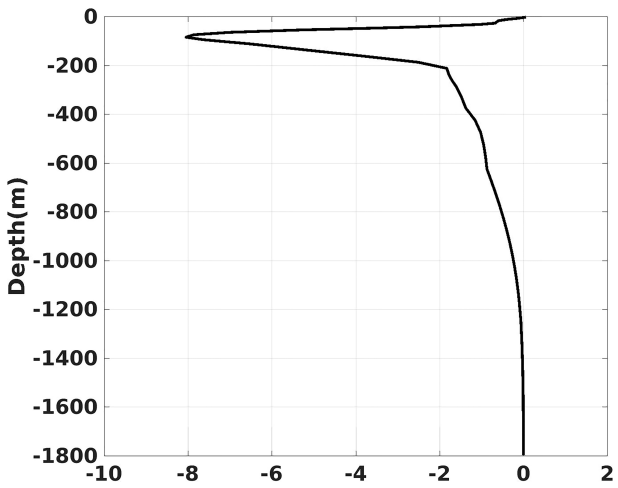


FIGURE 6
The vertical structure function J(z) of potential density of the implanted 3D eddy.

ocean. In the real ocean, the mesoscale eddies are not always upright, and may have a "lean" in the most dynamic regions (Poulsen et al., 2019). The decomposition that treats the mesoscale eddies as an upright body is not perfect, from which the constructed 3D eddy is only a proxy of the real one. Therefore, the effect of the VISTEMP for lean eddies may be limited. Actually, the predicted eddy using the VISTEMP (Figures 8F, I) also shows a lean in vertical structure, suggesting that the upright vertical structure of the implanted 3D eddy could be adjusted to a leaning one that is close to the "true" state with model integration.

The impact of the estimate biases of the implanted 3D warm eddy vertical structure on the mesoscale eddy prediction

As shown in Figures 8A, B, the eddy track predicted by ED_BOG_W_2J has larger errors after 15 days than that by ED_BOG_W, which could be due to errors in the predicted movement speed (Figures 8C, D). Regarding to the other feature variables, including ECSLA, RV, SSEKE, amplitude, size, ROS, ED_BOG_2J generally overestimates them compared with ED_BOG_W. For instance, the absolute value of RV from ED_BOG_2J, which is a key measurement for the strength of an eddy, is much larger than that from ED_BOG_W, indicating a stronger eddy predicted by ED_BOG_W _2J; moreover, the absolute value of RV decreases with forecast time in both ED_BOG_W _2J and ED_BOG_W, suggesting that the eddy is dissipated gradually (Figure 8F).

The results show that the key to a good performance in eddy prediction by the VISTEMP is the accuracy of retrieving the implanted 3D eddy from SLA, which mainly depends on the accuracy of $J(z)$ calculation. In the ideal experiment, the accuracy can be guaranteed because the $J(z)$ can be easily fitted to the vertical structure of the "true" 3D eddy in the OSSEs by adjusting the unknown parameters in the theoretical formulas. However, the $J(z)$ for eddies in the real ocean can only be obtained through a composite analysis based on the historical satellite-derived SLA and T/S profiles (e.g. Argo T/S profile), in which the accuracy of $J(z)$ may be somewhat reduced. Therefore, how to make the bias in the $J(z)$ calculation within an acceptable range is the key to the successful application of the VISTEMP in the real mesoscale eddy prediction.

Summary

In this study, a vortex-implanted initialization scheme for the mesoscale eddy prediction (VISTEMP) is developed. With VISTEMP, a 3D eddy is implanted into the model initial field by replacing the original current field with the one derived from SLA, generating a new initial field that provides a more accurate description of mesoscale eddy for prediction. A set of OSSEs based on the idealized model setting are conducted to evaluate the effect of the VISTEMP on the prediction of mesoscale eddies. The results show that the application of the VISTEMP improves

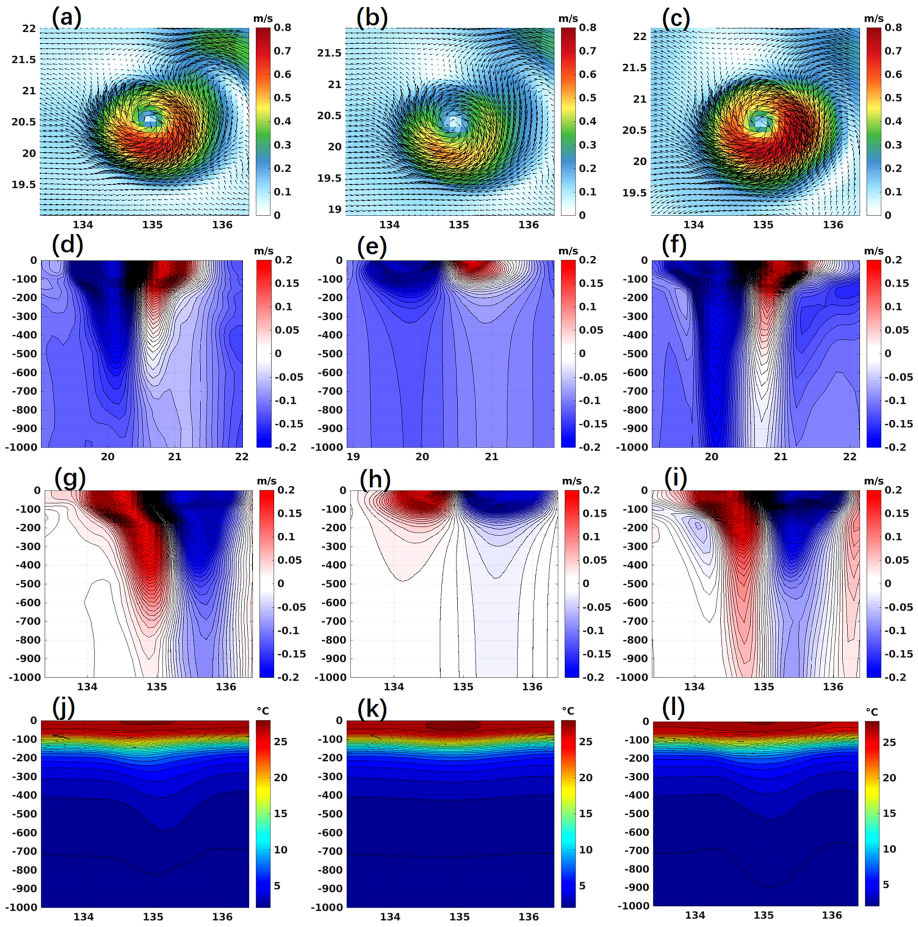


FIGURE 7
The **(A–C)** sea surface current, eddy center-crossed section of **(D–F)** u, **(G–I)** v component of current and **(J–L)** temperature on 10th prediction day from **(A, D, G, J)** "true" 3D warm eddy, **(B, E, H, K)** DASLA_W and **(C, F, I, L)** ED_BOG_W of the idealized experiment, in which the color and arrows in **(A–C)** represent the current speed and direction, respectively.

TABLE 3 The 30-day mean prediction error of the eddy feature variables from experiment DASLA_W, ED_BOG_W and ED_BOG_2J.

Experiment	DASLA_W	ED_BOG_W	ED_BOG_W_2J
track(km)	20.18	8.25	17.39
ECSLA(m)	0.04	0.02	0.05
RV($10^{-6}s^{-1}$)	4.53	1.24	4.74
SSEKE($10^{8}m^{4}s^{-2}$)	1.71	1.34	5.67
amplitude(m)	0.02	0.01	0.04
size(km)	4.92	4.96	10.60
ROS(ms^{-1})	0.12	0.04	0.17
mu(ms^{-1})	0.03	0.03	0.03
mv(ms^{-1})	0.02	0.02	0.02

The metrics in the table include track, eddy center SLA (ECSLA), 0-500m averaged relative vorticity (RV), sea surface eddy kinetic energy (SSEKE), amplitude, rotation speed (ROS), size, and zonal/meridional moving speeds (mu and mv).

TABLE 4 The same as Table 3 except for the DASLA_C and ED_BOG_C.

Experiment	DASLA_C	ED_BOG_C
track(km)	16.95	12.84
ECSLA(m)	0.06	0.04
RV($10^{-6}s^{-1}$)	5.23	2.19
SSEKE($10^{8}m^{4}s^{-2}$)	6.01	1.56
amplitude(m)	0.04	0.01
size(km)	5.31	2.80
ROS(ms^{-1})	0.21	0.07
mu(ms^{-1})	0.02	0.02
mv(ms^{-1})	0.04	0.02

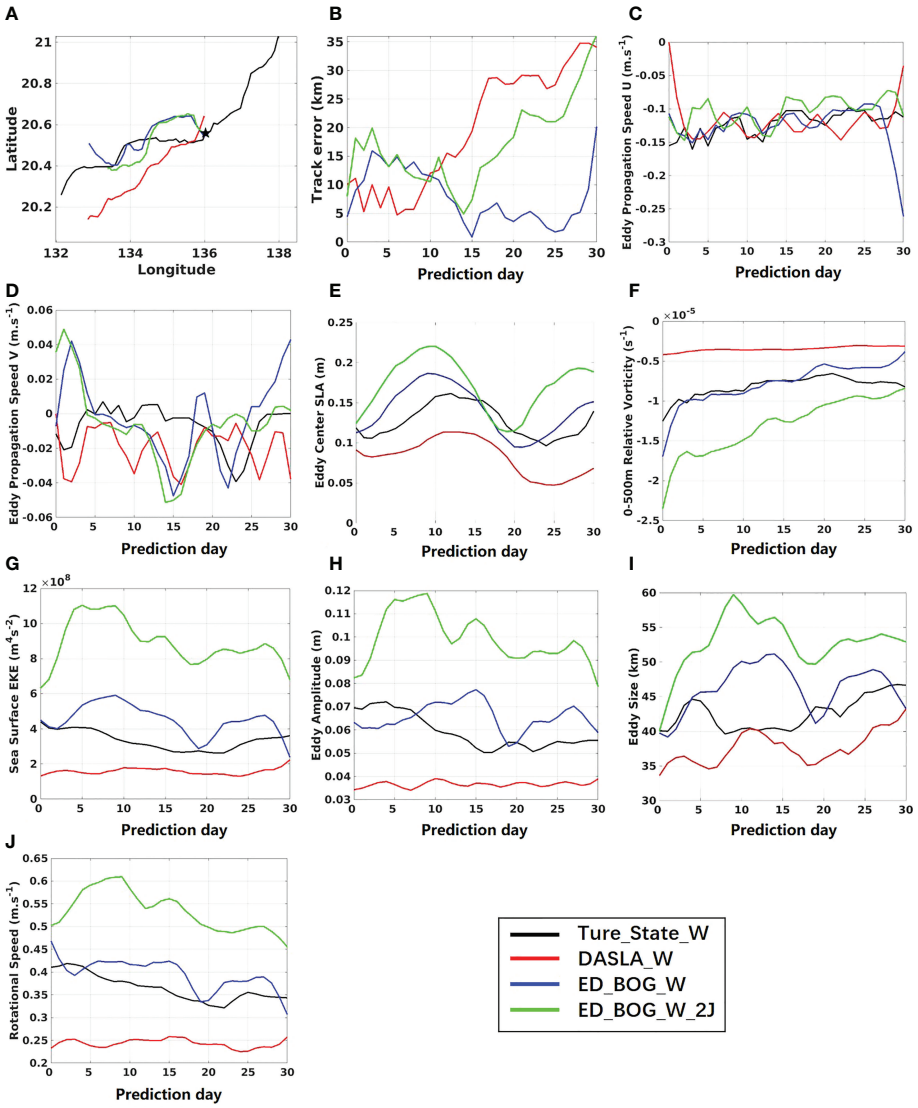


FIGURE 8
The **(A)** track, time series of **(B)** track prediction error, **(C)** mu, **(D)** mv, **(E)** strength, **(F)** RV, **(G)** SSEKE, **(H)** amplitude, **(I)** scale and **(J)** ROS of predicted eddy from "true" warm eddy, DASLA_W, ED_BOG_W and ED_BOG_2J during the prediction period.

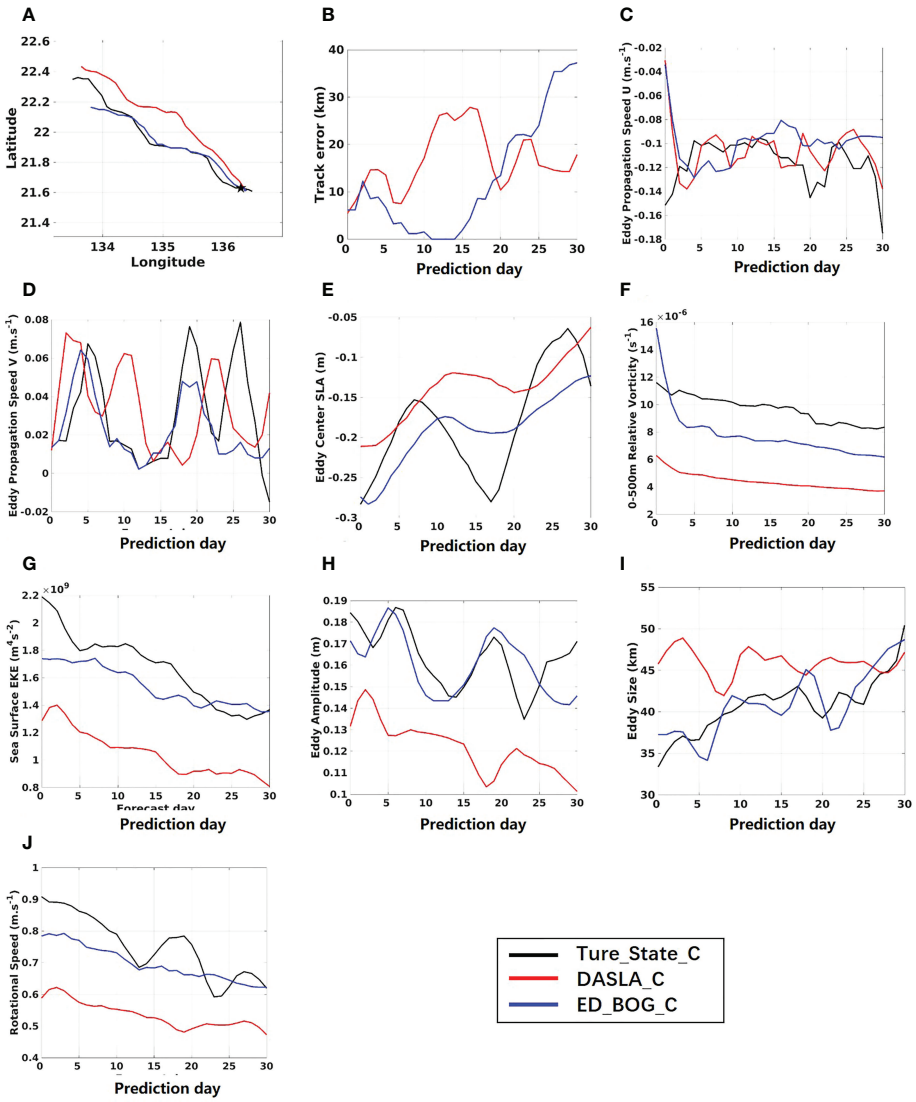


FIGURE 9
The **(A)** track, time series of **(B)** track prediction error, **(C)** mu, **(D)** mv, **(E)** strength, **(F)** RV, **(G)** SSEKE, **(H)** amplitude, **(I)** scale and **(J)** ROS of predicted eddy from "true" cold eddy, DASLA_C and ED_BOG_C during the prediction period.

the prediction for mesoscale eddies in terms of track, ECSLA, vertical structure and so on, as compared to the experiment that only assimilates SLA. The results of OSSEs also indicate that the improvement of eddy prediction is largely influenced by the estimate biases of the vertical structure in the construction of the 3D eddy.

This study provides an innovative method for the mesoscale eddies' prediction, which could have great potential application in operational services of the marine environments. However, more experiments and analysis need to be carried out before the practical application, such as the OSSEs for real simulation and the application in real hindcast. This will be our work in the future.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

YZ developed the vortex-implanted initialization scheme, designed and conducted the forecast experiments, validated and analyzed the experiment results, wrote the manuscript. SP provided the original idea, offered suggestions, polished up the

manuscript. YL offered suggestions. All authors contributed to the article and approved the submitted version.

Funding

This work was jointly supported by the National Natural Science Foundation of China (Grant Nos. U20A20105, 41931182, 41890851 and U21A6001), Guangdong Special Support Program (Grant No.2019BT2H594), Guangxi Key Laboratory of Marine Disaster in the Beibu Gulf, Beibu Gulf University (Grant No.2021KF01), the Strategic Priority Research Program of the Chinese Academy of Sciences (Grant No. XDA19060503), and the Independent Research Project Program of State Key Laboratory of Tropical Oceanography under contract (Grant No.LTOZZ2004). The authors gratefully acknowledge the use of the HPCC at the South China Sea Institute of Oceanology, Chinese Academy of Sciences.

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Arnold, C. P., and Dey, C. H. (1986). Observing-systems simulation experiments: past, present, and future. *Bull. Am. Meteorol. Soc.* 67, 687–695. doi: 10.1175/1520-0477(1986)067<0687:OSSEPP>2.0.CO;2

Chaigneau, A., and Pizarro, O. (2005). Eddy characteristics in the eastern South pacific. *J. Geophys. Res. Oceans* 110 (C8), 207–229. doi: 10.1029/2004JC002815

Chaigneau, A., and Pizarro, O. (2011). Eddy characteristics in the eastern South pacific. *J.Geophys. Res. Oceans* 110 (C8), 207–229. doi: 10.1029/2004JC002815.

Chassignet, E. P., Hurlburt, H. E., Metzger, E. J., Smedstad, O. M., Cummings, J. A., Halliwell, G. R., et al. (2009). US GODAE: Global ocean prediction with the HYbrid coordinate ocean model (HYCOM). *Oceanography* 22 (2), 64–75. doi: 10.5670/oceanog.2009.39

Chelton, D. B., Freilich, M. H., and Esbensen, S. K. (2000). Satellite observations of the wind jets off the pacific coast of central america. part I: Case studies and statistical characteristics, Mon. *Weather Rev.* 128, 1993–2018. doi: 10.1175/1520-0493(2000)128<1993:SOOTWJ>2.0.CO;2

Chelton, D. B., Schlax, M. G., and Samelson, R. M. (2011). Global observations of nonlinear mesoscale eddies. *Prog. Oceanogr.* 91 (2), 167–216. doi: 10.1016/j.pocean.2011.01.002

Chelton, D. B., Schlax, M. G., Samelson, R. M., and Szoeke, R. A. D. (2007). Global observations of large oceanic eddies. *Geophys. Res. Lett.* 34 (15), 87–101. doi: 10.1029/2007GL030812

Daley, R. (1993). *Atmospheric data analysis* (New York: Cambridge University Press).

Debreu, L. P., Marchesiello, P., Penven, P., and Cambon, G. (2012). Two way nesting in split explicit ocean models: Algorithms, implementation and validation. *Ocean Model.* 49-50, 1–21. doi: 10.1016/j.ocemod.2012.03.003

Dong, C., Lin, X., Liu, Y., Nencioli, F., Chao, Y., Guan, Y., et al. (2012). Three-dimensional oceanic eddy analysis in the southern California bight from a numerical product. *J. Geophys. Res.* 117, C00H14. doi: 10.1029/2011JC007354

Flierl, G. R. (1987). Isolated eddy models in geophysics. *Ann. Rev. Fluid Mech.* 19, 493–530. doi: 10.1146/annurev.fl.19.010187.002425

Graham, A. (1981). *Kronecker Products and matrix calculus with applications* (Chichester, UK: Ellis Horwood).

Halo, I., Penven, P., Backeberg, B. C., Ansorge, I., Shillington, F. A., and Roman, R. (2014). Mesoscale eddy variability in the southern extension of the East Madagascar current: Seasonal cycle, energy conversion terms, and eddy mean properties. *J. Geophys. Res. Ocean.* 119 (10), 7324–7356. doi: 10.1002/2014JC009820

Isern-Fontanet, J. E., Font, J., Garcia-Ladona, E., Emelianov, M., Millot, C., and Taupier-Letage, I. (2004). Spatial structure of anticyclonic eddies in the Algerian basin (Mediterranean Sea) analyzed using the okubo–Weiss parameter. *Deep-Sea Res. II* 51, 3009–3028. doi: 10.1016/j.dsr2.2004.09.013

Isern-Fontanet, J. E., Garcia-Ladona, E., and Font, J. (2003). Identification of marine eddies from altimetric maps. *J. Atmos. Ocean. Technol.* 20, 772–778. doi: 10.1175/1520-0426(2003)20<772:IOMEFA>2.0.CO;2

Isern-Fontanet, J. E., Garcia-Ladona, E., and Font, J. (2006). Vortices of the Mediterranean Sea: An altimetric perspective. *J. Phys. Oceanogr.* 36, 87–103. doi: 10.1175/JPO2826.1

Kloosterziel, R. C. (1990). On the large-time asymptotics of the diffusion equation of infinite domains. *J. Eng. Math.* 24, 213–236. doi: 10.1007/BF00058467

Kurihara, Y., Bender, M. A., and Ross, R. J. (1993). An initialization scheme of hurricane models by vortex specification. *Mon. Weather Rev.* 121 (7), 2030. doi: 10.1175/1520-0493(1993)121<2030:AISOHM>2.0.CO;2

Large, W. G., McWilliams, J. C., and Doney, S. C. (1994). Oceanic vertical mixing: A review and a model with a nonlocal boundary layer parameterization. *Rev. Geophys.* 32 (4), 63–403. doi: 10.1029/94RG01872

Lee, H., Chang, P. H., Kang, K., Kang, H. S., and Kim, Y. (2018). Assessment of ocean surface current forecasts from high resolution global seasonal forecast system version 5. *Ocean Polar Res.* 40 (3), 99–114. doi: 10.4217/OPR.2018.40.3.099

Liang, J. H., Mcwilliams, J. C., Kurian, J., Colas, F., Wang, P., and Uchiyama, Y. (2012). Mesoscale variability in the northern tropical pacific: Forcing mechanisms and eddy properties. *J. Geophys. Res. Atmos.* 117 (C7), 165–165. doi: 10.1029/2012JC008008

Li, Z. J., Chao, Y., Mcwilliams, J. C., and IDE, K. (2008a). A three-dimensional variational data assimilation scheme for the regional ocean modeling system. *J. Atmos. Ocean. Technol.* 25, 2074–2090. doi: 10.1175/2008JTECHO594.1

Li, Z. J., Chao, Y., Mcwilliams, J. C., and IDE, K. (2008b). A three-dimensional variational data assimilation scheme for the regional ocean modeling system: Implementation and basic experiments. *J. Geophys. Res.* 113, C05002. doi: 10.1029/2006JC004042

Liu, Y., Dong, C. M., and Guan, Y. P. (2012). Eddy analysis in the subtropical zonal band of the north pacific ocean. *Deep Sea Res. Part I: Oceanogr. Res. Pap.* 68, 54–67. doi: 10.1016/j.dsr.2012.06.001

Martin, M. J., Hines, A., and Bell, M. J. (2007). Data assimilation in the FOAM operational short-range ocean forecasting system: a description of the scheme and its impact. *Q. J. R. Meteorol. Soc.* 133, 981–995. doi: 10.1002/qj.74

Masutani, M., Woollen, J. S., Lord, S. J., Emmitt, G. D., and Kleespies, T. J. (2010). Observing system simulation experiments at the national centers for environmental prediction. *J. Geophys. Res.: Atmos.* 115, D07101. doi: 10.1029/2009JD012528.

Morrow, R., Birol, F., Griffin, D., and Sudre, J. (2004). Divergent pathways of cyclonic and anti-cyclonic ocean eddies. *Geophys. Res. Lett.* 31 (31), 357–370. doi: 10.1029/2004GL020974

Orlanski, I. (1976). A simple boundary-condition for unbounded hyperbolic flows. *J. Comput. Phys.* 21 (3), 251–269. doi: 10.1016/0021-9991(76)90023-1

Parrish, D. F., and Derber, J. C. (1992). The national meteorological center's spectral statistical-interpolation analysis system. *Mon. Weather Rev.* 120 (8), 1747–1763. doi: 10.1175/1520-0493(1992)120<1747:TNMCSS>2.0.CO;2

Peng, S. Q., Zhu, Y. H., Li, Z. J., Li, Y. N., Xie, Q., Liu, S. J., et al. (2019). Improving the real-time marine forecasting of the northern south China Sea by assimilation of gliderobserved T/S profiles. *Sci. Rep* 9, 17845, . doi: 10.1038/s41598-019-54241-8

Penven, P., Debreu, L., Marchesiello, P., and McWilliams, J. C. (2006). Evaluation and application of the ROMS 1-way embedding procedure to the central California upwelling system. *Ocean Model.* 12, 157–187. doi: 10.1016/j.ocemod.2005.05.002

Poulsen, M. B., Jochum, M., Maddison, J. R., Marshall, D., and Nuterman, R. (2019). A geometric interpretation of southern ocean eddy form stress. *J. Phys. Oceanogr.* 49 (10), 2553–2070. doi: 10.1175/JPO-D-18-0220.1

Raymond, W. H., and Kuo, H. L. (1984). A radiation boundary-condition for multi-dimensional flows. *Q. J. R. Meteorol. Soc.* 110, 535–551. doi: 10.1002/qj.49711046414

Sandhya, K. G., Murty, P. L. N., Deshmukh, A. N., Nair, T. M. B., and Shenoi, S. S. C. (2018). An operational wave forecasting system for the east coast of India. *Estuar. Coast. Shelf Sci.* 202, 114–124. doi: 10.1016/j.ecss.2017.12.010

Shchepetkin, A. F., and McWilliams, J. C. (2003). A method for computing horizontal pressure-gradient force in an oceanic model with a nonaligned vertical coordinate. *J. Geophys. Res.: Ocean.* 108 (C3), 35.1–35.4. doi: 10.1029/2001JC001047

Shchepetkin, A. F., and McWilliams, J. C. (2005). The regional oceanic modeling system (ROMS): a split explicit, free surface, topography following coordinate oceanic model. *Ocean Model.* 9 (4), 347–404. doi: 10.1016/j.ocemod.2004.08.002

Shriver, J. F., Hurlburt, H. E., Smedstad, O. M., Wallcraft, A. J., and Rhodes, R. C. (2007). 1/32° real-time global ocean prediction and value-added over 1/16° resolution. *J. Mar. Syst.* 65, 3–26. doi: 10.1016/j.jmarsys.2005.11.021

Usui, N., Lshizaki, S., Fujii, Y., Tsujino, H., Yasuda, T., and Kamachi, M. (2006). Meteorological research institute multivariate ocean variational estimation (MOVE) system: Some early results. *Adv. Space Res.* 37, 806–822. doi: 10.1016/j.asr.2005.09.022

Wang, G., Su, J. L., and Qi, Y. Q. (2005). Advances in studying mesoscale eddies in south China Sea. *Adv. Earth Sci.* 20 (8), 882–886. doi: 10.11867/J.ISSN.1001-8166.2005.08.0882.

Wang, H., Liu, N., Li, B. X., and Li, X. (2014). An Overview of Ocean Predictability and Ocean Ensemble Forecast. *Advances in Earth Science* 29 (11), 1212–1225 (In Chinese).

Xu, L. X., Xie, S. P., McClean, J. L., Liu, Q. Y., and Sasaki, H. (2016). Observing mesoscale eddy effects on mode-water subduction and transport in the north pacific. *Nat. Commun.* 7, 10505. doi: 10.1038/ncomms10505

Yang, G., Wang, F., Li, Y., and Lin, P. (2013). Mesoscale eddies in the northwestern subtropical pacific ocean: statistical characteristics and three-dimensional structures. *J. Geophys. Res.* 118, 1–20. doi: 10.1002/jgrc.20164

Zhang, Z. W., Tian, J., Qiu, B., Zhao, W., Chang, P., Wu, D., et al. (2016). Observed 3D structure, generation, and dissipation of oceanic mesoscale eddies in the south China Sea. *Sci. Rep.* 6, 24349. doi: 10.1038/srep24349

Zhang, Z. G., Zhang, Y., Wang, W., and Huang, R. X. (2013). Universal structure of mesoscale eddies in the ocean. *Geophys. Res. Lett.* 40, 3677–3681. doi: 10.1002/grl.50736

Zhu, Y. H., Li, Y. N., and Peng, S. Q. (2020). The track and accompanying Sea wave forecasts of the supertyphoon mangkhut, (2018) by a real-time regional forecast system. *J. Atmos. Ocean. Technol.* 37 (11), 2075–2084. doi: 10.1175/JTECH-D-19-0196.1

Appendix

1) The definition of mesoscale eddies' feature variables.

The feature variables of mesoscale eddies in this study are defined as below. a) Track: the moving track of the eddy represented by the longitude and latitude of an eddy center; b) Radius (km): The radius of a circle with equal area surrounded by the edge of an eddy; c) Rotation speed (ms^{-1}): The maximum of the averaged speed along each closed streamline of an eddy (Chelton et al., 2011); d) Sea surface eddy kinetic energy (m^4s^{-2}): sea surface eddy kinetic energy of the mesoscale eddy SSEKE, which is represented as

$$\mathrm{SSEKE} = \frac{S*\left(u^2 + v^2\right)}{2} \tag{1}$$

In which S is the area of the mesoscale eddy at sea surface; e) Amplitude (m): The height difference between the eddy center and the averaged SSH along the eddy edge (Chelton et al., 2011); f) Size (km): The mean distance between the eddy center and the streamline where the rotation speed located (Chelton et al., 2011); g) Relative vorticity (s^{-1}): sea surface relative vorticity Ω, which is represented as

$$\Omega = \frac{\partial v'}{\partial x} - \frac{\partial u'}{\partial y} \tag{2}$$

2) The detection of mesoscale eddies.

The detection method of mesoscale eddies used in this study is the Okubo-Weiss (OW, Chelton et al., 2007) method. The OW method define the area where the parameter W less than $2e\times10^{-12}s^{-2}$ as the area of the eddy (Chaigneau et al., 2011), in which the parameter W is defined as (2006; Isern-Fontanet et al., 2003):

$$W = 4\left(u_x^2 + v_x u_y\right) \tag{3}$$

Several thresholds are defined for the eddy detection: The maximum and minimum radiuses of an eddy is 400km and 30km, respectively. The minimum amplitude is 0.02m and the minimum life period is 7 days.

The eddy should be numbered after detection. Penven (2006) proposes a dimensionless parameter that represents the similarity of eddies at different times. For two eddies e_1, e_2, the similarity parameter is defined as

$$X_{e_1e_2} = \sqrt{\left(\frac{\Delta L}{L_0}\right)^2 + \left(\frac{\Delta R}{R_0}\right)^2 + \left(\frac{\Delta \xi}{\xi_0}\right)^2 + \left(\frac{\Delta Z}{Z_0}\right)^2 + \left(\frac{\Delta A}{A_0}\right)^2} \tag{4}$$

In which L_0 , R_0 , ξ_0 , Z_0 and A_0 represent the feature distance, feature radius, feature relative vorticity, feature mean SSH and feature amplitude, which are set to 15km, 100km, $10^{-5}s^{-1}$, 0.1m and 0.1m, respectively. Δ represents the difference of feature variable between e_1, e_2. There are two rules in the determination of the eddy number: a) The two eddies with the minimum $X_{e_1}e_2$ are detected as the same eddy at different time; b) The moving speed of the eddy e_1/e_2 in a) should not beyond $0.3ms^{-1}$.

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 08 December 2022
DOI 10.3389/fmars.2022.1011347

Check for updates

OPEN ACCESS

EDITED BY
JInyu Sheng,
Dalhousie University, Canada

REVIEWED BY
Jing Tao,
Bedford Institute of Oceanography (BIO), Canada
Junbiao Tu,
Tongji University, China
Gaolei Cheng,
South China Sea Institute of Oceanology (CAS), China

*CORRESPONDENCE
Jie Shi
shijie@ouc.edu.cn
Xinyan Mao
maoxinyan@ouc.edu.cn

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

RECEIVED 04 August 2022
ACCEPTED 22 November 2022
PUBLISHED 08 December 2022

CITATION
Geng H, Peng H, Shi J, Mao X and Lv X (2022) Application of interpolation methodology with dynamical constraint to the suspended particulate matter in the Liaodong Bay. *Front. Mar. Sci.* 9:1011347. doi: 10.3389/fmars.2022.1011347

COPYRIGHT
© 2022 Geng, Peng, Shi, Mao and Lv. 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(s) 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.

Application of interpolation methodology with dynamical constraint to the suspended particulate matter in the Liaodong Bay

Haoyi Geng[1,2], Haohong Peng[3,4], Jie Shi[1,2*], Xinyan Mao[3,4*] and Xianqing Lv[3,4]

[1]Key Laboratory of Marine Environment and Ecology, Ministry of Education of China, Ocean University of China, Qingdao, China, [2]Laboratory for Marine Ecology and Environmental Sciences, Qingdao National Laboratory for Marine Science and Technology, Qingdao, China, [3]Key Laboratory of Physical Oceanography, Ministry of Education of China, Ocean University of China, Qingdao, China, [4]College of Oceanic and Atmospheric Sciences, Ocean University of China, Qingdao, China

Introduction: Suspended Particulate Matter (SPM) influences the primary production and the distributions of pollutants in the ocean. Besides, the regulation mechanisms of SPM in the Liaodong Bay were complicated.

Method: To analyze the distributions and influencing factors of SPM, based on the adjoint assimilation method, an interpolation method with dynamical constraint was established in the Liaodong Bay.

Result: In two ideal experiments, the cost function, Mean Absolute Error (MAE) and Normalized Mean Error (NME) all had reduced by more than 90%, which proved the accuracy of the interpolation method. Based on conventional observations of SPM, the distributions of dynamically constrained, Kriging and radial basis function (RBF) interpolations in March, May, August and October of 2015 were obtained.

Discussion: The cross-validation was carried out to compare the dynamically constrained interpolation and the unconstrained interpolations. Among seven unconstrained interpolation methods, the averaged MAE of RBF interpolation was the lowest, which was 10.976 mg/L. The averaged MAE of dynamically constrained interpolation was 7.703 mg/L, reduced by 29.8% compared with the RBF interpolation. It was indicated that RBF interpolation was the most accurate among the seven unconstrained interpolations and dynamically constrained interpolation was more accurate than unconstrained interpolations at the observation stations. The distributions of dynamically constrained and RBF interpolations were compared with Korean Geostationary Ocean Color Imager (GOCI) satellite-derived distributions of SPM concentrations in the Liaodong Bay. Fully considering the influences of the hydrodynamic processes, the dynamically constrained interpolation

provided distributions more consistent with the satellite-derived distributions. However, due to the lack of observations in some areas and ignoring the influences of currents, some high values of SPM concentration were not captured by the distributions of RBF interpolation. Moreover, in accordance with the results of dynamically constrained interpolation, it was found that the SPM concentrations in the bay were affected by the SPM discharge from the Liao River Basin.

KEYWORDS

material transport model, adjoint assimilation model, interpolation with dynamical constraint, cross-validation, suspended particulate matter

1 Introduction

Suspended Particulate Matter (SPM) has a profound impact on the biogeochemical processes in the ocean. The transport and transformation of nutrients in seawater are influenced by SPM (Volpe et al., 2011). Besides, SPM determines the optical indexes of seawater such as chroma and transparency. Hence, phytoplankton can be affected by SPM. It was found that phytoplankton productivity in turbid ocean areas was less than in the adjacent ocean (Cloern, 1987) and SPM concentration was negatively correlated with phytoplankton density (He et al., 2017) in previous studies. Therefore, the primary production in ocean is influenced by the distributions of SPM concentrations. During transferring along the currents, SPM adsorbs and releases the pollutants such as heavy metals, thereby effecting the distributions of marine pollutants. The pollutants can be transported from the coastal area to open sea along with SPM, and have potentially harmful effects on marine organisms (Yao et al., 2016). In addition, the sedimentation and resuspension processes of substances are also affected by SPM. These processes have an important impact on the coastal sea area (D'Sa et al., 2006).

The distributions of SPM in seawater are complicated and will be affected by many factors. Hydrodynamic processes will influence the distributions of SPM. Walker and Hammack (2000) found that strong northerly winds in winter could increase the concentration of SPM in the water column by five times in the northern Gulf of Mexico, and a large turbid plume was produced under the influence of wind-drive horizontal currents and resuspension. Estuarine discharge is a critical factor causing the variations of SPM distributions. Based on the remotely sensed data, Li et al. (2019) proved that due to the variations of SPM discharge from the Yellow River estuary, the SPM concentrations in Laizhou Bay and the southern Bohai Sea will be significantly higher in spring and winter than in summer and autumn. Moreover, the atmospheric deposition, the fragmentation process and the flocculation process in coastal systems can affect the concentration of SPM in seawater (Maerz et al., 2016; Xie et al., 2022; Fettweis et al., 2014). Therefore, it is necessary to study the distributions and the influencing factors of SPM in marine biogeochemical research.

The Bohai Sea is the largest inland sea in China (Figure 1A). The Liaodong Bay is one of the three most important bays of the Bohai Sea which is located in the northern part of the Bohai Sea (Figure 1B). The Liaodong Bay is surrounded by Hebei and Liaoning provinces on three sides. The line from Qinhuangdao of Hebei Province to Laotieshan Cape of Liaodong Peninsula is the boundary between the Liaodong Bay and the central area of the Bohai Sea (Lan et al., 2016). The estuary area (40.6°N-41.2° N, 121.5°E-122.2°E) is located in the northeast of the Liaodong Bay, including estuaries of Liaohe River, Daliao River and Daling River, which is the most important SPM source in the Liaodong Bay (Yan et al., 2020). As the second largest river in the Bohai Sea, the Liaohe River transports a high load of fine-grained sediments to the Liaodong Bay, which increases the SPM concentration in the bay (Jiang et al., 2004). In addition, the Yellow River discharges a large number of SPM every year. Under the influence of ocean currents, the water with high SPM concentration was transported into the Liaodong Bay through the central Bohai Sea from Laizhou Bay (Liu et al., 2007). The distributions of SPM in the Liaodong Bay are also influenced by the circulation. The detrital fine-grained sediments in the north of the Liaodong Bay were transported mainly *via* the surface currents (Dou et al., 2014). Moreover, the resuspension process also contributes to the formation of high SPM concentration in the Liaodong Bay (Jiang et al., 2004).

In order to study the spatial and temporal variations and regulation mechanisms of SPM in the Liaodong Bay, *in-situ* observation is an essential method to acquire the SPM concentrations. However, observations of SPM in the Liaodong Bay are spatially discrete. Hence, there is a demand of the continuously spatial distributions of SPM concentrations

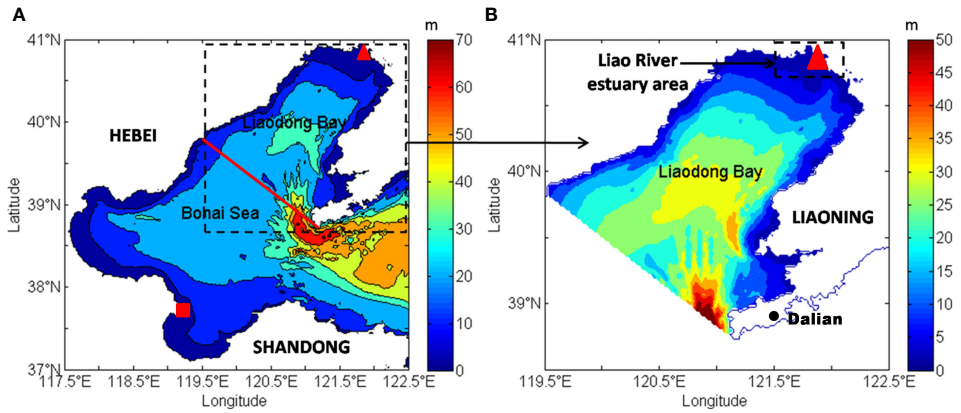


FIGURE 1
Geographical locations and bathymetry of the Bohai Sea and the Liaodong Bay. **(A)** shows the Bohai Sea, in which the red solid line is the boundary between the Liaodong Bay and the central Bohai Sea. The northwest end of the boundary is Qinhuangdao and the southeast end of the boundary is the Laotieshan Cape. The red triangle points out the estuary of the Liao River and the red square stands for the estuary of the Yellow River. The Liaodong Bay is selected using the black dotted line. **(B)** shows the study area, which is the Liaodong Bay. The red triangle still represents the estuary of the Liao River. The land on northwest side is Hebei Province and the land on the southeast side is Liaoning Province. Dalian city is located at the east of the mouth of the Liaodong Bay. The depth of the Liaodong Bay is deep in the central area and shallow around. The deepest region is near the mouth of the bay and the depth is more than 50m.

in different observation months through interpolation (Davis, 1975). In previous studies, various interpolation methods had been used to acquire the continuously spatial distributions, including Kriging interpolation, radial basis function (RBF) interpolation and Cressman interpolation and so on. Bourgain and Gascard (2011) spatially interpolated the observations collected in the central Arctic basin from 1997 to 2008 using Kriging interpolation. The interannual change of halocline was discussed based on the interpolated distributions on the profile. Wang et al. (2014) developed a new RBF interpolation-based method to obtain a sub-pixel map. Through three visual and quantitative assessments on reversion experiments of remote sensing images, the accuracy of RBF interpolation-based method was verified. Injan et al. (2021) used the Cressman interpolation to initialize the model with an analysis product based on observations, and a Cressman Initialized Ensemble Intermediate Coupled Model was established for more accurate Sea surface temperature (SST) analysis and prediction. However, it is difficult for the unconstrained interpolation methods to provide accurate interpolated distributions in the regions lacking observations. In addition, too many observations may be taken into calculation during the interpolation, which may make the interpolated distributions tend to be average, resulting in large errors. Therefore, in previous studies, researchers optimized the unconstrained interpolation methods to improve the accuracy. Based on support vector machine, Wang et al. (2008) improved Kriging interpolation. The sea surface salinity and height data was obtained in the ocean area missing observations using the improved method. Wang et al. (2016) predicted chlorophyll concentrations based on least square support vector regression and RBF neural network.

Nevertheless, the optimizations of unconstrained interpolations only improved the mathematical methods, while the effects of hydrodynamic processes on interpolation were not taken into account. Since the adjoint assimilation method have been widely used to reverse the open boundary conditions (Pan et al., 2017 and Chen et al., 2014), optimize the initial field (Peng and Xie, 2006) and optimize the control parameters of marine ecosystem dynamical models (Qi et al., 2011; Li et al., 2013 and Goldberg and Heimbach, 2013), researchers start to use interpolation method with dynamical constraint based on adjoint assimilation method in marine research. Wu et al. (2021) adjusted the parameters in Ekman model based on the cubic spline interpolation method with the adjoint assimilation model, and obtained the optimized wind stress resistance coefficient. Zheng et al. (2020) applied the dynamically constrained interpolation method to the Bohai, Yellow, and East China Seas. Utilizing the time and space information of the observations, the constrained interpolation provided precise M_2 cotidal charts. Considering the influences of wind on the distributions of $PM_{2.5}$, Li et al. (2021) used the $PM_{2.5}$ transport model as a dynamical constraint, and obtained high-accuracy national-scale distributions of $PM_{2.5}$ in China from August to November 2014. Based on the adjoint assimilation method and taking the sediment transport model as the constraint condition, Mao et al. (2018) used a dynamically constrained interpolation method to obtain the SPM concentrations in the Bohai Sea. However, unconstrained interpolations weren't used in practical applications due to insufficient observations, and the higher

accuracy of dynamically constrained interpolation than unconstrained interpolation had not been proved.

In this study, an interpolation method with dynamical constraint based on adjoint assimilation model was established in the Liaodong Bay first of all. Two ideal experiments were carried out to validate the accuracy of the interpolation. Then, the observational SPM concentrations in March, May, August and October of 2015 in the Liaodong Bay were spatially interpolated using the dynamically constrained, Kriging and RBF interpolation methods. After that, seven unconstrained interpolation methods including Kriging and RBF interpolations were compared using cross-validation to find a more accurate unconstrained interpolation method. Whether the accuracy of dynamically constrained interpolation was improved compared with unconstrained interpolation methods would be verified by cross-validation. Then, satellite-derived distributions of SPM concentrations in the Liaodong Bay were used in the validation of dynamically constrained interpolation. In addition, the interpolated distributions were analyzed to find the controlling factors of temporal and spatial variations of SPM concentrations.

2 Data and methods

2.1 Observations of SPM concentrations in the Liaodong Bay

The observation data used in this study came from conventional monitoring in March, May, August and October of 2015 in the Liaodong Bay. SPM concentrations were calculated after sampling and filtering in situ. In this study, only SPM observations of the sea surface were used. The quantities of observations in March, May, August and October were 72, 49, 52 and 54 respectively. The distributions of the observational SPM concentrations were shown in Figure 2. Most of the monitoring stations were located in the coastal areas. In this study, the observed SPM concentrations were used in both dynamically constrained interpolation method and unconstrained interpolation method. The observations of August were used for cross-validation of all the interpolation methods. The results of the dynamically constrained interpolation were used to discuss the spatio-temporal variations.

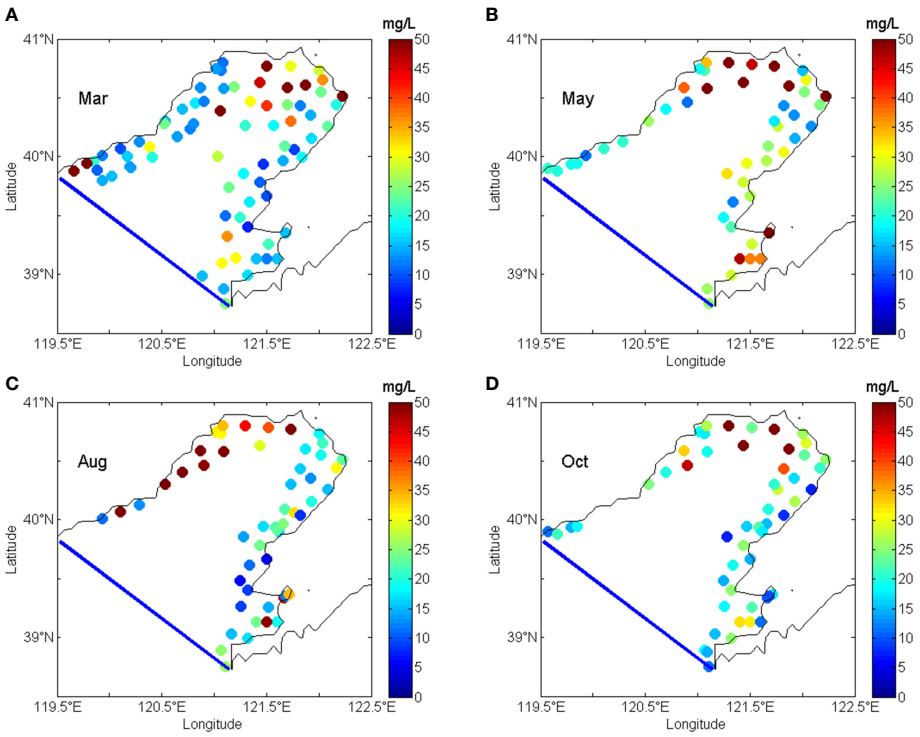


FIGURE 2
Distributions of observations in the Liaodong Bay. **(A–D)** show the observations in March, May, August and October, respectively. The blue line is the boundary of the Liaodong Bay. The dots exhibit the monitoring stations. The color of the dots indicates SPM concentration.

2.2 Satellite data of SPM

The satellite data of SPM used in this study was obtained from Korean Geostationary Ocean Color Imager (GOCI) satellite and downloaded from Korea Ocean Satellite Center (http://kosc.kiost.ac.kr/gociSearch/list.nm?menuCd=50&lang=en&url=gociSearch). The satellite provided SPM concentrations in March, May, August and October of 2015 in the Liaodong Bay. GOCI satellite provided gridded output eight times a day with a spatial resolution higher than 1 km. Monthly averaged distributions of SPM in the Liaodong Bay can be obtained from GOCI dataset, and were used to verify the results of dynamically constrained interpolation method in this study.

2.3 Hydrodynamic field of the Liaodong Bay

The hydrodynamic fields used in this study were the results of a hydrodynamic model based on the Princeton Ocean Model (POM). The domain of the hydrodynamic model included the Bohai Sea and the Yellow Sea (34.5°-41°N, 117.5°-127.5°E). The horizontal resolution was 10'in both latitude and longitude, and there were 21 sigma levels in the vertical direction. The open boundary was set at 34.5°N. Four main tidal components (M_2, S_2, K_1 and O_1) were implemented along the open boundary. The Ocean Circulation and Climate Advanced Model (OCCAM) was used to provide the distribution of the climatic circulation in each layer at the open boundary. The initial conditions of temperature and salinity were obtained from NODC (Levitus) World Ocean Atlas 2001. The hydrodynamic model was driven by the wind field. The model was run for one year after spin-up, and the three-dimensional results of currents and water temperature were stored every half an hour. Wang et al. (2013) had verified the hydrodynamic field and used it in the simulation of the Bohai Sea. The Liaodong Bay part of the hydrodynamic flow fields was used in this study. The hydrodynamic fields were spatial interpolated to the grid of our model.

2.4 Interpolation methods without dynamical constraint

2.4.1 Kriging interpolation

The observation data of SPM concentrations recorded as a standard format (x_i,y_i,z_i), i =1,2,3, … , N, where N is the number of observation data. $(x_i,y_i,)$is the rectangular coordinate converted from the longitude and latitude of the observation station, and z_i is the observed SPM concentration. The interpolated SPM concentration z_j at a given point (x_j,y_j) can be expressed as

$$z_j = \sum_{i=1}^{N} \beta_i \cdot z_i \tag{1}$$

where $\beta_i$is the interpolation coefficient of Kriging interpolation. Based on different variogram model, β_i can be calculated differently. Kriging interpolation was carried out based on linear variogram model in this study.

2.4.2 Radial basis function interpolation

RBF interpolation has been widely used in spatial interpolation of observation data. The preprocessing of observation data is consistent with Kriging interpolation, and the interpolated SPM concentration at a given point (x_j,y_j) can be expressed as

$$z_j = c_0 + c_1 x + c_2 y + \sum_{i=1}^{i=N} \lambda_i \cdot \varphi(d_{ij}) \tag{2}$$

where d_{ij} is the distance from the given point (x_j,y_j) to observation (x_j,y_j);$\varphi(d_{i,j})$ is the basis function; c_0, c_1, c_2 and λ_i are constant coefficients. The following five RBFs are usually used, 1) Linear, $\varphi(d_{ij})=d_{ij}$, 2) Cubic, $\varphi(d_{ij})=d_{ij}^3$, 3) Thin plate spline, $\varphi(d_{ij})=d_{ij}^2 ln(d_{ij}+1)$, 4) Gaussian, $\varphi(d_{ij})=exp(-0.5\cdot d_{ij}^2/\sigma^2)$, 5) Multi-quadrics, $\varphi(d_{ij}) = \sqrt{1 + d_{ij}^2/\sigma^2}$. RBF interpolation was carried out based on linear RBF in this study.

2.5 Interpolation method with dynamical constraint based on adjoint assimilation method

Adjoint assimilation is an effective method of data assimilation. It combines the Variation Principle with the Optimization Control Theory, and converts the physical problem to the mathematical problem of finding the minimum value (Sasaki, 1970; Lu and Zhang, 2006). It takes the equations, initial conditions and boundary conditions of model as constraint conditions, and minimizes the cost function representing the errors between the observations and the simulations. Adjoint assimilation method includes three parts: forward model, backward adjoint model and optimization scheme. The adjoint model is derived from Lagrange Multiplier Method. The governing equation of the time-dependent marine ecological model is assumed to be

$$\frac{\partial x}{\partial t} = f(x,c) \tag{3}$$

where x is the state variable of the ecological model; c is the control variable; f represents the nonlinear vector operator; t is the time variable. Lagrange function is defined as

$$L(x,c,\lambda) = J(x,c) + \langle \lambda \cdot G(x,c) \rangle \tag{4}$$

where $G(x,c)=\frac{\partial x}{\partial t}-f(x,c)$; J(x,c) is cost function;<-> is the inner product defined on a Hilbert space.

Assuming that there are observations of the state variable x on Ω×T, where Ω means special scale and T means temporal scale, the cost function can be expressed as

$$J(x,c)=\frac{1}{2}\int_{\Omega\times T}(x-x_{obs})^2 dwdt \tag{5}$$

Therefore, the constrained minimum problem is transformed into an unconstrained problem about x^*, c^* and λ^* . Consequently, determining the stagnation point of cost function $J(x,c)$ under the restriction of strong constraint condition $G(x,c)=0$ is equivalent to determining the stagnation point of Lagrange function with respect to state variable x, control variable c and Lagrange multiplier λ . The equations representing stagnation point of Lagrange function are also called Euler-Lagrange equations of constrained minimum value problem. Euler-Lagrange optimal conditions (optimal x^*, c^* and λ^*) can be determined by the following equations

$$\frac{\partial L}{\partial x}(x^*,c^*,\lambda^*)=0 \tag{6}$$

$$\frac{\partial L}{\partial c}(x^*,c^*,\lambda^*)=0 \tag{7}$$

$$\frac{\partial L}{\partial \lambda}(x^*,c^*,\lambda^*)=0 \tag{8}$$

Equation (7) corresponds to the original model equation. Equation (6), the adjoint equation of equation (7), is a set of equations about Lagrange multiplier. Equation (8) is the gradient expression of cost function J with respect to control variable c. Based on equation (6) and Equation (7), the gradient in equation (8) can be calculated. Since cost function declines in the inverse direction of its gradient with respect to control variable, the direction to optimize the control variable can be determined.

Considering the convection and diffusion processes of SPM and the hydrodynamic conditions of the Liaodong Bay, the material transport model in the Liaodong Bay was built at first and the governing equation was

$$\frac{\partial c}{\partial t}+u\frac{\partial c}{\partial x}+v\frac{\partial c}{\partial y}+w\frac{\partial c}{\partial z}=\frac{\partial}{\partial x}\left(A_H\frac{\partial c}{\partial x}\right)+\frac{\partial}{\partial y}\left(A_H\frac{\partial c}{\partial y}\right)+\frac{\partial}{\partial z}\left(K_H\frac{\partial c}{\partial z}\right)-rc \tag{9}$$

where A_H was horizontal diffusion coefficient and K_H was vertical diffusion coefficient; c meant SPM concentration; u, v and w were the flow velocity in the x, y and z directions, respectively; r was degradation rate of SPM and r=0 considering SPM as conserved substance in the Liaodong Bay.

The adjoint equation of Equation (9) was

$$-\frac{\partial c^*}{\partial t}-\frac{\partial}{\partial z}\left(K_H\frac{\partial c^*}{\partial z}\right)=\frac{\partial uc^*}{\partial x}+\frac{\partial vc^*}{\partial y}+\frac{\partial wc^*}{\partial z}+\frac{\partial}{\partial x}\left(A_H\frac{\partial c^*}{\partial x}\right)+\frac{\partial}{\partial y}\left(A_H\frac{\partial c^*}{\partial y}\right)-K_c(c-\overline{c}) \tag{10}$$

where c^* was adjoint variable of SPM concentration c.

The gradient of the cost function with respect to the SPM concentrations at the initial moment was

$$\frac{\partial J}{\partial c^1}=\left(\frac{\partial c^*}{\partial t}\right)^1+\left(\frac{\partial uc^*}{\partial x}\right)^1+\left(\frac{\partial vc^*}{\partial y}\right)^1+\left(\frac{\partial wc^*}{\partial z}\right)^1+\frac{\partial}{\partial x}\left(A_H\frac{\partial c^*}{\partial x}\right)^1+\frac{\partial}{\partial y}\left(A_H\frac{\partial c^*}{\partial y}\right)^1+\frac{\partial}{\partial z}\left(K_H\frac{\partial c^*}{\partial z}\right)^1 \tag{11}$$

where superscript 1 denoted the first interation step in the calculation process.

The differential formats of Equations (9) and (10) were expressed as

$$\begin{aligned}&\frac{c^{l+1}_{i,j,k}-c^{l}_{i,j,k}}{\Delta t}-\left[\frac{K_H\left(c^{l+1}_{i,j,k+1}-c^{l+1}_{i,j,k}\right)}{\Delta z_{k+1/2}\cdot\Delta z_{k+1}}-\frac{K_H\left(c^{l+1}_{i,j,k}-c^{l+1}_{i,j,k-1}\right)}{\Delta z_{k+1/2}\cdot\Delta z_{k}}\right]=\\&-\frac{u^{l}_{i,j,k}\left(c^{l}_{i+1,j,k}-c^{l}_{i-1,j,k}\right)}{2\Delta x_j}-\frac{v^{l}_{i,j,k}\left(c^{l}_{i,j+1,k}-c^{l}_{i,j-1,k}\right)}{2\Delta y}-\frac{w^{l}_{i,j,k}\left(c^{l}_{i,j,k+1}-c^{l}_{i,j,k-1}\right)}{2\Delta z_k}\\&+\left[\frac{A_H\left(c^{l}_{i+1,j,k}-c^{l}_{i,j,k}\right)}{\Delta x_{j+1/2}\cdot\Delta x_{j+1}}-\frac{A_H\left(c^{l}_{i,j,k}-c^{l}_{i-1,j,k}\right)}{\Delta x_{j+1/2}\cdot\Delta x_{j}}\right]\\&+\left[\frac{A_H\left(c^{l}_{i,j+1,k}-c^{l}_{i,j,k}\right)}{(\Delta y)^2}-\frac{A_H\left(c^{l}_{i,j,k}-c^{l}_{i,j-1,k}\right)}{(\Delta y)^2}\right]\end{aligned} \tag{12}$$

$$\begin{aligned}&\frac{c^{*l-1}_{i,j,k}-c^{*l}_{i,j,k}}{\Delta t}-\left[\frac{K_H\left(c^{*l-1}_{i,j,k+1}-c^{*l-1}_{i,j,k}\right)}{\Delta z_{k+1/2}\cdot\Delta z_{k+1}}-\frac{K_H\left(c^{*l-1}_{i,j,k}-c^{*l-1}_{i,j,k-1}\right)}{\Delta z_{k+1/2}\cdot\Delta z_{k}}\right]\\&=\frac{\left(u^{l}_{i+1,j,k}c^{*l}_{i+1,j,k}-u^{l}_{i-1,j,k}c^{*l}_{i-1,j,k}\right)}{2\Delta x_j}+\frac{\left(v^{l}_{i,j+1,k}c^{*l}_{i,j+1,k}-v^{l}_{i,j-1,k}c^{*l}_{i,j-1,k}\right)}{2\Delta y}\\&+\frac{\left(w^{l}_{i,j,k+1}c^{*l}_{i,j,k+1}-w^{l}_{i,j,k-1}c^{*l}_{i,j,k-1}\right)}{2\Delta z_k}\\&+\left[\frac{A_H\left(c^{*l}_{i+1,j,k}-c^{*l}_{i,j,k}\right)}{\Delta x_{j+1/2}\cdot\Delta x_{j+1}}-\frac{A_H\left(c^{*l}_{i,j,k}-c^{*l}_{i-1,j,k}\right)}{\Delta x_{j+1/2}\cdot\Delta x_{j}}\right]\\&+\left[\frac{A_H\left(c^{*l}_{i,j+1,k}-c^{*l}_{i,j,k}\right)}{(\Delta y)^2}-\frac{A_H\left(c^{*l}_{i,j,k}-c^{*l}_{i,j-1,k}\right)}{(\Delta y)^2}\right]-K_c\left(c^{l}_{i,j,k}-\overline{c^{l}_{i,j,k}}\right)\end{aligned} \tag{13}$$

The computational domain was 38.5°N-41°N, 119.5°E-122.5°E (the Liaodong Bay) with a 4′ × 4′ horizontal resolution. There were 6 layers in vertical profiles. The depth of each layer from top to bottom was 10m, 20m, 30m, 50m, 75m, and 100m, respectively. The computing time was 30 days and the time step was set to be 6 h. The hydrodynamic field was provided by POM, as described above. In order to improve the simulation accuracy of the adjoint assimilation model, the independent grids were selected every 4 common grids. Only the SPM

concentrations of these independent grids needed to be optimized while those of other grid can be calculated by Cressman interpolation. The influencing radius of Cressman interpolation was 1.2 times of the distance between adjacent independent grids. Assimilation stopped when the preset ending condition was satisfied, and the monthly averaged distribution of the model results was taken as the result of interpolation with dynamical constraint.

2.5.1 Reversion of given distribution in ideal experiments

In order to verify the accuracy of the interpolation method with dynamical constraint, two ideal experiments were carried out. An initial distribution was given in the Liaodong Bay. Then the forward model was run for 30 days to obtain distributions at every time step. Idealized observations were picked up from the model-generated concentrations according to the following principle: the sampling locations were the same as the locations of monitoring stations. The adjoint assimilation model was used to optimize the distribution, and the monthly averaged result of the adjoint assimilation model was taken as the result of the interpolation with dynamical constraint. Compared with the given distribution, Mean Absolute Error (MAE) and Normalized Mean Error (NME) can be calculated. The accuracy of the dynamically constrained interpolation method was evaluated by the decrease of the simplified cost function, MAE and NME.

Two ideal experiments were carried out respectively:

1) Assume that the initial distribution of concentrations showed a parabolic surface with downward convex, and the concentration at any grid can be calculated by formula (14). The vertex of the parabolic surface was located at the geometric center of the Liaodong Bay (39.97°N, 121.02°E). The curvature and the maximum value of the parabolic surface were adjusted to make sure the minimum value of the parabolic surface was 10.0 mg/L, which was the SPM background value of the Liaodong Bay and the average value of the parabolic surface was 32.3 mg/L, which was the average value of the observational SPM concentrations.
2) Assume that the initial distribution of concentrations increased uniformly with the distance from the mouth of the Liaodong Bay, and the concentration at any grid can be calculated by formula (15).The minimum value of the given distribution of concentrations was 10.0 mg/L, which was the SPM background value of the Liaodong Bay. The constant values were adjusted to make sure that the gradient of the given concentration was perpendicular to the boundary line between the Liaodong Bay and the central Bohai Sea. The value of the gradient was adjusted to make sure the average value of the given distribution was 32.3 mg/L, which was the average value of the observational SPM concentrations.

$$-25\times\left[(lon(i)-121.02)^2+(lat(j)-39.97)^2\right]+45.3 \tag{14}$$

$$18\times\frac{\left|-0.6743\times lon(i)-lat(j)+120.4112\right|}{\sqrt{(-0.6743)^2+1}}+10.0 \tag{15}$$

where $lon(i)$ indicated the longitude and $lat(j)$ indicated the latitude. The index triplet (i, j) was a pointer to certain grid in the given initial field. We guessed that the initial concentration of the model at any grid was 10.0 mg/L and adjusted the distributions using adjoint assimilation model.

2.5.2 Dynamically constrained interpolation in practical applications

The concentrations of SPM in the Liaodong Bay in March, May, August and October 2015 were interpolated with dynamical constraint. The initial concentration at any grid was set as the average value of the observations. With the simulation of adjoint assimilation model for 30 days, the distributions of the SPM concentrations can be obtained, meanwhile MAEs and NMEs at the grid of observation stations can be calculated.

2.6 Cross-validation

Cross-validation is an effective method to evaluate the results of interpolations (Robinson and Metternicht, 2006; Hofierka et al., 2007; Wise, 2011; Etherington, 2020). The leave-one-out cross-validation method was adopted in this study. All observations were randomly and averagely divided into N groups. The N-1 groups were used in spatial interpolation to obtain the SPM concentration at any grid. Since the remaining 1 group had not been used in interpolation, it can be used as test data to check the interpolated distribution. MAEs and NMEs were calculated by combining the interpolated distribution with test data. Their calculation formulas were

$$MAE=\frac{1}{M}\sum_{i=1}^{M}\left|O_{\mathrm{int}erp}-O_{obs}\right| \tag{16}$$

$$NME=\frac{1}{M}\sum_{i=1}^{M}\frac{\left|O_{\mathrm{int}erp}-O_{obs}\right|}{O_{obs}} \tag{17}$$

where O_{obs} was SPM observation in the test data; O_{interp} was interpolated SPM concentration, and the location of O_{interp} was the same as the location of O_{obs}; M was the quantity of test data. The experiment was repeated N times, hence each group of observations was used as test data by turns. The averaged MAE

and NME of the N times were calculated as the indexes to evaluate interpolation methods. A 10-fold cross-validation method was adopted in this study (N=10).

3 Result

3.1 Application of the interpolation method with dynamical constraint in two ideal experiments

Two ideal experiments were carried out to testify the feasibility and validity of the interpolation methodology with dynamical constraint. The initial distribution, the idealized observations and the evaluation indexes in the ideal experiment were described in Section 2.5.1.

3.1.1 Ideal experiment I: Distribution of concentrations shows a parabolic surface

In Ideal Experiment I, the given initial distribution of SPM concentrations in the Liaodong Bay showed a parabolic surface with downward convex. SPM concentrations had the maximum value of 45.3 mg/L in the central Liaodong Bay and decreased outward. The low values appeared around the coastal areas, and the minimum value of 10.0 mg/L occurred in the top and the mouth of the bay. The concentration of SPM at any grid can be calculated by formula (14). The interpolation method with dynamical constraint was used to calculate the interpolated distribution. The given initial and the interpolated distributions of SPM concentrations were compared in Figure 3. It was obvious that the distribution of the dynamically constrained interpolation was consistent with the given initial distribution, which had the maximum concentration in the central Liaodong Bay and decreased in all direction. The interpolated distribution of SPM concentrations also showed a parabolic surface with downward convex analogously.

To quantitatively compare the interpolated and observed concentrations, MAE and NME were calculated by formula (16) and formula (17). The initial values of MAE and NME and the final values of MAE and NME after 100 calculating steps of the adjoint assimilation were compared in Table 1. The final values of MAE and NME were 1.003 mg/L and 0.043, and decreased by 94.6% and 92.9%, respectively. The final NME was 0.043, which indicated that the final errors of the interpolated results accounted for only 4.3% of the observations. The reduction ratio of the simplified cost function was also shown in Table 1. The cost function decreased by 99.8%. These high reduction ratios suggested that dynamically constrained interpolation based on the adjoint assimilation model greatly reduced the differences between the interpolated distribution and the observed distribution during the interpolation process. An accurate distribution of SPM concentrations was obtained using the interpolation method with dynamical constraint.

3.1.2 Ideal experiment II: Distribution of concentrations increases uniformly

In Ideal Experiment II, the given initial distribution exhibited a uniform increase of SPM concentrations from the mouth to the top of the bay. SPM concentrations had the lowest value of 10.0 mg/L along the mouth and the highest value of 45.6 mg/L around the top of the Liaodong Bay. The concentration of SPM at any grid can be calculated by formula (15).The interpolation method with dynamical constraint was used to obtain the interpolated distribution. The given initial distribution and the interpolated distribution of SPM concentrations were compared in Figure 4. It suggested that the distribution of the dynamically constrained interpolation

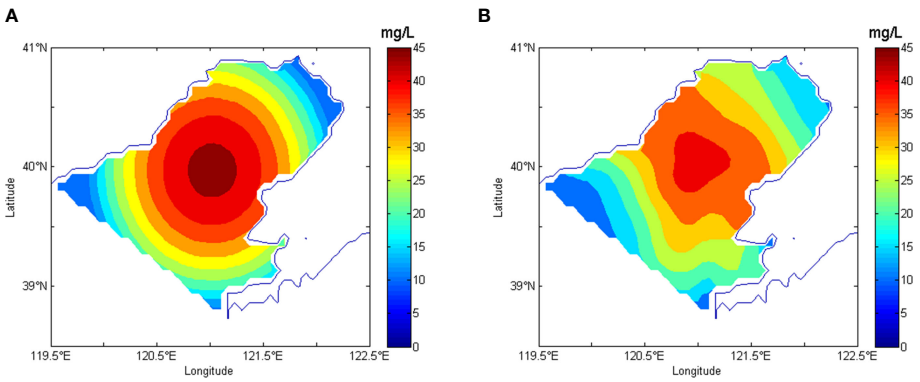


FIGURE 3
Interpolated results of given initial distribution of SPM that shows a parabolic surface with downward convex. **(A)** is the given distribution in Ideal Experiment I. **(B)** is interpolated distribution derived from dynamically constrained interpolation. The observations using in the interpolation were picked up from the model-generated concentrations based on given distributions.

TABLE 1 The cost function, MAE and NME of Ideal Experiment I.

	Cost function	MAE(mg/L)	NME
Initial value	1.0000	18.404	0.603
Final value	0.0019	1.003	0.043
Reduction ratio	99.8%	94.6%	92.9%

The decrease of cost function, MAE and NME was used to evaluate the accuracy of the dynamically constrained interpolation when observations showed a parabolic surface with downward convex. The initial values were acquired at the first calculating step of the adjoint assimilation model, and the final values were at the last calculating step. A large reduction ratio means that the interpolated distribution was close to the observations.

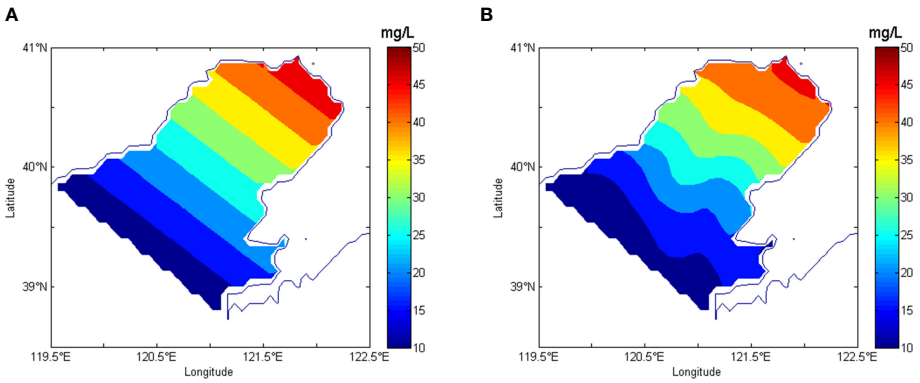


FIGURE 4
Interpolated results of given initial distribution of SPM that increases uniformly with the distance from the mouth. **(A)** is the given initial distribution in Ideal Experiment II. **(B)** is interpolated distribution derived from dynamically constrained interpolation. The observations using in the interpolation were picked up from the model-generated concentrations based on given distributions.

exhibited high agreement with the given initial distribution. SPM concentrations had the minimum value along the mouth of the Liaodong Bay and then increased gradually to the maximum value near the top.

The simplified cost function decreased by 99.9% (Table 2). The final values of MAE and NME were 0.924 mg/L and 0.036, which decreased by 95.1% and 93.8% respectively. The final NME was 0.036, which indicated that the final errors of the interpolated results accounted for only 3.6% of the observations. It was obvious that the dynamically constrained interpolation based on the adjoint assimilation model greatly reduced the differences between the interpolated and the observed distributions. An accurate distribution of SPM concentrations can be obtained using interpolation method with dynamical constraint in Ideal Experiment II.

Furthermore, the results of the two ideal experiments were compared with the similar ideal experiments carried out in previous studies. Wang et al. (2013) set four ideal experiments respectively in the Bohai Sea. The given distributions showed a parabolic surface with upward convex, a parabolic surface with downward convex, a conical surface with upward convex and a conical surface with downward convex in four ideal experiments respectively. The cost function decreased by 96.3%, 91.7%, 95.6% and 90.9%, respectively. The MAE decreased by 88.7%, 86.9%, 92.5% and 92.1%, respectively. The NME decreased by 85.2%, 87.4%, 86.5% and 87.1%, respectively. Huang et al. (2021) set a similar ideal experiment in the Laizhou Bay. The concentrations of petroleum hydrocarbon pollutants decreased exponentially in the given distribution of ideal experiment. The MAE decreased by 88.40%. Their models exhibited high accuracy during

TABLE 2 The cost function, MAE and NME of Ideal Experiment II.

	Cost function	MAE(mg/L)	NME
Initial value	1.0000	19.037	0.585
Final value	0.0009	0.924	0.036
Reduction ratio	99.9%	95.1%	93.8%

The decrease of cost function, MAE and NME was used to evaluate the accuracy of the dynamically constrained interpolation when observations increased uniformly. The calculation and function of cost function, MAE and NME were similar to those in Ideal Experiment I.

simulation. In this study, the reduction ratios of the cost functions, MAE and NME were larger, which proved that the errors were smaller.

3.2 Application of the interpolation method with dynamical constraint in practical experiments

SPM observations in March, May, August and October 2015 were used for dynamically constrained interpolation. Most of the observation stations were located in the coastal area, while the observation stations in the central Liaodong Bay were rare (Figure 2). In March of 2015 (Figure 2A), a high value area of the observations appeared near the top of the Liaodong Bay. There were two high value regions of observations on the northwest and the southeast side of the mouth. The maximum value of 101.8 mg/L occurred at the top of the Liaodong Bay. The SPM observations along the northwest and southeast coastal area were low. In May (Figure 2B), the SPM observations around the top of the bay were still high. Another high value area of observations was in the semi-enclosed bay near the Dalian city. The maximum value of the observations was 547.0 mg/L, which was observed near the top. In August (Figure 2C), the SPM observations on the northwest side of the Liaodong Bay were much higher than those on the southeast side. The observations near the Dalian city were still high. The maximum value of 124.0 mg/L occurred on the northwest side of the Liaodong Bay. In October (Figure 2D), there was still a high value region of observations in the top area. The maximum value was observed here, which was 464.0 mg/L. The observations near the Dalian city were also high.

Based on the SPM observations of four months, the distributions of SPM were derived from dynamically constrained interpolation, which were shown in Figure 5. The general trend of the distributions was that the concentration was high in the top and north of the Liaodong Bay, and low in the southeast of the Liaodong Bay, which was basically consistent with the distributions of observations (Figure 2). In March of 2015 (Figure 5A), an area of high SPM concentrations appeared in the northeast part of the bay, with the values higher than 40 mg/L. A band of high values extended from northeast to southwest along the central bay, and then turned to the east area of the mouth of the bay. Another area with high values appeared in the west end of the mouth. Along the southeast and northwest coastline of the bay, there were areas with low

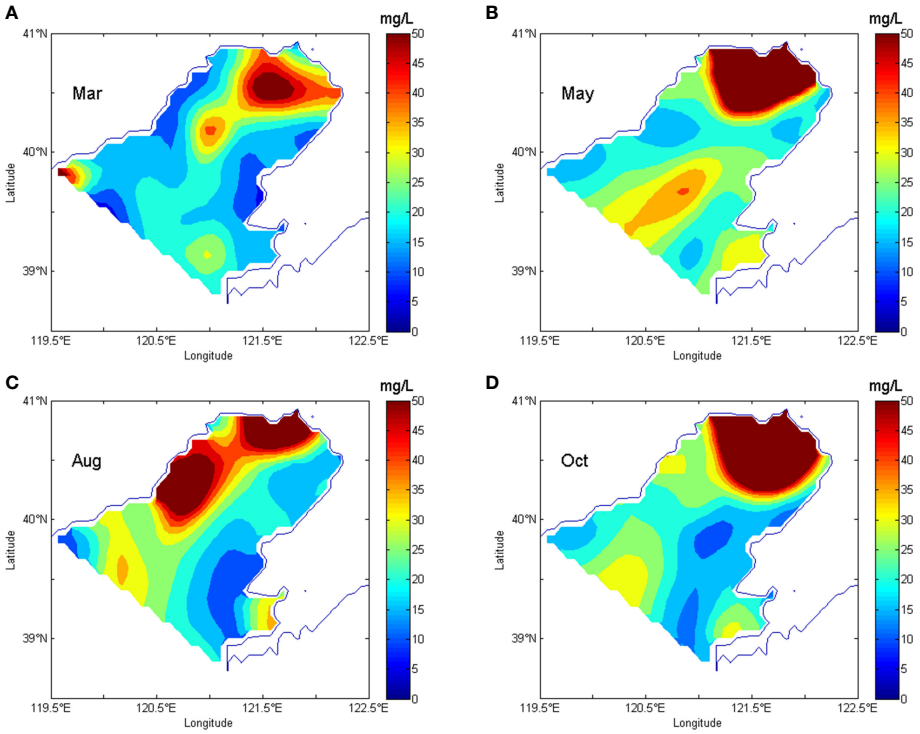


FIGURE 5
Interpolated distributions with dynamical constraint of SPM concentrations in the Liaodong Bay. **(A)** is the distribution obtained by dynamically constrained interpolation based on SPM observations in March 2015. **(B–D)** are the interpolated distributions of May, August and October 2015, respectively.

concentrations, with values lower than 15 mg/L. In May (Figure 5B), the high SPM concentrations area in the northeast part of the bay was enlarged, and the gradient of SPM concentrations was also increased. The concentration of the area was higher than 50 mg/L. A band of high concentrations extending from the central bay to the central area of bay mouth occurred. The SPM concentrations in the semi-enclosed bay near the Dalian city were higher than 30 mg/L, forming a high value area. The low concentration areas along the southeast and northwest coastline were still there, while the concentration increased. In August (Figure 5C), the high-values band across the central Liaodong Bay, appearing in March and May, translated to the northwest coastline. From the top to the mouth of the bay, the concentrations of the three high value areas on the band were greater than 50.0 mg/L, 50.0 mg/L and 30.0 mg/L, respectively. The high value areas near the Dalian city still appeared, with the maximum values higher than 35.0 mg/L. The SPM concentrations of the southeast part of the Liaodong Bay were lower than those of the northwest part of the bay. There were two low value areas occurred here, with concentrations lower than 15 mg/L and 10 mg/L respectively. In October (Figure 5D), the SPM concentrations in the northeast part of the bay were higher than 50 mg/L, forming a high value area with high value of gradient of SPM concentrations. The other high value area was found in the northwest area of the mouth, with SPM concentrations higher than 25 mg/L. The semi-enclosed bay in the southeast of the mouth of the Liaodong Bay remained a high value area, with SPM concentrations higher than 20 mg/L. A band of low values extended along the southeast coastline, with concentrations lower than 15 mg/L.

3.3 Application of Kriging interpolation and RBF interpolation in practical experiments

The interpolated distribution of SPM concentrations in the Liaodong Bay had been obtained by interpolation method with dynamical constraint in Section 3.2. The interpolation method without dynamical constraint such as Kriging interpolation and RBF interpolation were also widely used in marine research. In this part, to verify the superiority of interpolation method with dynamical constraint over interpolation method without dynamical constraint, the spatial interpolation of SPM concentrations using Kriging method and RBF method was carried out respectively. Kriging interpolation and RBF interpolation were used to interpolate SPM concentration data observed in the Liaodong Bay in 2015. The interpolated distributions of SPM concentrations using Kriging interpolation and RBF interpolation were shown in Figures 6, 7 respectively.

The interpolated distributions of Kriging interpolation, RBF interpolation and dynamically constrained interpolation were roughly similar. The high value and low value areas appearing in the interpolated results of the three interpolation methods were generally consistent with the observed distributions of SPM concentrations (Figure 2). As shown in Figures 5A, 6A, 7A, a banded high value region of SPM concentrations appeared in the central area of the Liaodong Bay. The high value area of SPM concentrations occurred in the top area of the bay, with the concentrations greater than 50 mg/L. The low values of SPM concentrations lower than 5 mg/L were in the coastal area at the northwest and southeast of Liaodong Bay. Figures 5C, 6C, 7C indicated that the interpolated results of the three interpolation methods all had a high value area of SPM concentrations in the top of the Liaodong Bay, with the concentrations greater than 50 mg/L. The low value of SPM concentrations lower than 15 mg/L appeared in the southeast of the Liaodong Bay. The interpolated distributions of the three interpolation methods all had a high value around the top of the Liaodong Bay in March and October, with the concentrations of SPM greater than 50 mg/L (Figures 5B, D, 6B, D, 7B, D).

Some high value regions in the interpolated distributions of dynamically constrained interpolation didn't appear in the interpolated distributions of Kriging interpolation and RBF interpolation, such as the high value region in the northwest area of the mouth of the Liaodong Bay in October. The accuracy and the authenticity of interpolation with dynamical constraint should be tested.

4 Discussion

4.1 Cross validations of the interpolation methods

The 10-fold cross-validation was used to evaluate and compare the results of unconstrained and dynamically constrained interpolation methods, and the averaged MAEs of each method were considered to be the evaluation index (Table 3). Beside Kriging and RBF interpolation, five more unconstrained interpolation methods were also carried out in this section to further test the accuracy of different unconstrained interpolation methods. They were the Cressman interpolation with the influencing radius of 0.40°, the Cressman interpolation with influencing radius of 0.75°, the minimum curvature interpolation, the nearest neighbor value interpolation and the inverse-distance weighting interpolation.

Since the rand-size relationship of MAEs of these interpolation methods in the four months were the same, the results of August were described as below. Among the seven unconstrained interpolation methods, the averaged MAE of RBF interpolation was the minimum, which was 10.976 mg/L. The averaged MAE of the minimum curvature interpolation was the maximum, followed by the inverse-distance weighting

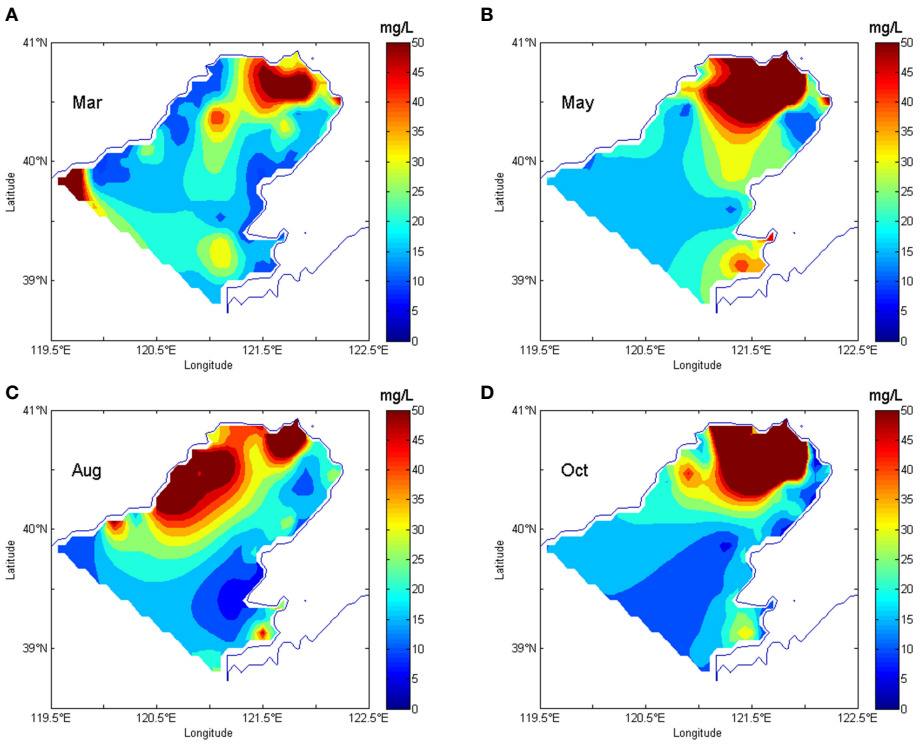


FIGURE 6
Interpolated distributions of SPM concentrations in the Liaodong Bay using Kriging interpolation. **(A)** is the distribution obtained by Kriging interpolation based on SPM observations in March 2015. **(B–D)** are the interpolated distributions of May, August and October 2015, respectively.

interpolation. The averaged MAEs of the nearest neighbor value interpolation, the Cressman interpolation with influencing radius of 0.40° and the Cressman interpolation with influencing radius of 0.75° was 10.990, 12.025 and 12.179 mg/L, which were lower than the averaged MAE of Kriging interpolation which was 13.240 mg/L. It was indicated that the Cressman and nearest neighbor value interpolations were more accurate than the Kriging interpolation. However, the distributions of these three interpolation methods weren't reasonable. The interpolated distribution of SPM concentrations of Cressman interpolation with influencing radius of 0.40° was incompletely interpolated since the influencing radius of 0.40° was not large enough to guarantee the results at every grid can be calculate by the observations. The interpolated distribution of Cressman interpolation with influencing radius of 0.75° was excessively smooth because the influencing radius was too large and too many observations were taken into calculation. The distribution derived from the nearest neighbor value interpolation was also incompletely interpolated and the concentration gradients at some grids were infinite. Kriging interpolation was more accurate than these three interpolation methods because the interpolated distribution of SPM concentrations was more reasonable. RBF interpolation was the most accurate among the unconstrained interpolation methods, followed by Kriging interpolation. Among the seven unconstrained interpolation methods, RBF and Kriging behaved much better.

When using the interpolation method with dynamically constraint, the averaged MAE and NME of all observation stations were 7.703 mg/L and 0.277 respectively. The averaged MAE of dynamically constrained interpolation method was 29.8%-44.5% lower than that of the seven unconstrained interpolation methods. The averaged MAE and NME of Kriging interpolation were 13.240 mg/L and 0.349 respectively, and those of RBF interpolation were 10.976 mg/L and 0.304 respectively. It was suggested that the averaged MAE of the dynamically constrained interpolation method was 41.8% and 29.8% smaller than those of Kringing and RBF, while the averaged NME was 20.6% and 8.8% smaller respectively.

The results of cross-validations indicated a better agreement between the distribution from the dynamically constrained interpolation and observations. However, the cross-validation only considered the errors between the interpolation results and the observations at the monitoring stations. In the areas where there was lack of observations, whether the interpolated distributions were reasonable and true should be further tested using more observations.

4.2 Influences of hydrodynamic processes on the interpolation of SPM concentrations

In Section 4.1, it was proved that RBF and Kriging interpolations were the two most accurate methods among the seven unconstrained interpolation methods, and the dynamically constrained interpolation was more accurate than RBF and Kriging interpolation. However, whether the interpolated distributions were consistent with the observations was more important. In this section, the influences of the currents (Figure 8) on the interpolated distributions were discussed, and the distributions of dynamically constrained and unconstrained interpolations were compared with the satellite-derived distributions obtained from GOCI (Figure 9). The satellite-derived SPM concentrations (Figure 9) were smaller than the *in-situ* observations (Figure 2) and the interpolated concentrations (Figures 5–7) based on the *in-situ* observations. The satellite data is frequently missing, because of the cloud cover, the variation of irradiances and the effect of sensor spatial resolution. Therefore, compared with the *in-situ* observations, the satellite-derived concentrations will underestimate the SPM concentration (Wielicki and Parker, 1992; Eleveld et al., 2014; Jia et al., 2021 and Mei et al., 2014). However, the underestimation will not affect the trend of satellite-derived SPM distributions and the satellite-derived distributions can be used to verify the interpolated distributions (Xu et al., 2020). Since the distributions of Kriging interpolation was similar to those of RBF interpolation and the results of cross-validation proved that RBF interpolation was more accurate, only the distributions of RBF interpolation was used to be compared with the distributions of dynamically constrained interpolation in this section.

TABLE 3 Cross-validation results of eight interpolation methods.

Interpolation methods	MAE (mg/L)
Minimum curvature interpolation	13.891
Inverse-distance weighting interpolation	13.316
Kriging interpolation	13.240
Cressman interpolation (0.75°radius)	12.179
Cressman interpolation (0.40°radius)	12.025
Nearest neighbor value interpolation	10.990
RBF interpolation	10.976
Dynamically constrained interpolation	7.703

Seven unconstrained interpolation methods and dynamically constrained interpolation method were compared by cross-validation. The MAEs represented the error between the interpolated results and the observations.

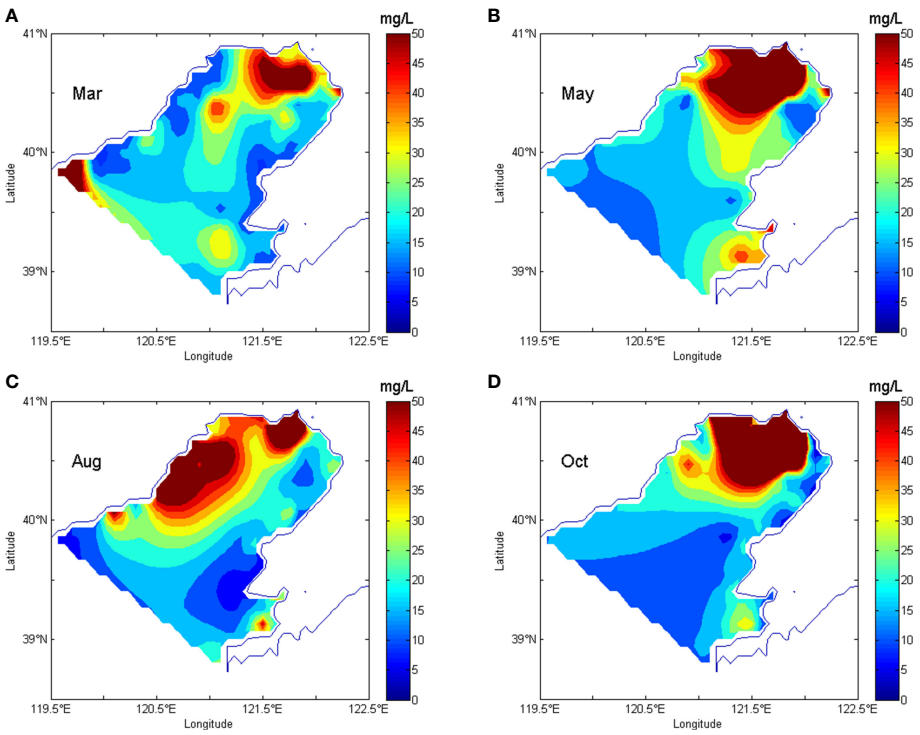


FIGURE 7
Interpolated distributions of SPM concentrations in the Liaodong Bay using RBF interpolation. **(A)** is the distribution obtained by RBF interpolation based on SPM observations in March 2015. **(B–D)** are the interpolated distributions of May, August and October 2015, respectively.

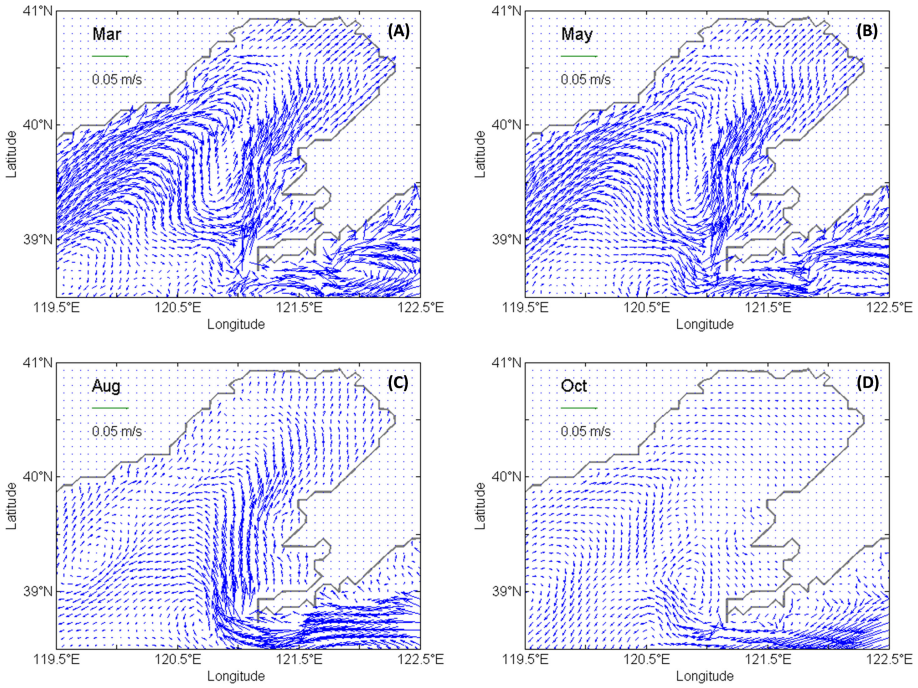


FIGURE 8
Monthly averaged currents field in the Liaodong Bay. The hydrodynamic fields were calculated by POM and had been validated by Wang et al. (2013). The Liaodong Bay part of the hydrodynamic field was interpolated to the grid of our model. **(A–D)** show the monthly averaged currents fields in March, May, August and October 2015, respectively.

In March of 2015, the distribution derived from dynamically constrained interpolation (Figure 5A) was similar to the GOCI-derived distribution (Figure 9A). Water flowed into the Liaodong Bay along the northwest and the southeast coastline, and flowed out of the Liaodong Bay through the central area (Figure 8A). Water with low SPM concentration was transported along the coastline from the central Bohai Sea to the top of the Liaodong Bay. Meanwhile, water with high SPM concentration was transported through the central Liaodong Bay from the top to the mouth of the Liaodong Bay. Consequently, in the distributions from dynamically constrained interpolation and satellite, the low values of SPM concentration appeared in the coastal areas of the Liaodong Bay and a band of high concentration extended from the top through the center to the mouth of the bay. Although the influences of currents weren't considered in the RBF interpolation process, the distribution from RBF interpolation (Figure 7A) was also consistent with the satellite-derived distribution, because the quantity of observations was large and the observation stations covered a large enough area in March.

In May, the currents were similar to those in March (Figure 8B). The currents brought the water with low SPM concentration into the bay from the central Bohai Sea to the top of the bay along the coastline, and brought the water with high SPM concentration out of the bay from the top to the mouth through the center of the bay. In the distribution of the dynamically constrained interpolation (Figure 5B), a high value area appeared around the top of the bay. The other high value area occurred in the central area of the mouth of the Liaodong Bay. In the satellite-derived distribution (Figure 9B), a same high value area appeared at the top of the bay. There was also a high value area in the bay mouth area, but it was closer to the southeast side of the bay than in the distribution of the dynamically constrained interpolation. Although the location of the area in the bay mouth (Figure 5B) was not totally consistent with that derived by satellite data (Figure 9B), the results of dynamically constrained interpolation successfully reproduced the area with high SPM concentrations in the bay mouth. However, the high values of SPM concentration in the central region of the mouth was not captured in the distribution of the RBF interpolation (Figure 7B), because of not considering the influences of hydrodynamic processes and lack of observations.

In August, the inflow at the northwest side of the Liaodong Bay became weak. Water flowed into the Liaodong Bay along the southeast side and out of the Liaodong Bay along the northwest side (Figure 8C). Water with low SPM concentration was transported along the southeast side into the Liaodong Bay and water with high SPM concentration was taken out of the Liaodong Bay along the northwest side. Therefore, the low values of SPM concentration appeared in the southeast Liaodong Bay, while the high values appeared in the northwest. The distribution of the

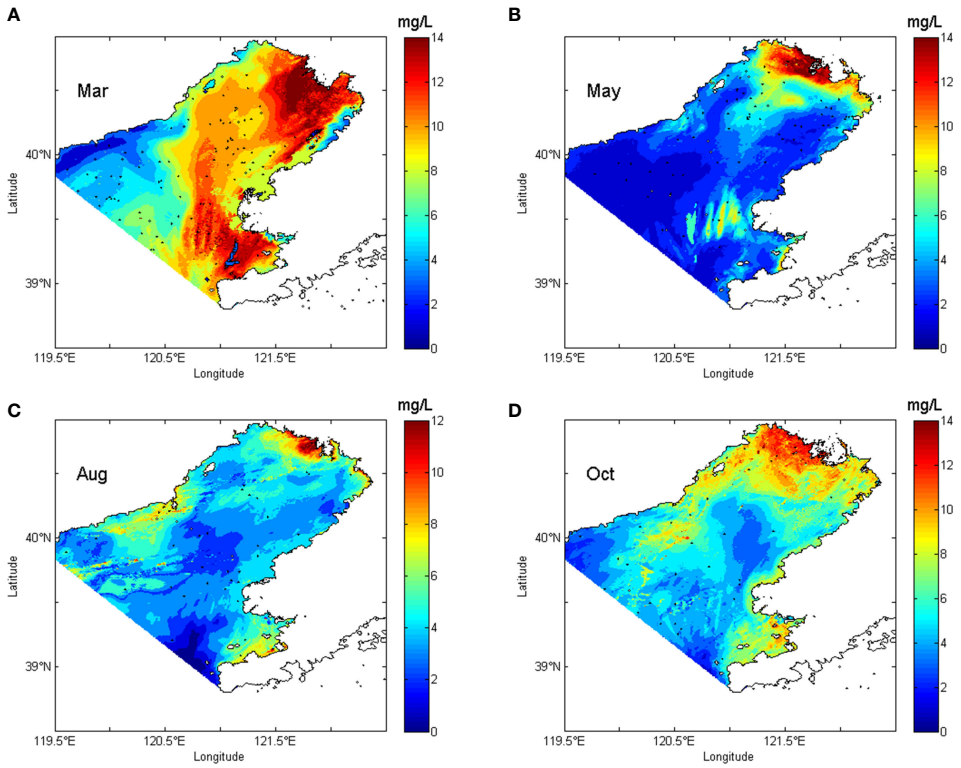


FIGURE 9
Monthly averaged SPM concentrations in the Liaodong Bay obtained from satellite. Satellite observations were obtained from GOCI. **(A)** shows the derived monthly averaged distribution of SPM concentrations in March 2015. **(B–D)** are the distributions of May, August and October 2015, respectively.

dynamically constrained interpolation (Figure 5C) was similar to the satellite-derived distribution (Figure 9C). The SPM concentrations in the northwest part of the Liaodong Bay were higher than those in the southeast. In the distribution of the dynamically constrained interpolation, a high value area occurred in the northwest of the mouth of the Liaodong Bay (Figure 5C). The same high values of SPM concentration also appeared in the SPM distribution obtained from GOCI satellite (Figure 9C). However, the same high value region didn't appear in the distribution of RBF interpolation (Figure 7C). There was a lack of observations in this area (Figure 2C). Besides, the influences of currents were not considered in the RBF interpolation. As a result, the high values here were not reflected in the distribution of RBF interpolation.

In October, the currents in the Liaodong Bay (Figure 8D) were similar to those in August but much weaker. Water flowed into the Liaodong Bay along the southeast coastline and out of the bay along the northwest coastline slowly. The SPM concentrations in the northwest of the bay were higher than those in the southeast of the bay. There were two high value areas found in the distribution of the dynamically constrained interpolation (Figure 5D) and GOCI-derived distribution (Figure 9D). The maximum SPM concentration occurred in the region near the top of the bay and the gradient of SPM concentration was obvious. Besides, a high value area was found in the northwest of the mouth of the Liaodong Bay. The high values of SPM concentration near the top was also reflected in the distribution of RBF interpolation (Figure 7D), because the currents here were weak and the observations here were abundant (Figure 2D). However, the high values around the mouth of the Liaodong Bay was not captured by the interpolated distribution using RBF interpolation (Figure 7D), because the influences of currents were not taken into consider and the observations here were rare (Figure 2D).

In this part, we discussed the influence of hydrodynamic processes on the interpolated distribution with dynamical constraint, and verified the interpolated results (Figure 5) with SPM distributions from satellite (Figure 9). Dynamically constrained interpolation can provide the distributions of SPM concentrations more consistent with the observations than the unconstrained interpolation methods. Besides, dynamically constrained interpolation can fully consider the influence of hydrodynamic processes, and provided the interpolated concentration similar to the actual in the area missing observations.

4.3 Factors controlling the temporal and spatial variations of SPM concentrations in the Liaodong Bay

Based on the SPM observations of March, May, August and October 2015 in the Liaodong Bay and interpolated SPM distributions obtained from dynamically constrained interpolation, the influencing factors of temporal and spatial variations of SPM concentrations were discussed in this part. The average value of SPM concentrations observed in March was 23.6 mg/L, the maximum value was 101.8 mg/L, and the minimum value was 7.4 mg/L. The average value of SPM concentrations observed in May was 42.0 mg/L, the maximum value was 547.0 mg/L, and the minimum value was 11.0 mg/L. The average value of SPM concentrations observed in August was 28.0 mg/L, the maximum value was 124.0 mg/L, and the minimum value was 5.6 mg/L. The average value of SPM concentrations observed in October was 39 mg/L, the maximum value was 464 mg/L, and the minimum value was 7.2 mg/L. The results of spatial interpolation with dynamical constraint indicated that the averaged SPM concentration of May was the highest, which was consistent with the observations. The maximum value of SPM concentration of May was located near the estuary area. Besides, the maximum SPM concentration of May was also the highest.

SPM concentrations in the Liaodong Bay were greatly affected by river's run-off SPM inputs. Based on the River Sediment Bulletin of China published by the Ministry of Water Resources of the People's Republic of China (http://www.mwr.gov.cn), the river discharge and sediment discharge in the estuary of the Liao River Basin were analyzed. The river discharge in March, May, August and October 2015 was 60, 220, 100 and 70 million m3, respectively. The sediment discharge in March, May, August and October 2015 was 16, 68, 16 and 3 thousand tons, respectively. The concentration of sediment in the water discharged from the estuary of Liao River Basin in March, May, August and October 2015 was 266.7, 309.1, 160.0 and 42.9 mg/L, respectively. The sediment discharge and concentration of sediment were both the highest in May. It is presumed that the increase of the river run-off SPM inputs in May leaded to the increase of the SPM concentration in the Liaodong Bay. This indicates that the SPM discharge of the estuary area may be an important reason for the variations of SPM concentration in the Liaodong Bay.

The currents in the Liaodong Bay also had influences on the distributions of SPM concentrations, which were discussed in Section 4.2. The SPM concentration tended to be high in the area currents flowing from the top of the bay out of the bay. Low concentrations tended to occur in the area currents flowing from the outer bay to the inner bay. There was a permanent northeast flow in the southeast side of Liaodong Bay, bring fresh water from the central Bohai Sea into the Liaodong Bay, which made the SPM concentration in the southeast side of Liaodong Bay low throughout the year. In March and May, the currents flowed from the top of the bay out of the bay through the central Liaodong Bay, and the value of the SPM concentration of the central Liaodong Bay was higher than other areas. In August and October, the currents flowed from the top of the bay out of the bay along the northwest Liaodong Bay, and the high value of SPM concentration occurred there.

5 Conclusion

In this study, an interpolation method with dynamical constraint was established to interpolate the SPM concentrations. The results of dynamically constrained interpolation were optimized iteratively with the adjoint assimilation method. In two ideal experiments, the final NMEs after 100 calculating steps were 0.043 and 0.036, which means the final errors of the interpolated results accounted for only 4.3% and 3.6% of the observations. The results proved the accuracy of dynamically constrained interpolation.

In the cross-validation experiments, the averaged MAE of dynamically constrained interpolation was 7.703 mg/L, which was 29.8%-44.5% lower than that of other unconstrained interpolation methods. Compared with unconstrained interpolation methods, dynamically constrained interpolation provided more consistent interpolated SPM concentrations at the observation stations. In addition, the averaged MAE of RBF interpolation method was 10.976 mg/L, which was the lowest among the seven unconstrained interpolation methods. RBF interpolation method was the best choice among unconstrained interpolation methods when dynamically constrained interpolation cannot be used due to the lack of dynamically constrained conditions.

The interpolated distributions were compared with the distributions obtained from satellite. The GOCI satellite-derived distributions of SPM were similar to the distributions of interpolation method with dynamical constraint. However, because of the lack of observations in some areas of Liaodong Bay and not taking the hydrodynamic processed into consider, there were many high value regions weren't captured by the RBF interpolation distributions, which existed in the satellite distributions. It was revealed that the hydrodynamic processes had important influence on the distributions of SPM, which could be used to improve the accuracy of interpolations.

The determining factors of the temporal and spatial variations of SPM concentrations in the Liaodong Bay were analyzed. The SPM concentration in the Liaodong Bay was highest in May due to the influence of the change of river SPM discharge in the Liao River Basin. Under the influence of the flow field, the SPM concentration in the northwest and the central Liaodong Bay was higher than that in the southeast Liaodong Bay.

Data availability statement

The original contributions presented in the study are included in the article/supplementary materials. Further inquiries can be directed to the corresponding author.

Author contributions

Conceptualization, JS. Methodology, JS. Software, HG. Validation, JS and XM. Formal analysis, HG and HP. Investigation, XM and HP. Data curation, HP. Writing—original draft preparation, HG. Writing—review and editing, JS and HG. Supervision, XL. Project administration, XL. and XM. All authors contributed to the article and approved the submitted version.

Funding

This paper was supported by the National Natural Science Foundation of China (U1806214 and 42176022).

Acknowledgments

We are grateful to all those who helped us work out our problems during this study. We express our heartfelt gratitude to all the group members for their helpful assistance and advice.

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Bourgain, P., and Gascard, J. C. (2011). The Arctic ocean halocline and its interannual variability from 1997 to 2008. *Deep Sea Res. Part I: Oceanographic Res. Papers* 58 (7), 745–756. doi: 10.1016/j.dsr.2011.05.001

Chen, H., Cao, A., Zhang, J., Miao, C., and Lv, X. (2014). Estimation of spatially varying open boundary conditions for a numerical internal tidal model with adjoint method. *Mathematics Comput. Simulation* 97, 14–38. doi: 10.1016/j.matcom.2013.08.005

Cloern, J. E. (1987). Turbidity as a control on phytoplankton biomass and productivity in estuaries. *Continental shelf Res.* 7 (11-12), 1367–1381. doi: 10.1016/0278-4343(87)90042-2

Davis, P. J. (1975). *Interpolation and approximation* (New York: Dover Publications (reprint)). doi: 10.1112/jlms/s1-39.1.568

Dou, Y., Li, J., Zhao, J., Wei, H., Yang, S., Bai, F., et al. (2014). Clay mineral distributions in surface sediments of the liaodong bay, bohai Sea and surrounding river sediments: Sources and transport patterns. *Continental Shelf Res.* 73, 72–82. doi: 10.1016/j.csr.2013.11.023

D'Sa, E. J., Miller, R. L., and Castillo, C. D. (2006). Bio-optical properties and ocean color algorithms for coastal waters influenced by the Mississippi river during a coldfront. *Appl. Op.* 45, 7410–7428. doi: 10.1364/AO.45.007410

Eleveld, M. A., van der Wal, D., and Van Kessel, T. (2014). Estuarine suspended particulate matter concentrations from sun-synchronous satellite remote sensing: Tidal and meteorological effects and biases. *Remote Sens. Environ.* 143, 204–215. doi: 10.1016/j.rse.2013.12.019

Etherington, T. R. (2020). Discrete natural neighbour interpolation with uncertainty using cross-validation error-distance fields. *PeerJ Comput. Sci.* 6, e282. doi: 10.7717/PEERJ-CS.282

Fettweis, M., Baeye, M., Zande, D. V. D., Eynde, D. V. D., and Lee, B. J. (2014). Seasonality of floc strength in the southern north Sea. *J. Geophysical Research: Oceans* 119, 1911–1926. doi: 10.1002/2013jc009750

Goldberg, D. N., and Heimbach, P. (2013). Parameter and state estimation with a time-dependent adjoint marine ice sheet model. *Cryosphere* 7 (6), 1659–1678. doi: 10.5194/tc-7-1659-2013

He, Q., Qiu, Y., Liu, H., Sun, X., Kang, L., Cao, L., et al. (2017). New insights into the impacts of suspended particulate matter on phytoplankton density in a tributary of the three gorges reservoir, China. *Sci. Rep.* 7 (1), 1–11. doi: 10.1038/s41598-017-13235-0

Hofierka, J., Cebecauer, T., and Šúri, M. (2007). "Optimisation of interpolation parameters using cross-validation," in *Digital terrain modelling* (Berlin, Heidelberg: Springer), 67–82. doi: 10.1007/978-3-540-36731-4_3

Huang, S., Han, H., Li, X., Song, D., Shi, W., Zhang, S., et al. (2021). Inversion of the degradation coefficient of petroleum hydrocarbon pollutants in laizhou bay. *J. Mar. Sci. Eng.* 9 (6), 655. doi: 10.3390/jmse9060655

Injan, S., Wangwongchai, A., Humphries, U., Khan, A., and Yusuf, A. (2021). Reinitializing Sea surface temperature in the ensemble intermediate coupled model for improved forecasts. *Axioms* 10 (3), 189. doi: 10.3390/axioms10030189

Jia, H., Ma, X., Yu, F., and Quaas, J. (2021). Significant underestimation of radiative forcing by aerosol–cloud interactions derived from satellite-based methods. *Nat. Commun.* 12 (1), 1–11. doi: 10.1038/s41467-021-23888-1

Jiang, W., Pohlmann, T., Sun, J., and Starke, A. (2004). SPM transport in the bohai Sea: field experiments and numerical modelling. *J. Mar. Syst.* 44 (3-4), 175–188. doi: 10.1016/j.jmarsys.2003.09.009

Lan, X. H., Meng, X. J., and Hou, F. H. (2016). Element geochemistry of sediments in liaodong bay of the bohai Sea. *Mar. Geo. Fron.* 32, 48–53. doi: 10.16028/j.1009-2722.2016.05007

Li, P., Ke, Y., Bai, J., Zhang, S., Chen, M., and Zhou, D. (2019). Spatiotemporal dynamics of suspended particulate matter in the yellow river estuary, China during the past two decades based on time-series landsat and sentinel-2 data. *Mar. pollut. Bull.* 149, 110518. doi: 10.1016/j.marpolbul.2019.110518

Liu, J. G., Li, A. C., and Chen, M. H. (2007). Geochemical characteristics of holocene argillaceous sediments in the bohai Sea. *Geo. Chem.* 36, 559–568. doi: 10.1631/jzus.2007.A1858

Li, X., Wang, C., Fan, W., and Lv, X. (2013). Optimization of the spatio-temporal parameters in a dynamical marine ecosystem model based on the adjoint assimilation. *Math. Problems Eng.* 2013, 1–12. doi: 10.1155/2013/373540

Li, N., Xu, J., and Lv, X. (2021). Application of dynamically constrained interpolation methodology in simulating national-scale spatial distribution of $PM_{2.5}$ concentrations in China. *Atmosphere* 12 (2), 272. doi: 10.3390/atmos12020272

Lu, X., and Zhang, J. (2006). Numerical study on spatially varying bottom friction coefficient of a 2D tidal model with adjoint method. *Continental shelf Res.* 26 (16), 1905–1923. doi: 10.1016/j.csr.2006.06.007

Maerz, J., Hofmeister, R., Lee, E., Grwe, U., Riethmüller, R., and Kai, W. W. (2016). Evidence for a maximum of sinking velocities of suspended particulate matter in a coastal transition zone. *Biogeosciences Discussions* 13, 4863–4876. doi: 10.5194/bg-13-4863-2016

Mao, X., Wang, D., Zhang, J., Bian, C., and Lv, X. (2018). Dynamically constrained interpolation of the sparsely observed suspended sediment concentrations in both space and time: A case study in the bohai Sea. *J. Atmospheric Oceanic Technol.* 35 (5), 1151–1167. doi: 10.1175/JTECH-D-17-0149.1

Mei, Y., Anagnostou, E. N., Nikolopoulos, E. I., and Borga, M. (2014). Error analysis of satellite precipitation products in mountainous basins. *J. Hydrometeorology* 15 (5), 1778–1793. doi: 10.1175/JHM-D-13-0194.1

Pan, H., Guo, Z., and Lv, X. (2017). Inversion of tidal open boundary conditions of the M_2 constituent in the bohai and yellow seas. *J. Atmospheric Oceanic Technol.* 34 (8), 1661–1672. doi: 10.1175/JTECH-D-16-0238.1

Peng, S. Q., and Xie, L. (2006). Effect of determining initial conditions by four-dimensional variational data assimilation on storm surge forecasting. *Ocean Model.* 14 (1-2), 1–18. doi: 10.1016/j.ocemod.2006.03.005

Qi, P., Wang, C., Li, X., and Lv, X. (2011). Numerical study on spatially varying control parameters of a marine ecosystem dynamical model with adjoint method. *Acta Oceanologica Sin.* 30 (1), 7–14. doi: 10.1007/s13131-011-0085-8

Robinson, T. P., and Metternicht, G. (2006). Testing the performance of spatial interpolation techniques for mapping soil properties. *Comput. Electron. Agric.* 50 (2), 97–108. doi: 10.1016/j.compag.2005.07.003

Sasaki, Y. (1970). Some basic formalisms in numerical variational analysis. *Monthly Weather Rev.* 98 (12), 875–883. doi: 10.1175/1520-0493(1970)098<0875:SBFINV>2.3.CO;2

Volpe, V., Silvestri, S., and Marani, M. (2011). Remote sensing retrieval of suspended sediment concentration in shallow waters. *Remote Sens. Environ.* 115 (1), 44–54. doi: 10.1175/1520-0493(1970)098<0875:SBFINV>2.3.CO;2

Walker, N. D., and Hammack, A. B. (2000). Impacts of winter storms on circulation and sediment transport: Atchafalaya-vermilion bay region, Louisiana, USA. *J. Coast. Res.* 16, 996–1010. doi: 10.2307/4300118

Wang, C., Li, X., and Lv, X. (2013). Numerical study on initial field of pollution in the bohai Sea with an adjoint method. *Math. Problems Eng.* 4, 389–405. doi: 10.1155/2013/104591

Wang, Q., Shi, W., and Atkinson, P. M. (2014). Sub-Pixel mapping of remote sensing images based on radial basis function interpolation. *ISPRS J. Photogrammetry Remote Sens.* 92, 1–15. doi: 10.1016/j.isprsjprs.2014.02.012

Wang, X., Wang, G., and Zhang, X. (2016). "Prediction of chlorophyll-a content using hybrid model of least squares support vector regression and radial basis function neural networks," in *2016 Sixth International Conference on Information Science and Technology (ICIST)*, Dalian, China, 6–8 May 2016, 366–371.

Wang, H. Z., Zhang, R., Liu, K. F., Liu, W., Wang, G. H., and Li, N. (2008). Improved kriging interpolation based on support vector machine and its application in oceanic missing data recovery. *Int. Conf. Comput. Sci. Software Engineering.* 4, 726–729. doi: 10.1109/CSSE.2008.924

Wielicki, B. A., and Parker, L. (1992). On the determination of cloud cover from satellite sensors: The effect of sensor spatial resolution. *J. Geophysical Research: Atmospheres* 97 (D12), 12799–12823. doi: 10.1029/92JD01061

Wise, S. (2011). Cross-validation as a means of investigating DEM interpolation error. *Comput. Geosciences* 37 (8), 978–991. doi: 10.1016/j.cageo.2010.12.002

Wu, X., Xu, M., Gao, Y., and Lv, X. (2021). A scheme for estimating time-varying wind stress drag coefficient in the ekman model with adjoint assimilation. *J. Mar. Sci. Eng.* 9 (11), 1220. doi: 10.3390/jmse9111220

Xie, L., Gao, X., Liu, Y., Yang, B., Yuan, H., Li, X., et al. (2022). Atmospheric deposition as a direct source of particulate organic carbon in region coastal surface seawater: Evidence from stable carbon and nitrogen isotope analysis. *Sci. Total Environ.* 854, 14. doi: 10.1016/j.scitotenv.2022.158540

Xu, L., Yang, D., Greenwood, J., Feng, X., Gao, G., Qi, J., et al. (2020). Riverine and oceanic nutrients govern different algal bloom domain near the changjiang estuary in summer. *J. Geophysical Res. Biogeosci.* 125 (10), e2020JG005727. doi: 10.1029/2020JG005727

Yan, X. L., Zheng, H., and Zhao, X. H. (2020). Identification of heavy metal pollution sources and health risk assessment in the northern estuary of the liaodong bay. *J. Environ. Sci.* 40, 3028–3039. doi: 10.13671/j.hjkxxb.2020.0118

Yao, Q., Wang, X., Jian, H., Chen, H., and Yu, Z. (2016). Behavior of suspended particles in the changjiang estuary: Size distribution and trace metal contamination. *Mar. pollut. Bull.* 103 (1-2), 159–167. doi: 10.1016/j.marpolbul.2015.12.026

Zheng, J., Mao, X., Lv, X., and Jiang, W. (2020). The M_2 cotidal chart in the bohai, yellow, and East China seas from dynamically constrained interpolation. *J. Atmospheric Oceanic Technol.* 37 (7), 1219–1229. doi: 10.1175/JTECH-D-19-0212.1

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 13 January 2023
DOI 10.3389/fmars.2022.1084159

Check for updates

OPEN ACCESS

EDITED BY
Yasumasa Miyazawa,
Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Japan

REVIEWED BY
Jicai Zhang,
Zhejiang University, China
Shuhei Masuda,
Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Japan

*CORRESPONDENCE
Wei Li
liwei1978@tju.edu.cn
Guijun Han
guijun_han@tju.edu.cn

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

RECEIVED 30 October 2022
ACCEPTED 19 December 2022
PUBLISHED 13 January 2023

CITATION
Liang K, Li W, Han G, Wang X and Qiu X (2023) An operational improvement of A-4DEnVar and its application to the estimation of the spatially varying bottom friction coefficients of the M_2 constituent in the Bohai and Yellow seas. *Front. Mar. Sci.* 9:1084159. doi: 10.3389/fmars.2022.1084159

COPYRIGHT
© 2023 Liang, Li, Han, Wang and Qiu. 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(s) 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.

An operational improvement of A-4DEnVar and its application to the estimation of the spatially varying bottom friction coefficients of the M_2 constituent in the Bohai and Yellow seas

Kangzhuang Liang[1,2], Wei Li[1,2,3]*, Guijun Han[3]*, Xuan Wang[3] and Xiaobin Qiu[1,2]

[1]Tianjin Key Laboratory for Oceanic Meteorology, Tianjin, China, [2]Tianjin Institute of Meteorological Science, Tianjin, China, [3]School of Marine Science and Technology, Tianjin University, Tianjin, China

The analytical four-dimensional ensemble variational (A-4DEnVar) data assimilation scheme inherits the advantages of the conventional four-dimensional variational (4D-Var) data assimilation scheme and removes the adjoint model. However, compatible operational improvements such as the reduction of the computational costs and the localization method should be considered when it is used in realistic systems. In this paper, the computational complexity of calculating the inverse of background error covariance (the **B** matrix) is decreased by a precondition transform method, i.e., introducing a new state variable whose product with the **B** matrix is the original state variable to be optimized in the cost function. Furthermore, an independent point (IP) scheme is combined to construct an implicit localization method and further decreases the computational cost. Based on the Princeton Ocean Model with the generalized coordinate system (POMgcs), the operational improved A-4DEnVar is applied to optimize the spatially varying bottom friction coefficients (BFCs) of the M_2 constituent in the Bohai and Yellow seas. A twin experiment with idealized observations is designed to validate the effectiveness of the proposed method. In practical experiments, with no more than 10 IPs, the algorithm can assimilate observations from the National Astronomical Observatory (NAO) dataset and obtain a good simulation. The experimental performances increase with the increase of either the IPs or observations, which indicates the efficacy of the proposed algorithm.

KEYWORDS

A-4DEnVar, hybrid, bottom friction coefficients, Bohai and Yellow Seas, localization

1 Introduction

Modern numerical weather prediction (NWP) and/or ocean prediction suffer from the uncertainty of initial conditions and parameters in numerical models. Data assimilation methods are widely implemented to provide better initial fields and parameters by blending numerical models with information extracted from a variety of observations. The four-dimensional variational data assimilation method (4D-Var, Courtier and Talagrand, 1987; Talagrand and Courtier, 1987; Rabier et al., 2000; Stammer et al., 2002; Rawlins et al., 2007 and Wunsch and Heimbach, 2013) and ensemble Kalman filter (EnKF) are two major advanced data assimilation methods (Evensen, 1994; Evensen, 2004). They have been successfully applied in many operational centers, and both of them have great characteristics (Kalnay et al., 2007; Buehner et al., 2010a; Buehner et al., 2010b). As a non-sequential data assimilation method, 4D-Var can assimilate asynchronous observations so that the analysis results from 4D-Var are continuous in a long data assimilation window. Developed from the Kalman filter and extended Kalman filter, EnKF benefits from the estimation of a flow-dependent analysis error covariance. However, the two well-accepted powerful data assimilation schemes are also hampered by their deficiencies. On the one hand, a defection for 4D-Var is its high dependency on the adjoint model (of the forecast model) that usually costs a lot of human resources and is difficult to maintain or transplant to other systems. Furthermore, the lack of flow-dependent background error covariance (denoted as $\boldsymbol{B}$) reduces the performance of 4D-Var. On the other hand, by considering observations sequentially, the EnKF can hardly assimilate all observations spread in a long data assimilation window simultaneously. The discontinuous analysis fields are not consistent with the natural facts.

A suite of cutting-edge hybrid data assimilation methods has been proposed to possess the advantages of 4D-Var and EnKF and avoid their disadvantages. En4DVar and 4DEnVar methods are both outstanding schemes of them (Goodliff et al., 2015; Lorenc et al., 2015; Liu and Xue, 2016). The En4DVar constructs the $\boldsymbol{B}$ matrix as a linear combination of a static part with a flow-dependent part estimated from ensemble members (Buehner, 2005). Since it retains the structure of 4D-Var, the adjoint model is also necessary for the En4DVar scheme. The 4DEnVar is devoted to not only introduce a flow-dependent $\boldsymbol{B}$ matrix into the form of 4D-Var but also remove the adjoint model. To this end, the 4DEnVar utilizes some ensemble perturbation evolutions to approximate the tangent linear model. Liu et al. (2008; 2009) and Liu and Xiao (2013) proposed the primary formulation of the 4DEnVar in which the analysis field is a linear combination of the ensemble mean with observation perturbations. Other studies like Tian et al. (2018) noticed the shortcomings of the 4DEnVar schemes that the ensemble members and thereby the values of the adjoint model are unchanged in iteration loops. To improve the performance, Tian et al. (2018) proposed a series of non-linear least squares En4DVar (NLSi-En4DVar) methods. Especially, the fifth variant (NLS5-En4DVar) is a so-called iterative improvement scheme. NLS5-En4DVar is composed of inner loops and outer loops in which the conventional 4DEnVar and update of ensemble members are implemented, respectively. A multigrid scheme is also developed in Zhang and Tian (2018). The ill-posedness of the inverse problem caused by the increase of the number of variables to be estimated or optimized should be eliminated by reducing the dimensions of variables or introducing useful regularizations. Nowadays, Tian and Zhang (2019a); Tian and Zhang (2019b) studied the generation of ensemble members and expanded the cost function with a regularization (Tian et al., 2020). The most significant superiority of the 4DEnVar schemes beyond others is that it inherits the performance of the conventional 4D-Var and avoids the usage of the adjoint model. As assumed explicitly or implicitly in the above studies, to estimate a reasonable adjoint model, the ensemble perturbations need to be as small as possible. Large perturbations are not compatible with the framework of the 4DEnVar schemes. However, in these above works, the ensemble covariance is set to be the same as the $\boldsymbol{B}$ matrix like EnKF does. To overcome this drawback, Liang et al. (2021) constructed the ensemble covariance by multiplying a small scaling factor (denoted as μ) to the $\boldsymbol{B}$ matrix. It is demonstrated that only if the scaling factor is less than 10^{-6} for the Lorenz-63 model, the value of the adjoint model and the gradients estimated by ensemble members would be close enough to the value obtained from the conventional 4D-Var. Otherwise, the assertion that 4DEnVar is equivalent to the conventional 4D-Var only holds for short data assimilation windows where the non-linearity is limited. Compared to other schemes, A-4DEnVar proposed by Liang et al. (2021) is a more rigorous framework. The superiority of the method has been proven not only in theory but also in ideal twin experiments, especially with long data assimilation windows.

Beyond these advantages, A-4DEnVar lacks of discussions in realistic models. In practice, the dimensions of the state variables in operational systems are often large. For ensemble-based methods, operational improvements and/or modifications are necessary to avoid the high computational cost of the inverse of the $\boldsymbol{B}$ matrix and restrain the spurious correlations caused by insufficient ensemble members. For the first issue, to remove the inverse of the $\boldsymbol{B}$ matrix, the usual process is introducing a new state variable, denoted as $\boldsymbol{\omega}$, for example, and preconditioning the cost function using the transformation $\tilde{\boldsymbol{x}} = \boldsymbol{B}^{1/2}\boldsymbol{\omega}$ (Oddo et al., 2016; Storto et al., 2018) or $\tilde{\boldsymbol{x}} = \boldsymbol{B}\boldsymbol{\omega}$, where $\tilde{\boldsymbol{x}}$ is the state variables or perturbations in the physical space. For the second issue, a straightforward way to eliminate the spurious correlations is to modify the background error covariance matrix through a localization operator or function. Some improvements and variants are also studied. Liu et al. (2009) proposed a reasonable localization approach through multiplying ensemble members by the first M empirical

orthogonal functions (EOFs) of the Schur operator. Although the approach is well accepted in ensemble-based algorithms, the necessary computational cost increases because it increases the ensemble size by M times. Despite the above methods, some simple and convenient strategies are enlightening. One of these methods, the so-called independent point (IP) or feature point (FP) scheme, is proposed in the studies of Zhang and Lu (2008); Zhang and Lu (2010) and Zhang et al. (2011) and is used in Zhang and Wang (2014); Guo et al. (2017), and Qian et al. (2021) to optimize the open-boundary conditions of tidal currents and/or bottom friction coefficients (BFCs) of the tidal constituent. In this scheme, only the values of several selected nodes are optimized, and the value of other rest nodes is calculated by Cressman linear interpolation. As they demonstrated, both the dimension of the parameter space and the ill-posedness of inverse problems are reduced. Apparently, the IPs scheme is easy to implement, especially for the estimation of spatially or temporally distributed parameters. Pan et al. (2017) discussed the IP interpolation methods in detail. While the conventional adjoint method is implemented in the above studies, a natural question is how to integrate this convenient approach into A-4DEnVar and exploit its ability on some realistic operational systems.

We address the above questions in this paper. We not only constructed an operational improvement variant of A-4DEnVar with a precondition transformation but also combined it with the IP scheme and explored its potential for the BFC estimation problems. At the beginning of the paper, the preconditioning transformation is introduced to A-4DEnVar to avoid the inverse of the $\boldsymbol{B}$ matrix. The algorithm is decomposed as inner and outer loops to further improve its ease of implementation. Following the derivation and combining with the IP scheme, we estimate the spatially varying BFCs using the newly improved variant. The estimation is based on the external mode of the Princeton Ocean Model general coordinate system (POMgcs, Ezer and Mellor, 2004), which is, in fact, a two-dimensional shallow water equation. The effectiveness and good performance of A-4DEnVar are validated by a twin experiment with some known and randomly chosen BFCs. A practical experiment in which the sea-level evaluation observations are generated from the NAO (Matsumoto et al., 2000) tidal harmonic constant datasets (https://www.miz.nao.ac.jp/staffs/nao99/data/nao99Jb.tar.gz) is implemented to investigate the distribution of BFCs in the Bohai and Yellow seas.

The remainder of this paper is organized as follows. In section 2, the formulations and derivations of the operational improved variant algorithm are described in detail. Section 3 introduces the numerical model and datasets. The data assimilation experiment setups and their results are shown in section 4. The last section lists the discussions and conclusions of this study.

2 Methodology

2.1 Preconditioning of the cost function

For the conventional 4D-Var scheme, the cost function is defined as

$$J = \frac{1}{2}(\boldsymbol{x}_0 - \boldsymbol{x}_b)^T \boldsymbol{B}^{-1}(\boldsymbol{x}_0 - \boldsymbol{x}_b) + \frac{1}{2}(\boldsymbol{\lambda} - \boldsymbol{\lambda}_b)^T \boldsymbol{C}^{-1}(\boldsymbol{\lambda}_0 - \boldsymbol{\lambda}_b) + \frac{1}{2}\sum_i [\mathscr{H}_i \mathscr{M}_i(\boldsymbol{x}_0, \boldsymbol{\lambda}) - \boldsymbol{y}_i]^T \boldsymbol{R}_i^{-1} [\mathscr{H}_i \mathscr{M}_i(\boldsymbol{x}_0, \boldsymbol{\lambda}) - \boldsymbol{y}_i] \quad (1)$$

where $\boldsymbol{x}_0$ and $\boldsymbol{\lambda}$ are the initial conditions and the parameters in dynamic models. $\boldsymbol{x}_\mathrm{b}$ and $\boldsymbol{B}$ are the background initial conditions and the corresponding background error covariance. Here, *a priori* or background parameter $\boldsymbol{\lambda}_b$ and error covariance $\boldsymbol{C}$ are also included for the combined estimation problem, but they are not always explicit in practice. The dynamic model evolving from t_0 to t_i is denoted as $\mathscr{M}_i =$. The observation operator at time step t_i is $\mathscr{H}_i$. $\boldsymbol{y}_i$ and $\boldsymbol{R}_i$ are the observations and their error covariances, respectively. Although the cost function mentioned above contains the estimation of the initial conditions and parameters simultaneously, it is convenient to reconstruct it to the estimation problems in which only initial condition or parameters are considered. However, to simplify the descriptions below, we consider Equation 1 directly and without any modifications for both initial condition estimation and parameter estimation problems.

The cost function in A-4DEnVar is similar with that in conventional 4D-Var except for the expression of independent variables that are composed of an approximation of the truth (denoted as $\boldsymbol{x}_0^*$ and $\boldsymbol{\lambda}^*$) and their perturbations (denoted as $\tilde{\boldsymbol{x}}_0$ and $\tilde{\boldsymbol{\lambda}}$) to better illustrate an iterative algorithm. That is

$$J = \frac{1}{2}\left(\tilde{\boldsymbol{x}}_0 + \boldsymbol{x}_0^* - \boldsymbol{x}_b\right)^T \boldsymbol{B}^{-1}\left(\tilde{\boldsymbol{x}}_0 + \boldsymbol{x}_0^* - \boldsymbol{x}_b\right) + \frac{1}{2}\left(\tilde{\boldsymbol{\lambda}} + \boldsymbol{\lambda}^* - \boldsymbol{\lambda}_b\right)^T C^{-1}\left(\tilde{\boldsymbol{\lambda}} + \boldsymbol{\lambda}^* - \boldsymbol{\lambda}_b\right) + \frac{1}{2}\sum_i \left[\mathscr{H}_i \mathscr{M}_i\left(\tilde{\boldsymbol{x}}_0 + \boldsymbol{x}_0^*, \tilde{\boldsymbol{\lambda}} + \boldsymbol{\lambda}^*\right) - \boldsymbol{y}_i\right]^T \boldsymbol{R}_i^{-1}\left[\mathscr{H}_i \mathscr{M}_i\left(\tilde{\boldsymbol{x}}_0 + \boldsymbol{x}_0^*, \tilde{\boldsymbol{\lambda}} + \boldsymbol{\lambda}^*\right) - \boldsymbol{y}_i\right] \quad (2)$$

If there are no precondition processes, a straight way to obtain the minimum of the cost function is to calculate its gradient as done in the original A-4DEnVar. In the initial optimization problems, for the fixed $\boldsymbol{x}_0^*$, the gradient of the cost function with respect to perturbations are

$$\nabla J(\tilde{\boldsymbol{x}}_0) = \boldsymbol{B}^{-1}\left(\tilde{\boldsymbol{x}}_0 + \boldsymbol{x}_0^* - \boldsymbol{x}_b\right) + \sum_i \left(\boldsymbol{M}_i|_{\boldsymbol{x}_0^*}\right)^T \boldsymbol{H}_i^T \boldsymbol{R}_i^{-1}\left[\mathscr{H}_i \mathscr{M}_i\left(\boldsymbol{x}_0^* + \tilde{\boldsymbol{x}}_0, \boldsymbol{\lambda}\right) - \boldsymbol{y}_i\right] \quad (3)$$

Due to the large dimensions of the background error covariances, accurate calculations for the inverse matrix $\boldsymbol{B}^{-1}$ require a lot of computational resources. If the localizations are not considered, an elegant way is to estimate them with the covariances of several (often less than 100) ensemble members so that the inverse is, in fact, replaced with the pseudo-inverse

matrix [see the appendix of Liang et al. (2021) for details)]. However, localizations play a significant role in the construction of the background error covariances and are often irreplaceable in operational systems to eliminate the long-distance pseudo-correlations. The elegant way would introduce some unexpected errors into gradient. Similar questions also exist in parameter estimations if the dimension of parameters is large. A convenient way is to remove the inverse matrix in Equations 2 and 3 through a precondition process as described below.

To avoid the calculation of the inverse of $\boldsymbol{B}$, one often preconditions the independent variable using $\tilde{\boldsymbol{x}}_0 = \boldsymbol{B}\boldsymbol{\omega}$. The cost function is hence expressed as

$$J(\boldsymbol{\omega}) = \frac{1}{2}\left(\boldsymbol{B}\boldsymbol{\omega} + \boldsymbol{x}_0^* - \boldsymbol{x}_b\right)^T \boldsymbol{B}^{-1}\left(\boldsymbol{B}\boldsymbol{\omega} + \boldsymbol{x}_0^* - \boldsymbol{x}_b\right) + \frac{1}{2}(\boldsymbol{\lambda} - \boldsymbol{\lambda}_b)^T \boldsymbol{C}^{-1}(\boldsymbol{\lambda}_0 - \boldsymbol{\lambda}_b) + \frac{1}{2}\sum_i \left[\mathscr{H}_i \mathscr{M}_i\left(\boldsymbol{x}_0^* + \boldsymbol{B}\boldsymbol{\omega}, \boldsymbol{\lambda}\right) - \boldsymbol{y}_i\right]^T \boldsymbol{R}_i^{-1}\left[\mathscr{H}_i \mathscr{M}_i\left(\boldsymbol{x}_0^* + \boldsymbol{B}\boldsymbol{\omega}, \boldsymbol{\lambda}\right) - \boldsymbol{y}_i\right] \tag{4}$$

In fact, the terms corresponding to the parameters are treated as constants in the initial condition estimation problem. The gradient of the cost function is

$$\nabla J(\boldsymbol{\omega}) = \boldsymbol{B}\boldsymbol{\omega} - \left(\boldsymbol{x}_b - \boldsymbol{x}_0^*\right) + \sum_i \left(B\boldsymbol{M}_i\big|_{\boldsymbol{x}_0^*}\right)^T H_i^T R_i^{-1}\left[\mathscr{H}_i \mathscr{M}_i\left(\boldsymbol{x}_0^* + \boldsymbol{B}\boldsymbol{\omega}, \boldsymbol{\lambda}\right) - \boldsymbol{y}_i\right] \tag{5}$$

where $\boldsymbol{H}_i$ denotes the tangent linear model of the observation operator $\mathscr{H}_i$. It should be noted that the value of tangent linear model $\boldsymbol{M}_i\big|_{\boldsymbol{x}_0^*}$ and, hence, the structure of cost function depends on $\boldsymbol{x}_0^*$, the prior approximation of the truth. Thus, to achieve the minimization of the cost function in Equation 1, $\boldsymbol{x}_0^*$ should be updated iteratively.

Since $\tilde{\boldsymbol{x}}_0$ is a small perturbation, the Taylor expansions of the dynamic model operators indicate that

$$\mathscr{M}_i\left(\boldsymbol{x}_0^* + \tilde{\boldsymbol{x}}\right) \approx \mathscr{M}_i\left(\boldsymbol{x}_0^*\right) + \boldsymbol{M}_i\big|_{\boldsymbol{x}_0^*}\tilde{\boldsymbol{x}}_0 \tag{6}$$

or it can be further simplified as

$$\tilde{\boldsymbol{x}}_i = \boldsymbol{M}_i\big|_{\boldsymbol{x}_0^*}\tilde{\boldsymbol{x}}_0 \tag{7}$$

in which $\boldsymbol{M}_i\big|_{\boldsymbol{x}_0^*}$ is the tangent linear model of $\mathscr{M}_i$. Equation 7 provides an efficient way to estimate the tangent linear model. The processes are similar to what was mentioned in Liang et al. (2021). Once there are N sufficient samples of $\tilde{\boldsymbol{x}}_0$ and $\tilde{\boldsymbol{x}}_i$, denoted as $\tilde{\boldsymbol{X}}_0 = (\tilde{\boldsymbol{x}}_0^1, \tilde{\boldsymbol{x}}_0^2, \cdots, \tilde{\boldsymbol{x}}_0^N)$and $\tilde{\boldsymbol{X}}_i = (\tilde{\boldsymbol{x}}_i^1, \tilde{\boldsymbol{x}}_i^2, \cdots, \tilde{\boldsymbol{x}}_i^N)$, respectively, the tangent linear model satisfies

$$\tilde{\boldsymbol{X}}_i = \boldsymbol{M}_i\big|_{\boldsymbol{x}_0^*}\tilde{\boldsymbol{x}}_0 \tag{8}$$

The least squares estimation of the tangent linear model is

$$\hat{\boldsymbol{M}}_i\big|_{\boldsymbol{x}_0^*} = \left[\frac{1}{N}\tilde{\boldsymbol{X}}_0\tilde{\boldsymbol{X}}_0^T\right]^{-1}\frac{1}{N}\tilde{\boldsymbol{X}}_i\tilde{\boldsymbol{X}}_0^T \tag{9}$$

if the dimension of state variables is not large and $\tilde{\boldsymbol{X}}_0\tilde{\boldsymbol{X}}_0^T$ is full rank. It is clear that for different $\tilde{\boldsymbol{X}}_0$ (although the $\tilde{\boldsymbol{X}}_i$ is different in this situation), Equations 8 and 9 are satisfied. Thus, the generation of ensemble does not influence the estimation of the tangent linear models and adjoint models. Otherwise,

$$\hat{\boldsymbol{M}}_i\big|_{\boldsymbol{x}_0^*} = \left[\frac{1}{N}\tilde{\boldsymbol{X}}_0^T\tilde{\boldsymbol{X}}_0\right]^{-1}\frac{1}{N}\tilde{\boldsymbol{X}}_0^T\tilde{\boldsymbol{X}}_i = \tilde{\boldsymbol{X}}_0^+\tilde{\boldsymbol{X}}_i \tag{10}$$

where a pseudo-inverse of $\tilde{\boldsymbol{X}}_0$ denoted as $\tilde{\boldsymbol{X}}_0^+ = [\tilde{\boldsymbol{X}}_0^T\tilde{\boldsymbol{X}}_0]^{-1}\tilde{\boldsymbol{X}}_0^T$ is adopted. In this situation, the estimations of tangent linear models are different with the difference of space generated by ensemble members. The widely used EOFs should be used to construct reasonable ensembles. However, due to the fact that the IP scheme is adopted in this paper, this situation is avoided actually. Substituting the $\hat{\boldsymbol{M}}_i\big|_{\boldsymbol{x}_0^*}$ into Equation 5 yields the estimation of the gradient

$$\nabla J(\boldsymbol{\omega}) = \boldsymbol{B}\boldsymbol{\omega} - \left(\boldsymbol{x}_b - \boldsymbol{x}_0^*\right) + \sum_i \left(B\hat{\boldsymbol{M}}_i\big|_{\boldsymbol{x}_0^*}\right)^T H_i^T R_i^{-1}\left[\mathscr{H}_i \mathscr{M}_i\left(\boldsymbol{x}_0^* + \boldsymbol{B}\boldsymbol{\omega}, \boldsymbol{\lambda}\right) - \boldsymbol{y}_i\right] \tag{11}$$

Equation 11 is the gradient of cost functions with respect to the precondition variables in the initial condition optimization problems. The construction details of $B\hat{\boldsymbol{M}}_i\big|_{\boldsymbol{x}_0^*}$ are related to the generation of the ensemble members, which will be discussed in the next subsections.

For the parameter estimation, the processes and derivations are similar as what was derived in the initial condition estimation except for some details. Set parameters are composed of prior components and perturbation components. That is

$$\boldsymbol{\lambda} = \boldsymbol{\lambda}^* + \tilde{\boldsymbol{\lambda}} \tag{12}$$

and, again, assume that

$$\tilde{\boldsymbol{\lambda}} = C\boldsymbol{\omega} \tag{13}$$

The parameter perturbations have

$$\mathscr{M}_i\left(\boldsymbol{x}_0, \boldsymbol{\lambda}^* + \tilde{\boldsymbol{\lambda}}\right) \approx \mathscr{M}_i\left(\boldsymbol{x}_0, \boldsymbol{\lambda}^*\right) + \boldsymbol{M}_i\big|_{\boldsymbol{\lambda}^*}\tilde{\boldsymbol{\lambda}} \tag{14}$$

where the tangent linear model is about the parameter variables λ^*.

Let

$\tilde{\boldsymbol{\lambda}} = (\tilde{\boldsymbol{\lambda}}^1, \tilde{\boldsymbol{\lambda}}^2, \cdots, \tilde{\boldsymbol{\lambda}}^N)$ be a matrix with N samples of $\tilde{\boldsymbol{\lambda}}$ as its columns and their corresponding state variable perturbations also be composed of a matrix named as $\tilde{\boldsymbol{X}}_i$. Then, the estimation of

$M_i|_{\lambda^*}$ is

$$\hat{\boldsymbol{M}}_i\big|_{\boldsymbol{\lambda}^*} = \left(\frac{1}{N}\tilde{\boldsymbol{\lambda}}\tilde{\boldsymbol{\lambda}}^T\right)^{-1}\frac{1}{N}\tilde{\boldsymbol{X}}_i\tilde{\boldsymbol{\lambda}}^T \tag{15}$$

when the $\tilde{\boldsymbol{\lambda}}\tilde{\boldsymbol{\lambda}}^T$ is full rank. Otherwise, the estimation is

$$\hat{\boldsymbol{M}}_i\big|_{\boldsymbol{\lambda}^*} = \left[\frac{1}{N}\tilde{\boldsymbol{\lambda}}^T\tilde{\boldsymbol{\lambda}}\right]^{-1}\frac{1}{N}\tilde{\boldsymbol{\lambda}}^T\tilde{\boldsymbol{X}}_i = \tilde{\boldsymbol{\lambda}}^+\tilde{\boldsymbol{X}}_i \tag{16}$$

where $\tilde{\boldsymbol{\lambda}}^+$ is the pseudo-inverse of $\tilde{\boldsymbol{\lambda}}$.

The cost function for parameter estimation is expressed as

$$J(\omega) = \frac{1}{2}(\boldsymbol{x}_0 - \boldsymbol{x}_b)^T B^{-1}(\boldsymbol{x}_0 - \boldsymbol{x}_b) + \frac{1}{2}\left(\boldsymbol{\lambda}^* + C\omega - \boldsymbol{\lambda}_b\right)^T C^{-1}\left(\boldsymbol{\lambda}^* + C\omega - \boldsymbol{\lambda}_b\right) + \frac{1}{2}\sum_i \left[\mathscr{H}_i \mathscr{M}_i(\boldsymbol{x}_0, \boldsymbol{\lambda}^* + C\omega) - \boldsymbol{y}_i\right]^T R_i^{-1}\left[\mathscr{H}_i \mathscr{M}_i(\boldsymbol{x}_0, \boldsymbol{\lambda}^* + C\omega) - \boldsymbol{y}_i\right] \tag{17}$$

The gradient with respect to $\boldsymbol{\omega}$ is

$$\nabla J(\omega) = C\omega - (\boldsymbol{\lambda}_b - \boldsymbol{\lambda}^*) + \sum_i \left(C\hat{\boldsymbol{M}}_i|_{\boldsymbol{\lambda}^*}\right)^T H_i^T R_i^{-1} R_i^{-1} \left[\mathscr{H}_i \mathscr{M}_i(\boldsymbol{x}_0, \boldsymbol{\lambda}^* + C\omega) - y_i\right] \quad (18)$$

Again, the generation of ensemble members will be discussed in the next sections.

Since the gradients are obtained, combining with a suitable optimization method [e.g., the linear research gradient descent method, the Newton descent method, and the limited-memory quasi-Newton method (L-BFGS) method], the minimization of the cost function can be calculated.

2.2 Generation of the ensemble perturbations

In subsection 2.1, we presented the framework of A-4DEnVar with precondition variables. The remaining question is how to generate reasonable ensemble perturbations in practice. To ensure an accurate estimation of tangent linear models in Equations 10 and 16, the basic assumption is that the magnitudes of perturbations should be much smaller than that of independent variables. It must be mentioned that the errors generated by a probability density distribution with the covariance $\boldsymbol{B}$ (for initial estimation) or $\boldsymbol{C}$. (for parameter estimation) are usually not available. However, for initial conditions, it is presented above

$$B\hat{\boldsymbol{M}}_i|_{\boldsymbol{x}_0^\star} = B\left[\frac{1}{N}\tilde{\boldsymbol{X}}_0\tilde{\boldsymbol{X}}_0^T\right]^{-1}\frac{1}{N}\tilde{\boldsymbol{X}}_i\tilde{\boldsymbol{X}}_0^T \quad (19)$$

or

$$B\hat{\boldsymbol{M}}_i|_{\boldsymbol{x}_0^\star} = \mathbf{B}\tilde{\mathbf{X}}_\mathbf{0}^+\tilde{\mathbf{X}}_\mathbf{i} \quad (20)$$

and, for parameters,

$$C\hat{\boldsymbol{M}}_i|_{\boldsymbol{\lambda}^\star} = C\left(\frac{1}{N}\tilde{\boldsymbol{\lambda}}\tilde{\boldsymbol{\lambda}}^T\right)^{-1}\frac{1}{N}\tilde{\boldsymbol{X}}_i\tilde{\boldsymbol{\lambda}}^T \quad (21)$$

or

$$C\hat{\boldsymbol{M}}_i|_{\boldsymbol{\lambda}^\star} = C\tilde{\boldsymbol{\lambda}}^+\tilde{\boldsymbol{X}}_i \quad (22)$$

If setting the ensemble perturbation covariances to be

$$\frac{1}{N}\tilde{\boldsymbol{X}}_0\tilde{\boldsymbol{X}}_i^T \approx \mu B \quad (23)$$

and

$$\frac{1}{N}\boldsymbol{\lambda}\boldsymbol{\lambda}^T \approx \mu C \quad (24)$$

The inverse terms in Equations 19–21 are eliminated. For initial condition estimation, the gradient is

$$\nabla J = \boldsymbol{B}\boldsymbol{\omega} - \left(\boldsymbol{x}_b - \boldsymbol{x}_0^\star\right) + \sum_i \frac{1}{\mu N}\tilde{\boldsymbol{X}}_0\tilde{\boldsymbol{X}}_i^T H_i^T R_i^{-1}\left[\mathscr{H}_i\mathscr{M}_i\left(\boldsymbol{x}_0^\star + \boldsymbol{B}\boldsymbol{\omega}, \boldsymbol{\lambda}\right) - y_i\right] \quad (25)$$

For parameter estimation, the gradient is

$$\nabla J = C\omega - \left(\boldsymbol{\lambda}_b - \boldsymbol{\lambda}^*\right) + \sum_i \frac{1}{\mu N}\tilde{\boldsymbol{\lambda}}\tilde{\boldsymbol{X}}_i^T H_i^T R_i^{-1} R_i^{-1}\left[\mathscr{H}_i\mathscr{M}_i(\boldsymbol{x}_0, \boldsymbol{\lambda}^* + C\omega) - y_i\right] \quad (26)$$

Note that the gradients in Equations 25 and 26 are calculated through ensemble perturbations under the assumption that $\boldsymbol{x}_0^\star$ and $\lambda^\star$ are fixed. Thus, once the convergence of prediction variable $\boldsymbol{\omega}$ is achieved, $\boldsymbol{x}_0^\star$ and $\lambda^\star$ should be updated. The complete algorithm is composed of two parts, i.e., the iterations of $\boldsymbol{\omega}$ and the updates of $\boldsymbol{x}_0^\star$ and/or $\lambda^\star$. A further development is to leave the iterations into outer loops in which several inner loops are included. The steps for the operational improved A-4DEnVar are described below in Figure 1.

In the outer loops, the ensemble perturbations are generated based on the $\mu\boldsymbol{B}$ (and/or $\mu\boldsymbol{C}$), and then, they are added to the approximation state variables to obtain the initial conditions or parameters of the ensemble members. The ensemble initial field and parameters are reconstructed in the outer loops of each cycle. The main computational costs for the integrations of ensemble members are only needed in the outer loops,

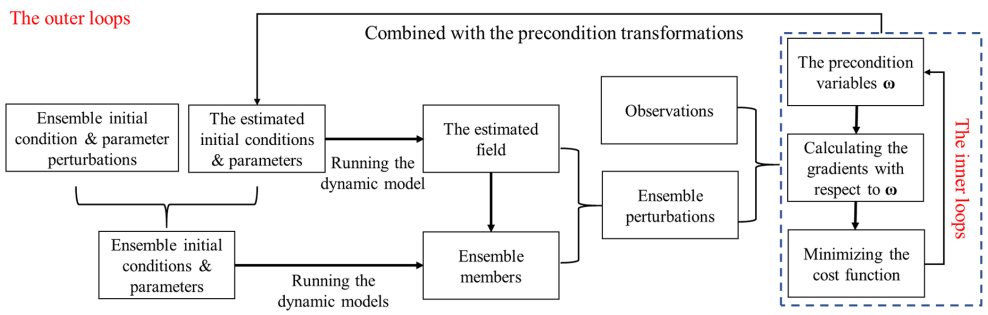


FIGURE 1
The flowchart of the operational improved analytical four-dimensional ensemble variational (A-4DEnVar).

whereas the inner loops focus on the iteration of precondition state variables.

2.3 The independent point scheme

The basic assumption of IP scheme is that the features or characters of the variables to be optimized (such as spatially distributed BFCs) can be determined by the variables located at some key points (Wang et al., 2021). In data assimilation, only these key variables are optimized directly, whereas other general variables are optimized by using interpolation methods (such as Cressman method and uniform or nonuniform spline interpolation methods) based on these key variables. The IP scheme restricts estimation problems from the variable space to the IP space. In practice, the key points are often chosen to be far away from each other so that they are considered to be independent. The background error covariance in the IP space is thus simplified as a diagonal matrix. Meanwhile, the interpolation coefficients are determined by the distances between points; hence, the adjacent points have similar values, obviously. It can be seen that the IP scheme not only reduces the dimension complexity but also introduces an implicit localization in data assimilation schemes.

If the Cressman method (Cressman, 1959) is adopted, it can be described as

$$F_g = \sum_{i=1}^{N} W_{gi} F_i \tag{27}$$

$$W_{gi} = \begin{cases} \frac{\left(R^2 - d_{gi}^2\right)}{\left(R^2 + d_{gi}^2\right)} & d_{gi} < R \\ 0 & d_{gi} > R \end{cases} \tag{28}$$

where F_g denotes the variable located at a general point, F_i is the variable located at a key point and/or IP, and W_{gi} is the Cressman coefficient that is determined by the distance between the general point and the IP i.e., d_{gi}, and a predefined parameter R.

3 The numerical model

3.1 Governing equations of forward model

Here, the external mode of the Princeton Ocean Model (POM, Blumberg and Mellor, 1987) is used to simulate tides and tidal currents. The attributes of the POM model include the sigma-coordinate system, curvilinear orthogonal coordinates with an Arakawa C differencing scheme in the horizontal grid, explicit horizontal time differencing, and implicit vertical time differencing. The free surface and split time step method is adopted in which the two-dimensional external mode uses a short time step for a high-accuracy surface simulation and the three-dimensional internal mode uses a long-time step to reduce the computational cost. Based on the framework of POM, the POMgcs is extended to a mixed coordinate system to make the z-coordinate and sigma-coordinate compatible.

After the variables are integrated in the original three-dimensional model from the bottom to the top, the governing equations in the external mode (in z-coordinate) for BFC estimation is

$$\begin{cases} \frac{\partial \eta}{\partial t} + \nabla_h \cdot \left[(\eta + H) \vec{V}_h \right] = 0 \\ \frac{\partial (\eta + H) \vec{V}_h}{\partial t} + \nabla_h \cdot \left[(\eta + H) \vec{V}_h \vec{V}_h \right] + \nabla_h \cdot \int_{-H}^{\eta} \left(\vec{v}'_h \vec{v}'_h \right) dz + f(\eta + H) \cdot \hat{k} \times \vec{V}_h = \\ -g(\eta + H) \nabla_h (\eta - \bar{\eta}) - \frac{1}{\rho_0} \nabla_h \int_{-H}^{\eta} \left(\int_z^{\eta} \tilde{\rho} g dz \right) dz + \frac{1}{\rho_0} \left(\int_z^{\eta} \tilde{\rho} g dz \right) | z = -H \nabla_h H \\ - \vec{\tau}_{top} + \vec{\tau}_{bottom} \end{cases} \tag{29}$$

where $\nabla_h = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right)$ is a Hamilton operator in the horizontal directions. The meanings of other symbols are listed in Table 1.

The bottom friction is calculated by

$$\vec{\tau}_{bottom} = \left(\tau_{bx}, \tau_{by} \right) = \rho_0 C_d \left| \vec{V}_h \right| \vec{V}_h \tag{30}$$

where C_d is the BFCs to be optimized.

3.2 Model setting

The model domain is the Bohai and Yellow seas, which covers from 35°N to 41°N in latitudes and 117.5°E to 127.25°E in

TABLE 1 The symbols in the external mode of the Princeton Ocean Model with the generalized coordinate system.

Symbols	Variables	Units
t	Time	s
η	Sea surface height	m
H	Depth	m
$\vec{V}_h$	Average current velocity	$m\ s^{-1}$
$\vec{v}'_h$	Current perturbation	$m\ s^{-1}$
f	Coriolis parameter	s^{-1}
g	Gravity	$m\ s^{-2}$
ρ_o	Average density	$kg\ m^{-3}$
$\tilde{\rho}$	Density perturbations	$kg\ m^{-3}$
$\vec{\tau}_{top}$	Surface friction	$m^2\ s^{-2}$
$\vec{\tau}_{bottom}$	Bottom friction	$m^2\ s^{-2}$

longitudes. The gridded topography is interpolated from the ETOPO5 datasets. It is shown (in Figure 2) that the depths are all less than 85 m, such that the tide is a basic and major motion in this area. The spatial resolutions are all 1/12° in the east and north directions. In POMgcs, the integration time step is 4 s, which is much enough to simulate the M_2 constituent.

A static initial condition with zero sea surface height and current velocity is used to spin up the model. Due to the fact that the external mode of POMgcs (a shallow water model) is adopted, the temperature and salinity in initial and open-boundary conditions are constant. The open-boundary conditions including the tide (sea-level height evolution) and tidal currents. Both of them are fixed in the BFC optimization experiments. The open-boundary tide sea-level heights are calculated through the tidal harmonic constants from the same datasets that used to generate the observations, i.e., the NAO datasets. However, because the NAO datasets do not provide the harmonic constants of tidal currents, we have to generate tidal currents from the harmonic constants of TPXO7 (Egbert and Erofeeva, 2002) datasets. To eliminate the misfits of tide and tidal currents, we assimilated the same observations with a fixed-background BFC (that is 0.0025 as used in below experiments) value to provide the optimized open-boundary tidal currents before the inversion of BFCs. The data assimilation method we used is also the operational improved A-4DEnVar. The fixed value is consistent with or close to what was used in other studies. In Yao et al. (2012), the open-boundary conditions of the M_2 tidal constituent in the Bohai and the Yellow seas are optimized with the BFCs 0.003. Fan et al. (2019) demonstrated that the value of BFCs in the East China Shelf Seas varies from 0.001 to 0.003. Wang et al. (2021) set the initial BFCs to 0.002 in their experiments. Following these studies, we think that 0.0025 is a reasonable and acceptable value in our study. Moreover, the forcing terms from the sea surface such as wind stress, sensible and/or latent heat flux, and precipitation are not included. The Coriolis parameter takes the local value. The model ran from the 1st to the 7th of January in 2000 in which the first 6 days were used to spin up. All observations are the sea-level heights of the M_2 tidal constituent. After that, data assimilation based on A-4DenVar is implemented.

3.3 The locations of IPs and observation stations

The factors that influence the BFC coefficients include topography, water depth, and sediments. Regardless of the attributes of these factors in detail, the distributions of BFCs ought to be continuous and smooth in most regions. The trends and/or characters of the distributions are able to be described through several control points (namely, IPs mentioned above) as indicated in studies. We followed these studies and conducted a twin experiment (denoted as EXP 1) with 10 IPs and a suite of practical experiments (denoted as EXP 2 to EXP 5) with 1, 4, 7, and 10 IPs, respectively. With 10 IPs, we also tested the impacts of observations on the data assimilation performances (EXP 1 and 5). In the last suite of experiments (EXP 6 and EXP 7), density IP stations evenly distributed on an 1°×1° grids are used so that smaller Cressman parameters would be acceptable.

As the main purpose of this paper is to show an operational implementation for A-4DEnVar, we do not discuss how to design a very strict optimal solution for IP locations. Empirically, the selection principle of IPs is that they should be distributed evenly and cover the entire simulation area as much as possible. In EXP 1 to EXP 5, the locations of these points are artificial and presented in Figure 2 (marked with yellow crosses). The observations are calculated from either a

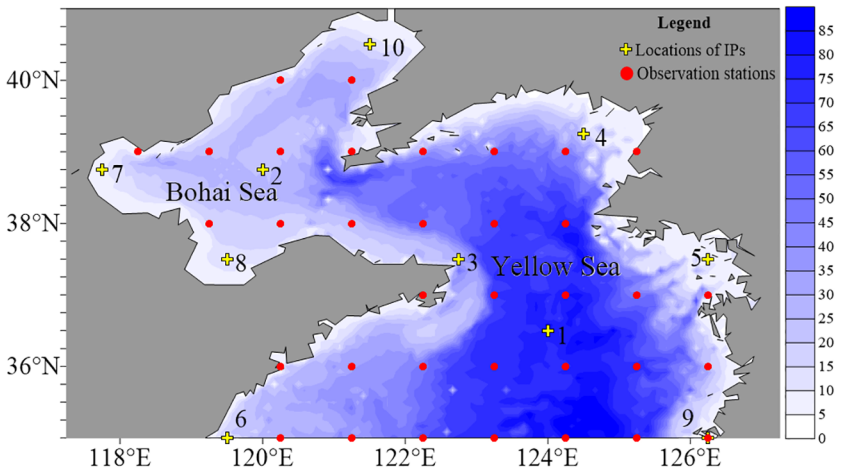


FIGURE 2
The locations of independent points (IPs; marked with yellow crosses) and the observation stations in EXP 1 to EXP 5 (marked with red dots) for bottom friction coefficient (BFC) optimization.

twin model (in EXP 1) or the tidal constants of the NAO model (in EXP 2 to EXP 5). The observations are located on stations that gradually increase from 1°×1° (for both the EXP 1 and EXP 2), 1/2°×1/2°,1/4°×1/4° and 1/12°×1/12° in EXP 1 and 5 (the station locations are not shown here). In EXP 6 and EXP 7, the observational stations are only on the 1/12°×1/12° grids. Furthermore, at each station, the simulated observation lasts for 1 day at a frequency of 1 h. The Cressman parameter is set to 5° and fixed so that the whole simulation domain is covered even with only one IP in EXP 1 to EXP 5, whereas it is set to be 2.5° and 1.5° in EXP 6 and EXP 7, respectively.

4 Numerical experiments

4.1 The twin experiments (EXP 1)

The twin experiments are conducted to evaluate the feasibility of A-4DEnVar, and the implementations are expressed in the following steps. First, a suite of random values varying from 1×10^{-4} to 3×10^{-3} are assigned to the BFCs of the 10 IPs. Second, with these values, the BFCs of the general locations are calculated by the Cressman interpolation method. Starting from the static initial condition and open-boundary conditions mentioned in the above sections, the external mode is integrated for 7 days. In the last 24 h, observations without noise are provided from the simulated "true" sea-level heights at the observation stations. Finally, the operational improved A-4DEnVar is implemented for data assimilation.

Due to the fact that the perfect observations are assimilated in the twin experiment, the optimizations should trust the observations as much as possible. To this end, the algorithm starts from an initial iteration BFC value that is 0.0025 and the background error covariance terms of the cost function are not included. The ensemble size is set to be 10, which is equal to the dimension of state variables, and the factor μ is fixed to 10^{-6}. The stopping criteria for inner loops in each outer loop is important to ensure the coverages of the algorithm. However, the ensemble members and hence the value of tangent linear models and the adjoint models are only updated in the outer loop; it is not necessary to achieve a very exact global minimum in every inner loop where the value of the adjoint models are fixed. Given the balance between effectiveness and the computational cost, empirically, it is convenient to limit the inner loops less than five times in this paper. The outer loop iterations are terminated once the value of the cost function does not decrease significantly.

Figure 3 shows the changes of the cost function value with inner iterations when 1°×1° observational grids are employed. Here, the zero IPs or the zero iterations mean the result without data assimilation. It is clear that, with the iterations, the value of the cost function decreases continually. The experiment results indicate that the model after data assimilation is indeed

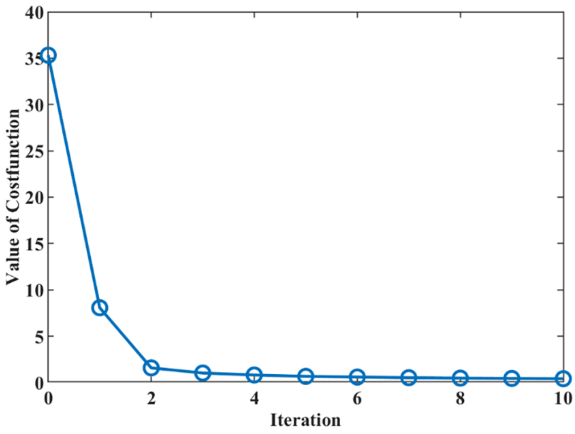


FIGURE 3
The cost function value decreases with iterations in the twin experiment with perfect observations on the 1°×1° grids.

restrained by the observations. With only one outer loop containing five inner loops, the cost function decreases by more than 77% of the initial cost function value. After 10 iterations, the algorithm converges to a local minimum of the cost function. At the end of the iterations, the cost function value is only 1.2% of the initial cost function value and is much close to zero. However, because the influences of BFCs to the evolutions of sea-level heights are complicated, it is hard to achieve the global minimum for any gradient-based optimization algorithm.

The distributions of the BFCs are presented in Figure 4 of which Figure 4A is the truth, Figures 4B to E are the experiment results with increasing observational grids, i.e., 1°×1°, 1/2°×1/2°, 1/4°×1/4° and 1/12°×1/12°, respectively, and Figure 4F shows the difference between the results in Figures 4E and 4A. These patterns from EXP 1 are much similar with each other and the experiment result performance is slightly better with the increases of observations. With either sparse or density observations, the proposed algorithm converges to a local minimum of the cost function. Compared with the true BFCs, all the absolute values of these misfits are under 4.5×10^{-4} which is much less than the values of BFCs. Specifically, in most areas, the BFC values from experiment results are very close to the truth but are larger in the west of the Yellow seas, near the open boundary, and are smaller in the north of Bohai seas.

4.2 The practical experiments with different IPs (EXP 2 to EXP 5)

As mentioned above, we designed four experiments (denoted as EXP 2 to EXP 5) with 1, 4, 7, and 10 IPs, respectively, to estimate the BFCs in Bohai and Yellow seas. The settings of practical experiments are similar to that of the twin experiment except for the number of IPs and the generation

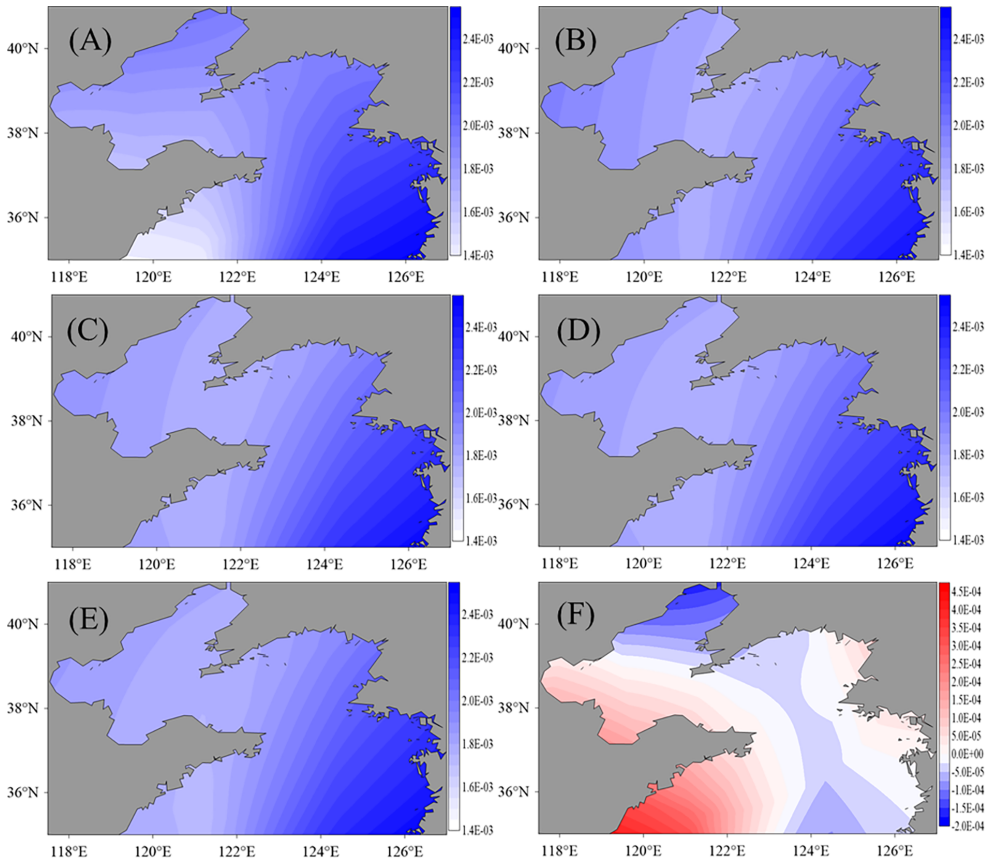


FIGURE 4
Distributions of the true BFCs **(A)** and that of experiment results with the 1°×1° **(B)**, 1/2°×1/2° **(C)**, 1/4°×1/4° **(D)**, and 1/12°×1/12° **(E)** observational grids. **(F)**is the difference between **(E)** and **(A)**.

of observations. The locations of these IPs are the same as what was mentioned in the twin experiments, whereas the observations are calculated from the harmonic constants of the M_2 tidal constituent derived from the NAO tide model dataset on their corresponding observation stations.

A great difference between the twin experiment and the practical experiments is that not only the uncertainty from BFCs but also that from other model parameters should be considered simultaneously. However, how to distinguish these uncertainties is complex and beyond the scope of this paper. Instead of discussing them in detail, we assume that these uncertainties are reflected through the observations and the observational errors are Gaussian white noise with a standard deviation of 0.1 m. In the cost function, the background values of BFCs are 0.0025 with a standard deviation of 0.001 (dimensionless). The iteration processes in the practical experiments are set to be the same with that in the twin experiments.

For the experiment with 1°×1° observational grids, Figure 5A shows the impact of the number of IPs on the converged cost function values in EXP 2 to EXP 5. What should be mentioned is that, influenced by other uncertainties in the dynamic model, the cost function values at the beginning and the end of the iterations are both larger than that in the twin experiments. The cost function value decreases by approximately 14% when only one IP is used. The trend that the minimum values of the cost function obtained by A-4DEnVar decreases with the increase of the IPs indicates that the distributions of BFCs are better described with more IPs. When the number of IPs is 7 (in EXP 4), the optimization effect tends to be stable. After that, increasing the number of IPs would not significantly improve the performance. Figure 5B presents the cost function values changes with iterations while 10 IPs are used. Like what we see in the twin experiment, the cost function values also decrease rapidly. In the first five inner iterations, the algorithm almost converges with the fixed ensemble members. After that, once the ensemble members are updated in the outer loop, cost function values decrease a rather big step and then converges again in the inner loops. When 10 inner iterations are implemented, cost function values decrease by approximately 29.6%.

The amplitudes and phases of M_2 tidal constituents are also calculated by the harmonic analysis process. In Figure 6, the cotidal chart from the study of Wang et al. (2021, Figure 6A) and

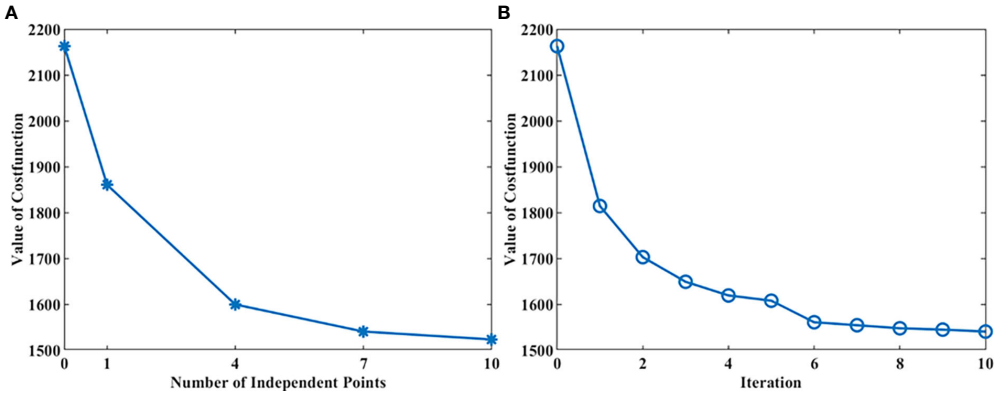


FIGURE 5
The left **(A)** shows the impact of the number of IPs on the cost function values when EXP 2 to EXP 5 converged. The right **(B)** shows the impact of the iterations on the cost function values when 10 IPs are used in EXP. 4.

the NAO dataset (Figure 6B) are shown, respectively. Compared to them, it is clear that three amphidromic points in the simulated area and the results we obtained are consistent with those from other studies.

The cotidal chart differences between the control model results ("CTL" in figures) before data assimilation and that from the NAO model are shown in Figure 7A (for the amplitudes denoted as "H") and (B) (for the phases denoted as "g"). Figures 7C, D are similar with those in Figures 7A, B but for the results after data assimilation ("DA" in figures). Comparing Figures 7A with C, it is shown that the amplitudes in the north of the Yellow Sea and the middle of the Bohai Sea are much

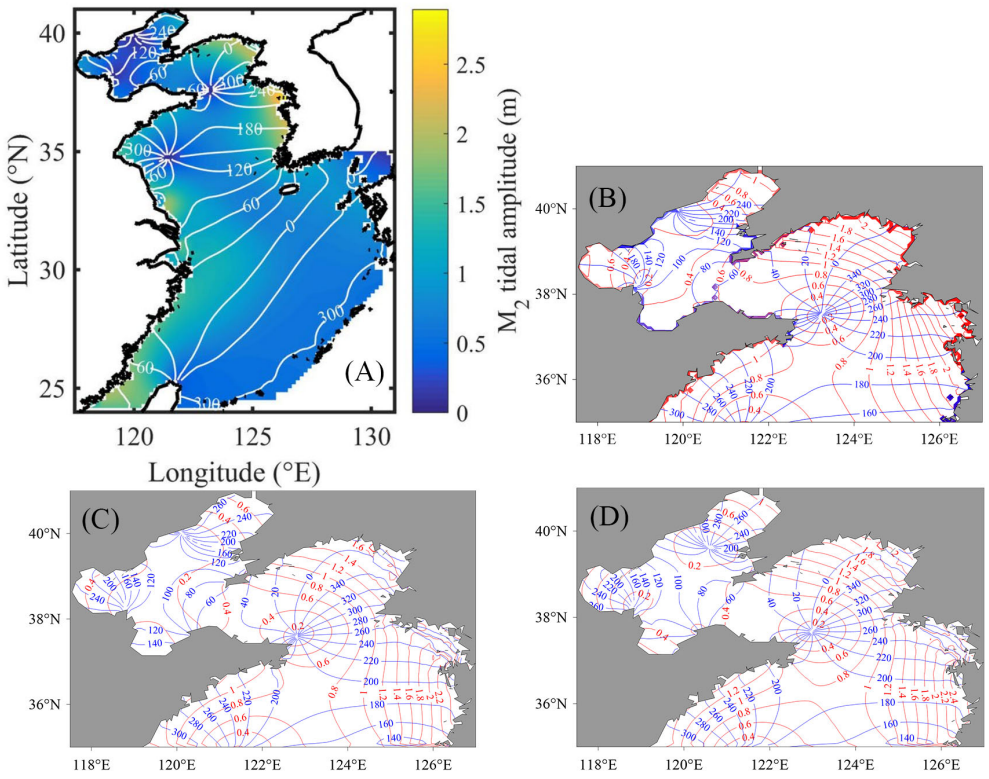


FIGURE 6
(A, B) are the cotidal charts from Wang et al. (2021) and the National Astronomical Observatory (NAO) data, respectively. **(C, D)** are the cotidal charts before and after data assimilation.

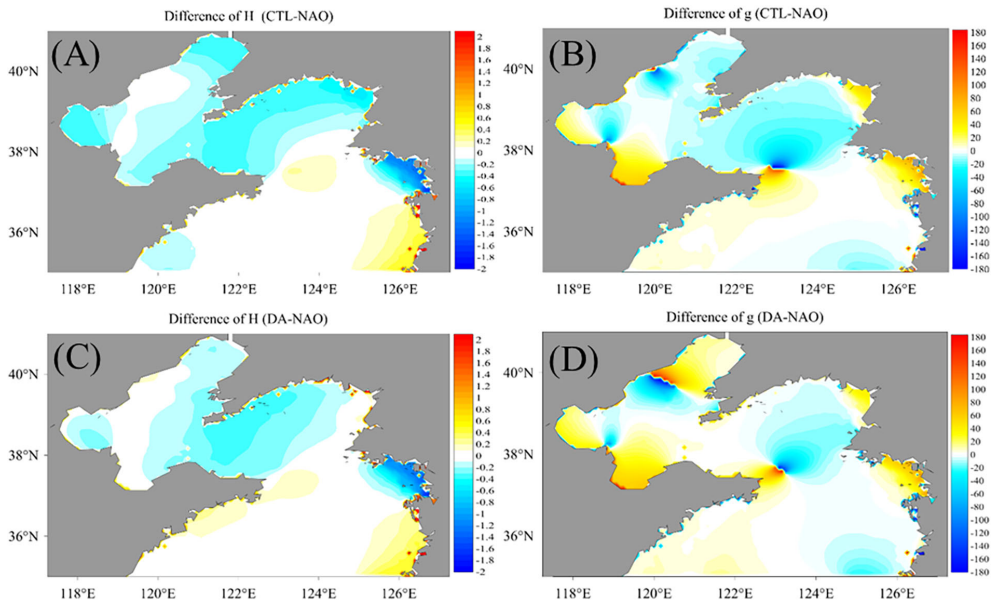


FIGURE 7
(A, B) are the differences of amplitudes and phases between the control model and NAO datasets, respectively. **(C, D)** are similar with **(A, B)** but for the differences between data assimilation results and NAO datasets.

improved. For the phases, great improvements are presented in the north of the Yellow seas around the amphidromic point near (123.5°*E*, 37.5°*N*) and the middle of the Bohai Sea. The misfits around (120°*E*, 40°*N*) in the north of the Bohai Sea are caused by the different locations of the amphidromic points from the results after data assimilation and NAO datasets. Considering the amplitudes in these areas are almost zeros, although the phases are not close enough to that from the NAO datasets, the sea-level evolutions are, in fact, much closer.

Specifically, Figure 8 shows the data assimilation results in EXP 5 with 10 IPs and different observation densities. Again, consistent with the results in the twin experiment in the above subsections, the patterns are much similar with each other.

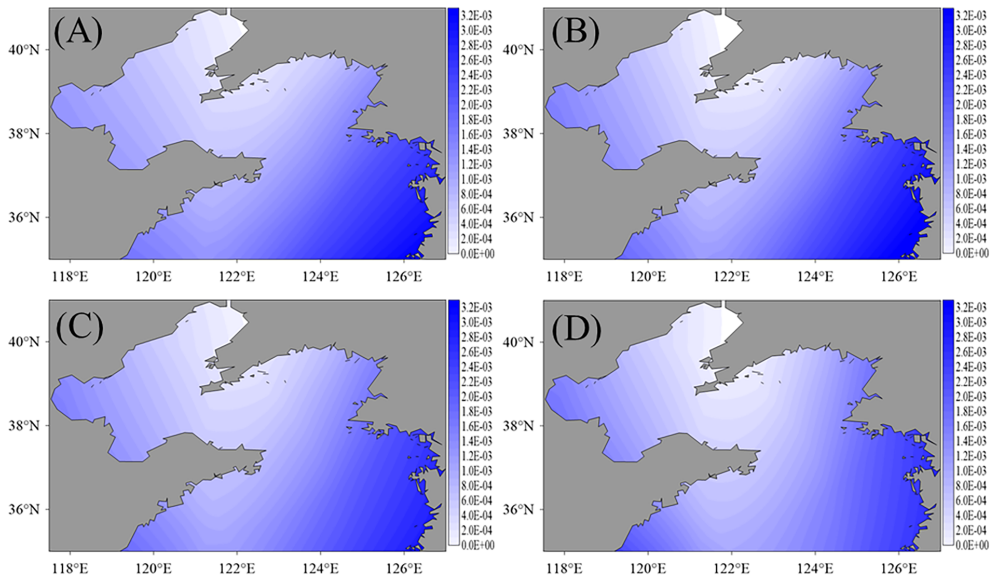


FIGURE 8
The BFC values after data assimilation from EXP 2 **(A)**, EXP 3 **(B)**, EXP 4 **(C)** and EXP 5 **(D)** with different observational grids.

The distributions of BFCs obtained here are much smooth. This is mainly because only 10 IPs and a larger Cressman interpolation parameter are used. For the experiment result in Figure 8A, for example, the average BFC value in the simulation areas is 1.323×10^{-3} and the maximum and minimum values are 3.179×10^{-3} and 9.276×10^{-5}, respectively. For the same longitude, the values of BFCs in the north are less, and, for the same latitude, the values near the middle of the area are less than that in the east or west areas.

4.3 The practical experiments with different Cressman parameters (EXP 6 and EXP 7)

The last suite of experiments consists of EXP 6 and EXP 7. For both of them, 35 IPs and the 1/12°×1/12° observational grids are used. The optimization processes for them are the same as those in the above subsections. The only differences between them are the Cressman parameters, which are 2.5° and 1.5° in EXP 6 and EXP 7 as mentioned in subsection 3.3.

The experiment results are shown in Figure 9. A larger Cressman parameter might introduce some compensations within the BFC values of IPs. Thus, in the above subsections, the patterns are smoother and the optimization process is more stable than what we obtained here, and, in Figure 9, the pattern from EXP 6 is smoother than that from EXP 7. In Figure 9A, the minimum values of BFCs are located near (37°*N*, 123°*E*) in the Yellow seas and the north of the Bohai seas. The maximum values are approximately 5.487×10^{-3} and distributed along the east shores in the Yellow seas. In Figure 9B, with a smaller Cressman parameter, the distribution of the BFCs shows more local features and details. The maximum values are also distributed along the east shores in the Yellow seas, whereas the minimum values appear at (37°*N*, 123°*E*).

A straight intuitive comparison of the data assimilation results is provided through the cotidal charts from EXP 6 and EXP 7 in Figures 10A, B. The differences between the NAO dataset and EXP 6 [amplitudes in (Figure 10C) and phase in (Figure 10D)] are also presented. Compared to Figures 7A, C, it is clear that the differences of amplitudes are further decreased in the Bohai seas and the north of the Yellow seas. A slight increase occurs in the west of the Yellow seas. The differences of phases are not changed significantly here.

How to choose the best Cressman parameter is a complex issue to be discussed. The basic principle is to ensure that the interpolation result can cover the full simulation domain. Furthermore, if the stability of data assimilation processes can be guaranteed, one should choose the parameter so that the cost function convergences to a lower value. In Lu and Zhang (2006), the Cressman parameter is set to 2° for a 1° IP grid. In this paper, the cotidal charts from EXP 6 and EXP 7 are much similar to each other and the difference between their cost function values after convergences is less than 2% of its initial value. Based on this evidence, we think that both of the two parameters are acceptable. A more objective and accurate approach might be discussed in the future.

5 Summary

A-4DEnVar is one of the most cutting-edge hybrid data assimilation schemes. In early studies, both theoretical and experimental studies based on ideal chaos dynamic models have shown its equivalence to the conventional 4D-Var and avoided the adjoint models. As we all know, there are still several issues to be explored if one wants to apply the theory and algorithm ideas into an operational system.

This paper is a bridge from the theoretical algorithm to practical operational developments. To reduce the computational cost caused by the inverse of the $\boldsymbol{B}$ matrix, precondition variables are introduced. The algorithm is further designed as an outer–inner loop structure. The ensemble members and hence the estimated corresponding adjoint model values in 4D-Var are updated only in the outer loops so that too-frequent updates of the ensemble members are avoided.

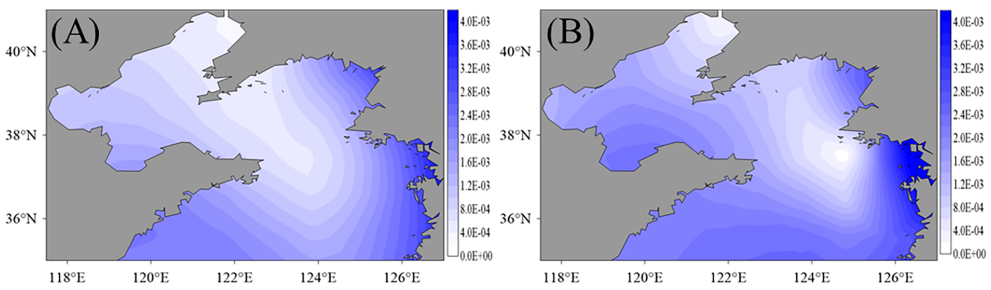


FIGURE 9
The BFC distributions after data assimilation with the same 35 IPs, 1/12°×1/12° observational grids but with Cressman parameter 2.5° **(A)** in EXP 6 and 1.5° **(B)** EXP 7.

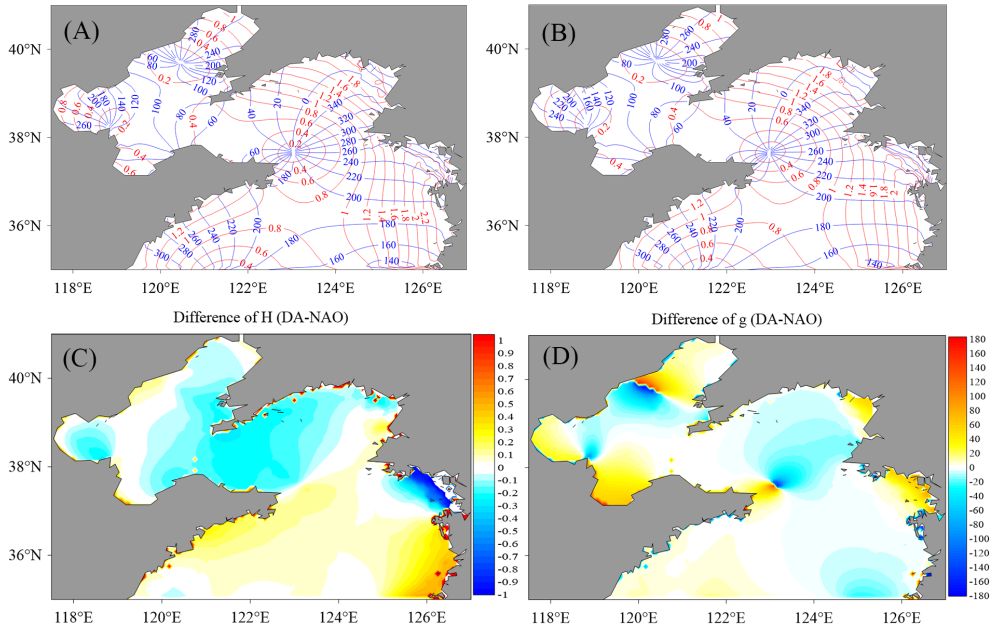


FIGURE 10
The cotidal chart from EXP 6 **(A)** and EXP 7 **(B)** and the differences between EXP 6 and NAO for amplitudes **(C)** and phases **(D)**.

The inner loops focus on the gradient-based optimization of the precondition variables. The local minimum of the precondition variables is combined to the increments of state variables in outer loops.

The IP scheme is, in fact, an implicit localization method to construct an empirical ***B*** matrix. In practice, the values of the IPs are optimized using data assimilation, and the values of other general points are calculated through interpolation methods such as the Cressman method, linear method, and spline interpolation method. Due to the fact that only IPs are explicitly optimized, the IP scheme can reduce the freedom degree and the complicity of the data assimilation process. In the framework of the operational improved A-4DEnVar, the combination of independent schemes makes it possible to restrict the optimization problems on the IP space. Considering that the number of IPs is much less than that of original variables, the ill-posedness is, in fact, limited. Moreover, the IP scheme also introduces an implicit localization that increases the stability of the algorithm.

We constructed the twin experiments and practical experiments to validate the performance of the improvement variant of the A-4DEnVar on the optimization of the BFCs in the Bohai and Yellow seas. In the twin experiment, the perfect observations are used and the cost function decreases rapidly. In the practical experiment, the observations calculated from the tidal harmonic constants of the NAO tide model are used. Compared with the cotidal chart from the original NAO datasets, both the tidal heights and lag phases are improved. After assimilation, the BFCs of the open-sea areas are larger than that of the coastal areas for the same latitude, and, for the same longitude, the BFCs of the Bohai Sea are slightly smaller than that of the Yellow Sea. It is concluded that the operational improved variant of A-4DEnVar works well in the spatially distributed parameter estimation problems.

We demonstrate that, if without the operational improvement, it is a difficult task for the original A-4DEnVar to optimize the BFCs in such vast areas. On the one hand, to construct a full-rank background error covariance and ensure a proper estimation of the adjoint model, the original A-4DEnVar has to generate more than 4,000 (which is the number of BFCs in the POMgcs model) ensemble members. It is clear that the burden of computation cost is too high if the IP scheme is not considered. On the other hand, the local correlations of these BFCs are hard to construct if the Cressman interpolation (or other interpolation skill) method is not considered. It would be a trouble to calculate the inverse of such a big background error covariance matrix, too. In summary, although better improvements might be proposed for A-4DEnVar in the future, in this stage, the IP scheme is necessary and cannot be replaced or ignored in the experiments.

In addition to the issues above, there are some points that are beyond the scope of this paper, but they should be studied in our future studies. On the one hand, this paper explores the performance of A-4DEnVar on the parameter estimation with an implicit localization method, but the proposed algorithm should be widely validated in the initial condition field estimation with large freedom degrees. On the other hand, only the M_2 tide constituent is mentioned in this paper.

However, many tidal constituents are combined with each other in the real world. It is worthwhile to apply the improved scheme to the inversion of BFCs under a more realistic model.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Author contributions

KL and WL had the conceptualization and methodology; GH and XW validated the result; KL wrote the manuscript and WL, GH, XW, and XQ reviewed it. All authors have read and agreed to the published version of the manuscript.

Funding

This research was funded by grants from the National Key Research and Development Program (2022YFC3104804 and 2021YFC3101501) and the National Natural Science Foundation (41876014) of China.

Acknowledgments

The authors acknowledge valuable discussions with Hanyu Liu, Song Hu, and Ru Wang from the School of Marine Science and Technology of Tianjin University.

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Blumberg, A. F., and Mellor, G. L. (1987). "A description of a three-dimensional coastal ocean circulation model," in *Three-dimensional coastal ocean models, coastal estuarine studies 4*. Ed. N. S. Heaps (Washington: American Geophysical Union), 1–16.

Buehner, M. (2005). Ensemble-derived stationary and flow-dependent background-error covariances: Evaluation in a quasi-operational NWP setting. *Q.J.R. Meteorol. Soc* 131 (607), 1013–1043. doi: 10.1256/qj.04.15

Buehner, M., Houtekamer, P. L., Charette, C., Mitchell, H. L., and He, B. (2010a). Intercomparison of variational data assimilation and the ensemble kalman filter for global deterministic NWP. part I: Description and single-observation experiments. *Mon. Wea. Rev.* 138 (5), 1550–1566. doi: 10.1175/2009MWR3157.1

Buehner, M., Houtekamer, P. L., Charette, C., Mitchell, H. L., and He, B. (2010b). Intercomparison of variational data assimilation and the ensemble kalman filter for global deterministic NWP. part II: One-month experiments with real observations. *Mon. Wea. Rev.* 138 (5), 1567–1586. doi: 10.1175/2009MWR3158.1

Courtier, P., and Talagrand, O. (1987). Variational assimilation of meteorological observations with the adjoint vorticity equation II: Numerical results. *Q.J.R. Meteorol. Soc* 113, 1329–1347. doi: 10.1002/qj.49711347813

Cressman, G. P. (1959). An operational objective analysis system. *Monthly Weather Rev.* 87, 10, 367–374. doi: 10.1175/1520-0493(1959)087<0367:AOOAS>2.0.CO;2

Egbert, G. D., and Erofeeva, S. Y. (2002). Efficient inverse modeling of barotropic ocean tides. *J. Atmospheric Oceanic Technology*. 19, 183–204. doi: 10.1175/1520-0426(2002)019<0183:EIMOBO>2.0.CO;2

Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte-Carlo methods to forecast error statistics. *J. Geophys. Res.* 99, 10143–10162. doi: 10.1029/94JC00572

Evensen, G. (2004). Sampling strategies and square root analysis schemes for the EnKF. *Ocean Dynam.* 54, 539–560. doi: 10.1007/s10236-004-0099-2

Ezer, T., and Mellor, G. L. (2004). A generalized coordinate ocean model and a comparison of the bottom boundary layer dynamics in terrain-following and in z-level grids. *Ocean Modelling.* 6 (3-4), 379–403. doi: 10.1016/S1463-5003(03)00026-X

Fan, R., Zhao, L., Lu, Y., Nie, H., and Wei, H. (2019). Impacts of currents and waves on bottom drag coefficient in the East China shelf seas. *J. Geophysical Research: Oceans* 124 (11), 7344–7354. doi: 10.1029/2019jc015097

Goodliff, M., Amezcua, J., and Van Leeuwen, P. J. (2015). Comparing hybrid data assimilation methods on the Lorenz 1963 model with increasing non-linearity. *Tellus A* 67 (1). doi: 10.3402/tellusa.v67.26928

Guo, Z., Pan, H. D., Fan, W., and Lv, X. Q. (2017). Application of surface spline interpolation in inversion of bottom friction coefficients. *J. Atmos. Ocean. Technol.* 34 (9), 2021–2028. doi: 10.1175/JTECH-D-17-0012.1

Kalnay, E., Li, H., Miyoshi, T., Yang, S. C., and Ballabrera, J. (2007). 4D-var or ensemble kalman filter? *Tellus A.* 59 (5), 758–773. doi: 10.1111/j.1600-0870.2007.00261.x

Liang, K., Li, W., Han, G., Shao, Q., Zhang, X., Zhang, L., et al. (2021). An analytical four-dimensional ensemble-variational data assimilation scheme. *J. Adv. Modeling Earth Syst.* 13, e2020MS002314. doi: 10.1029/2020MS002314

Liu, C., and Xiao, Q. (2013). An ensemble-based four-dimensional variational data assimilation scheme. part III: Antarctic applications with advanced research WRF using real data. *Mon. Wea. Rev.* 141 (8), 2721–2739. doi: 10.1175/MWR-D-12-00130.1

Liu, C., Xiao, Q., and Wang, B. (2008). An ensemble-based four-dimensional variational data assimilation scheme. part I: Technical formulation and preliminary test. *Mon. Wea. Rev.* 136 (9), 3363–3373. doi: 10.1175/2008MWR2312.1

Liu, C., Xiao, Q., and Wang, B. (2009). An ensemble-based four-dimensional variational data assimilation scheme. part II: Observing system simulation experiments with advanced research WRF (ARW). *Mon. Wea. Rev.* 137 (5), 1687–1704. doi: 10.1175/2008MWR2699.1

Liu, C. S., and Xue, M. (2016). Relationships among four-dimensional hybrid ensemble-variational data assimilation algorithms with full and approximate ensemble covariance localization. *Mon. Wea. Rev.* 144, 591–606. doi: 10.1175/MWR-D-15-0203.1

Lorenc, A. C., Bowler, N. E., Clayton, A. M., Pring, S. R., and Fairbairn, D. (2015). Comparison of hybrid-4DEnVar and hybrid-4DVar data assimilation methods for global NWP. *Mon. Wea. Rev.* 143 (1), 212–229. doi: 10.1175/MWR-D-14-00195.1

Lu, X. Q., and Zhang, J. C. (2006). Numerical study on spatially varying bottom friction coefficient of a 2D tidal model with adjoint method. *Continental Shelf Res.* 26 1905–1923. doi: 10.1016/j.csr.2006.06.007

Matsumoto, K., Takanezawa, T., and Ooe, M. (2000). Ocean tide models developed by assimilating TOPEX/POSEIDON altimeter data into hydrodynamical model: A global model and a regional model around Japan. *J. Oceanography.* 56, 567–581. doi: 10.1023/A:1011157212596

Oddo, P., Storto, A., Dobricic, S., Russo, A., Lewis, C., Onken, R., et al. (2016). A hybrid variational-ensemble data assimilation scheme with systematic error correction for limited-area ocean models. *Ocean Sci.* 12 (5), 1137–1153. doi: 10.5194/os-12-1137-2016

Pan, H., Guo, Z., and Lv, X. (2017). Inversion of Tidal Open Boundary Conditions of the M2 Constituent in the Bohai and Yellow Seas. *J. Atmospheric and Oceanic Tech.* 34 1661–1672. doi: 10.1175/JTECHD-16-0238.1

Qian, S. H., Wang, D. S., Zhang, J. C., and Li, C. Y. (2021). Adjoint estimation and interpretation of spatially varying bottom friction coefficients of the M2 tide for a tidal model in the bohai, yellow and East China seas with multi-mission satellite observations. *Ocean Modelling.* 161. doi: 10.1016/j.ocemod.2021.101783

Rabier, F., Järvinen, H., Klinker, E., Mahfouf, J.-F., and Simmons, A. (2000). The ECMWF operational implementation of four-dimensional variational assimilation. I: Experimental results with simplified physics. *Q.J.R. Meteorol. Soc* 126, 1143–1170. doi: 10.1002/qj.49712656415

Rawlins, F., Ballard, S. P., Bovis, K. J., Clayton, A. M., Li, D., Inverarity, G. W., et al. (2007). The met office global four-dimensional variational data assimilation scheme. *Q.J.R. Meteorol. Soc* 133, 347–362. doi: 10.1002/qj.32

Stammer, D., Wunsch, C., Giering, R., Eckert, C., Heimbach, P., Marotzke, J., et al. (2002). Global ocean circulation during 1992–1997, estimated from ocean observations and a general circulation model. *J. Geophys. Res.* 107 (C9), 3118. doi: 10.1029/2001JC000888

Storto, A., Oddo, P., Cipollone, A., Mirouze, I., and Lemieux, B. (2018). Extending an oceanographic variational scheme to allow for affordable hybrid and four-dimensional data assimilation. *Ocean Modelling.* 128, 67–86. doi: 10.1016/j.ocemod.2018.06.005

Talagrand, O., and Courtier, P. (1987). Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory. *Q.J.R. Meteorol. Soc* 113, 1311–1328. doi: 10.1002/qj.49711347812

Tian, X., Han, R., and Zhang, H. (2020). An adjoint-free alternating direction method for four-dimensional variational data assimilation with multiple parameter tikhonov regularization. *Earth Space Sci.* 7, e2020EA001307. doi: 10.1029/2020EA001307

Tian, X., and Zhang, H. (2019a). A big data-driven nonlinear least squares four-dimensional variational data assimilation method: Theoretical formulation and conceptual evaluation. *Earth Space Sci.* 6, 1771–1774. doi: 10.1029/2019EA000735

Tian, X., and Zhang, H. (2019b). Implementation of a modified adaptive covariance inflation scheme for the big data-driven NLS-4DVar algorithm. *Earth Space Sci.* 6, 2593–2604. doi: 10.1029/2019EA000963

Tian, X., Zhang, H., Feng, X., and Xie, Y. F. (2018). Nonlinear least squares En4DVar to 4DEnVar methods for data assimilation: Formulation, analysis, and preliminary evaluation. *Mon. Wea. Rev.* 146, 77–93. doi: 10.1175/MWR-D-17-0050.1

Wang, D., Zhang, J., and Wang, Y. p. (2021). Estimation of bottom friction coefficient in multi-constituent tidal models using the adjoint method: Temporal variations and spatial distributions. *J. Geophysical Res. Oceans* 126, e2020JC016949. doi: 10.1029/2020JC016949

Wunsch, C., and Heimbach, P. (2013). Dynamically and kinematically consistent global ocean circulation and ice state estimates. *Int. Geophysics* 103, 553–579. doi: 10.1016/B978-0-12-391851-2.00021-0

Yao, Z., He, R., Bao, X., Wu, D. X., and Song, J. (2012). M2 tidal dynamics in bohai and yellow seas: a hybrid data assimilative modeling study. *Ocean Dynamics* 62, 753–769. doi: 10.1007/s10236-011-0517-1

Zhang, J. C., and Lu, X. Q. (2008). Parameter estimation for a three-dimensional numerical barotropic tidal model with adjoint method. *Int. J. Numer. Methods Fluids* 57, 47–92. doi: 10.1002/fld.1620

Zhang, J. C., and Lu, X. Q. (2010). Inversion of three-dimensional tidal currents in marginal seas by assimilating satellite altimetry. *Comput. Methods Appl. Mech. Eng.* 199, 3125–3136. doi: 10.1016/j.cma.2010.06.014

Zhang, J. C., Lu, X. Q., Wang, P., and Wang, Y. P. (2011). Study on linear and nonlinear bottom friction parameterizations for regional tidal models using data assimilation. *Continental Shelf Res.* 31, 6, 555–573. doi: 10.1016/j.csr.2010.12.011

Zhang, H., and Tian, X. (2018). A multigrid nonlinear least squares four-dimensional variational data assimilation scheme with the advanced research weather research and forecasting model. *J. Geophys. Res. Atmos.* 123, 5116–5129. doi: 10.1029/2017JD027529

Zhang, J., and Wang, Y. P. (2014). A method of inversion of periodic open boundary conditions in two-dimensional tidal models. *Comput. Methods Appl. Mech. Engrg.* 275, 20–38. doi: 10.1016/j.cma.2014.02.020

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 16 January 2023
DOI 10.3389/fmars.2022.1073377

Check for updates

OPEN ACCESS

EDITED BY
Shiqiu Peng,
State Key Laboratory of Tropical Oceanography, Chinese Academy of Sciences, China

REVIEWED BY
Tao Song,
China University of Petroleum, China
Lin Bo,
National Marine Environmental Forecasting Center, China
Guangjun Xu,
Guangdong Ocean University, China

*CORRESPONDENCE
Xuefeng Zhang
xuefeng.zhang@tju.edu.cn
Dianjun Zhang
zhangdianjun@tju.edu.cn

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

RECEIVED 18 October 2022
ACCEPTED 28 December 2022
PUBLISHED 16 January 2023

CITATION
Miao Y, Zhang X, Li Y, Zhang L and Zhang D (2023) Monthly extended ocean predictions based on a convolutional neural network *via* the transfer learning method. *Front. Mar. Sci.* 9:1073377. doi: 10.3389/fmars.2022.1073377

COPYRIGHT
© 2023 Miao, Zhang, Li, Zhang and Zhang. 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(s) 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.

Monthly extended ocean predictions based on a convolutional neural network *via* the transfer learning method

Yonglan Miao[1], Xuefeng Zhang[1*], Yunbo Li[2], Lianxin Zhang[3] and Dianjun Zhang[1*]

[1]School of Marine Science and Technology, Tianjin University, Tianjin, China, [2]Hydrology and environmental protection department, Army 91001 of Chinese People's Liberation Army, Beijing, China, [3]Key Laboratory of Marine Environmental Information Technology, National Marine Data and Information Service, Ministry of Natural Resources, Tianjin, China

Sea surface temperature anomalies (SSTAs) and sea surface height anomalies (SSHAs) are indispensable parts of scientific research, such as mesoscale eddy, current, ocean-atmosphere interaction and so on. Nowadays, extended-range predictions of ocean dynamics, especially in SSTA and SSHA, can provide daily prediction services in the range of 30 days, which bridges the gap between synoptic-scale weather forecasts and monthly average scale climate predictions. However, the forecast efficiency of extended range remains problematic. With the development of ocean reanalysis and satellite remote sensing products, large amounts datasets provide an unprecedented opportunity to use big data for the extended range prediction of ocean dynamics. In this study, a hybrid model, combing convolutional neural network (CNN) model with transfer learning (TL), was established to predict SSTA and SSHA at monthly scales, which makes full use of these data resources that arise from delayed gridding reanalysis products and real-time satellite remote sensing observations. The proposed model, where both ocean and atmosphere reanalysis datasets serve as the pretraining dataset and the satellite remote sensing observations are employed for fine-tuning based on the transfer learning (TL) method, can effectively capture the evolving spatial characteristics of SSTAs and SSHAs with low prediction errors over the 30 days range. When the forecast lead time is 30 days, the root means square errors for the SSTAs and SSHAs model results are 0.32°C and 0.027 m in the South China Sea, respectively, indicating that this model has not only satisfactory prediction performance but also offers great potential for practical operational applications in improving the skill of extended-range predictions.

KEYWORDS

extended-range forecast, sea surface temperature anomaly, sea surface height anomaly, remote sensing, convolutional neural network, transfer learning

1 Introduction

Sea surface temperature anomalies (SSTAs) and sea surface height anomalies (SSHAs), which play a crucial role in ocean dynamic processes, such as mesoscale eddy, current. Moreover, they are also important indicators to evaluate the ocean-atmosphere interaction phenomena (such as El Niño/Southern Oscillation, tropical storms, etc.). Therefore, it is significant to predict SSTA and SSHA accurately. However, the prediction of SSTA and SSHA still faces great challenges due to the rapid changes of the ocean and the nonlinear dynamics of complex factors. Generally, common SSTA and SSHA forecast methods consist of numerical and conventional statistical methods, which are two main methods of predicting marine variables. Due to the prediction uncertainty caused by errors related to the initial conditions, boundary conditions and various discretization/parameterization approximations, the numerical prediction model is substantially limited in terms of improving the prediction time (Peng and Xie, 2006; Hervieux et al., 2019). Since Charney et al. (1950) made the first numerical 24 hour weather prediction in 1950, the numerical methods have made significant progress in prediction skills and temporal range. However, according to nonlinear predictability theory, the limit to the daily numerical weather prediction range is approximately 2 weeks (Lorenz, 1963) since the information regarding the initial parameterization of the ocean field cannot remain stationary for long due to nonlinear chaos. At present, numerical ocean models can predict 7-10 days at the synoptic scale, and most of the numerical models mainly focus on short- and medium-range or seasonal forecast systems. Although some numerical models can be used for extended-range prediction, such as the National Centers for Environmental Prediction (NCEP) Climate Forecast System (CFSv2) and the European Centre for Medium-Range Weather Forecasts (ECMWF) Variable Resolution Ensemble Prediction System for monthly predictions (VarEPS-monthly), they still suffer from lower prediction skill (Saha et al., 2014; Nageswararao et al., 2022). The theoretical limit of the validity of numerical prediction is 2 weeks, and the correlation coefficient of the predictions is less than 0.5.

With the increased adoption and development of artificial intelligence (AI), new deep learning based statistical prediction methods now outperform the general statistical methods in terms of predicting marine dynamics. At present, various deep learning neural networks have emerged, such as convolutional neural networks (CNNs) (Karpathy et al., 2014; Kim, 2014; LeCun et al., 2015; Shelhamer et al., 2017; Kohler and Langer, 2020), long short term memory networks (LSTMs) (Hochreiter and Schmidhuber, 1997; Chen et al., 2022), convolutional long short term memory networks (ConvLSTM) (Shi et al., 2015; Tong et al., 2022), and transformers (Vaswani et al., 2017; Wang Y et al., 2022). As one of the most popular models, CNNs have the advantages of offering 1) powerful self-learning ability, 2) high processing efficiency for multiple-dimensional data, and 3) self-adaptability (Krizhevsky et al., 2012; Oquab et al., 2014; LeCun et al., 2015). These models can be potentially beneficial in geoscience studies and have been successfully used in object detection (Salberg, 2015; Liu et al., 2016; Long et al., 2017; Zhao et al., 2019; Santana et al., 2022), classification (Castelluccio et al., 2015; Luus et al., 2015; Chen et al., 2016; Masoumi, 2021), extreme weather prediction (Gorricha et al., 2013; Zhuang and Ding, 2016; Castangia et al., 2023), etc. In addition, these models are also used for predicting marine variables. Braakmann-Folgmann et al. (2017) combined CNN and recurrent neural network (RNN) models to predict sea level anomalies (SLA) and analyze the spatiotemporal evolution of the northern and central Pacific Ocean. Han et al. (2019) utilized SST, SSH, and sea surface salinity (SSS) to predict subsurface temperature (ST) based on a CNN model. Wang G et al. (2022) combined ensemble empirical mode decomposition (EEMD) with ConvLSTM to construct a hybrid model to predict sea level anomalies (SLA). It is also worthwhile to note that CNN model was used to predict the El Niño/Southern Oscillation (ENSO) for 1.5 years by Ham et al. (2019), which was a substantial achievement for ENSO predictions. The function of a CNN model is to extract hierarchical characteristics from the input data through a convolution filter, which makes the model suitable for extracting spatial features from marine meteorology data. Meanwhile, the model also offers superior performances in terms of time series analysis by inputting continuous time series data. Based on the ability of CNNs to learn from gridded data and spatiotemporal features, it is a suitable tool for the prediction of SSTA and SSHA in this study. In addition, transfer learning (TL) is also a popular technology that has been successfully applied in research, which was proposed by Pan to solve the problem of limited training samples (Pan and Yang, 2010). The goal of TL methods is to transfer knowledge learned in cases of sufficient source data to target domains consisting of less data. Additionally, this method can solve similar difficult tasks by fine tuning a pretrained model. TL has been actively applied in many studies. For instance, Ham et al. (2019) applied the transfer learning technique to train a CNN model by utilizing CMIP5 outputs and then used SODA data to retrain the model on the basis of the former trained weights.

Currently, large amounts of relatively stable and mature ocean-atmosphere reanalysis data can be acquired easily. Ocean (Atmosphere) reanalysis gridded datasets are able to reproduce historical oceanic (atmospheric) states by combining oceanic (atmospheric) observations from multiple sources with a state-of-the-art numerical ocean (atmosphere) model using robust data assimilation techniques. The development of reanalysis products has provided an unprecedented golden opportunity for deep learning to explore time series statistical predictions methods (Song et al., 2021). With the development of numerical models and the increase in grid resolution, as well as the improvement of data assimilation skills, long sequential and higher quality reanalysis data products have begun to emerge to serve as indicators of global/local climate and ecological change. However, there is a gap with the gridded reanalysis data between the short term and monthly extended-range prediction owing to the absent of the real-time reanalysis products, which means that it is inconvenient to directly use the reanalysis data as indicators of initial conditions for extended-range predictions of ocean elements in time. Fortunately, real-time and/or quasi-real-time satellite remote sensing observations of the Earth's resources over the past several decades have made notable contributions in monitoring and understanding oceanic and atmospheric variability at both global and regional scales. The use of ocean and atmosphere reanalysis datasets as the pretraining datasets within neural network frameworks, followed by TL-based fine tuning with satellite remote sensing observations, can be expected to improve the skill of extended-range predictions of ocean elements to some degree.

The structure of this study is as follows. Section 2 introduces the study area and data preparation, as well as the CNNTL method. Then, the experimental results are presented and discussed in section 3. Based on those results, section 4 provides a summary to discuss the contribution of this work and future research.

2 Materials and methods

2.1 Study area

As shown in Figure 1, the area focused in this study is the South China Sea (SCS), located in the Western Pacific Ocean (5°-24.75°N, 105°-124.75°E). This region is connected to the Indian Ocean, Sulu Sea, Pacific Ocean, and East China Sea through numerous straits. The SCS is a semienclosed marginal sea featuring complex ocean dynamic processes, such as many mesoscale ocean eddies, multiple circulation systems, internal waves and other ocean conditions due to complex submarine topography and a large north-south span (Wang et al., 2012; Hu et al., 2014). Moreover, it is also a typical monsoon area located in the middle of the world's largest source of oceanic heat, the Asian-Australian monsoon region. Monsoons can lead to complex thermodynamics and kinetics in the upper SCS. Thus, this is a sensitive area where significant ocean-atmosphere interactions frequently occur that play a crucial role in changes to global and regional climate. Therefore, the ocean-atmosphere processes in the region can have significant impacts on economies, fisheries, and regional transportation.

2.2 Data

Considering the significance of oceanic predictions in the SCS, the goals of this study involve making extended-range predictions of SSTA and SSHA using convolutional neural networks and transfer learning technology (CNNTL) that are based on remote sensing observations. Moreover, this study also takes the sea surface dynamic processes into account, using the wind speed along with SSTA as input variables to predict SSTA. Both ocean and atmosphere reanalysis data and remote sensing data are used in this study. The reanalysis datasets shown in Table 1, downloaded from Copernicus Marine and Environment Monitoring Service (CMEMS, download from https://resources.marine.copernicus.eu/products) and European Centre for Medium-Range Weather Forecasts (ECMWF, download from https://www.ecmwf.int/en/forecasts/datasets/browse-reanalysis-datasets), are used for the pretraining. These data range from 1 January 1993 to 31 December 2018. Due to data availability, the satellite remote sensing data are restricted to a 3 year period, from 1 January 2018 to 31 December 2020, and are divided into a training set (from 2018 to 2019) and a testing set (2020) for the transfer learning model training. The spatial resolution is the same as that of the reanalysis data, which is 0.25°×0.25°. The data were extracted from an area located at 5°N to 24.75°N and 105°E to 124.75°E; thus, the total number of grids was 80×80. These blended satellite products were used to build and test this deep learning model. The SSH satellite observations are mainly built by combining multiple satellite altimeter missions (Jason-3, Sentinel-3A, HY-2A, Saral/AltiKa, Cryosat-2, Jason-2, Jason-1, T/P, ENVISAT, GFO, ERS1/2) processed by the DUACS multimission altimeter data processing system. SST satellite observations are multiproduct ensembles produced by the GHRSST multiproduct ensemble (GMPE) system at the Met Office. In addition, the 6-hourly blended wind speed is from the WindSat radiometer onboard the Coriolis satellite.

To eliminate the influence of climate variability on model training, the daily SSHA and SSTA values are calculated based on reanalysis data to train the pretraining model by using the daily values minus the average over the 26 years between 1993 and 2018. In addition, the 6-hourly blended wind speed data was used to calculate daily means.

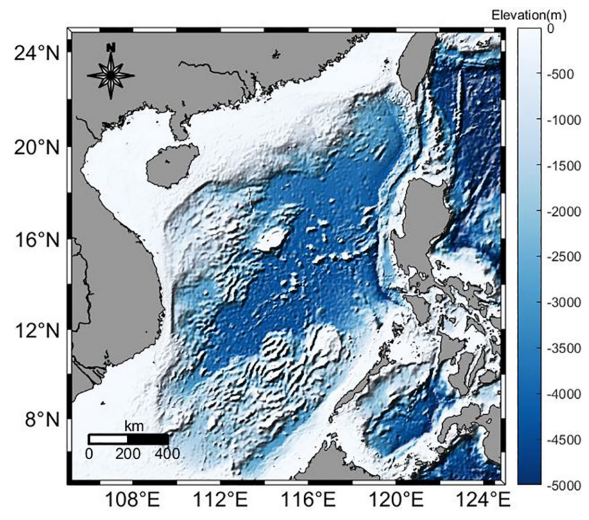


FIGURE 1
The study area.

2.3 CNNTL model

Convolutional neural networks have been adopted to conduct studies successfully since LeCun et al. proposed them in 1998 (LeCun et al., 1998). Generally, CNNs mainly contain an input layer, convolution layer, pooling layer, and fully connected layer (FC). Figure 2 shows an example of a CNN. In the image classification field, the input layer represents the pixel matrix of an image as two dimensional or three-dimensional tensors to input the network. The introduction of convolution and pooling layers makes it superior to traditional neural networks. The convolution layer extracts feature from the network model through a convolution operation, which is a weighted evaluation process that involves sliding convolution filters.

As one of the most crucial parts of CNNs, the pooling layer aims to reduce the resolution for further layers and controls overfitting. Mean pooling and max pooling are two main methods. Due to special linear changes in the convolution operation, an activation function layer must be adopted to increase the network nonlinearity before transferring the results to the pooling layer. The common activation functions are sigmoid, ReLU, tanh, Elu, and so on. In addition, a fully connected layer is used to acquire outputs, which integrates the extracted features from the convolution and pooling processes. To

TABLE 1 Description of reanalysis and satellite remote sensing data.

Data Type	Variables	Sources	Date	Spatial Resolution	Temporal Resolution
Reanalysis	SSH	CMEMS	1993.01.01-2018.12.31	0.25°×0.25°	daily
	SST				
	wind speed	ECMWF			
Satellite remote sensing	SSH	CMEMS	2018.01.01-2020.12.31		
	SST				
	wind speed				6-hourly

accelerate the convergence speed of the gradient descent algorithm and establish a robust model, it is necessary to normalize the training data to a mean of 0 and a variance of 1 with normalization before model training.

In this study, a CNNTL model was constructed for extended-range scale (30 days) predictions of SSHA and SSTA in the SCS. The structure of the CNNTL model is shown in Figure 3. It is mainly composed of 10 convolution layers, 2 pooling layers, 1 fully connected layer and 2 transposed convolution layers. There are 300 convolution kernels for the first two convolution layers with a size of 5×5. The other layers have 30 convolution kernels with sizes of 5×5 (C3 to C5 layer) and 3×3 (C6 to C10 layer). In addition, there is a batch normalization layer after each convolution layer to solve internal covariate shifts in neural network training. This approach was proposed by Ioffe and Szegedy (2015) to improve the generalization ability and training speed of the network. The convolution is a linear process that is difficult to solve linearly inseparable and complex problems in reality. Therefore, to increase the network's nonlinearity, the activation function, as an effective nonlinear method, is generally used after the convolution layer. The rectified linear unit (ReLU) is a popular activation function that has been widely used in recent studies due to its fast calculation time. The ReLU is defined by selecting the max value between the input data and 0, which leads to the problem of dying neurons, the formula is as follows: $\varphi(z)=\max\{0, z\}$. However, the normalized data in this study may have negative values, which indicates that the ReLU is not suitable for using with the data in this study. To avoid this problem, an exponential linear unit (ELU) was used as the activation function here, which was introduced by Clevert in 2016. Generally, pooling layers are added to process feature mapping results obtained through convolution operations. These layers summarize the eigenvalues of a position and adjacent positions as the value of this new position. Therefore, this method can reduce the resolution for further layers and avoid overfitting. This study used maximum pooling as a downsampling method to achieve resolution reduction with a size of 2×2, and hence, the feature map dimension can be reduced by half. In addition, the dropout layers are also used to avoid overfitting, which randomly discards the neurons in the training neural network. In this CNNTL model, they are set behind two max pooling layers and the flattened layer with a probability of 0.5. To acquire the same dimension for the output matrix as the input, transposed convolution, an upsampling method, plays an important role in neural networks. This is set to two layers with 30 kernels, and the kernel sizes are 3×3 and 5×5.

The CNNTL consists of two main parts. First, a CNN model was trained by reanalysis data to predict SSTA and SSHA at extended ranges (30 time steps, or days). Second, the transfer learning models were retrained by remote sensing data on the basis of the CNN model. The framework of this proposed method is demonstrated in Figure 4.

The SSTA time series for the previous 14 and current time steps (T-14 to T, i.e., the 1st to 15th steps) along with wind speed data for the future 10 time steps (from T to T+10, i.e., the 1st to 10th steps) were used to predict 30 days of future SSTAs (T+1 to T+30, i.e., the 1st to 30th steps). Notably, the wind speed is the future rather than historical data, and the reasons for this selection are as follows: the

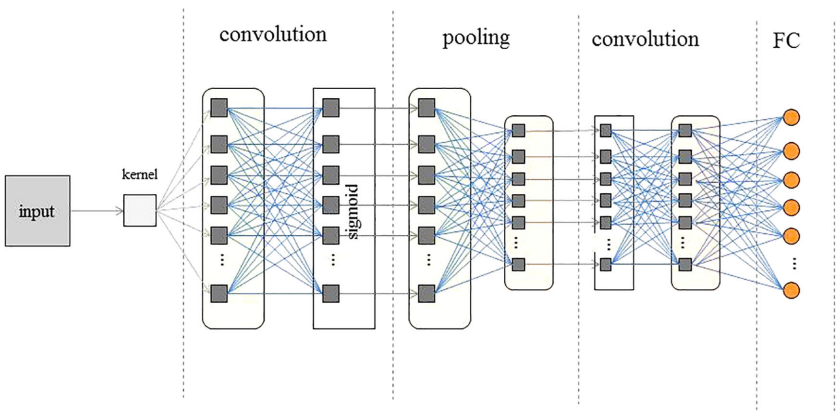


FIGURE 2
An example structure of a CNN.

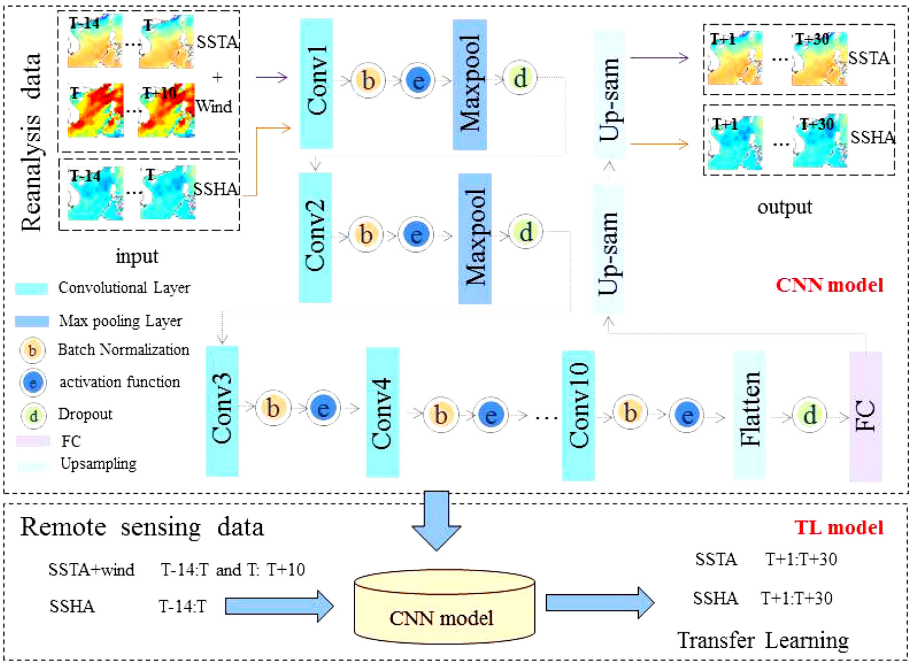


FIGURE 3
Structure of the CNNTL model. This model is composed of ten convolutional layers, ten BN layers, two max pooling layers, one fully connected layer and two upsampling layers.

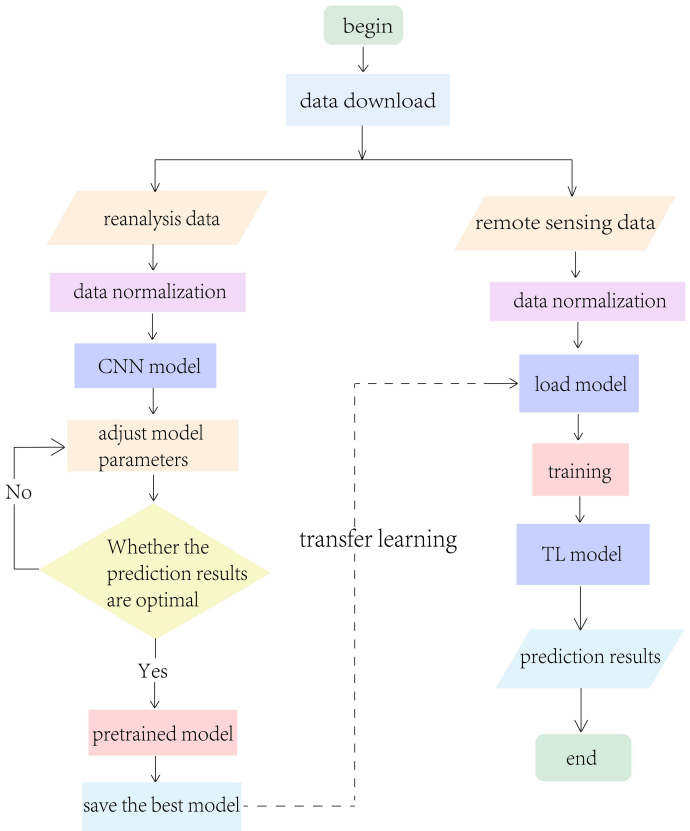


FIGURE 4
Framework of the proposed prediction model.

wind stress has an obvious correlation with SSTA. It can mix sea water, affect oceanic dynamic processes, and influence latent heat flux by accelerating evaporation, which leads to changes in SSTA. Therefore, the wind speed can play an important role in SSTA prediction. In addition, if this model is adopted for actual applications, the wind speed forecast products (e.g., the ECMWF 10-day wind speed forecast products) can be acquired as input data. In contrast, the SSHA is mainly influenced by quasi-geostrophic currents. Thus, only SSHA was used as an input variable in this study to predict the monthly extended SSHAs. The SSHA time series for the previous 14 and current time steps (i.e., T-14 to T, i.e., the 1st to 15th steps) were used to predict the SSHA of the following 30 time increments. Therefore, the full range of the SSTA (SSHA) time dimension is 25 (15). In this study, "T" indicates the current time. Here, the length of the original reanalysis data sequence was 9490, where the sliding window was set to 1 day. Therefore, 9445 samples were acquired to train the CNN model, and then it was divided into a training dataset and validation dataset, 70% and 30%, respectively. The length of the remote sensing training data sequence was 730 from 2018 to 2019. The data from 2020 were selected as the test dataset. The original remote sensing data were also processed as reanalysis data.

Parameter selection is important for neural networks, such as the loss function and optimizer. The loss function used in this study is the mean squared error (MSE), which intuitively reflects the model's training quality according to the difference between the training and validation phases. The smaller value of the loss function is, the smaller the deviation between the results obtained by the model and the real value is, that is, the model is more accurate. The optimizer is used to adjust the parameters to reduce the value of the loss function. The Adam optimizer, a deep neural network method for the adaptive learning rate, is used in this study. It dynamically adjusts the learning rate of each parameter using first- and second-order moment estimators of the gradient. Furthermore, the batch size and epoch size are also crucial parameters to reflect the speed of the model convergence and fitting degree, which are set to 128 and 10, respectively.

2.4 Prediction performance evaluation

In this study, the correlation coefficient (CC) and root mean squared error (RMSE) are metrics used to evaluate the performance of the CNNTL model. The CC reflects the degree of linear correlation between variables. As a common measure of the difference between values, the value of the RMSE is usually the metric used to reflect model performance. The smaller RMSE is, the smaller the prediction difference is, and the better the performance of the model is. Specifically, the two metrics are defined as follows:

$$CC = \frac{\sum_{n=1}^{N}((X_n - \overline{X})(Y_n - \overline{Y}))}{\sqrt{\sum_{n=1}^{N}(X_n - \overline{X})2\sum_{n=1}^{N}(Y_n - \overline{Y}_n)^2}} \quad (1)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(X_n - Y_n)^2} \quad (2)$$

where N denotes the number of samples and X and Y denote the true matrix and prediction matrix, respectively. Meanwhile, the temporal trends of CC and RMSE are calculated sample by sample at spatial grids during the forecast lead time (30 days).

3 Results and discussions

3.1 Accuracy during the test period

This study constructed a CNNTL model based on transfer learning using satellite remote sensing data to predict the monthly scale extended SSTA and SSHA in the SCS. To better express the degree of accuracy for the extended range scale predictions from the model, a time series of regionally averaged CCs and RMSEs were calculated among the test samples and presented in Figure 5 and Table 2. In general, the CCs and RMSEs of SSTA (SSHA) are stable, fluctuating between 0.6 to 0.79 (0.8 to 0.89) and 0.22°C to 0.32°C (0.020 m to 0.027 m), respectively. Figure 5 shows that the CCs (RMSEs) of the dataset gradually decrease (increase) with an increase in time. When the lead time is 30 days, the CC of SSTA (SSHA) exceeds 0.5 (0.7), indicating that the extended monthly predictions of the CNNTL model is ideal.

Furthermore, the spatial distributions of average CC and RMSE of the CNNTL model are given in Figure 6. The average CCs and RMSEs of the prediction made from the 2020 test dataset were calculated grid by grid from daily data over an extended range scale (30 days). As shown in Figure 6A, the CCs of the SSTAs are mainly above 0.6 over most areas of the SCS, showing the availability of the CNNTL model in extended-range prediction. As the predictions progressed, the CCs decreased significantly, indicating that the predictions became less stable as time increased. The overall RMSEs of the SSTAs (Figure 6C) during the prediction are mostly between 0.2-0.5°C, except in the coastal sea region and the south-central SCS. The RMSE in the central sea basin increases distinctly with increasing number of forecast lead days. For the SSHA, CCs are mainly above 0.8 (Figure 6B), and RMSEs (Figure 6D) are mainly 0.01-0.04 m during the prediction. Unlike the SSTAs, the CCs of the SSHA showed greater stability as

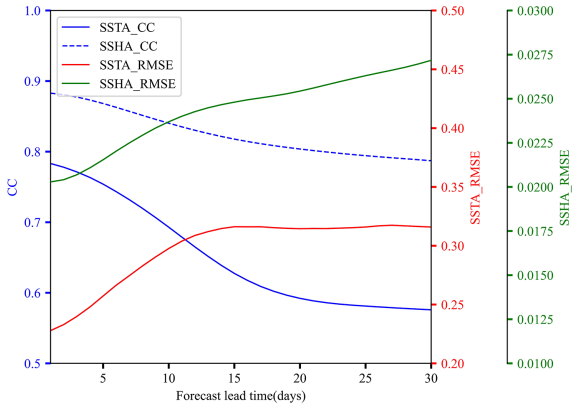


FIGURE 5
Accuracy of the CNNTL model during the forecast lead time (30 days). Temporal trend of CC (blue solid line: SSTA, blue dotted line: SSHA) and RMSE (red solid line: SSTA, unit: °C; green solid line: SSHA, unit: m).

TABLE 2 The experimental results of 10-, 20-, and 30-day forecasts.

	10-Day		20-Day		30-Day	
	CC	RMSE	CC	RMSE	CC	RMSE
SSTA	0.6930	0.2976°C	0.5921	0.3145°C	0.5758	0.3159°C
SSHA	0.8403	0.0237 m	0.8037	0.0254 m	0.7871	0.0272 m

prediction time increases, indicating the strong prediction ability of SSHA in the SCS. The areas in the northeast along with the north to the west of the SCS have larger RMSEs than the others, which is because these areas have complicated dynamic processes that influence the variation SSHA, more details of which are discussed in section 3.4.

3.2 SSTA and SSHA spatial-temporal evolution patterns

Figure 7 shows an example (from 14 June to 13 July 2020) of SSTA spatial maps of satellite observations, predictions, and differences at different prediction times with an interval of 5 days. A strong similarity between satellite observations and predictions can be seen in this figure, which indicates that the CNNTL model has good prediction ability. However, the spatial characteristics between them are slightly different. From Figure 7A, there is mainly a positive anomaly in most areas of the SCS, while the coastal marine area presents a negative anomaly. The SSTA prediction results shown in Figure 7B are consistent with the observations in the most areas of SCS. However, the location and intensity of the negative and positive anomalies have large disagreements with the observations during the prediction lead days, especially along the northern coast. Moreover, from Figure 7A, it can be seen that there was a distinct high value center in the southeast Vietnam, especially when the lead time is 25-days and 30-days. However, it is not obvious in the prediction maps (Figure 7B). The reason may be that this area is the appearance of Vietnam cold eddy which was largely depended by the wind speed. If the wind stress is weaker, the cold eddy may not significant, which makes SSTA higher than the normal. Though the CNNTL model was embedded in wind speed, it only contained future 10 days, which played a minor role in extended-scale prediction. In addition, the spatial distributions differences between prediction and observation are shown in Figure 7C. The differences have no obvious changes during prediction time. The higher values are mainly focus on the Beibu Gulf, southeast Vietnam where the complex dynamic processes occur frequently. The prediction results are overestimated in the northern SCS. However, they are underestimated in most areas of the SCS.

Compared to SSTA prediction, the CNNTL model can capture the spatial-temporal distribution of SSHA more accurately (shown in Figure 8). Figure 8A shows that there was an obvious dipole double eddy structure in the eastern of Vietnam during the period from June 14 to July 13, 2020. In addition, there were obvious warm eddies in the eastern part of Taiwan Island (labeled eddy "a" in Figure 8A-Day 5, the same below) and the eastern part of Luzon Strait (eddy "b"), and cold eddies in the Luzon Strait (eddy "c"), southeastern part of Vietnam (eddy "d") and western part of Luzon Island (eddy "e" and "f"). Figure 8B shows the prediction results of CNNTL model, it can be seen that there was a strong similarity between satellite observations and the CNNTL predictions in terms of the overall pattern and the characterization of eddies. Specifically, the warm eddies in the eastern Luzon Strait (eddy "a" and "b"), the southwest of

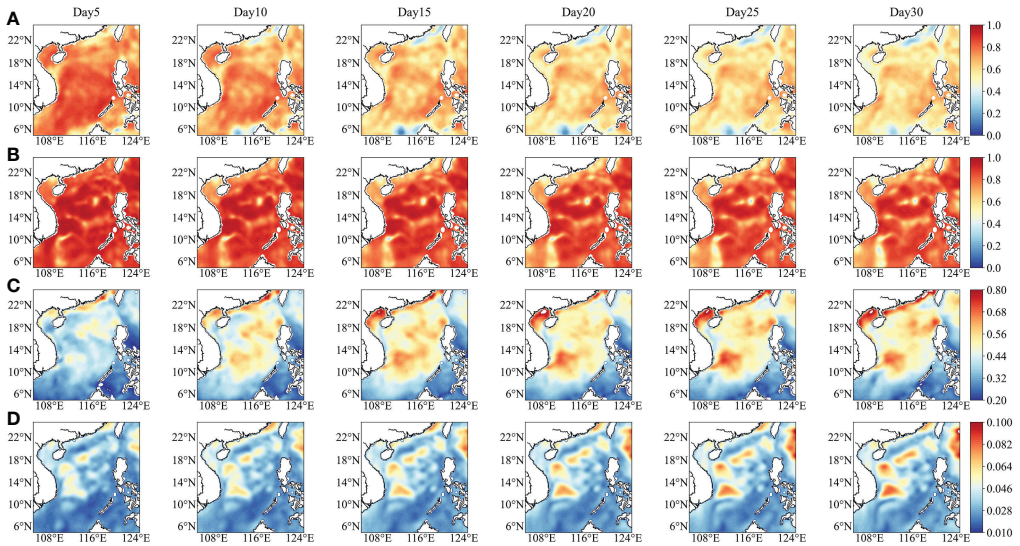


FIGURE 6
Spatial distribution maps of CC and RMSE on 5, 10, 15, 20, 25, and 30 days. **(A)** SSTA CC, **(B)** SSHA CC, **(C)** SSTA RMSE, unit: °C, **(D)** SSHA RMSE, unit: m.

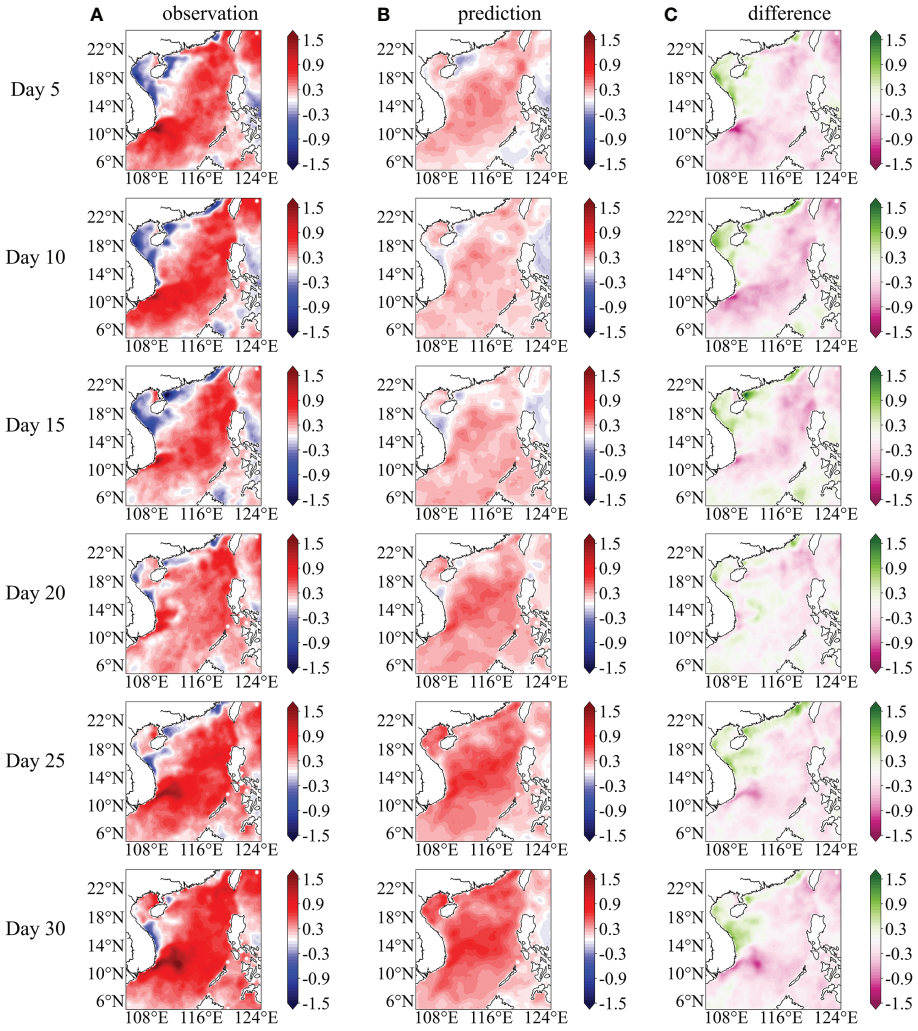


FIGURE 7
Observation **(A)**, prediction **(B)**, and difference **(C)** snapshots of sea surface temperature anomaly (SSTA, °C) prediction for 1–30 days (interval of 5 days), corresponding to June 14 to July 13, 2020.

Taiwan Island (eddy "g"), the eastern and southeastern Indo-China Peninsula (eddy "h" and "i") were well captured during the prediction interval. Although the predicted intensity was relatively weaker compared with the satellite observations, the trend of gradual attenuation of two warm eddies in eastern and southeastern Vietnam (eddy "h" and "i") during the period of June 14 to July 13 was well captured. The cold eddy in southeast Vietnam (eddy "d") developed gradually during this period. However, this temporal evolution trend was not captured by CNNTL. The model can sufficiently capture the locations of the cold eddies in eastern and western Luzon Strait (eddy "c" and "j") and the eastern Indo-China Peninsula (eddy "k"), but the prediction intensity is weaker than the actual observations. As the prediction time increases, the spatio-temporal prediction patterns of the CNNTL model can also consistent with the observations, which demonstrates the high performance offered by the model for extended range scale predictions. Figure 8C shows spatial differences of SSHA, it can be seen that the difference displayed nonuniform patterns. The difference values are higher in the locations of mesoscale eddies. According to Figures 8A, B, the areas of warm eddies (the northern and central regions of the SCS) are underestimated, while the areas of cold eddies are overestimated. This pattern remains almost unchanged as the predication time increases, which was consisted with the patterns of the warm eddies and cold eddies. The main reason for this may be that SSHA is mainly affected by the quasi-geostrophic current, showing significant seasonal and interannual variation characteristics and changes slowly over the extended-range scale.

Figures 9 and 10 shows the evolutionary characters of satellite observation and prediction SSHA during prediction lead time (from June 14 to July 13, 2020) at the 21.5°N section and 118°E section, respectively. From the Figure 9, it showed that there was a cold eddy in the eastern of the Luzon Strait, and this pattern can be well captured by the CNNTL model. With the prediction lead time increasing, this cold eddy gradually attenuated firstly, then became stronger and presented a trend of westward motion. Fortunately, the westward motion trend of the CNNTL model prediction was in good agreement with those satellite observations. However, the prediction strength was much weaker than the actual observation. The reason for this may be that influenced by Kuroshio extension and complex topography prominently, the dynamic processes of this area are

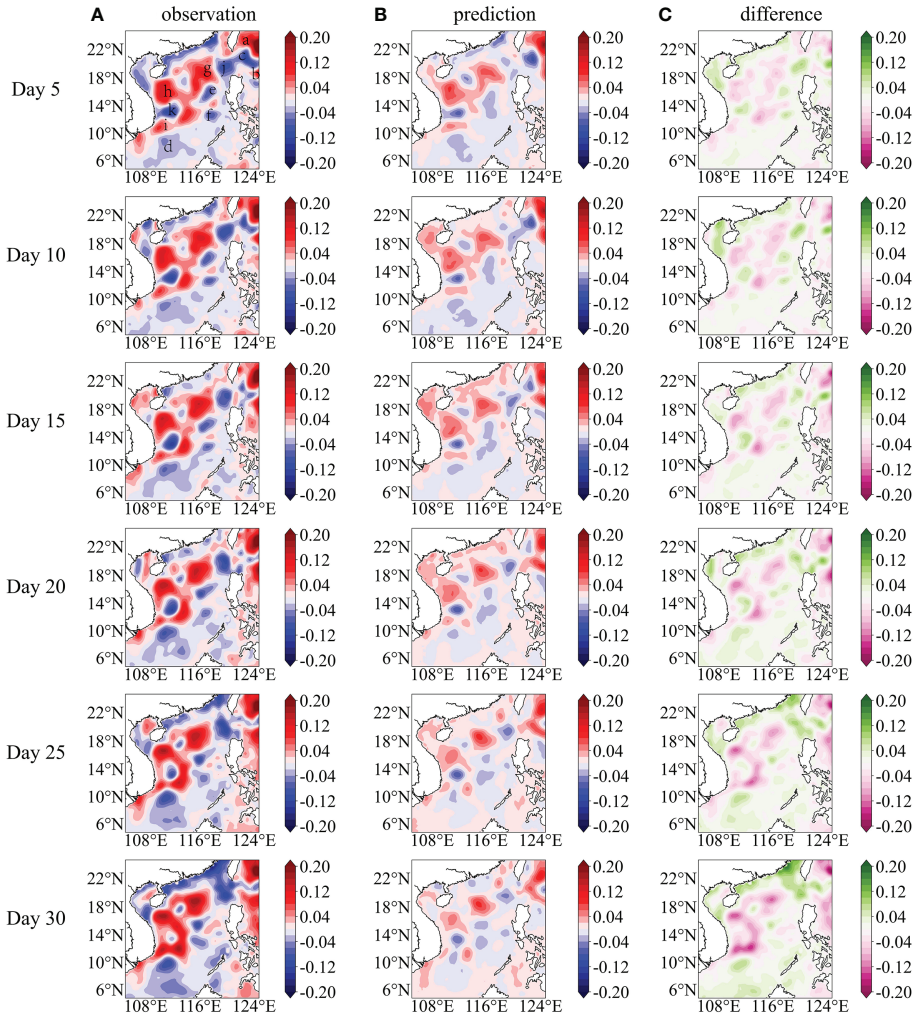


FIGURE 8
Observation **(A)**, prediction **(B)**, and difference **(C)** snapshots of sea surface height anomaly (SSHA, m) prediction for 1–30 days (interval of 5 days), corresponding to June 14 to July 13, 2020.

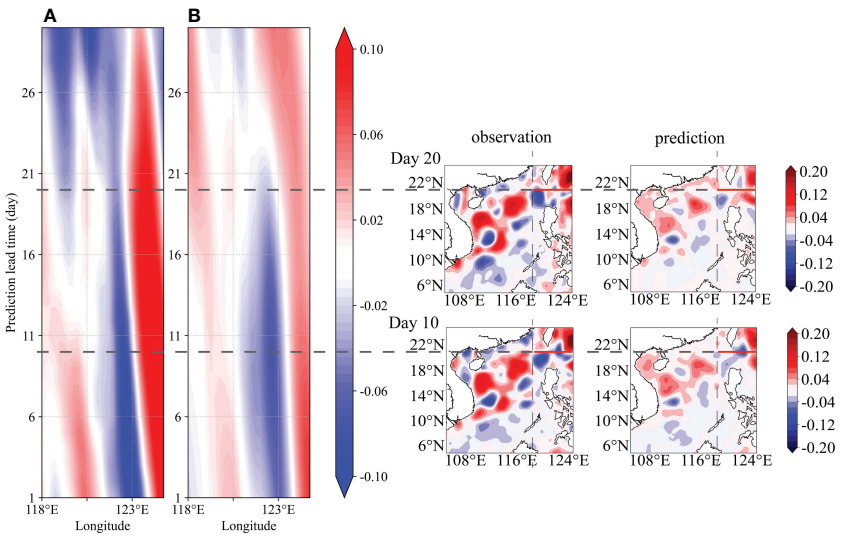


FIGURE 9
Longitude-time maps at the 21.5°N section. **(A)** observation, **(B)** prediction. The red solid line is the longitude range (118°E-124.75°E) of the section.

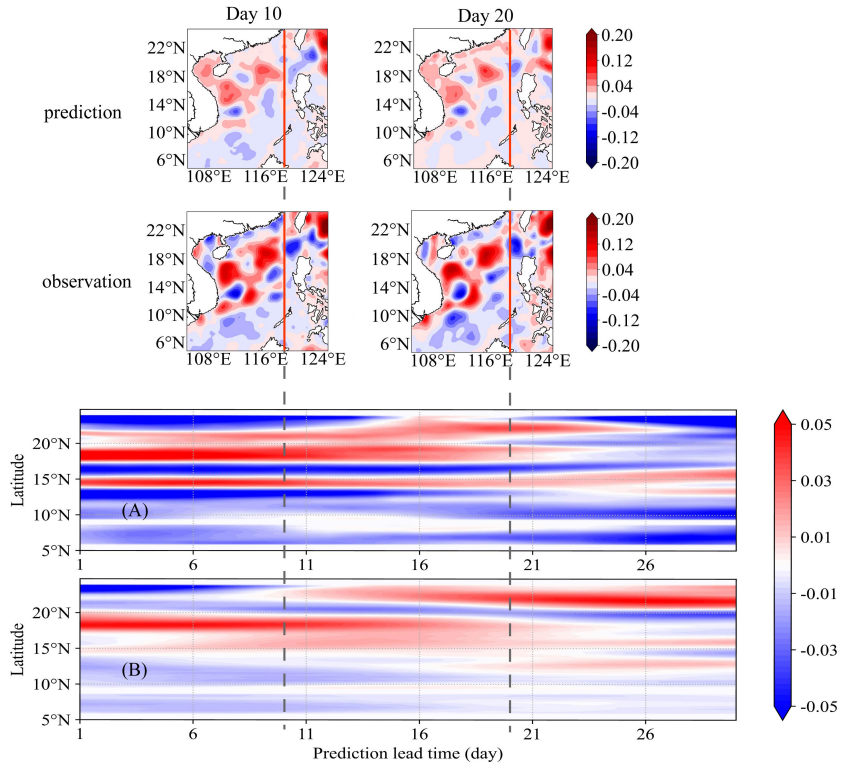


FIGURE 10
Latitude-time maps at the 118°E section. **(A)** observation, **(B)** prediction. The red solid line is the latitude range (5°N-24.75°N) of the section.

extremely complicated, which makes the CNNTL model predict difficultly. From the evolutionary maps at the 118°E section shown in Figure 10, the overall patterns of prediction was consistent with observation well within the 10 days prediction windows. But this pattern was not last longer, the prediction strength was much weaker than observation after 10 days. Specially, the cold eddy at approximately 16°N presented a trend of northward motion. The CNNTL model cannot capture this trend too well. The reason may be that Kuroshio intrudes onto the SCS in summer through Luzon Strait in large scale, contributing to the Luzon Strait cold eddy moved northward. And the strength of Kuroshio intruding is influenced by many factors, making the dynamic processes of this area more complex, which makes it more difficult for the model to predict.

Statistical histograms of the SSTA differences and SSHA differences between the predictions and observations over forecast lead days in 2020 are displayed in Figure 11. It can be seen that the statistics of this model have a lower bias and there is a higher proportion of SSTA (SSHA) differences within a range of ±0.5°C (± 0.035 m).

3.3 Comparison with other models

The comparison of the CNNTL model with the transformer and ConvLSTM model was further quantified. Taking the prediction of SSHA as an example, the RMSEs are shown in Figure 12. It showed that the prediction error of the CNNTL model was stable in the range 0.02-0.03 m within extended-range, with an average RMSE of 0.024 m. Compared with the CNNTL model, the RMSE from the transformer model was less stable and increases significantly with prediction days increasing. It had small error in the initial first day, indicating its suitability for a very short-term prediction. However, it does not preform very well for extended-scale prediction. The average RMSE of transformer model was 0.061 m during the prediction lead time. Besides, the prediction error of ConvLSTM model was more stable than the transformer, but it had higher RMSE during prediction window than the CNNTL model, with the average RMSE of 0.037 m. The RMSEs of tansformer, ConvLSTM, and CNNTL at the end of the prediction window were approximately 0.087 m, 0.041 m, and 0.027 m, respectively. This indicates that the CNNTL model is better than those other two models and has outstanding performance for extended-scale prediction. Notably, both the transformer model and ConvLSTM model used in this study are the basic network models without using other tricks. From the view of the model advancement, both the transformer and ConvLSTM models are more sophisticated than the easy-to-use CNN model, which usually contain the more parameters that need to be fine-tuned to avoid the overfitting and underfitting, thus more data samples are needed to perform the training process for further improving the forecast accuracy.

3.4 Discussion

Based on the above results, when the forecast time is 30 days, the RMSEs of these model predictions for SSTA and SSHA are approximately 0.32°C and 0.027 m, indicating that the CNNTL model performs satisfactorily on an extended range scale. The SCS is one of the most complicated dynamic oceanic areas in the world,

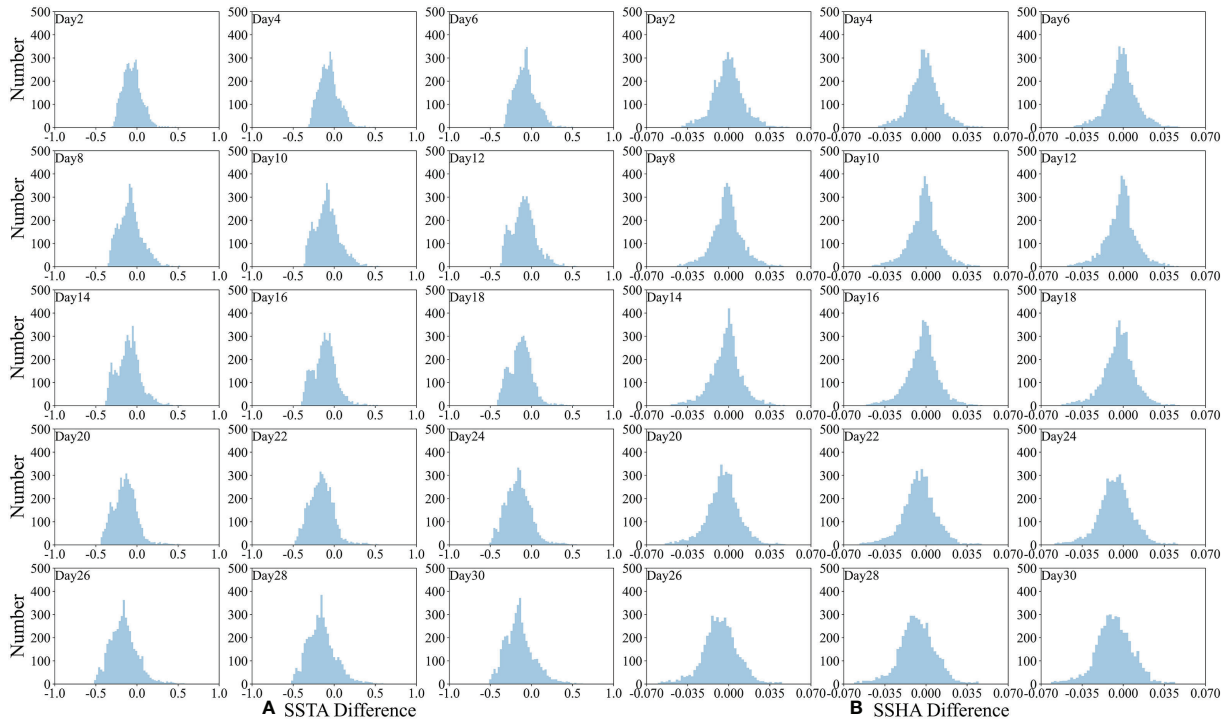


FIGURE 11
Statistics of the SSTA differences (°C) and SSHA differences (m) between the predictions and the observations in 2020. **(A)** Shows the SSTA differences (°C) and **(B)** shows the SSHA differences (m) with leading times from 1 to 30 days.

having nonlinear and chaotic hydrodynamic processes. From the spatial distribution maps of the RMSEs, the performances of areas are quite different. The RMSEs of the SSTAs and SSHAs are higher in the northern SCS and southeast of Taiwan Island. This is because the area has strong nonlinear dynamical processes, which bring about strong impacts on the CNNTL model prediction and hence lead to a higher RMSE in this area. The predictions are very similar to the satellite observations in terms of the overall pattern and the characteristics, although the changes are not captured well in some areas. From the difference maps of the spatial distribution, the differences in SSTA obviously change in the central sea basin with increasing forecast days. The reasons for these results may be as follows: in principle, the wind stress, which disturbed the sea surface water, can intensify the instability of the SSTAs; in addition, SSTAs are also influenced by other various factors, such as local advection and heat flux, while this CNNTL model only considers wind stress without other factors. Therefore, it limits the performance of this model. The differences in SSHA are higher in the northern part of the central basin. In the SCS, mesoscale eddies are quite active and mainly focus on southwestern Taiwan Island, northwestern and southwestern Luzon Island, and the open sea of Vietnam. They are mainly caused by the following two reasons. First, wind stress forces the upper layer of seawater to move, which can lead to Ekman pumping through divergence and convergence, influencing the eddy kinetic energy (EKE) and then contributing to the strength of mesoscale eddies. Second, the Kuroshio intrusion and baroclinic instability of the background current can change the distributions of SSHAs, which mainly appear in the north of the SCS. Although the accuracy of CNNTL is influenced by those factors, it also offers good results for SSTA and SSHA prediction.

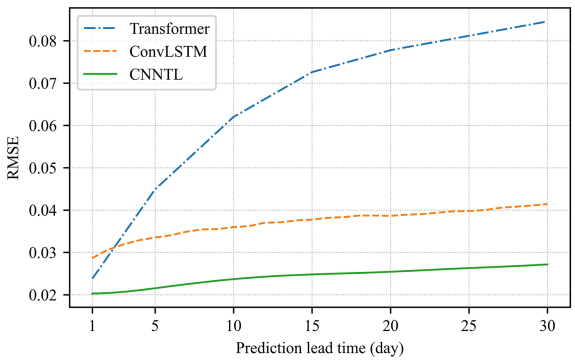


FIGURE 12
The root mean square error (RMSE, unit: m) of SSHA. The CNNTL model was compared with a transformer model (blue dot dash line), a convolutional long short term memory networks (ConvLSTM) model (orange dashed line), and CNNTL (green solid line).

4 Conclusions

Currently, limited by nonlinear chaos predictability, substantial difficulties exist in developing realistic numerical prediction models that operate over longer temporal ranges. However, the rules influencing complicated oceanic processes are hidden within large volumes of spatiotemporal data and can be revealed. Therefore, the growing availability of reanalysis and satellite sensing data makes powerful deep learning technology a promising alternative for

predictions and can circumvent certain temporal restrictions. Based on this, a hybrid statistical predication model (CNNTL) for extended predictions of SSTA and SSHA at monthly scales was adopted in this study. This model accurately predicted the spatiotemporal variations in the SSTA and SSHA that are consistent with the satellite observations. For a forecast time of 30 days, the CCs of the model forecasts for SSTA and SSHA were approximately 0.58 and 0.79, respectively. The RMSEs were 0.32°C and 0.027 m, respectively, which are much smaller than those of transformer model and ConvLSTM model. The forecast assuracy of the more robust models are expected to be further improved throuth increasing the data samples and/or using the fine-tune skills. The spatial distribution of the CCs and RMSEs demonstrate that the RMSEs of SSTA are mainly between 0.2-0.5°C during the course of the predication, and the CCs are mainly above 0.6. Except for some sensitive areas that have complex dynamic processes, the CCs and RMSEs of the SSHA are approximately above 0.8 and within 0.05 m, respectively. To further evaluate the model's performance, this study also analyzed differences in the predication and the satellite observations. For the SSTA results, the positive anomalies were mainly distributed in the northern SCS. For the SSHA, influenced by wind stress, the Kuroshio intrusion and the baroclinic instability of background current, it is obvious that the differences are higher in the central basin, where mesoscale eddies frequently appear. Although influenced by these factors, the CNNTL model demonstrated remarkable performances not only in the temporal trend but also the spatial distribution, indicating that it has sufficient capacity for monthly scale extended predictions. Moreover, the latitude and longitude section results showed that this model can capture the eddy evolutionary accurately. Although the CNNTL model improves the skill of the extended SSTA and SSHA predictions, it can be further improved for applications in the near future. First, the model can be used as a more advanced method to predict monthly scale SSTAs and SSHAs, such as by applying the self-attention mechanism. Second, more factors can be included to more accurately forecast SSTAs and SSHAs and to explore its stability under extreme weather conditions.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding authors.

Author contributions

XZ and YM conceived and designed the experiments. YM conducted the data analysis, model validation, writing, and visualization. XZ, YL, DZ, and LZ conducted the formal analysis, review, editing, and supervision. All authors contributed to the article and approved the submitted version.

Funding

This study was funded by the National Key R&D Program of China (2021YFC2803003) and the Open Fund Project of Key Laboratory of Marine Environmental Information Technology, Ministry of Natural Resources of the People's Republic of China.

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Braakmann-Folgmann, A., Roscher, R., Wenzel, S., Uebbing, B., and Kusche, J. (2017). Sea Level anomaly prediction using recurrent neural networks. *arXiv*. doi: 10.48550/arXiv.1710.07099

Castangia, M., Grajales, L. M. M., Aliberti, A., Rossi, C., Macii, A., Macii, E., et al. (2023). Transformer neural networks for interpretable flood forecasting. *Environ. Modell. Software* 160, 105581. doi: 10.1016/j.envsoft.2022.105581

Castelluccio, M., Poggi, G., Sansone, C., and Verdoliva, L. (2015). Land use classification in remote sensing images by convolutional neural networks. *arXiv*. doi: 10.48550/arXiv.1508.00092

Charney, J. G., Fjortoft, R., and Neumann, V. (1950). Numerical integration of the barotropic vorticity equation. *Tellus* 2, 237–254. doi: 10.3402/tellusa.v2i4.8607

Chen, Y., Jiang, H., Li, C., Jia, X., and Ghamisi, P. (2016). Deep feature extraction and classification of hyperspectral images based on convolutional neural networks. *IEEE Trans. Geosci. Remote Sens.* 54, 6232–6251. doi: 10.1109/TGRS.2016.2584107

Chen, J., Wang, X., and Xu, X. (2022). GC-LSTM: graph convolution embedded LSTM for dynamic network link prediction. *Appl. Intell.* 52, 7513–7528. doi: 10.1007/s10489-021-02518-9

Clevert, D. A., Thomas, U., and Sepp, H. (2016). Fast and accurate deep network learning by exponential linear units (ELUs). *arXiv*. doi: 10.48550/arXiv.1511.07289

Gorricha, J., Lobo, V., and Costa, A. C. (2013). A framework for exploratory analysis of extreme weather events using geostatistical procedures and 3d self-organizing maps. *Int. J. Adv. Intelligent Systems* 6, 16–26.

Ham, Y. G., Kim, J. H., and Luo, J. J. (2019). Deep learning for multi-year ENSO forecasts. *Nature* 573, 568–572. doi: 10.1038/s41586-019-1559-7

Han, M., Feng, Y., Zhao, X., Sun, C., Hong, F., and Liu, C. (2019). A convolutional neural network using surface data to predict subsurface temperatures in the pacific ocean. *IEEE Access* 7, 172816–172829. doi: 10.1109/ACCESS.2019.2955957

Hervieux, G., Alexander, M. A., and Stock, C. A. (2019). More reliable coastal SST forecasts from the north American multimodel ensemble. *Clim. Dyn.* 53, 7153–7168. doi: 10.1007/s00382-017-3652-7

Hochreiter, S., and Schmidhuber, J. (1997). Long short-term memory. *Neural Comput.* 9, 1735–1780. doi: 10.1162/neco.1997.9.8.1735

Hu, W., Wu, R., and Liu, Y. (2014). Relation of the south China Sea precipitation variability to tropical indo-pacific SST anomalies during spring-to-summer transition. *J. Climate* 27, 5451–5467. doi: 10.1175/JCLI-D-14-00089.1

Ioffe, S., and Szegedy, C. (2015). Batch normalization: accelerating deep network training by reducing internal covariate shift. *arXiv*. doi: 10.48550/arXiv.1502.03167

Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R., and Fei-Fei, L. (2014). "Large-Scale video classification with convolutional neural networks," in *2014 IEEE*

Conference on Computer Vision and Pattern Recognition. 1725–1732. doi: 10.1109/CVPR.2014.223

Kim, Y. (2014). Convolutional neural networks for sentence classification. *arXiv*. doi: 10.48550/arXiv.1408.5882

Kohler, M., and Langer, S. (2020). Statistical theory for image classification using deep convolutional neural networks with cross-entropy loss. *arXiv*. doi: 10.48550/arXiv.2011.13602

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Commun. ACM* 60, 1097–1105. doi: 10.1145/3065386

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature* 521, 436–444. doi: 10.1038/nature14539

LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *P. IEEE* 86, 2278–2324. doi: 10.1109/5.726791

Liu, Y., Racah, E., Correa, J., and Collins, W. (2016). Application of deep convolutional neural networks for detecting extreme weather in climate datasets. *arXiv*. doi: 10.48550/arXiv.1605.01156

Long, Y., Gong, Y., Xiao, Z., and Liu, Q. (2017). Accurate object localization in remote sensing images based on convolutional neural networks. *IEEE Trans. Geosci. Remote Sens.* 55, 2486–2498. doi: 10.1109/TGRS.2016.2645610

Lorenz, E. N. (1963). Deterministic nonperiodic flow. *J. Atmos. Sci.* 20, 130–141. doi: 10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2

Luus, F. P., Salmon, B. P., Van den Bergh, F., and Maharaj, B. T. J. (2015). Multiview deep learning for land-use classification. *IEEE Geosci. Remote Sens. Lett.* 12, 2448–2452. doi: 10.1109/LGRS.2015.2483680

Masoumi, M. (2021). Ocean data classification using unsupervised machine learning: Planning for hybrid wave-wind offshore energy devices. *Ocean. Eng.* 219, 108387. doi: 10.1016/j.oceaneng.2020.108387

Nageswararao, M. M., Zhu, Y., and Tallapragada, V. (2022). Prediction skill of GEFSv12 for southwest summer monsoon rainfall and associated extreme rainfall events on extended range scale over India. *Weather Forecast* 37, 1135–1156. doi: 10.1175/WAF-D-21-0184.1

Oquab, M., Bottou, L., Laptev, I., and Sivic, J. (2014). "Learning and transferring mid-level image representations using convolutional neural networks," in *Proc. IEEE Conf. on Computer Vision And Pattern Recognition*. 1717–1724. doi: 10.1109/CVPR.2014.222

Pan, S. J., and Yang, Q. (2010). A survey on transfer learning. *IEEE T. Knowl. Data En.* 22, 1345–1359. doi: 10.1109/TKDE.2009.191

Peng, S. Q., and Xie, L. (2006). Effect of determining initial conditions by four-dimensional variational data assimilation on storm surge forecasting. *Ocean Model.* 14, 1–18. doi: 10.1016/j.ocemod.2006.03.005

Saha, S., Moorthi, S., and Wu, X. (2014). The NCEP climate forecast system version 2. *J. Climate* 27, 2185–2208. doi: 10.1175/JCLI-D-12-00823.1

Salberg, A.-B. (2015). "Detection of seals in remote sensing images using features extracted from deep convolutional neural networks. 2015 IEEE int," in *Geoscience and remote sensing symp* (Milan, Italy: IEEE), 1893–1896. doi: 10.1109/IGARSS.2015.7326163

Santana, O. J., Hernández-Sosa, D., and Smith, R. N. (2022). Oceanic mesoscale eddy detection and convolutional neural network complexity. *Int. J. Appl. Earth Obs* 113, 102973. doi: 10.1016/j.jag.2022.102973

Shelhamer, E., Long, J., and Darrell, T. (2017). Fully convolutional networks for semantic segmentation. *IEEE T. Pattern. Anal.* 39, 640–651. doi: 10.1109/TPAMI.2016.2572683

Shi, X. J., Chen, Z. R., Wang, H., Yeung, D. Y., Wong, W. K., and Woo, W. (2015). Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *arXiv*. doi: 10.48550/arXiv.1506.04214

Song, T., Han, N., and Zhu, Y. (2021). Application of deep learning technique to the sea surface height prediction in the south China Sea. *Acta Oceanol. Sin.* 40, 68–76. doi: 10.1007/s13131-021-1735-0

Tong, B., Wang, X., Fu, J. Y., Chan, P. W., and He, Y. C. (2022). Short-term prediction of the intensity and track of tropical cyclone *via* ConvLSTM model. *J. Wind Eng. Ind. Aerod.* 226, 105026. doi: 10.1016/j.jweia.2022.105026

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., and Gomez, A. N. (2017). Attention is all you need. *arXiv*. doi: 10.48550/arXiv.1706.03762

Wang, G. H., Li, J. X., Wang, C. Z., and Yan, Y. W. (2012). Interactions among the winter monsoon, ocean eddy and ocean thermal front in the south China Sea. *J. Geophys. Res.* 117, C08002. doi: 10.1029/2012JC008007

Wang, Y., Sun, H., Wang, X., Zhang, B., Li, C., Xin, Y., et al. (2022). MAFormer: A transformer network with multi-scale attention fusion for visual recognition. *arXiv*. doi: 10.48550/arXiv.2209.01620

Wang, G., Wang, X., Wu, X., Liu, K., Qi, Y., Sun, C., et al. (2022). A hybrid multivariate deep learning network for multistep ahead Sea level anomaly forecasting. *J. Atmos. Ocean. Tech.* 39, 285–301. doi: 10.1175/JTECH-D-21-0043.1

Zhao, J., Guo, W., and Zhang, Z. (2019). A coupled convolutional neural network for small and densely clustered ship detection in SAR images. *Sci. China Inf. Sci.* 62, 42301. doi: 10.1007/s11432-017-9405-6

Zhuang, W., and Ding, W. (2016). Long-lead prediction of extreme precipitation cluster *via* a spatiotemporal convolutional neural network. *Proc. Sixth Int. Workshop Climate Informatics Boulder CO NCAR NCAR/TN-529+PROC*, 128–131. doi: 10.5065/D6K072N6

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 31 January 2023
DOI 10.3389/fmars.2023.1021650

Check for updates

OPEN ACCESS

EDITED BY
JInyu Sheng,
Dalhousie University, Canada

REVIEWED BY
Yanling Wu,
Hohai University, China
Bozena Wojtasiewicz,
Oceans and Atmosphere (CSIRO), Australia

*CORRESPONDENCE
Florent Gasparin
florent.gasparin@legos.obs-mip.fr

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

RECEIVED 17 August 2022
ACCEPTED 16 January 2023
PUBLISHED 31 January 2023

CITATION
Gasparin F, Lellouche J-M, Cravatte SE, Ruggiero G, Rohith B, Le Traon PY and Rémy E (2023) On the control of spatial and temporal oceanic scales by existing and future observing systems: An observing system simulation experiment approach.
Front. Mar. Sci. 10:1021650.
doi: 10.3389/fmars.2023.1021650

COPYRIGHT
© 2023 Gasparin, Lellouche, Cravatte, Ruggiero, Rohith, Le Traon and Rémy. 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(s) 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.

On the control of spatial and temporal oceanic scales by existing and future observing systems: An observing system simulation experiment approach

Florent Gasparin[1,2]*, Jean-Michael Lellouche[1], Sophie E. Cravatte[3], Giovanni Ruggiero[1], B. Rohith[1,2], Pierre Yves Le Traon[1] and Elisabeth Rémy[1]

[1]Mercator Ocean International, Toulouse, France, [2]Université de Toulouse, LEGOS (IRD/UPS/CNES/CNRS), Toulouse, France, [3]Université de Toulouse, LEGOS (IRD/UPS/CNES/CNRS), Nouméa, New Caledonia

Ocean monitoring and forecasting systems combine information from ocean observations and numerical models through advanced data assimilation techniques. They are essential to monitor and report on past, present and future oceanic conditions. However, given the continuous development of oceanic models and data assimilation techniques in addition to the increased diversity of assimilated platforms, it becomes more and more difficult to establish how information from observations is used, and to determine the utility and relevance of a change of the global ocean observing system on ocean analyses. Here, a series of observing system simulation experiments (OSSE), which consist in simulating synthetic observations from a realistic simulation to be subsequently assimilated in an experimental analysis system, was performed. An original multiscale approach is then used to investigate (i) the impact of various observing system components by distinguishing between satellites and *in situ* (Argo floats and tropical moorings), and (ii) the impact of recommended changes in observing systems, in particular the impact of Argo floats doubling and enhancements of tropical moorings, on the fidelity of ocean analyses. This multiscale approach is key to better understand how observing system components, with their distinct sampling characteristics, help to constrain physical processes. The study demonstrates the ability of the analysis system to represent 40-80% of the temperature variance at mesoscale (20-30% for salinity), and more than 80% for larger scales. Satellite information, mostly through altimetric data, strongly constrains mesoscale variability, while the impact of *in situ* temperature and salinity profiles are essential to constrain large scale variability. It is also shown that future enhancements of Argo and tropical mooring arrays observations will likely be beneficial to ocean analyses at both intermediate and large scales, with a higher impact for salinity-related quantities. This work provides a better understanding on the respective role of major satellite and *in situ* observing system components in the integrated ocean observing system.

KEYWORDS

Argo floats, tropical moored buoys, ocean analysis systems, observing system simulation experiment (OSSE), impact studies, Ocean monitoring and forecasting systems

1 Introduction

As the ocean plays a fundamental role in regulating climate variability, it has been recognized in the 1980-90's that systematic ocean observations are essential to understand and monitor the changing climate of the Earth (e.g., Fu et al., 1994; McPhaden et al., 1998). Initially focused on capturing oceanic variability at large spatial McPhaden et al., 1998, the scope of sustained ocean observations is now expanded to serve diverse end-users, with multi-scale sampling and multi-disciplinary needs (Moltmann et al., 2019). These observations are integrated in the global ocean observing system (GOOS), which includes data stream from satellites and *in situ* platforms. These strong efforts have been essential for the development of ocean models and data assimilation methods allowing to validate and optimize numerical simulations (Smith, 1993). Given these different sources of ocean information from satellites, *in situ* platforms and assimilative models, the Global Ocean Data Assimilation Experiment (GODAE) has strongly supported the implementation of global ocean analysis and forecasting capabilities for operational oceanography (Cummings et al., 2009; Bell et al., 2015). The objective was to build operational systems able to provide to the scientific and broader communities the most accurate estimates of essential physical oceanic variables.

Over the last three decades, significant progress has been made in model developments and data assimilation techniques (Moore et al., 2019), and the diversity of assimilated *in situ* platforms is steadily increasing (Tanhua et al., 2019). Yet, given the complexity of operational systems, it is not currently possible to easily establish the efficiency of the various observations in constraining the ocean state, and to determine how information from observations is used. It also becomes more and more difficult to evaluate the influence of a change in the observing system on ocean analyses, which would be essential to determine the utility and relevance of such changes from the integrated ocean observing system perspective.

Numerous studies based on numerical experiments have investigated the impact of existing or future *in situ* observations in ocean analysis and forecasting systems (Fujii et al., 2019). For instance, the complementarity between tropical mooring, Argo and altimetry data has been demonstrated for global ocean analysis (Balmaseda et al., 2007; Turpin et al., 2016) and seasonal forecasting (Balmaseda and Anderson, 2009; Balmaseda et al., 2013; Fujii et al., 2015). Other studies have also focused on specific regions, like the Tropical Pacific (e.g., Zhu et al., 2021), the Australian coast (e.g., Jones et al., 2012; Aydogdu et al., 2016), or the abyssal Ocean (e.g., Gasparin et al., 2020; Levin et al., 2021). However, a large part of impact studies was dedicated to existing *in situ* observations and/or did not consider the integrated value of the global ocean observing system (e.g., no assimilation of altimetry; Fujii et al., 2019). In addition, usual evaluation metrics, mostly based on box-averaged statistics, make it difficult to separate observation impacts depending on specific space and temporal scales.

The purpose of the present study is thus to analyze the impact of *in situ* observations on constraining oceanic analyses/reanalyses based on Observing System Simulation Experiments (OSSEs), considering both satellites and *in situ* observations and their complementarity or redundancy. A first objective is to disentangle the added value of *in situ* observations when satellite data are already assimilated, while a second objective is to determine their respective ability to constrain specific temporal and spatial scales. Special attention is given to the evaluation of international recommendations for the Argo array and the tropical moored array. Our results will illustrate the need for continuing assessments and improvements of data assimilation techniques for *in situ* observations and should pave the way for future operational systems improvements to increase benefits of ocean observations.

The paper is organized as follows. In the next section, the methodology based on numerical experiments is described. In Section 3, the added value of the main components of the ocean observing system is presented. Section 4 discusses potential outcomes of enhancements of the *in situ* observing system in western boundary current regions and in tropical basins. Discussion and conclusion are provided in Section 5.

2 Data and methodology

The present study is based on a comparison of a series of numerical experiments, called observing system simulation experiments (OSSEs), in which different designs of ocean observations have been assimilated. The present work follows the OSSEs best practices proposed by Halliwell et al. (2014) as much as possible. The OSSEs ingredients are (i) an unconstrained simulation, named the Nature Run, assumed to provide a good representation of the "true" ocean variability over the space and time scales of interest, (ii) a set of synthetic realistic observations simulating different observing system designs, generated from the Nature Run, and (iii) a global experimental system assimilating the above synthetic observations.

2.1 Modelling components

The Nature Run corresponds to the free-running version (i.e., without data assimilation) of the GLORYS12 reanalysis at 1/12° horizontal resolution (Lellouche et al., 2021), called FREEGLORYS12. This unconstrained simulation has been developed at Mercator Ocean International and is based on the NEMO3.1 ocean model (Madec and The NEMO Team, 2008), using a 1/12° ORCA grid (horizontal resolution of 9 km at the equator, 7 km at mid-latitudes and 2 km near the poles). The ocean model is forced at the surface with the atmospheric fields from the ERA-Interim reanalysis produced by the European Centre for Medium-Range Weather Forecasts (ECMWF) (Dee et al., 2011). The Nature Run was initialized in October 1991, from the EN4 gridded fields of temperature and salinity (Good et al., 2013). Assuming that the ocean is initially at rest, the model physics then spins up a velocity field in balance with the density field after about 1 year. The Nature Run was run up until the end of 2017, during which the 2015-2017 period was used to generate synthetic observations.

The experimental analysis system used to perform OSSEs is based on the global Mercator Ocean operational system to be deployed in the Copernicus Marine Service (see Le Traon et al., 2019) portfolio by the end of 2022. The ocean model uses version 3.6 of NEMO (Madec

et al., 2017) with a ¼° ORCA grid type (Madec and Imbard, 1996), and is forced at the surface by the operational atmospheric fields from the ECMWF-Integrated Forecast System (ECMWF-IFS) with 3-h resolution. A coherent bulk formulation is derived from the IFS model (Brodeau et al., 2017), but no atmospheric pressure forcing is used. Moreover, the surface currents are not considered in the stress computation (absolute wind) as it was the case for the Nature Run. The ocean model uses an explicit barotropic mode solved by a split-explicit approach (Shchepetkin and McWilliams, 2005), a second order vertical mixing (k-espilon; Rodi, 1987) and a UBS scheme (Shchepetkin and McWilliams, 2008) for computing the horizontal momentum advection without addition of an explicit diffusion.

In addition to the ocean model, the assimilative system consists of a 3D-Var bias correction for the slowly evolving large-scale biases in temperature and salinity, and a local version of a reduced-order Kalman filter based on the Singular Evolutive Extended Kalman filter formulation (Brasseur and Verron, 2006). In practice, temperature and salinity observations are selected depending on the innovation value, defined as the observation minus model forecast equivalent. For the 3D-Var corrections, innovations are considered on a temporal window of 1 month (i.e., at a given 7-day cycle and the three previous cycles) and on spatial window of the order of 400–500 km in order to map large-scale temperature and salinity corrections. For the assimilation of the SEEK filter, the analysis at a given point is based on surrounded innovations determined by spatial and temporal correlation scales, ranging from 50 to 450 km in the zonal direction, from 50 to 200 km in the meridional direction and from 3 to 15 days. From the innovations and specified observation errors, the SEEK filter generates a localized analysis increment, which is a linear combination of short-scale anomalies from a statistical ensemble representative of the forecast error covariances (Lellouche et al., 2013). The 3D-Var correction and the SEEK increment are applied progressively using the incremental analysis update (IAU) method (another tendency term added in the model prognostic equations), to avoid model shock every week due to the imbalance between the analysis increments and the model physics (Bloom et al., 1996; Benkiran and Greiner, 2008).

Main features of this assimilation system have already been described in Lellouche et al. (2013, 2018). However, a main update has been included for the OSSEs and is related to the use of a 4D analysis (Benkiran et al., 2021) allowing an improvement in the spatiotemporal continuity of mesoscale structures. Note that, unlike Lellouche et al. (2018), no mean dynamic topography is used for referencing the altimetric sea level anomaly, since the total sea surface height is directly assimilated. The system was initialized on January 07, 2015, using fields from a 4-yr spin-up run, and experiments were run up until the end of 2017.

2.2 Design experiments and synthetic data sets

A total of six global ocean experiments has been performed to disentangle the role of various ocean observations in constraining ocean model forecasts/analysis and demonstrate potential outcomes of observing system extensions (Table 1). By assimilating observing system components separately, we aim to identify the ability of observing arrays to constrain different range of spatial and temporal scales and highlight complementarity and redundancy of ocean observation information from an operational oceanography perspective. Each experiment is characterized by the data sets assimilated, which have been synthetically generated by subsampling the daily fields of the Nature Run at the space and time location of each observation. As the study investigates the role of *in situ* observations as part of the integrated ocean observing system, synthetic data sets include both satellite and *in situ* components, with only one satellite configuration used (which is assumed to be close to current satellite constellation).

The common synthetic satellite observations consist in sea surface height (SSH) and sea surface temperature (SST) variables (Figure 1); synthetic sea ice concentration is not generated nor assimilated. The SSH data set is built from a constellation of the three nadir satellites

TABLE 1 Experiments performed in this study. Note that altimetry refers to the assimilation of SSH at the location of observed SLA.

Experiments	Assimilated observing arrays
FREE	No
ONLYSAT	Altimetry, SST
ONLYSITU	Argo_Nominal,Mooring_Nominal
NOMINAL	Altimetry, SST, Argo_Nominal, Mooring_Nominal
ENHANCED_AR	Altimetry, SST, Argo_Enhanced, Mooring_Nominal
ENHANCED_MO	Altimetry, SST, Argo_Nominal, Mooring_Enhanced
ENHANCED_ARMO	Altimetry, SST, Argo_Enhanced, Mooring_Enhanced

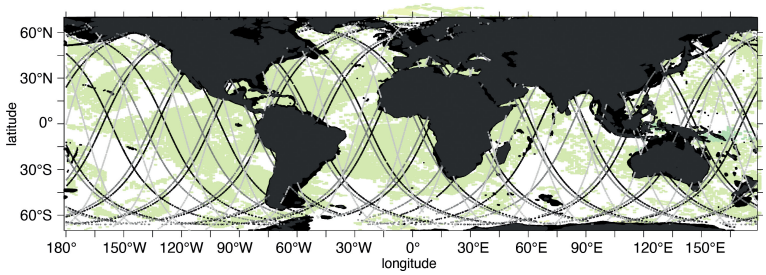


FIGURE 1
Configurations of synthetic SST (shading) and SSH (curves) observations for a given day. SST maps have been masked out following the ODYSSEA L3S product. SSH sampling corresponds to the orbitals of Jason-2, Jason-3 and Sentinel-3a altimeters.

Jason-2, Jason-3 and Sentinel-3a. Positions (longitude, latitude, time) are extracted from Copernicus Marine Service Sea Level TAC (Thematic Assembly Center) multi-mission along-track L3 altimeter products. Each satellite provides around 50,000 measurements per date (10-day repeat cycle and 13 orbits per day for Jason-2 and Jason-3; 27-day repeat cycle and 14 orbits per day for Sentinel-3a). The SST data set consists in daily maps obtained from the Copernicus Marine Service ODYSSEA multi-sensor L3S product. This product, consisting in a fusion of SST observations from multiple satellite sensors, daily, over a 0.1° resolution global grid, was used to mask regions without SST observations.

The synthetic *in situ* data sets consist of subsurface vertical profiles of temperature and salinity (T/S profiles) from two historical global *in situ* networks; the Argo global array of profiling floats (www.argo.net) and the Global Tropical Moored Buoy Array (GTMBA, www.pmel.noaa.gov/gtmba). Unlike satellite observations, synthetic *in situ* observations are built on idealized configurations to ease the interpretation of results. For each of those networks, two different designs are considered representing the current (NOMINAL) and enhanced (ENHANCED) arrays as follows (Figure 2).

- Argo-NOMINAL mimics the standard configuration and corresponds to one Argo float per 3°x3°x10-day square, sampling the 0-2000 m upper-ocean globally. Locations of T/S profiles are randomly distributed in a 3°x3° square, each square being sampled every 10 days. The day of the first 10-day cycle is randomly distributed in the first 10-day window and the space position is different for each 10-day cycle. This T/S configuration counts around 470 profiles per day from 3700 floats, with measurements located at the model vertical levels (including 22 levels within the upper 100 m, with 1-m resolution at the surface and 450-m resolution at the bottom).
- Argo-ENHANCED is based on the latter configuration, with one added float per 3°x3° square in highly energetic regions, i.e, in western boundary currents and in equatorial/tropical regions following international recommendations (Roemmich et al., 2019; Smith et al., 2019). As Argo-NOMINAL, the day of the first 10-day cycle of additional profiles is randomly distributed in the first 10-day window.
- Mooring-NOMINAL uses the position of tropical moorings during the 2020-2021 period (only two TRITON moorings in the western tropical Pacific). Vertical levels of T/S profiles are based on "standard instrumental depths" given by the GTMBA website. Eleven depth levels are located between the surface and 500 m, with 20-m resolution in the upper 150 m. Note that, unlike in the Indian and Atlantic Ocean, there are only temperature profiles in the Pacific.
- Mooring-ENHANCED is mostly characterized by an increased vertical resolution, which follows the recommendations of the tropical community (Foltz et al., 2019; Hermes et al., 2019; Smith et al., 2019), and by a reconfiguration of the Pacific array. Temperature sensors lie at 1 m, every 5 m from 5 to 30 m, every 10 m from 30 to 60 m, and with vertical resolution like present from 60 to 500 m (depending on the longitude in the basin). Salinity sensors are located at 1 m, every 5 m from 5 to 30 m, every 10 m from 30 to 80 m and at 100 m. The strong modification of the spatial distribution of tropical moorings in the Pacific follows the conclusions from the Tropical Pacific Observing System 2020 (Kessler et al., 2021).

To mimic the assimilation procedure of real observations within an operational system, synthetic data sets must deviate from the simulation in which they are assimilated, but also from the Nature

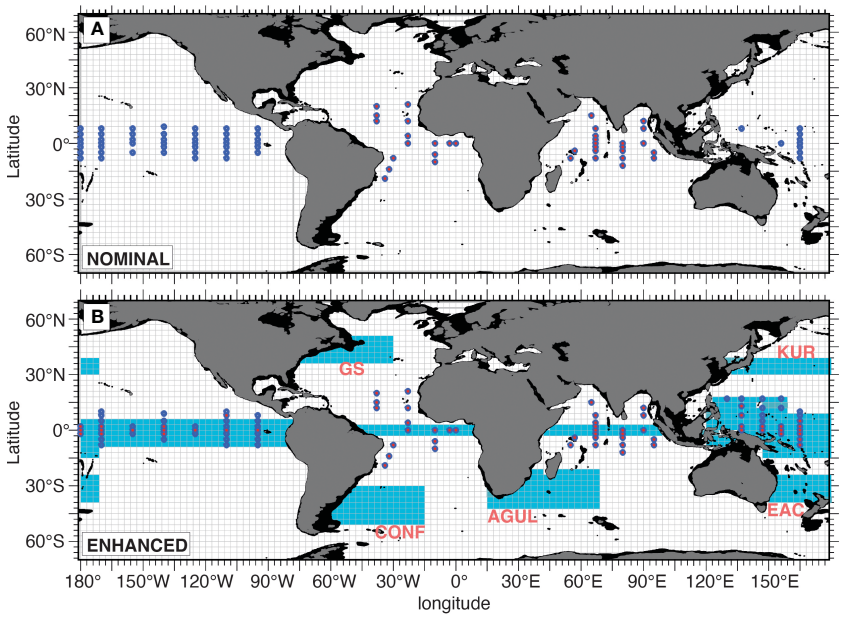


FIGURE 2
Nominal **(A)** and enhanced **(B)** configurations of synthetic temperature and salinity (T/S) profiles for Argo floats (shading) and tropical moorings (dots). Argo sampling corresponds to 1 and 2 floats per 3°x3°x10-day square in yellow and light blue, respectively. Subsurface T/S observations from moorings are indicated by blue (T) and red (S) dots.

Run realization: errors must be prescribed to the synthetic observations, as in the real system. We follow the methodology of Gasparin et al. (2019) to add errors of the synthetic satellites and *in situ* observations. It should be first noted that synthetic observations are generated from the Nature Run daily mean fields, which are at a higher horizontal resolution (at 1/12° resolution) than the experimental analysis system (at ¼° resolution) in order to consider variability in the synthetic observations associated with processes resolved in the 1/12° system, but not in the ¼° system. Observation error must include a representation error and an instrumental error. The representation error is generated to mimic unresolved or poorly resolved small-scale processes by the data assimilation scheme used in the OSSEs (e.g., internal waves), with horizontally and vertically correlated errors. For that, a time-shifting technique, usually used by the atmosphere community (e.g., Huang and Wang, 2018), generates weekly variability by randomly shifting the Nature Run fields by ± 3 days (following a uniform distribution, either 3 days before or 3 days after the given date). Finally, an instrumental error is added to each observation as an uncorrelated error following a Gaussian distribution with the standard deviation given by the instrumental accuracy (0.35°C for SST; 3 cm for SSH; 0.01°C/0.01 for Argo T/S profiles; and 0.02°C/0.02 for Mooring T/S profiles; Cabanes et al., 2013). We call "synthetic observation error" the sum (in variance) of these different of additional errors.

2.3 Experiments calibration

2.3.1 Synthetic observation error

As the reliability of OSSEs to correctly provide impact assessment partly lies in defining appropriate errors associated to synthetic observations, the representation error is first evaluated for the 100-m temperature and 10-m salinity by computing Root-Mean-Square difference between the original Nature Run fields and the 3-day shifted Nature Run fields. This representation error is of the order of O(0.2°C) for 100-m temperature and O(0.05) for 10-m salinity, with high variability in western boundary regions, tropics, and the Southern Ocean (Supplementary Figure S1). Both instrumental error and error due to the small-scale variability embedded in the 1/12° Nature Run not represented by the ¼° OSSE grid, are negligible in comparison to representation error (not shown), in agreement with Gasparin et al. (2019). The consistency of the synthetic observation errors is measured by comparing the synthetic observation error with observation error prescribed in the data assimilation system. A similar comparison is then carried out at three mooring locations in the equatorial Pacific (Supplementary Figure S2). Profiles of the amplitude of the representation error show a maximum at the thermocline level (~0.8°C), which is slightly higher than the observation error variance specified in the operational system (~0.5-0.8°C). These error estimates are also well-compared with the amplitude of the high-frequency variability of temperature (periods shorter than the 7-day assimilation window) based on the 10-min mooring time series. This indicates that the amplitude of the synthetic observation error, representing high-frequency variability, is thus realistic and in agreement with observation error variance specified in the operational system.

2.3.2 Residual error

To evaluate the good calibration of the experimental analysis system (see Halliwell et al., 2014), it is important to verify that the distance between the assimilated and non-assimilated runs is similar to that of the GLORYS12 reanalysis. If the NOMINAL observing system is comparable with the current observing system, the distance between NOMINAL and FREE experiments must be similar to that of GLORYS12 and its free version FREEGLORYS12. In Figures 3A, B the 100-m temperature and 10-m salinity RMS difference between NOMINAL and FREE experiments, zonally averaged, ranges between 1.0 and 1.5°C for temperature and 0.25 to 0.75 for salinity, and is similar to the RMS difference computed from the difference between GLORYS12 and FREEGLORYS12 (slightly higher for temperature). In addition, the distance of NOMINAL from the synthetic observations is compared with the distance of GLORYS12 from real

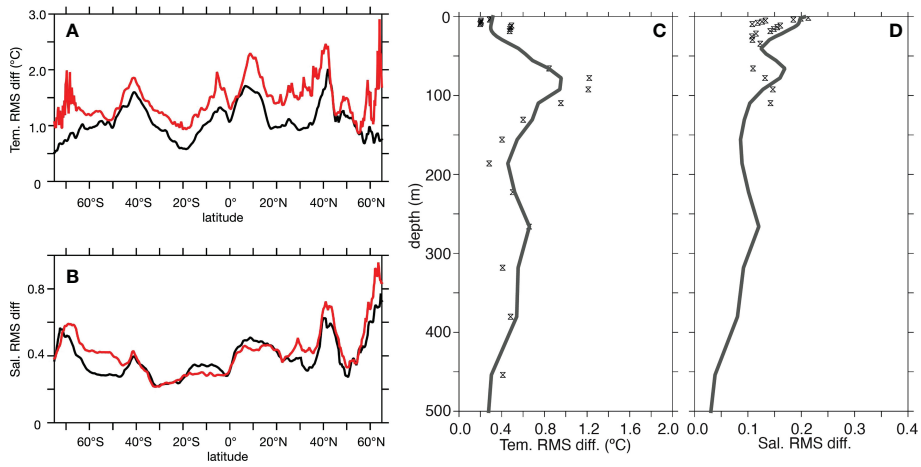


FIGURE 3
(A, B) 100-m temperature and 10-m salinity RMS difference, zonally averaged, between the free and assimilated simulations for the OSSE system (FREE and NOMINAL; black) and the reanalysis system (FREEGLORYS12 and GLORYS12; red). **(C, D)** Temperature and salinity RMS residuals (difference of OSSEs fields from the Nature Run fields) at 23°W, 0° (Atlantic) from the OSSE system (distance between NOMINAL and synthetic observations) and the GLORYS12 reanalysis.

observations at a mooring point in the equatorial Atlantic (Figures 3C, D). The similar shape of the profiles also provides a good confidence of the good calibration of the experimental analysis system and therefore a realistic behavior of OSSEs.

3 Spatial and temporal scales constrained by ocean observations

Given the sparse distribution of ocean observations, usual metrics in operational centers are mostly based on box-averaged statistics (Hernandez et al., 2009), making difficult to separate analysis skills according to spatial and temporal scales. To evaluate the ability of ocean observations to constrain ocean state estimates in the experimental analysis system, time series of sea surface steric height (relative to the bottom and referred in the following as steric height) from the FREE experiment, is first analyzed with separation of spatial and temporal scales. To ease the understanding of this global study, we choose to decompose signals and associated errors into three different space and time scales, although more complex techniques could have been used such as Ensemble Empirical Mode Decomposition (EEMD). Temporal and spatial scales shorter than 20 days and 100 km, respectively, are referred to as "small-scale variability". Small-scale variability, not resolved by the 7-day assimilation window and the ¼° horizontal grid of the eddy-permitting model, is isolated by applying a 1°x1°x20-day high-pass filter on the gridded fields. Note that this small-scale variability includes part of the mesoscale activity which cannot be resolved (and constrained by observations) given the experimental system (Cipollone et al., 2017; Yu et al., 2022). Using a similar filter than Roemmich and Gilson (2009), large-scale signals are obtained by applying a 9°x9°x100-day low-pass running mean filter to represent "large-scale variability". Finally, the difference between the 9°x9°x100-day and 1°x1°x20-day smoothed time series is referred to "intermediate-scale variability" to define processes such as mesoscale eddies (which can extent to 500-1000 km; Storer et al., 2022) and intraseasonal waves, with temporal and spatial scales of 20-100 days and around 100-1000 km, respectively. In addition, the term "residual error" refers to statistics based on the difference of OSSEs fields from the Nature Run fields. Note that our results do not depend sensitively on specific choice of scales separation, since other choices (e.g., 8°x8°x80-day, 10°x10°x100-day) yield similar results (not shown).

3.1 Variability amplitude for various scales

Figure 4A first shows the total variability of steric height from the FREE experiment. High variability regions are clearly identified in western boundary regions and in the Southern Ocean, with amplitude reaching more than 12 cm. These regions are characterized by instabilities of the strong mean flow generating meanders and eddies (Ducet and Le Traon, 2001). Moderate variability regions are seen in the tropical Indian Ocean and Pacific Ocean. Low variability regions are in the center of oceanic gyres. The standard deviation of steric height was then zonally averaged to show the latitude dependence of the total variability (Figure 4B), and spatial and temporal filters were applied to isolate the steric height variability at small, intermediate and large scales (Figure 4C). At latitudes of the high and moderate variability regions, intermediate variability represents almost 70% of the variability. At other latitudes, large-scale variability is equal or higher than intermediate variability. Small-scale variability, which is usually dominated by coherent vortices, fronts and filaments, represents a non-negligeable contribution to the total variability.

The amplitude of the small-, intermediate and large-scale variability is then compared to the residual error of the FREE experiment, estimated by the RMS difference from the Nature Run fields (Figure 5). The amplitude of the total residual error (including small-, intermediate- and large scales) is quite similar to the total variability of the SH signal, with slightly higher error amplitude at latitudes of high variability regions and slightly lower error amplitude in the tropical band (Figure 5A). Yet, the scale separation demonstrates that the amplitude of the residual error of the FREE experiment is differently distributed over scales than that of the variability (Figures 5B-D). High variability regions of the Northern Hemisphere have similar amplitude at intermediate and large scales, but the FREE residual error is mostly dominated by intermediate scales (mesoscale).

Residual errors are usually defined based on the total RMS difference between observations and analysis, without scale

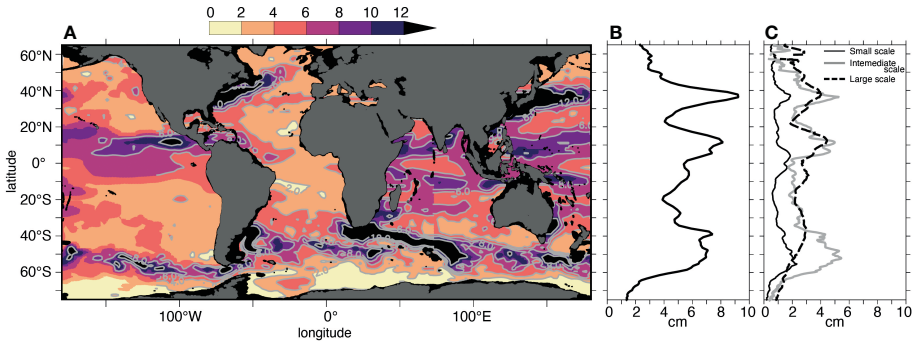


FIGURE 4
Standard deviation of the daily steric height (SH, cm) from the FREE experiment **(A)** spatial map, **(B, C)** zonal-average, black line). For comparison, zonally averaged standard deviation of the daily SH fields of the small scales (black line), intermediate (gray line) and large-scale variability (dashed line) are also shown.

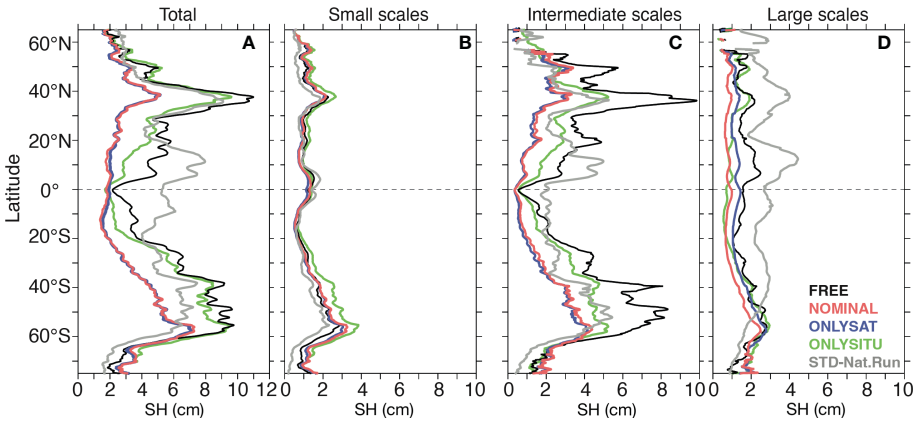


FIGURE 5
Zonally averaged steric height (SH, cm) RMS difference between the Nature Run and experiments (FREE, NOMINAL, ONLYSAT, ONLYSITU) for the total **(A)**, the small scales (smaller than 1°x1°x20-day, **B**), mesoscale (between 1°x1°x20-day and 9°x9°x90-day, **C**) and large variability (larger than 9°x9°x90-day, **D**). For comparison, the standard deviation of the Nature Run SH, zonally-averaged, is also shown (gray).

separation, and we clearly see the limit of this diagnostic here: it favors the dominant scales of the residual error but does not inform us on the processes that are, or not, constrained by the assimilation. This metric does not either weight the residual error amplitude according to the natural variability. We propose here to define the metric "percent of represented variance", calculated as one minus the proportion of the residual variance (e.g., the variance of the residual error divided by the signal variance), and to compute it for different scales. We argue it will allow better assessing the impact of the observing system components.

3.2 Impact of the various observing system components

Based on the comparison of several experiments assimilating separately (ONLYSAT, ONLYSITU) and conjointly satellites and *in situ* data sets (NOMINAL), the aim here is to disentangle the contribution of the satellite versus *in situ* components in improving the ocean state by separating impacts according to spatial and temporal scales. Note that the data assimilation system uses multi-variate approach meaning that SSH, temperature, salinity corrections are dynamically consistent. ONLYSAT assimilating SSH will provide information on SH, as well as ONLSITU assimilating temperature and salinity (indirectly SH).

3.2.1 Added value of satellites for mesoscale activity

Figure 5C shows the residual errors for steric height at intermediate scales. Even though the magnitude of residual error in ONLYSITU is lower than in FREE, it remains higher than the variability at almost all latitudes. In contrast, an important reduction of the residual error is seen in ONLYSAT, especially at latitudes of western boundary currents where the residual error decreases to 3-4 cm, in comparison to 8-10 cm in FREE. With a residual error at intermediate scales similar to that of ONLYSAT, NOMINAL benefits from satellite observations. This is not a surprise, as this is consistent with the scales of conventional one-dimensional nadir-looking altimeters having the ability to resolve wavelengths down to about 50–150 km depending on the specific satellite and geographic locations (Dufau et al., 2016; Ballarotta et al., 2019).

Sea surface height information provided by altimetry is a depth-integrated quantity and is closely related to steric height. Note that SSH and SH differ by the barotropic component (mass-related component) which is not considered in SH. A question arises on how the information provided by satellites is projected at depth. Some indications are given in Figures 6A, B, showing the globally averaged percentage of the Nature Run represented variance at intermediate scales for each experiment. First, for intermediate scales, the surface layer is better constrained by observations than the deeper layer, both for temperature and salinity, for all experiments. Then, satellite observation impacts (ONLYSAT) are clearly seen in both temperature and salinity variables at intermediate scales on the whole water column. In temperature, ONLYSAT leads to an improvement of 15% of variance at 500 m depth compared to FREE, to 50% at the surface. In salinity, the improvement is more modest: 5% of salinity variance at 500m depth to 30% at the surface. In comparison, the contribution of *in situ* observations (ONLYSITU) can represent up to half of the improvement seen in ONLYSAT at intermediate scales.

3.2.2 Added value of *in situ* observations for large-scale variability

At larger scales, residual errors in OSSEs are significantly lower than variability for all experiments (Figure 5D), including the unconstrained FREE experiment, likely due to the more predictable large-scale response of the ocean circulation to atmospheric forcing at low frequencies (Wunsch, 1998). Compared to FREE, the reduction of the residual error is more important in ONLYSITU than in ONLYSAT, suggesting that *in situ* observations provide a unique large-scale information to the analysis (similar residual error between ONLYSITU and NOMINAL). This is consistent with the characteristics of the Argo array which provides a global coverage of the upper ocean on broad spatial scales, O(1000 km), and on time scales of months and longer (Roemmich and Gilson 2009; Riser et al., 2016). The higher impact of *in situ* compared to satellites observations

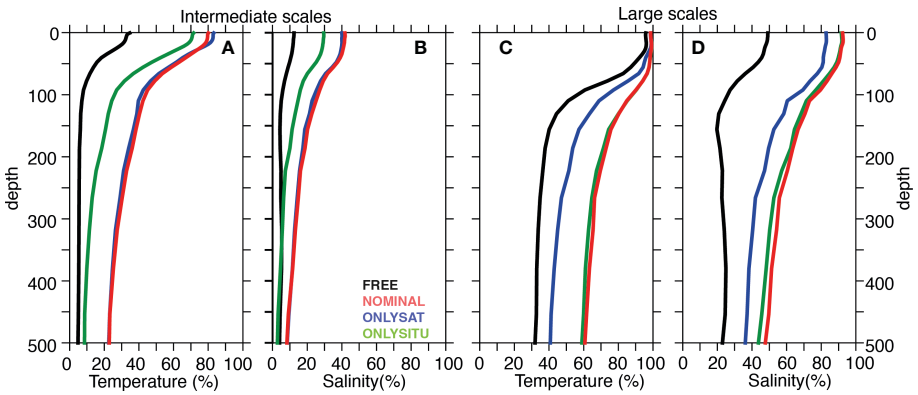


FIGURE 6
Globally averaged percentage of Nature Run represented variance for subsurface temperature and salinity at intermediate **(A, B)** and large scales **(C, D)** from the FREE, NOMINAL, ONLYSAT, ONLYSITU experiments. Timeseries have been filtered with running mean filters (1°x1°x20-day, 9°x9°x90-day) as explained in the text.

is confirmed in Figures 6C, D. In the FREE experiment, the percentage of represented variance of temperature is close to one at the surface and decreases to 30% in depth. *In situ* observations (ONLYSITU) improves the large-scale thermohaline stratification by increasing the percentage of represented variance by around 30% at depth for temperature, and up to 50% for salinity. In comparison, improvement from satellites observations (ONLYSAT) is lower, demonstrating a clear added value of *in situ* observations at large scales/low frequency.

The computation of statistics on residual errors by separating spatial and temporal scales highlights the evidence for the complementary role of satellites and *in situ* observations for constraining ocean analysis, with intermediate scales (mesoscale) mostly constrained by satellites and large scales by *in situ* observations. However, the reduction of analysis error at mesoscale is not null in ONLYSITU, as at large scales for ONLYSAT experiment, and the residual error reduction in NOMINAL is not directly related to the sum of ONLYSAT and ONLYSITU improvements. Both *in situ* and satellites observations provide a redundant information with regards to ocean analysis at a given spatial and temporal scale. This can explain the higher impact of *in situ* observations in ocean analysis when satellites observations are not assimilated (e.g., Zhu et al., 2021).

4 Potential outcomes of *in situ* observing system enhancements

The next objective of the present work is to evaluate potential outcomes of enhancing the *in situ* ocean observing system. One of the major evolutions of the *in situ* observing system recommended by the international community (Oceanobs'19 conference) is to double the number of Argo floats in western boundary currents and in equatorial regions and to enhance the vertical resolution of temperature and salinity measurements in the upper ocean layer on tropical moorings (Foltz et al., 2019; Smith et al., 2019; Hermès et al., 2019). Such evolutions have been evaluated based on three additional experiments (ENHANCED_AR, ENHANCED_MO, ENHANCED_AR_MO; Table 1).

4.1 Doubling Argo in western boundary currents

Western boundary currents are a fundamental element in the ocean circulation system given their impact on weather and climate both locally and remotely, on time scales from days to decades. Yet, it has been recognized that estimates of heat and freshwater contents have still large uncertainties due to insufficient sampling (e.g., Palmer et al., 2019, Todd et al., 2019). With high levels of mesoscale variability, enhancing Argo sampling in western boundary current is expected to reduce noise in tracking the temperature and salinity fields (Roemmich et al., 2019). The impact of doubling the number of Argo floats on constraining oceanic analyses will thus be assessed based on the experiment ENHANCED_AR, in which the number of Argo profiles has been doubled in western boundary current regions. For the assessment, we use integrated quantities as the ocean heat content (OHC) and the ocean freshwater content (OFC). The 0-700m OHC is computed based on the depth-integration of temperature anomaly from the 2016-2017 temporal mean multiplied by the heat capacity (3900 J/kg/°C) and ocean density of reference (1024 kg/m^3). Considering that about 3 cm of freshwater are needed to dilute 1 m of seawater by 1 psu, the OFC, expressed in meters, is the depth-integral of salinity anomaly multiplied by -0.03 (Gasparin and Roemmich, 2016).

Figure 7 shows the percentage of the OHC and OFC represented variance of the Nature Run for each experiment, area-averaged in the five main western boundary current regions (see Figure 1). OHC and OFC time series have been previously filtered to separate intermediate to large-scale variability. As expected, large-scale structures are better represented than intermediate scale features, with more than 90% of the OHC variance (Figure 7A) and more than 50% of the OFC variance (Figure 7C) represented by the NOMINAL experiment (in red). In agreement with previous results, *in situ* observations are generally more efficient than satellites observations in constraining the large-scale OHC and OFC variability (comparing blue and green bars in Figures 7A, C). This is more obvious for OFC. In addition, important differences between ONLYSAT (in blue) and ONLSITU (in green) experiments to represent both intermediate scales OHC and

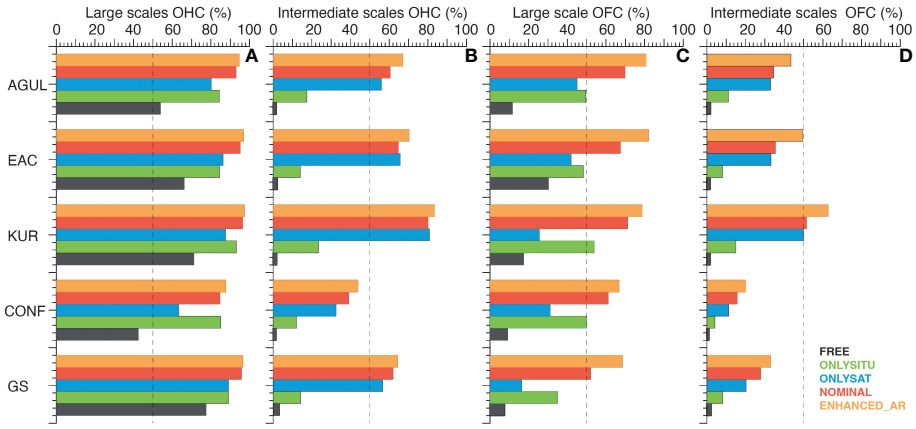


FIGURE 7
Percentage of the Nature Run represented variance, area-averaged in western boundary current regions, for 0-700 m Ocean Heat (OHC, **A, B**) and Freshwater Contents (OFC, **C, D**) at mesoscale and larger scales based on the FREE (black), NOMINAL (red), ONLYSAT (blue), ONLYSITU (green) and ENHANCED_AR (orange) experiments.

OFC confirm the key role of satellites observations to constrain mesoscale variability (Figures 7B, D). The complementarity of both observing system components is confirmed since the percentage of represented variance is systematically higher in NOMINAL than in ONLYSITU or ONLYSAT, especially in large-scale OFC where NOMINAL percentages of the Nature Run represented variance are 10 to 20% higher than that of ONLYSITU.

The added value of doubling the number of Argo floats is then assessed by comparing NOMINAL with ENHANCED_AR (in orange). Compared to NOMINAL, the slightly increased percentages of represented variance in ENHANCED_AR for both intermediate and large-scale OHC suggest that doubling Argo in western boundary current regions only has a limited impact on OHC. In contrast, substantial gain is seen for salinity (OFC) at both intermediate and large scales with improvements reaching more than 10% (e.g, Agulhas, EAC, Kuroshio regions). These results are qualitatively consistent with the reduction of the RMS error in the Gulf Stream from the multi-system approach of Gasparin et al. (2019).

4.2 Argo doubling and mooring enhancements in tropics

We now focus on the potential impacts of enhanced observing systems in the tropical oceans. Several extensions of the Argo and moored arrays have been proposed, both to help constraining the ocean state *via* their ingestion in data assimilation systems, and to help understanding critical processes not well sampled (Cravatte et al., 2016). These enhancements consist in a finer vertical resolution in the upper 100 meters, and an enhanced meridional resolution, with Argo doubling and more sensors on moorings in the upper ocean. The comparison, based on experiments in which Argo and moorings arrays have been enhanced separately or together (ENHANCED_AR, ENHANCED_MO and ENHENCED_AR_MO), aims at determining the relative and combined impacts of *in situ* enhancements. Results are shown for the Pacific Ocean, but conclusions can be extended to other tropical basins.

Figures 8A-D indicates the percentage of Nature Run represented variance for the 0-100m temperature and salinity at intermediate and large scales for NOMINAL and ENHANCED_AR_MO experiments, zonally averaged in the tropical Pacific. While NOMINAL can capture more than 70% of the Nature Run variance for temperature and salinity at both intermediate and large scales, only slight improvements are seen in ENHANCED_AR_MO (less than 5% higher than NOMINAL, area-averaged in the tropical Pacific). However, zonal averages mask a more complex behavior of the data assimilation system. In Figures 8E, F the difference of the percentage of the Nature Run represented variance at intermediate scales between ENHANCED_AR_MO and NOMINAL experiments for the 0-100 m layer indicates that *in situ* enhancements generally reduce residuals errors of temperature and salinity (Figures 8A, B), but improvement can reach 35% of the Nature Run variance of salinity in the western Pacific, while degradations are seen in some areas (higher residuals in ENHANCED_AR_MO than in NOMINAL).

The higher improvement in salinity results from the increased number of subsurface salinity measurements in ENHANCED_AR_MO (no salinity at mooring locations in NOMINAL), and since salinity is less constrained by altimetry as sea level variations are dominated by the thermosteric component (Storto et al., 2017). The degradation at some locations in ENHANCED_AR_MO might be due to several reasons. First, the number of temperature observations on moorings is decreased off-equator in ENHANCED_AR_MO compared to NOMINAL (See Figure 2). The percentage of variance, based on the ratio of error variance and signal variance, computed on the 2016-2017 period, underestimates *in situ* enhancement impacts due to the strong variability associated with the 2015-2016 El Nino. It is also noteworthy that data assimilation systems are built on subtle balances and conservation laws, data assimilation can also result in small degradations (e.g., Waters et al., 2017).

It is now possible to separate the effects of each component of the *in situ* observing system by comparing the percentage of Nature Run represented variance for temperature and salinity at intermediate and

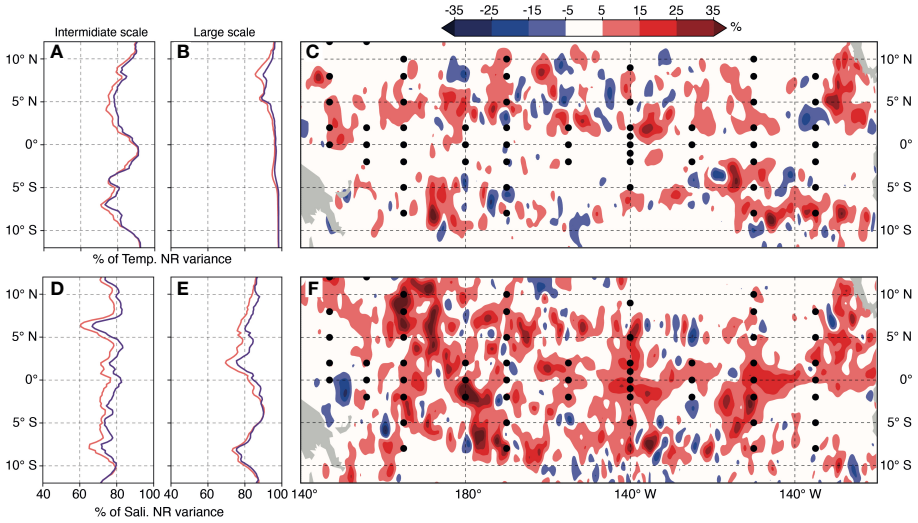


FIGURE 8
Zonally averaged percentage of Nature Run represented variance of NOMINAL (red) and ENHANCED_AR_MO (purple) experiments for the 0-100m temperature **(A, B)** and salinity **(D, E)** at intermediate and large scales. **(C, F)** ENHANCED_AR_MO-minus-NOMINAL difference of percentage of Nature Run represented variance at intermediate scales for the 0-100m **(C)** temperature and **(F)** salinity.

large scales, zonally averaged in the tropical Pacific, for ENHANCED_AR and ENHANCED_MO experiments (Figure 9). Interesting features are seen. First, temperature is not significantly improved at both intermediate and large scales, except off-equator where doubling Argo (ENHANCED_AR) increases the percentage of represented variance of 5% on average. Note that the percentage of variance is close to 90% in the equatorial band. The representation of the 0-100m salinity at intermediate scales is improved in both ENHANCED_AR and ENHANCED_MO. While doubling Argo benefits to all latitudes, mooring enhancements only provide a better estimate in the 4°S-4°N band. Such equator/off-equator differences are due to the fixed-point characteristics of moorings with a smaller number of salinity measurements in addition to shorter scales dynamics off-equator. At large scales, salinity benefits similarly from both Argo and moorings enhancements.

Similar diagnostics are applied to integrated quantities such as mixed layer depth (MLD) and barrier layer thickness (BLT) computed following the definition of de Boyer Montégut et al. (2004). Figure 10 shows the percentage of represented Nature Run variance at intermediate scales, zonally averaged in the western Pacific between 135°E and 155°E, for NOMINAL and the ENHANCEDs experiments. Different behaviors are seen. First, doubling Argo (ENHANCED_AR, yellow) systematically provides a better estimate than NOMINAL (red), while benefits from moorings do not occur at

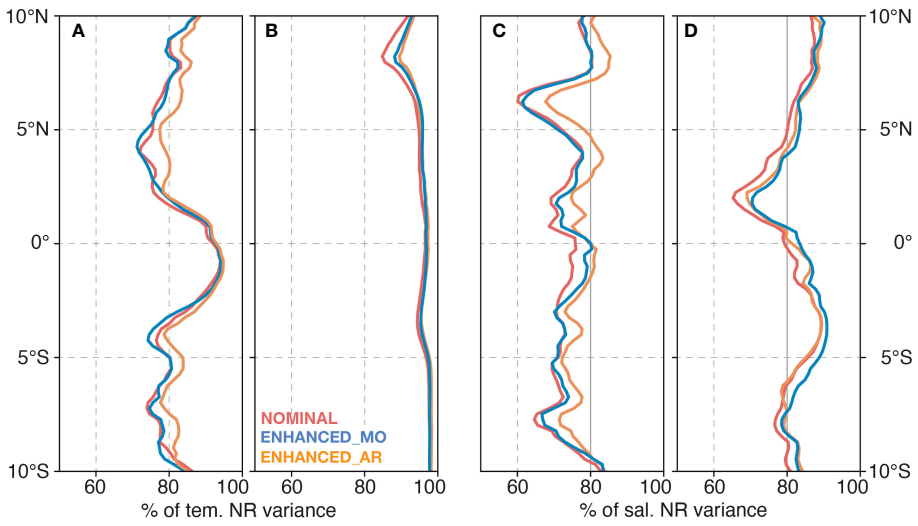


FIGURE 9
Percentage of Nature Run (NR) temperature **(A, B)** and salinity **(C, D)** represented variance at intermediate **(A, C)** and large **(B, D)** scales for NOMINAL (in red), ENHANCED_MO (blue), ENHANCED_AR (orange) experiments, zonally averaged in the tropical Pacific.

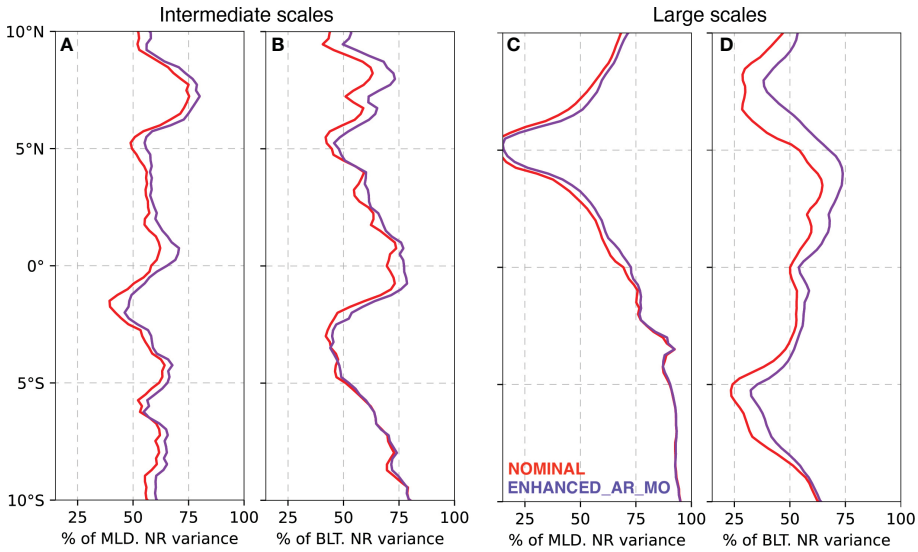


FIGURE 10
Percentage of Nature Run represented variance of Mixed Layer Depth and Barrier Layer Thickness at intermediate scales **(A, B)** and large scales **(C, D)** for the NOMINAL and ENHANCED_AR_MO experiments, zonally averaged in the western Paci!c (140E-180°E).

all latitudes. Even if ENHANCED_AR_MO often provides the best estimates, suggesting that both Argo and mooring data assimilation are complementary (south of 2°N), dedicated work is still necessary to make a better use of mooring observations (e.g., at 3°N).

Thus, *in situ* enhancements slightly improve the representation of temperature and salinity fields at both intermediate and large scales, with a strong regional dependency. A higher impact is seen in salinity (not directly constrained by altimetry), at both intermediate and large scales. However, as data assimilation techniques favor local impacts, moorings generally provide a highly accurate representation near the mooring points (at the thermocline level, Figure 11), while Argo gives a more uniformly accurate estimate across the basin (Figure 10A).

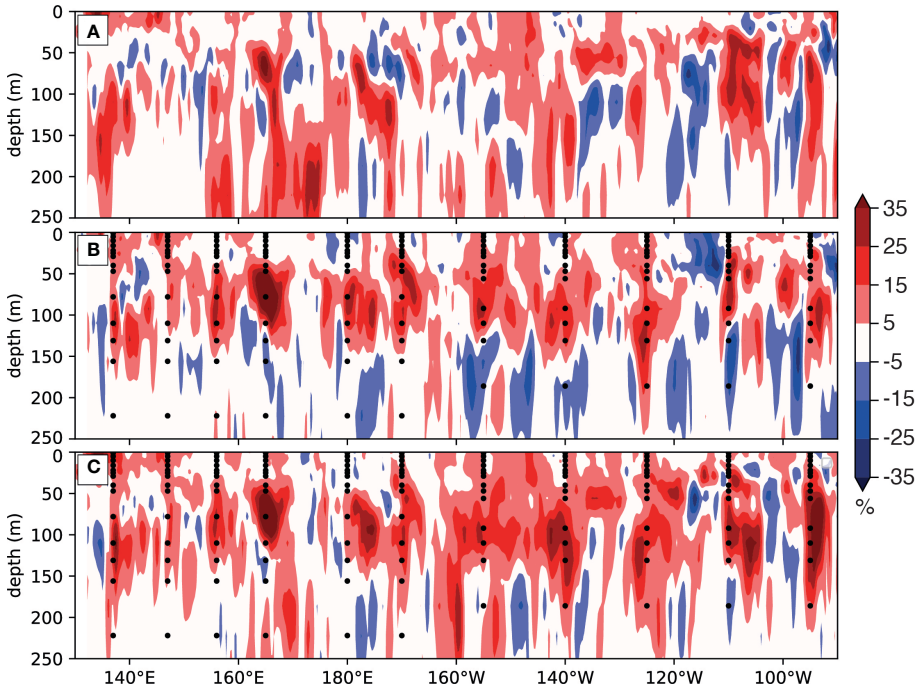


FIGURE 11
Difference of percentage of Nature Run (NR) salinity represented variance at intermediate scales from the NOMINAL experiment for **(A)** ENHANCED_AR, **(B)** ENHANCED_MO and **(C)** ENHANCED_AR_MO experiments along the equatorial Pacific. Red colors represent improvement of enhanced experiments compared to NOMINAL.

The strong complementarity between these two components is seen since both type of information benefit to the data assimilation system (Figure 11C), in agreement with previous studies (e.g., Gasparin et al., 2015; Zhu et al., 2021). It is important to mention that positive impacts of additional observations also result in some degradation (below 150m in Figure 11B). Further investigations are thus required to optimize the use of ocean observations and avoid spurious effects of data assimilation.

5 Discussion and conclusion

Based on a series of observing system simulation experiments (OSSE), a detailed description of the impacts of various ocean observations (from satellites and *in situ* networks) assimilated in a global data assimilation system is presented here. One of the important novelties here is the decomposition of the signals (both residual errors and percentage of represented variance) into specific temporal and spatial scales (i.e., small scales, intermediate scales and large scales). This allows a better understanding of the relative contributions of the observing system components in constraining various processes.

In general, mesoscale processes dominate residual errors in simulations with no data assimilation, especially in western boundary currents, while they are equivalent to large-scale residual errors in mid-latitude and low-latitude regions. Our experiments assimilating conjointly satellites and *in situ* data sets demonstrate the ability of an analysis system to represent 40-80% of the temperature variance at intermediate scales (20-30% for salinity), and more than 80% of the variance for large-scale variability. An important complementarity of satellites and *in situ* observations is shown; satellites information, mostly through altimetric data, strongly constrain mesoscale variability, while *in situ* data, mostly through Argo floats, provide a large-scale information. Each observing system component provides a substantial added value for ocean state estimates depending on the selected spatial and temporal scales. In addition to depth-integrated quantities, such as steric height, the added value of the observing system is clearly seen for subsurface temperature and salinity.

In addition to the evaluation of the current observing system design, our numerical experimental system also assessed the expected impact of future enhancements in observing systems, as recommended by GOOS. One of the main recommended enhancements is the doubling of the number of Argo floats in western boundary currents regions. It is demonstrated that the representation of the ocean freshwater content in these regions would be significantly improved, at both intermediate (mesoscale) and larger scales, whereas the representation of ocean heat content would be mostly improved at mesoscale. Enhancements of moorings observations are mostly through additional temperature and salinity sensors in the tropical Pacific (and a modification of the mooring locations in the tropical Pacific), and an increase in the vertical resolution in the surface layer in the Indian and Atlantic basins. In addition to substantial improvements at large scales, impacts of these enhancements are mostly seen locally as also seen in several studies (e.g., Gasparin et al., 2019; Zhu et al., 2021). Further developments of data assimilation systems might be needed to make a better use of these observations in models.

Unlike many other data assimilation systems, the benefits of *in situ* ocean observations are estimated here as a complementary information to satellite data sets (Fujii et al., 2019). The number of experiments, including simulations that assimilate observing system components separately, is a central point of this study and demonstrates the complexity of impact studies in a multivariate system. As expected, benefits of *in situ* observations are lower in the context of the integrated observing system. One limitation of this study is that results are obtained with a particular operational system. Efforts have been made to extract important messages that can likely be applied to other ocean analysis systems. However, it is noteworthy that observations impacts can be dependent on the data assimilation techniques and settings, and that similar experiments should be made for other systems to complement our results. More advanced techniques might hopefully increase the gain of a specific observing system component.

Finally, it is worth pointing out that the impact of ocean observations on analyses is addressed here only from the assimilation perspective. Ocean observations have multiple other benefits: they are key at the different steps of the development and qualification of an analysis system: they foster forecasting system advancement through a better understanding of the climate system; they allow validation and verification. All these indirect contributions are also very important for the analysis accuracy, albeit less visible.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

FG designed and wrote the manuscript. FG and J-ML generated synthetic observations, setted up the experimental system and performed experiments. FG conducted most of the analyses, with substantial contributions from J-ML, SC, GR and ER. Lastly, PL and BR helped in reading, editing figures and commenting on the paper. All authors contributed to the article and approved the submitted version.

Funding

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No.892626 (EUROSEA).

Acknowlegments

This study was conducted using Copernicus Marine Service Products. This paper used data collected and made freely available by programs that constitute the Global Ocean Observing System and the national programs that contribute to it (http://www.ioc-goos.org/), i.e., the Argo data collected and made freely available by the International Argo Project and the national programs that contribute

to it (http://argo.jcommops.org), the TAO-PIRATA-RAMA (PMEL-NOAA).

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2023.1021650/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Amplitude of errors (both representation and instrumental) associated to synthetic observations for 100-m temperature and 10-m salinity: **(A, C)** spatial maps, **(B, D)** zonally averaged (black line). As a comparison, zonally averaged prescribed error in the operational system is also shown in red line. Amplitude of synthetic observation errors (both representation and instrumental errors) in synthetic observations for 100-m temperature: **(A)** spatial maps, **(B)** zonally averaged (black line) and for 10-m salinity **(C)** spatial maps, **(B)** zonally averaged (black line). As a comparison, zonally averaged prescribed error in the operational system is also shown in red line in **(B)** and **(D)**.

SUPPLEMENTARY FIGURE 2
Amplitude of errors in synthetic temperature observations at three moorings locations in the equatorial Pacific (165°E, 140°W, 110°W). For comparison to the error estimate (black lines), prescribed errors in the operational system (red lines) and high-frequency variability from moorings (<7 days; blue crosses) are shown.

References

Aydogdu, A., Pinardi, N., Pistoi, J., Martinelli, M., Belardinelli, A., and Sparnocchia, S. (2016). Assimilation experiments for the fishery observing system in the Adriatic Sea. *J. Mar. Syst.* 162, 126–136. doi: 10.1016/j.jmarsys.2016.03.002

Ballarotta, M., Ubelmann, C., Pujol, M.-I., Taburet, G., Legeais, J.-F., Faugere, Y., et al. (2019). On the resolution of ocean altimetry maps. *Ocean Sci.* 156. doi: 10.5194/os-2018-156

Balmaseda, M., and Anderson, D. (2009). Impact of initialization strategies and observations on seasonal forecast skill. *Geophys. Res. Lett.* 36 (1), L01701. doi: 10.1029/2008GL035561

Balmaseda, M., Anderson, D., and Vidard, A. (2007). Impact of argo on analyses of the global ocean. *Geophysical Res. Lett.* 34 (16), L16605. doi: 10.1029/2007GL030452

Balmaseda, M. A., Trenberth, K. E., and Källén, E. (2013). Distinctive climate signals in reanalysis of global ocean heat content. *Geophysical Res. Lett.* 40 (9), 1754–1759. doi: 10.1002/grl.50382

Bell, M. J., Schiller, A., Pierre-Yves, Le T., Smith, N. R., Dombrowsky, E., and Wilmer-Becker, K. (2015). An introduction to GODAE OceanView. *J. Of Operational Oceanography* 8, S2–S11. doi: 10.1080/1755876X.2015.1022041

Benkiran, M., and Greiner, E. (2008). Impact of the incremental analysis updates on a real-time system of the north Atlantic ocean. *J. Atmos. Ocean. Technol.* 25, 2055–2073. doi: 10.1175/2008JTECHO537.1

Benkiran, M., Ruggiero, G., Greiner, E., Le Traon, P.-Y., Rémy, E., Lellouche, J. M., et al. (2021). Assessing the impact of the assimilation of SWOT observations in a global high-resolution analysis and forecasting system part 1: Methods. *Front. Mar. Sci.* 8. doi: 10.3389/fmars.2021.691955

Bloom, S. C., Takacs, L. L., Da Silva, A. M., and Ledvina, D. (1996). Data assimilation using incremental analysis updates. *Mon. Weather Rev.* 124, 1256–1271. doi: 10.1175/1520-0493(1996)124<1256:DAUIAU>2.0.CO;2

Brasseur, P., and Verron, J. (2006). The SEEK filter method for data assimilation in oceanography: a synthesis. *Ocean Dyn.* 56, 650–661. doi: 10.1007/s10236-006- 0080- 3

Brodeau, L., Barnier, B., Gulev, S. K., and Woods, C. (2017). Climatologically significant effects of some approximations in the bulk parameterizations of turbulent air–sea fluxes. *J. Phys. Oceanogr.* 47, 5–28. doi: 10.1175/jpo-d-16-0169.1

Cabanes, C., Grouazel, A., von Schuckmann, K., Hamon, M., Turpin, V., Coatanoan, C., et al. (2013). The CORA dataset: validation and diagnostics of insitu ocean temperature and salinity measurements. *Ocean Sci.* 9, 1–18. doi: 10.5194/os-9-1-2013

Cipollone, A., Masina, S., Storto, A., and Iovino, D. (2017). Benchmarking the mesoscale variability in global ocean eddy-permitting numerical systems. *Ocean Dynamics* 67 (10), 1313–1333. doi: 10.1007/s10236-017-1089-5

Cravatte, S., Kessler, W. S., Smith, N., Wijffels, S. E., Yu, L.Contributing Authors (2016). *Executive Summary. First report of TPOS 2020* (Salisbury: TPOS), GOOS-215, pp. i-xii. Available online at http://tpos2020.org/first-report/.

Cummings, J., Bertino, L., Brasseur, P., Fukumori, I., Kamachi, M., Martin, M. J., et al. (2009). Ocean data assimilation systems for GODAE. *Oceanography* 22 (3), 96–109. doi: 10.5670/oceanog.2009.69

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.* 109, C12003. doi: 10.1029/2004JC002378

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERAInterim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc* 137, 553–597. doi: 10.1002/ qj.828

Ducet, N., and Le Traon, P.-Y. (2001). A comparison of surface eddy kinetic energy and reynolds stresses in the gulf stream and the kuroshio current systems from merged TOPEX/Poseidon and ERS-1/2 altimetric data. *J. Geophys. Res.* 106, 16603–16622. doi: 10.1029/2000JC000205

Dufau, C., Orsztynowicz, M., Dibarboure, G., Morrow, R., and Le Traon, P. Y. (2016). Mesoscale resolution capability of altimetry: present & future. *J. Geophys. Res. Oceans* 121, 4910–4927. doi: 10.1002/2015JC010904

Foltz, G. R., Brandt, P., Richter, I., Rodríguez-Fonseca, B., Hernandez, F., Dengler, M., et al. (2019). The tropical atlantic observing system. *Front. Mar. Sci.* 6. doi: 10.3389/fmars.2019.00206

Fu, L. L., Christensen, E. J., Yamarone, C. A.Jr., Lefebvre, M., Menard, Y., Dorrer, M., et al. (1994). TOPEX/POSEIDON mission overview. *J. Geophysical Research: Oceans* 99 (C12), 24369–24381. doi: 10.1029/94JC01761

Fujii, Y., Ogawa, K., Brassington, G. B., Ando, K., Yasuda, T., and Kuragano, T. (2015). Evaluating the impacts of the tropical pacific observing system on the ocean analysis fields in the global ocean data assimilation system for operational seasonal forecasts in JMA. *J. Oper. Oceanogr.* 8, 25–39. doi: 10.1080/1755876X.2015.1014640

Fujii, Y., Rémy, E., Zuo, H., Oke, P., Halliwell, G., Gasparin, F., et al. (2019). Observing system evaluation based on ocean data assimilation and prediction systems: On-going challenges and a future vision for designing and supporting ocean observational networks. *Front. Mar. Sci.* 6, 417. doi: 10.3389/fmars.2019.00417

Gasparin, F., Guinehut, S., Mao, C., Mirouze, I., Rémy, E., King, R. R., et al. (2019). Requirements for an integrated *in situ* Atlantic ocean observing system from coordinated observing system simulation experiments. *Front. Mar. Sci.* 6, 83. doi: 10.3389/fmars.2019.00083

Gasparin, F., Hamon, M., Rémy, E., and Le Traon, P. Y. (2020). How deep argo will improve the deep ocean in an ocean reanalysis. *J. Climate* 33 (1), 77–94. doi: 10.1175/JCLI-D-19-0208.1

Gasparin, F., and Roemmich, D. (2016). The strong freshwater anomaly during the onset of the 2015/2016 El niño. *Geophysical Res. Lett.* 43 (12), 6452–6460. doi: 10.1002/2016GL069542

Gasparin, F., Roemmich, D., Gilson, J., and Cornuelle, B. (2015). Assessment of the upper-ocean observing system in the equatorial pacific: the role of argo in resolving intraseasonal to interannual variability. *J. Atmos. Oceanic. Technol.* 32, 1668–1688. doi: 10.1175/jtech-d-14-00218.1

Good, S. A., Martin, M. J., and Rayner, N. A. (2013). EN4: quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. *J. Geophys. Res.* 118, 6704–6716. doi: 10.1002/2013JC009067

Halliwell, G. R., Srinivasan, A., Kourafalou, V., Yang, H., Willey, D., Le Hénaff, M., et al. (2014). Rigorous evaluation of a fraternal twin ocean OSSE system for the open gulf of Mexico. *J. Atmospheric Oceanic Technol.* 31 (1), 105–130. doi: 10.1175/JTECH-D-13-00011.1

Hermes, J. C., Masumoto, Y., Beal, L. M., Roxy, M. K., Vialard, J., Andres, M., et al. (2019). A sustained ocean observing system in the Indian ocean for climate related scientific knowledge and societal needs. *Front. Mar. Sci.* 6. doi: 10.3389/fmars.2019.00355

Hernandez, F., Bertino, L., Brassington, G., Chassignet, E., Cummings, J., Davidson, F., et al. (2009). Validation and intercomparison studies within GODAE. *Oceanography* 22 (3), 128–143. doi: 10.5670/oceanog.2009.71

Huang, B., and Wang, X. (2018). On the use of cost-effective valid-time-shifting (VTS) method to increase ensemble size in the GFS hybrid 4DEnVar system. *Mon. Weather Rev.* 146, 2973–2998. doi: 10.1175/MWR-D-18-0009.1

Jones, E. M., Oke, P. R., Rizwi, F., and Murray, L. (2012). Assimilation of glider and mooring data into a coastal ocean model. *Ocean Model.* 47, 1–13. doi: 10.1016/j.ocemod.2011.12.009

Kessler, W. S., Cravatte, S.Lead Authors, 2021. "GOOS-268," in *Final report of TPOS 2020*, vol. 83. . Available at: https://tropicalpacific.org/tpos2020-project-archive/reports/.

Lellouche, J. M., Greiner, E., Bourdallé-Badie, R., Garric, G., Melet, A., Drévillon, M., et al. (2021). The Copernicus global 1/12 oceanic and sea ice GLORYS12 reanalysis. *Front. Earth Sci.* 9, 698876. doi: 10.3389/feart.2021.698876

Lellouche, J.-M., Greiner, E., Le Galloudec, O., Garric, G., Regnier, C., Drevillon, M., et al. (2018). Recent updates to the Copernicus marine service global ocean monitoring and forecasting real-time 1/12° high-resolution system. *Ocean Sci.* 14, 1093–1126. doi: 10.5194/os-14-1093-2018

Lellouche, J.-M., Le Galloudec, O., Drévillon, M., Régnier, C., Greiner, E., Garric, G., et al. (2013). Evaluation of global monitoring and forecasting systems at Mercator océan. *Ocean Sci.* 9, 57–81. doi: 10.5194/os-9-57-2013

Le Traon, P. Y., Reppucci, A., Alvarez Fanjul, E., Aouf, L., Behrens, A., Belmonte, M., et al. (2019). From observation to information and users: The Copernicus marine service perspective. *Front. Mar. Sci.* 6, 234. doi: 10.3389/fmars.2019.00234

Levin, J., Arango, H. G., Laughlin, B., Hunter, E., Wilkin, J., and Moore, A. M. (2021). Observation impacts on the mid-Atlantic bight front and cross-shelf transport in 4D-var ocean state estimates: Part II–the pioneer array. *Ocean Model.* 157, 101731. doi: 10.1016/j.ocemod.2020.101731

Madec, G., and The NEMO Team. (2008). NEMO ocean engine, version 3.0. Note du Pôle de modélisation de l'Institut Pierre-Simon Laplace, 27. Paris: Institut Pierre-Simon Laplace (IPSL), pp. 217. Available at: https://epic.awi.de/id/eprint/39698/1/NEMO_book_v6039.pdf

Madec, G., Bourdallé-Badie, R., Bouttier, P.-A., Bricaud, C., Bruciaferri, D., Calvert, D., et al. (2017). NEMO ocean engine Paris, France: *notes Du pôle de modélisation de l'institut Pierre-simon Laplace (IPSL)*. doi: 10.5281/zenodo.1472492

Madec, G., and Imbard, M. (1996). A global ocean mesh to overcome the north pole singularity. *Clim. Dynam.* 12, 381–388. doi: 10.1007/BF00211684

McPhaden, M. J., Busalacchi, A. J., Cheney, R., Donguy, J.-R., Gage, K. S., and Halpern, D. (1998). The tropical ocean-global atmosphere observing system: A decade of progress. *J. Geophys. Res.* 103, 14169–14240. doi: 10.1029/97JC02906

Moltmann, T., Turton, J., Zhang, H. M., Nolan, G., Gouldman, C., Griesbauer, L., et al. (2019). A global ocean observing system (GOOS), delivered through enhanced collaboration across regions, communities, and new technologies. *Front. Mar. Sci.* 6, 291. doi: 10.3389/fmars.2019.00291

Moore, A. M., Martin, M. J., Akella, S., Arango, H. G., Balmaseda, M., Bertino, L., et al. (2019). Synthesis of ocean observations using data assimilation for operational, real-time and reanalysis systems: A more complete picture of the state of the ocean. *Front. Mar. Sci.* 6, 90. doi: 10.3389/fmars.2019.00090

Palmer, M. D., Durack, P., Chidichimo, M. P., Church, J., Cravatte, S. E., Hill, K. L., et al. (2019). Adequacy of the ocean observation system for quantifying regional heat and freshwater storage and change. *Front. Mar. Sci.* 6. doi: 10.3389/fmars.2019.00416

Riser, S. C., Freeland, H. J., Roemmich, D., Wijffels, S., Troisi, A., Belbeoch, M., et al. (2016). Fifteen years of ocean observations with the global argo array. *Nat. Clim. Change* 6, 145–153. doi: 10.1038/nclimate2872

Rodi, W. (1987). Examples of calculation methods for flow andmixing in stratified fluids. *J. Geophys. Res.* 92, 5305–5328. doi: 10.1029/jc092ic05p05305

Roemmich, D., Alford, M. H., Claustre, H., Johnson, K., King, B., Moum, J., et al. (2019). On the future of argo: a global, full-depth, multi-disciplinary array. *Front. Mar. Sci.* 6. doi: 10.3389/fmars.2019.00439

Roemmich, D., and Gilson, J. (2009). The 2004–2008 mean and annual cycle of temperature, salinity, and steric height in the global ocean from the argo program. *Prog. Oceanogr.* 82, 81–100. doi: 10.1016/j.pocean.2009.03.004

Shchepetkin, A. F., and McWilliams, J. C. (2005). The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate ocean model. *Ocean Modell.* 9, 347–404. doi: 10.1016/j.ocemod.2004.08.002

Shchepetkin, A. F., and McWilliams, J. C. (2008). *Handbook of numerical analysis: Computational methods for the ocean and the atmosphere* Vol. 14 (Amsterdam: Elsevier Science), 121–183.

Smith, N. R. (1993). Ocean modeling in a global ocean observing system. *Rev. Geophysics* 31 (3), 281–317. doi: 10.1029/93RG00134

Smith, N. R. (1995). An improved system for tropical ocean sub-surface temperature analyses. *J. Atmos. Oceanic Technol.* 12, 850–870. doi: 10.1175/1520-0426(1995)012<0850:aisfto>2.0.co;2

Smith, N., Kessler, W. S., Cravatte, S., Sprintall, J., Wijffels, S., Cronin, M. F., et al. (2019). Tropical pacific observing system. *Front. Mar. Sci.* 6:31. doi: 10.3389/fmars.2019.00031

Storer, B. A., Buzzicotti, M., Khatri, H., Griffies, S. M., and Aluie, H.. (2022). Global energy spectrum of the general oceanic circulation. *Nature communications*, 13(1), 5314. doi: 10.1038/s41467-022-33031-3

Storto, A., Masina, S., Balmaseda, M., Guinehut, S., Xue, Y., Szekely, T., et al. (2017). Steric sea level variability, (1993–2010) in an ensemble of ocean reanalyses and objective analyses. *Climate Dynamics* 49 (3), 709–729. doi: 10.1007/s00382-015-2554-9

Tanhua, T., McCurdy, A., Fischer, A., Appeltans, W., Bax, N., Currie, K., et al. (2019). What we have learned from the framework for ocean observing: Evolution of the global ocean observing system. *Front. Mar. Sci.* 6, 471. doi: 10.3389/fmars.2019.00471

Todd, R. E., Chavez, F. P., Clayton, S., Cravatte, S. E., Goes, M. P., Graco, M. I., et al. (2019). Global perspectives on observing ocean boundary current systems. *Front. Mar. Sci.* 6.

Turpin, V., Remy, E., and Le Traon, P. Y. (2016). How essential are argo observations to constrain a global ocean data assimilation system? *Ocean Sci.* 12 (1), 257–274.

Waters, J., Bell, M. J., Martin, M. J., and Lea, D. J. (2017). Reducing ocean model imbalances in the equatorial region caused by data assimilation. *Q. J. R. Meteorol. Soc* 143, 195–208. doi: 10.1002/qj.2912

Wunsch, C. (1998). The work done by the wind on the oceanic general circulation. *J. Phys.Oceanogr.* 28, 2332–2340. doi: 10.1175/1520-0485(1998)028<2332:twdbtw>2.0.co;2

Yu, Y., Gille, S. T., Sandwell, D. T., and McAuley, J. (2022). Global mesoscale ocean variability from multiyear altimetry: An analysis of the influencing factors. *Artif. Intell. Earth Syst.* 1 (3), e210008. doi: 10.1175/AIES-D-21-0008.1

Zhu, J., Vernieres, G., Sluka, T., Flampouris, S., Kumar, A., Mehra, A., et al. (2021). Roles of TAO/TRITON and argo in tropical pacific observing system: An OSSE study for multiple time scale variability. *J. Climate* 34, 6797-6817. doi: 10.1175/JCLI-D-20-0951.1

Zuo, H., Balmaseda, M. A., Tietsche, S., Mogensen, K., and Mayer, M. (2019). The ECMWF operational ensemble reanalysis–analysis system for ocean and sea ice: a description of the system and assessment. *Ocean Sci.* 15, 779–808. doi: 10.5194/os-15-779-2019

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 01 February 2023
DOI 10.3389/fmars.2023.1085118

Check for updates

OPEN ACCESS

EDITED BY
Shiqiu Peng,
State Key Laboratory of Tropical Oceanography (CAS), China

REVIEWED BY
Xianqing Lv,
Ocean Unversity of China, China
Li Yineng,
South China Sea Institute of Oceanology, Chinese Academy of Sciences (CAS), China

*CORRESPONDENCE
Jicai Zhang
Jicai_Zhang@163.com

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

RECEIVED 31 October 2022
ACCEPTED 20 January 2023
PUBLISHED 01 February 2023

CITATION
Qian S, Du Y, Wei Z, Zhang J, Cheng J, Wang D and Wang YP (2023) Effects of spatial bottom friction parameterization scheme on the tidal dynamics in the macrotidal East China Seas. *Front. Mar. Sci.* 10:1085118. doi: 10.3389/fmars.2023.1085118

COPYRIGHT
© 2023 Qian, Du, Wei, Zhang, Cheng, Wang and Wang. 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(s) 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.

Effects of spatial bottom friction parameterization scheme on the tidal dynamics in the macrotidal East China Seas

Suhui Qian[1], Yunfei Du[1], Zilu Wei[1], Jicai Zhang[2]*, Jun Cheng[3], Daosheng Wang[4] and Ya Ping Wang[2]

[1]Institute of Physical Oceanography and Remote Sensing, Ocean College, Zhejiang University, Zhoushan, China, [2]State Key Laboratory of Estuarine and Coastal Research, East China Normal University, Shanghai, China, [3]Department of Environmental and Sustainability Sciences, Kean University, Union, NJ, United States, [4]Hubei Key Laboratory of Marine Geological Resources/College of Marine Science and Technology, China University of Geosciences, Wuhan, China

In this study, the effects of different bottom friction coefficient (BFC) parameterization schemes on the modelling of four principal tidal constituents (M_2, S_2, K_1, O_1 tides) in the macrotidal East China Seas were investigated by using a high-resolution model based on FVCOM (Finite Volume Community Ocean Model). The applied BFC schemes include: the empirical constant (EC-BFC), sediment-dependent form (SD-BFC), and spatial varying BFC obtained from adjoint data assimilation (SV-BFC). The comparisons between the simulated results and the observations from satellite altimeters and tidal gauge stations indicated that the SV-BFC scheme is superior to others. The locations of amphidromic points calculated with EC-BFC and SD-BFC were in the northwest of those from SV-BFC. The variations in tidal dynamics between different BFC schemes were closely related to the spatial distributions of BFCs, especially in high-valued BFC areas, e.g., the West Korea Bay, the South Yellow Sea, and the eastern coasts of Jiangsu, Zhejiang and Fujian provinces. The tidal energy flux transporting into Bohai and Yellow Seas increased under the SV-BFC scheme, while smaller tidal energy flux transporting from the Korea Strait was generated by SV-BFC as compared to those from EC-BFC and SD-BFC. The high-valued BFC areas in the SV-BFC scheme dissipated larger amounts of tidal energy, and the average values of Simpson-Hunter numbers were lower than those with the other two schemes. However, the values of Simpson-Hunter numbers increased in the West Korea Bay and Jianghua Bay with high-valued BFCs because of the decreasing current velocity under the headland-shaped topography.

KEYWORDS

bottom friction, tide, energy dissipation, mixing, East China Seas

1 Introduction

The most important sink of oceanic kinetic energy is the bottom boundary layer in shallow seas where intensive dissipation occurs (Munk and Wunsch, 1998; Blakely et al., 2022). The bottom boundary layer (BBL), the interface between the seabed and the overlying water column, is where exchanges of particles (Dyer and Soulsby, 1988; Brink, 2016), chemicals (Huettel et al., 2014), and organisms (Cowen and Sponaugle, 2009) take place. Frictional dissipation of energy and turbulent mixing of mass, momentum, and heat are rather significant in these regions, and thus, the BBL plays an important role in the oceanic momentum balance (McWilliams, 2006; Trowbridge and Lentz, 2018). The dissipation mechanisms for global tides include bottom friction dissipation and internal tide dissipation during the conversion from barotropic to baroclinic tides (Munk, 1997). Global tidal energy dissipation was evaluated to be equal to 3.7 terawatts, and nearly 2.8-3.1 terawatts was allocated to the dissipation in the turbulent BBL of marginal seas (Munk and Wunsch, 1998).

Bottom friction plays an important role in shelf flows, but knowledge about this term is still inadequate. The bottom friction coefficient (BFC) was introduced to parameterize the bottom friction term usually with a quadratic function of the near-bottom current velocity (Mofjeld, 1988; Guo and Yanagi, 1998). In early estimates of bottom friction dissipation by Taylor (1920) and Jeffreys (1921), the bottom friction dissipation depended on the product of the BFC and tidal velocity cubed. Field measurements ensure that BFC is not a universal constant (Cheng et al., 1999; Fan et al., 2019; Bo and Ralston, 2020), i.e. with temporal and spatial variations, and is believed to be one of the main uncertainties for the evaluation of tidal energy dissipation (Munk and Wunsch, 1998). Several methods were suggested to determine the BFC. Zhao et al. (1993) simulated the semidiurnal and diurnal tides and tidal currents in the whole East China Seas with different BFC in different subdomains. By using the depth-dependent form of BFC, Kang et al. (1998) carried out a fine grid tidal modeling experiment to study the tidal phenomena in the Yellow and East China Seas. Pringle et al. (2018) presented a semidata-informed method to estimate spatially varying BFC from seabed and physical properties of the flow. Blakely et al. (2022) used a depth-dependent Manning's coefficient to optimize the boundary layer friction parameters and estimate the boundary layer dissipation. They concluded that altering friction values in high-energy dissipation areas has significant basin-scale impacts on tidal results. In shallow coastal seas, BFC can be affected by multiple factors (Cheng et al., 1999; Fan et al., 2019; Bo and Ralston, 2020; Qian et al., 2021), which results in the spatiotemporal variations in BFC. The spatiotemporal distributions of BFC had been investigated extensively by parameter estimations based on data assimilation techniques (Das and Lardner, 1991; Lu and Zhang, 2006; Zhang et al., 2011; Gao et al., 2015; Qian et al., 2021; Wang et al., 2021). A more reasonable spatial distribution of BFC of the East China Seas was obtained by assimilating multi-missions satellite altimeter observations into an adjoint tidal model (Qian et al., 2021). The temporal and spatial variations in the estimated BFCs were significantly correlated with the current speed and water depth, which ultimately induce the erosion-deposition of sediments on the seabed (Wang et al., 2021).

Overall, different parameterization schemes of BFC have been used in previous studies (Zhao et al., 1993; Kang et al., 1998; Lee and Jung, 1999; Egbert and Erofeeva, 2002; Egbert et al., 2004; Wang et al., 2014; Pringle et al., 2018; Chu et al., 2021). Wang et al. (2014) investigated the effects of BFC schemes on single-tidal simulation. Chu et al. (2019) studied the sensitivities of modelling storm surge to BFC schemes, and they also investigated the effects of BFC schemes on the modelling of shallow-water tides and tidal duration asymmetry. However, so far there are few systematic comparisons among the different schemes of BFC on the estimation of energy flux, oceanic mixing, and bottom friction dissipation. The goal of this study is to investigate the effects of various schemes of BFC on the tidal dynamics in the macrotidal East China Seas. More specifically, the following tasks will be achieved. Firstly, a high-resolution unstructured model is developed based on FVCOM (Finite Volume Community Ocean Model) to simulate the four principal tidal constituents (M_2, S_2, K_1, and O_1) in the macrotidal East China Seas. Secondly, different schemes of BFC including empirical constant (EC-BFC), sediment-dependent form (SD-BFC), and spatial varying BFC obtained from data assimilation (SV-BFC) are compared. Furthermore, the variations in oceanic energy, mixing and bottom friction dissipation between different BFC schemes are discussed.

The paper is organized as follows. Section 2 introduces the methodology, including the model configuration and experiment settings, and model verifications. Section 3 describes the sensitivity analysis and model results. The discussions are arranged in Section 4. The conclusions are drawn in Section 5.

2 Methodology

2.1 Model development

FVCOM model, with a non-overlapping unstructured triangular grid ideally to resolve dynamics in irregular complex coastlines, is used in this study (Chen et al., 2003; Chen et al., 2007). The model solves the momentum and mass conservation equations in integral form by computing fluxes between nonoverlapping horizontal triangular control volumes. This numerical approach also provides optimal representations of mass, momentum, salinity, and heat conservation in coastal and estuarine regions with complex geometries.

The bathymetry and mesh of the computational domain are shown in Figure 1. The mesh includes parts of the East China Sea to minimize the influence of open boundaries. The land boundary is limited by the coastline, and open boundaries are located in the northern, southern, and eastern parts of the domain. The unstructured triangular grid of the computational domain consists of 29916 nodes and 57125 elements with a spatial resolution of 0.5 km for coastal zones and decreased resolution up to 20 km towards open sea boundaries. In addition, seven uniform σ layers are specified in the vertical profiles. The hydrodynamics in the study area is dominantly driven by currents induced by barotropic tide (Li et al., 2018; Wu et al., 2018). Four principal tidal constituents (M_2, S_2, K_1, and O_1) were used to generate the tidal elevations along the open boundaries. The astronomical tidal constituents along open boundaries were derived from TPXO 7.2 established by the University of Oregon (http://volkov.oce.orst.edu/tides/TPXO7.2.html). The high-resolution bathymetry data for the coastal areas adjacent to Zhejiang province and the Yangtze estuary were provided

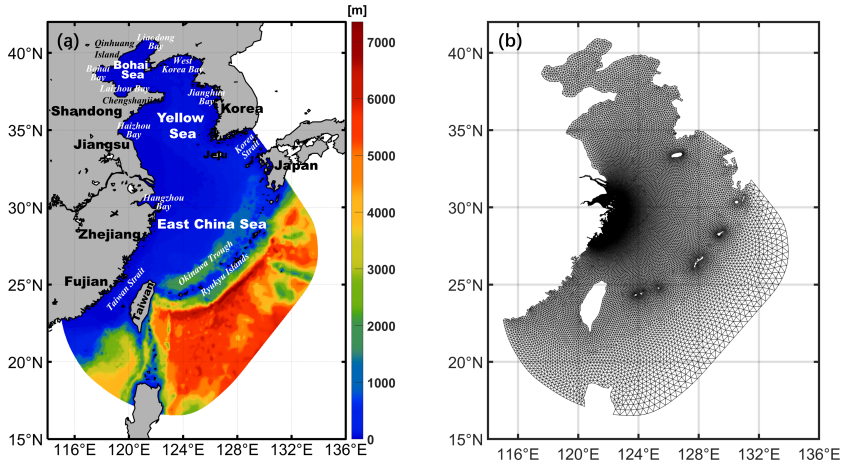


FIGURE 1
(A) Bathymetry of numerical model, **(B)** computational grid of study domain.

by the Ocean and Fisheries Bureau of Zhejiang Province. Bathymetry data for areas further offshore were obtained from Etopo1 (https://www.ngdc.noaa.gov/mgg/bathymetry/). All the bathymetry data were interpolated to the computational cells. The Mellor and Yamada level-2.5 turbulence closure scheme is adopted to parameterize the vertical mixing (Mellor and Yamada, 1982). The model used in this study is FVCOM v4.0, while the baroclinic effect is ignored. This model has been successfully applied to numerous estuaries and continental shelf areas (Chu et al., 2019; Zhang et al., 2021).

2.2 Numerical experiments

In this study, three BFC schemes are employed in the FVCOM to simulate the four principal tidal constituents. Generally, the BFC parameterizations can be summarized as: an empirical constant, different constants in different subdomains, depth-dependent form with Chezy coefficient, sediment form, and the spatial varying BFC obtained from data assimilation. Wang et al. (2014) concluded that the simulated M_2 tide in the East China Seas with the first three BFC schemes had larger discrepancies compared with field observations. Therefore, the employed BFC schemes in this study are described as follows.

EXP 1 (EC-BFC): The BFC is treated as a constant in the East China Seas, BFC = 0.0022. (Qian et al., 2021).

EXP 2 (SD-BFC): The BFCs are estimated by the semidata-informed method with the knowledge of seabed sediments and the physical properties of the flow (Pringle et al., 2018) (Figure 2A).

EXP 3 (SV-BFC): The spatially varying BFCs in the East China Seas were obtained by assimilating multi-mission satellite observations from TOPEX/Poseidon, Jason-1, and Jason-2 into a tidal model with an adjoint method (Qian et al., 2021) (Figure 2B).

The spatially varying bottom friction coefficient data of the SD-BFC scheme and SV-BFC scheme were interpolated into the computational cells. The model was launched with a cold start. With the assumption of zero heat flux, the simulation time of the model lasted from 1 February 2012 to 31 March 2012, and the modeling results from 17 March to 31 March (15 days) were used for analysis.

2.3 Model verification

Observations from altimeter cross points and tidal gauge stations are used to evaluate the model-simulated tidal constituents (Figure 3). The T_tide toolbox (Pawlowicz et al., 2002) was used to analyze the harmonic constants of two diurnal constituents (K_1 and O_1 tides) and

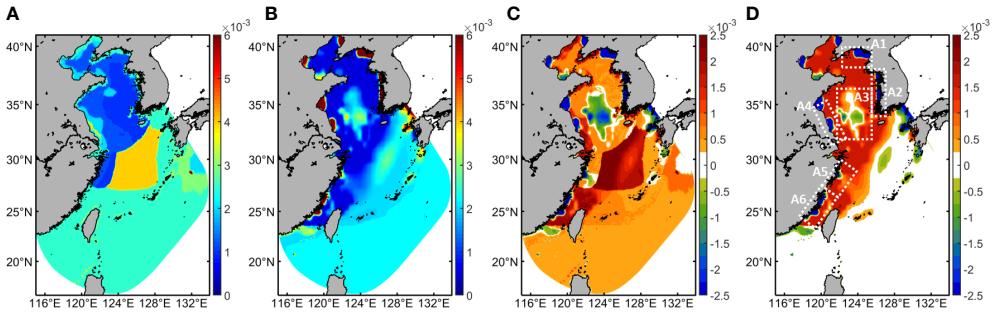


FIGURE 2
(A, B) Spatial distributions of BFCs of EXP 2 and EXP 3, respectively. **(C)** Differences between the BFCs of EXP 2 and those of EXP 3 ($BFC_{EXP2} - BFC_{EXP3}$). **(D)** Differences between the BFCs of EXP 1 and those of EXP 3 ($BFC_{EXP1} - BFC_{EXP3}$). White dashed boxes mark the areas of the Korea Bay (A1), the Jianghua Bay (A2), the middle South Yellow Sea (A3), and the eastern coasts of Jiangsu (A4), Zhejiang (A5) and Fujian provinces (A6).

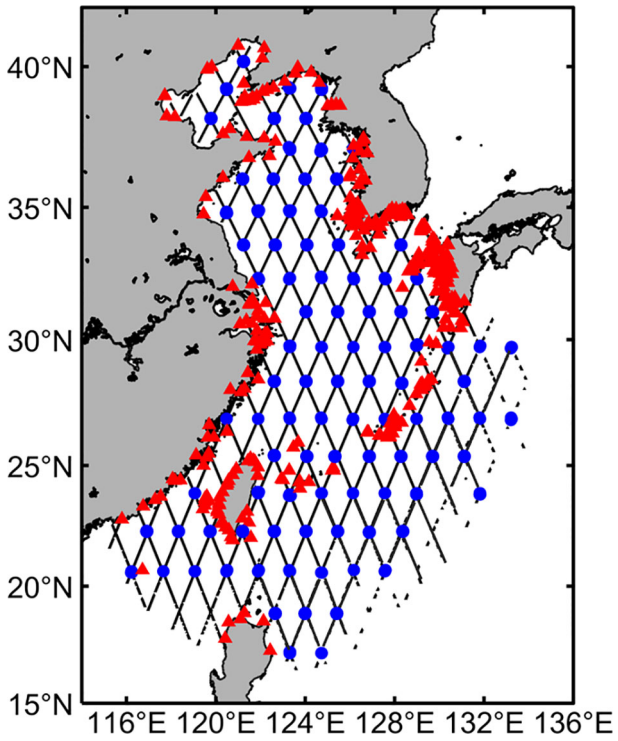


FIGURE 3
Locations of tidal gauge stations (red triangles) and cross points of altimeters (blue dots). The black solid lines show the altimeter tracks.

two semidiurnal constituents (M_2 and S_2 tides). To evaluate the effectiveness of the model, three skill parameters were calculated to quantify the differences between observations and simulations. The parameters were computed as follows (Wang et al., 2012; Chu et al., 2019; Chu et al., 2021; Blakely et al., 2022).

The correlation coefficien

$$CC = [\sum_{i=1}^{n}(X_m - \overline{X_m})(X_o - \overline{X_o})]/[\sum_{i=1}^{n}(X_m - \overline{X_m})^2\sum_{i=1}^{n}(X_o - \overline{X_o})^2]^{1/2} \quad (1)$$

where n is the number of the variable values; X_m and $\overline{X_m}$ are time-varying model results and time-averaged values, respectively; X_o and $\overline{X_o}$ are time-varying values of observed results and time-averaged values, respectively.

(2) The root-mean-square discrepancy was used to evaluate both the amplitude and phase lag of the error in one metric (Wang et al., 2012; Blakely et al., 2022)

$$DH = \sqrt{0.5\sum_{k}[(A_o^k)^2 + (A_m^k)^2 - 2A_o^kA_m^kcos(\theta_o^k - \theta_m^k)]} \quad (2)$$

$$\overline{DH} = \frac{1}{N}\sum_{j=1}^{N}DH(j) \quad (3)$$

where A and θ are the amplitude and phase lag of the kth constituent, o denotes observed amplitude/phase lag, and m denotes simulated amplitude/phase lag.

(3) The relative bias

$$RB = [\sum_{i=1}^{n}(X_m - X_o)^2/\sum_{i=1}^{n}|X_o|] \times 100\,\% \quad (4)$$

Figure 4 shows the comparisons between the simulated and observed amplitudes and phase lags of M_2, S_2, K_1, and O_1 tides from tidal gauge stations. The differences between simulated results and observations from tidal gauge stations are calculated and listed in Table 1. The correlation coefficients between the simulated amplitudes of M_2, S_2, K_1, and O_1 tides and observations from altimeter cross points in EXP 1-3 are larger than 0.92, 0.88, 0.88, and 0.91, respectively. The correlation coefficients between the simulated phase lags of M_2, S_2, K_1, and O_1 tides and observations from altimeter cross points in EXP 1-3 are larger than 0.89, 0.95, 0.95, and 0.85, respectively. For the amplitude, the relative bias of the M_2, S_2, K_1, and O_1 constituents between the observed and simulated harmonic constants in EXP 1-3 are smaller than 4.3%, 5.2%, 4.8%, and 4.9%, respectively. For the phase lag, the relative bias of M_2, S_2, K_1, and O_1 constituents in EXP 1-3 are smaller than 6.3%, 3.2%, 3.7%, and 6.2%, respectively. The $\overline{DH}$ between the simulated results and observations from satellite altimeters of M_2, S_2, K_1, and O_1 tides are (10.7, 15.7, 17.7, 17.4 cm) in EXP 1, (11.4, 15.7, 17.3, 18.0 cm) in EXP 2, and (10.5, 15.2, 17.2, 16.1 cm) in EXP 3. According to the calculations of the average absolute differences, correlation coefficients and root-mean-square discrepancies, the simulated results of EXP 3 fit the observations best, which has the smallest differences and highest correlation coefficient.

The tidal characteristics in the East China Seas are discussed based on the distributions of co-amplitude, co-phase, and tidal current ellipses of the four principal tidal constituents generated from the harmonic analysis of model results. As the simulated results from the SV-BFC scheme have the smallest root-mean-square discrepancy, the cotidal charts of M_2, S_2, K_1, and O_1 tides in EXP 3 are depicted in Figures 5A–D. The four principal tidal

10.3389/fmars.2023.1085118

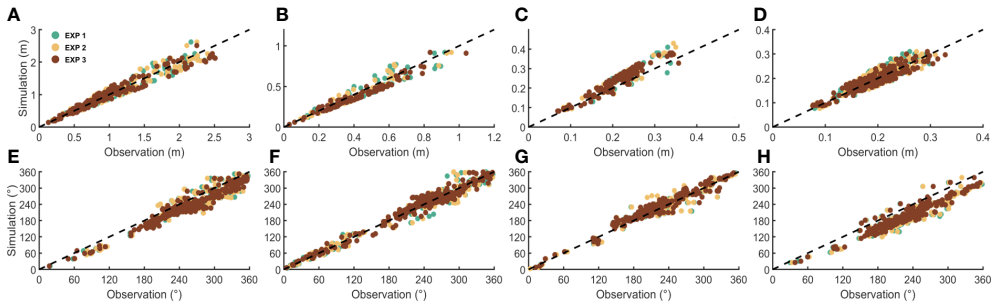


FIGURE 4
(A–D) Comparisons between the simulated and observed amplitudes of M_2, S_2, K_1, and O_1 tides from tidal gauge stations, respectively. Colors denote different BFC schemes. **(E–H)** Comparisons between the simulated and observed phase lags of M_2, S_2, K_1, and O_1 tides, respectively.

TABLE 1 The differences of tidal harmonic constants compared with observations from tidal gauge stations.

			EXP 1	EXP 2	EXP 3
Average absolute difference	M_2	AMP (cm)	4.3	4.1	3.7
		PHA (°)	12.1	11.6	10.9
	S_2	AMP (cm)	6.7	6.5	6.4
		PHA (°)	7.0	6.7	6.7
	K_1	AMP (cm)	5.7	5.7	5.6
		PHA (°)	7.3	7.3	7.2
	O_1	AMP (cm)	5.2	5.4	5.1
		PHA (°)	15.1	14.7	14.8
Root mean square discrepancy (cm)	M_2		9.9	10.5	9.9
	S_2		16.0	16.0	16.0
	K_1		20.6	21.7	20.7
	O_1		20.3	21.1	20.6
Correlation coefficient	M_2	AMP	0.94	0.94	0.96
		PHA	0.90	0.92	0.95
	S_2	AMP	0.96	0.97	0.97
		PHA	0.86	0.86	0.87
	K_1	AMP	0.94	0.95	0.96
		PHA	0.87	0.87	0.88
	O_1	AMP	0.93	0.94	0.94
		PHA	0.88	0.87	0.88

AMP denotes the difference in tidal amplitude, PHA denotes the difference in tidal phase lag.

constituents amplified in the coastal regions under the effect of shoaling and narrowing, and the study area is significantly dominated by semidiurnal tides. The simulated results show that the tidal currents propagating from the Pacific Ocean into the East China Sea are affected by coastal topography, and tidal currents mainly propagate as rotating waves in a counterclockwise direction. The co-amplitude lines of the semidiurnal tides show that the largest amplitude, about 2 m for M_2 tide and 1 m for S_2 tide, exists in the West Korea Bay and the eastern coasts of Zhejiang and Fujian provinces (Figures 5A, B). For diurnal tides, the largest amplitude (about 0.3 m) exists in the Liaodong Bay, the West Korea Bay and the eastern coasts of Zhejiang and Fujian provinces (Figures 5C, D). The variations in the maximum tidal elevation between different BFC schemes are depicted in Figures 6A, B. The average values of the maximum tidal elevation of EXP 1 are 0.26 m, 0.32 m, 0.24 m, and 0.29 m less than those of EXP 3 in the West Korea Bay and eastern coasts of Jiangsu, Zhejiang and Fujian provinces, respectively. However, the maximum tidal elevation of EXP 1 around Jeju island is about 0.10 m larger than that in EXP 3. The average values of the maximum tidal elevation of EXP 2 are 0.14 m, 0.27 m, 0.24 m, and

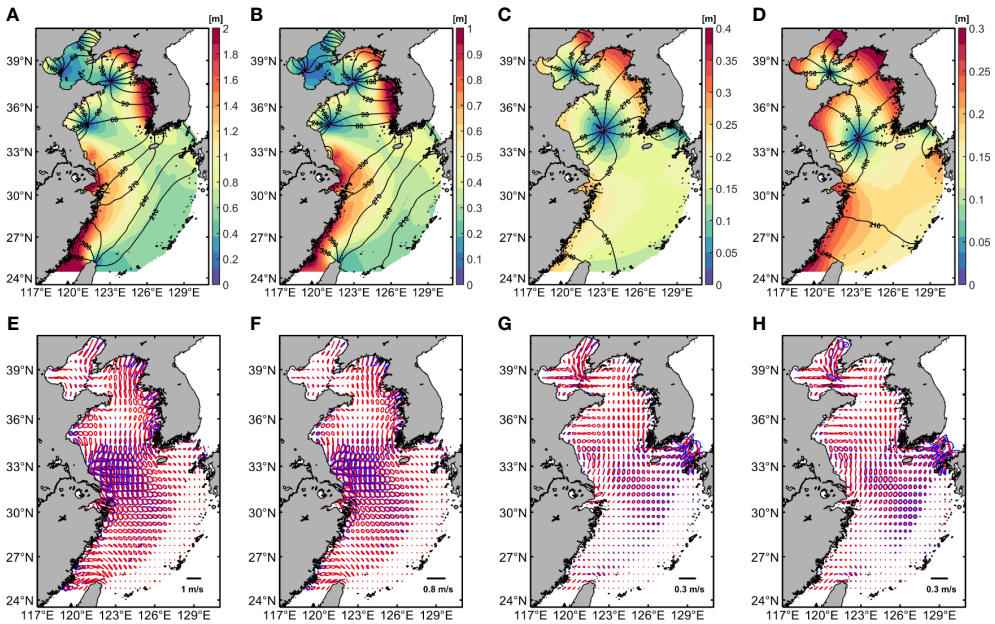


FIGURE 5
(A–D) Cotidal charts for M_2, S_2, K_1, and O_1 tides in EXP 3, respectively. Colormaps denote the magnitude of the co-amplitude (m). Contour lines are the co-phase lines (°). **(E–H)** Tidal current ellipses for M_2, S_2, K_1, and O_1 tides from the surface currents (blue) and bottom currents (red) in EXP 3, respectively.

0.35 m less than those of EXP 3 in the West Korea Bay and eastern coasts of Jiangsu, Zhejiang and Fujian provinces, respectively.

The co-phase lines show that there are four counterclockwise amphidromic systems for semidiurnal tides located in the Qinhuang Island, the Yellow River Estuary, the Chengshanjiao, and the Haizhou Bay, respectively, and two amphidromic systems for diurnal tides located in the Bohai Strait and middle of the South Yellow Sea (Figures 5A–D). The location of the amphidromic point is sensitive to the bottom friction, bottom topography, and coastlines (Fang et al., 1999). The locations of amphidromic points in this study coincide with previous studies (Fang et al., 2004; Zhu and Liu, 2012; Huang et al., 2017; Pringle et al., 2018). As the locations of the amphidromic points of M_2 and S_2 tides are almost similar, as well as the K_1 and O_1 tides, therefore only the comparisons of M_2 and K_1 tides are carried out. A comparison of the locations of the amphidromic points between different BFC schemes is depicted in Figure 7. The amphidromic systems of semidiurnal constituents in the Yellow River Estuary and Haizhou Bay are clustered. For the amphidromic points on the eastern coast of the Chengshanjiao, the amphidromic points of EXP 1 and EXP 2 are in the northwest of that from EXP 3, respectively. For diurnal constituents, the amphidromic points from different BFC schemes in the Bohai Strait are clustered. For those on the eastern coast of Jiangsu province, the amphidromic points of EXP 1 and EXP 2 are in the northwest of that from EXP 3.

The tidal current ellipses of tidal currents on the surface layer and bottom layer of EXP 3 are depicted in Figures 5E–H. The tidal current ellipses show the velocity vector tracks for a certain constituent, in which the major axis and minor axis correspond to the maximum and minimum tidal velocity of this constituent, respectively. Due to the complex topography of coastal seas, such as fjords, islands, and tidal sand ridges, the ellipticity of the tidal ellipse in coastal regions is larger, while that in the outer sea is smaller. The results show that the amplitude of the tidal current velocity of semidiurnal constituents is larger than that of diurnal constituents. The strong tidal currents of

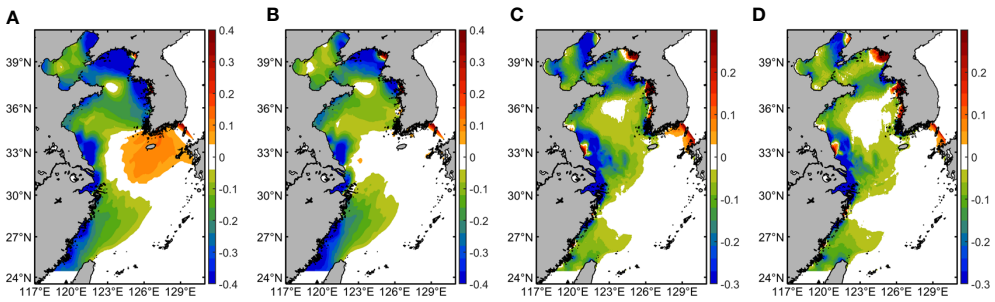


FIGURE 6
(A, B) Differences between the maximum tidal elevations (MTE) of EXP 1 and EXP 2 to those of EXP 3 ($MTE_{EXP1,2}$ - MTE_{EXP3}, unit: m). **(C, D)** Differences between the maximum bottom current velocity (MCV) of EXP 1 and EXP 2 to those of EXP 3 ($MCV_{EXP1,2}$ - MCV_{EXP3}, unit: m/s).

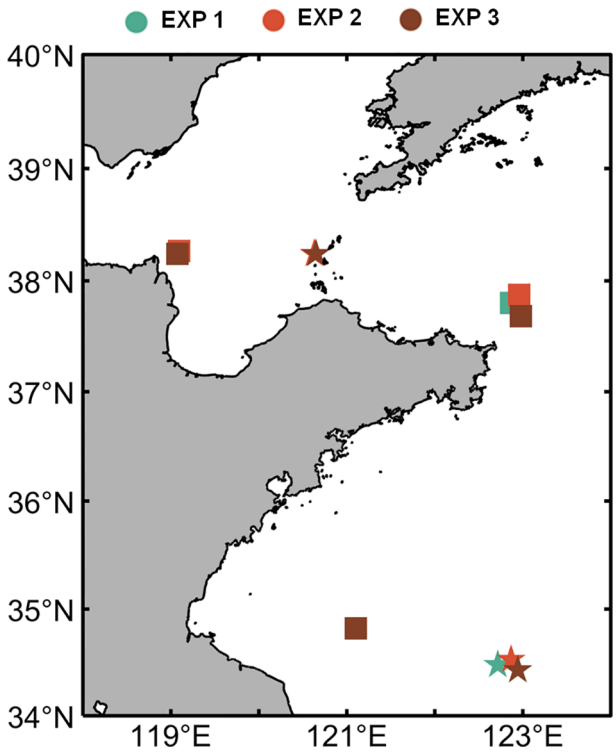


FIGURE 7
Locations of amphidromic points of M_2 tide (rectangles) and K_1 tide (pentagrams), respectively. Colors denote different experiments.

semidiurnal constituents are found in the West Korea Bay, the eastern coast of Jiangsu province, and the Taiwan Strait, while the strong currents of diurnal constituents are mainly found in the Bohai Strait and the Liaodong Bay. The distributions of tidal ellipses for bottom currents are similar to the surface currents except for the decreasing amplitudes of tidal current velocity. The variations in the maximum bottom current velocity between different BFC schemes are shown in Figures 6C, D. The average values of the maximum bottom current velocity of EXP 1 are 0.08 m/s, 0.23 m/s, 0.14 m/s, and 0.08 m/s less than those of EXP 3 in the West Korea Bay, and the eastern coasts of Jiangsu, Zhejiang and Fujian provinces respectively, while the values of maximum bottom current velocity in EXP 2 are 0.03 m/s, 0.21 m/s, 0.10 m/s, and 0.11 m/s less than those of EXP 3.

3 Results

3.1 Tidal energy flux

To understand the tidal dynamics and to explain the mechanisms of the processes of four principal tidal constituents, the distribution of tidal energy in the East China Seas is investigated in this section. Following Garrett (1975) and Egbert and Ray (2000), the expression for the tidal energy flux P is:

$$P = \rho g \langle U\zeta \rangle \tag{5}$$

where U is the volume transport vector, which equals velocity times water depth; ζ is sea surface elevation; the bracket $\langle\ \rangle$ denotes time average. To improve the accuracy of the estimation, the final results are the average values of 15 tidal cycles of the four principal tidal constituents.

The vectors of tidal energy flux in the East China Seas of EXP 3 are shown in Figures 8A–D. For semidiurnal constituents (Figures 8A, B), the tidal energy flux in the northwestern Pacific Sea is westward and divided into two parts: one branch crosses the Ryukyu Islands into the East China Sea, and the other branch crosses the Luzon Strait into the South China Sea. The semidiurnal tide entering the East China Sea from the Tokara Strait can fold to the northwest and continue to spread forward with progressive waves. When arriving on the western coast of Kyushu, Japan, it is divided into two branches: the west branch is the main branch, which continues to the northwest through the southwest side of Jeju Island and enters the Yellow Sea. The tidal energy flux entering the Yellow Sea mainly moves northward along the western coast of the Korean Peninsula, while part of the energy bends westward in the north Yellow Sea and enters the Bohai Sea through the Bohai Strait. The rest part turns back by the Shandong Peninsula and propagates southward along the coasts, forming a counterclockwise semidiurnal tidal wave system. Most of the semidiurnal tidal energy entering the East China Sea from the Ryukyu Islands is diverted into the Taiwan Strait by a counterclockwise rotation around the northern Taiwan island. The vectors of tidal energy flux of M_2 and S_2 constituents are similar, but the magnitude of S_2 tide is approximately one-fifth of the M_2 tide. For the diurnal constituents (Figures 8C, D), tidal energy flux from the Pacific Ocean can be divided into two parts: a small part diverts northwest into the East China Sea through the Tokara Strait; the rest is blocked by the topographical trench (Ryukyu trench), and continues to spread southwest along the Okinawa Trough, and enters

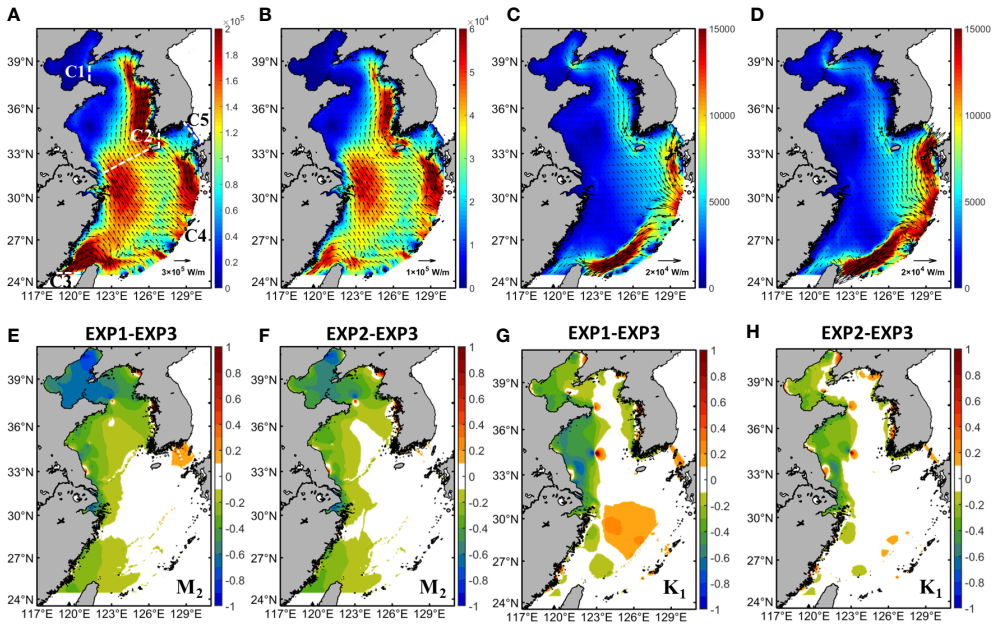


FIGURE 8
(A–D) Vectors of depth-averaged tidal energy flux for M_2, S_2, K_1, and O_1 tides in EXP 3, respectively. Colormaps denote the magnitude of tidal energy flux. **(E–H)** Relative differences of semidiurnal and diurnal tidal energy flux between EXP 1, EXP 2 to EXP 3, respectively.

the South China Sea through the Luzon Strait. The variations in the propagation of the tidal energy flux between semidiurnal and diurnal constituents are concluded: most of the semidiurnal energy flux transports into the East China Sea through the Ryukyu Islands, while the diurnal energy flux is obstructed by the topographical trench and mainly propagates into the South China Sea through the Luzon Strait, which also indicates that the study area is dominated by the semidiurnal constituents.

The tidal energy flux of semidiurnal constituents is significantly larger than that of diurnal tides in the Yellow Sea and East China Sea, while the magnitudes of tidal energy flux of semidiurnal and diurnal tides through the eastern Taiwan Island to the northern Tokara Strait are similar. The existence of islands and topography variations have a great influence on the spatial distribution of tidal energy flux. The variations in tidal energy flux between different BFC schemes are shown in Figures 8E–H. The results show that the tidal energy flux transporting into the Bohai Sea and the Yellow Sea in the EXP 3 are larger than those in EXP 1 and EXP 2. Large differences in semidiurnal tidal energy flux appear in the West Korea Bay, and the eastern coasts of Jiangsu, Zhejiang and Fujian provinces. The semidiurnal tidal energy flux of EXP 1 is 62.7% smaller than that of EXP 3 in the Bohai Sea, while the semidiurnal tidal energy flux of EXP 2 is 47.9% smaller than that of EXP 3. The semidiurnal tidal energy flux in EXP 1 is (20.2%, 17.1%, 32.8%, 15.5%, 20.1%) smaller than those in EXP 3 in the West Korea Bay, South Yellow Sea, and eastern coasts of Jiangsu, Zhejiang and Fujian provinces, respectively. The semidiurnal tidal energy flux in EXP 2 is (9.6%, 13.7%, 26.9%, 12.8%, 25.1%) smaller than those in EXP 3 in the West Korea Bay, South Yellow Sea, and eastern coasts of Jiangsu, Zhejiang and Fujian provinces, respectively. On the other hand, the diurnal tidal energy flux is greatly weaker than that of semidiurnal components in the East China Seas, and the large variations in diurnal tidal energy flux between different BFC schemes mainly appear in the Bohai Sea and the eastern coasts of Shandong and Jiangsu provinces. The area-averaged values of tidal energy flux of diurnal constituents in EXP 1 are 26.9% and 47.1% less than those in EXP 3 in the Bohai Sea and the eastern coasts of Shandong and Jiangsu provinces, while the area-averaged values of diurnal tidal energy flux of EXP 2 are 16.5% and 29.8% less than those in EXP 3 in the Bohai Sea and the eastern coasts of Shandong and Jiangsu provinces. However, the average values of diurnal tidal energy flux in the northwestern Ryukyu Island of EXP 1 are 14.6% larger than those in EXP 3, while the average values of diurnal tidal energy flux in the northwestern Ryukyu Island of EXP 2 are only 5.8% larger than those in EXP 3.

To quantify the transport of tidal energy flux in the East China Seas, five sections are selected (C1-C5, Figure 8A). Tidal energy flux through section C1 means energy from the Yellow Sea entering the Bohai Sea, section C2 for the East China Sea entering the Yellow Sea, section C3 for the Taiwan Strait entering the East China Sea, section C4 (from the eastern Taiwan Island to the northern Tokara Strait) for the Pacific Ocean entering the East China Sea, and section C5 for the Korea Strait entering the East China Sea. The specific values of tidal energy flux of the M_2, S_2, K_1, and O_1 tides propagating through the five sections are listed in Table 2. The statistic values of tidal energy flux through C1-C5 are consistent with previous studies (Fang et al., 2004; Li et al., 2005; Zhu and Liu, 2012; Chen and Cheng, 2020). Moreover, the variations in tidal energy flux between different BFC schemes across five sections are studied. For diurnal constituents, the tidal energy flux of EXP 1 and EXP 2 through C1, C2, C3, and C4 is smaller than that in EXP 3, while the tidal energy flux of EXP 1 and EXP 2 through C5 is larger than that in EXP 3. The absolute differences in tidal energy flux between EXP 1 and EXP 2 to EXP 3 are less than 0.2 GW, while the relative differences are less than 7.1%. For semidiurnal constituents, the tidal energy flux of EXP 1 and EXP

2 through C1, C2, C3, and C4 is smaller than that in EXP 3, while the tidal energy flux of EXP 1 and EXP 2 through C5 is larger than that in EXP 3. The statistic values show that there are large variations in C1 and C3, and the relative differences are small in C2, C4, and C5. The semidiurnal tidal energy flux in EXP 1 through C1 and C3 are 66.7% and 29.1% smaller than those in EXP 3, while the semidiurnal tidal energy flux in EXP 2 through C1 and C3 is 51.5% and 33.8% smaller than those in EXP 3. On the other hand, the semidiurnal tidal energy flux in EXP 1 through C2 and C4 is 7.6% and 5.6% smaller than those in EXP 3, while the semidiurnal tidal energy flux in EXP 2 through C2 and C4 is 8.8% and 6.9% smaller than those in EXP 3. The statistical values show that the spatial distribution of BFCs in EXP 3 results in the increasing tidal energy flux transporting into the Bohai and the Yellow Sea, while the tidal energy from the Korea Strait decreases.

3.2 Bottom friction dissipation

According to the calculations of tidal energy flux through the specific sections, the total tidal dissipation of the Bohai, Yellow and East China Seas is listed in Table 3. The statistical values and the spatial distribution of total tidal dissipation locate in a reasonable range, which corresponds well with previous studies (Zhao et al., 1993; Li et al., 2005; Zhu and Liu, 2012; Zhu et al., 2014; Wu et al., 2018; Chen and Cheng, 2020). The results show that semidiurnal tides mainly dissipate in the Yellow Sea, while diurnal tides mainly dissipate in the East China Sea. The variations in total dissipation of semidiurnal constituents are significant, while those of diurnal constituents are neglectable. The semidiurnal tidal energy dissipated more in EXP 3 compared with EXP 1 and EXP 2, while the total dissipations of diurnal tides in EXP 1-3 have few differences. The total tidal dissipation of semidiurnal tides of EXP 1 and EXP 2 in the Bohai and Yellow Sea is less than that in EXP 3. However, the total dissipation of semidiurnal constituents of EXP 1 and EXP 2 is larger than that in EXP 3 in the East China Sea. On the other hand, the variations in total dissipation of diurnal constituents between different BFC schemes in the Yellow and East China Seas are neglectable, while the total dissipation of diurnal constituents in the Bohai Sea of EXP 1 is smaller than that in EXP 3.

The tidal dissipation rate is estimated as a balance between the rate of working by tidal forces and the energy flux divergence (Egbert and Ray, 2000). The primary dissipation mechanisms for global tides are boundary layer dissipation and internal tide dissipation representing barotropic to baroclinic tidal conversion (Munk, 1997). For coastal areas, tidal dissipation is dominated by the bottom friction effect, and the expression for the bottom friction dissipation rate D is:

$$D = \rho C_d \langle \vec{u}\ 3 \rangle \tag{6}$$

where C_d is the bottom friction coefficient (BFC). The bottom friction dissipation of the four principal tidal constituents in EXP 3 is depicted in Figures 9A–D. The bottom friction dissipation of M_2 tide

TABLE 2 Statistic values of tidal energy flux cross specific sections in East China Seas.

		EXP 1	EXP 2	EXP 3	Zhu and Liu (2012)	Chen and Cheng (2020)
M_2 (GW)	C1	1.0	1.4	2.8	2.0	1.5
	C2	64.5	63.8	69.1	67.6	
	C3	26.5	24.7	37.3		
	C4	133.8	131.8	141.5	122.5	159.7
	C5	14.4	13.2	12.5		
S_2 (GW)	C1	0.1	0.2	0.5	0.2	0.1
	C2	17.8	17.5	20.0	11.7	
	C3	5.0	4.7	7.1		
	C4	35.5	35.1	37.8	20.9	33.9
	C5	5.0	4.6	4.4		
K_1 (GW)	C1	0.3	0.3	0.3	0.5	0.4
	C2	1.5	1.4	1.5	3.1	
	C3	0.6	0.6	0.7		
	C4	9.5	9.3	9.4	10.2	7.6
	C5	1.4	1.3	1.3		
O_1 (GW)	C1	0.2	0.3	0.3	0.3	0.2
	C2	1.3	1.3	1.4	1.6	
	C3	0.7	0.7	0.7		
	C4	14.7	14.7	15.2	7.4	4.6
	C5	1.9	1.8	1.8		

10.3389/fmars.2023.1085118

TABLE 3 The total tidal dissipation in the Yellow and East China Seas.

		EXP 1	EXP 2	EXP 3
Yellow Sea (C2-C1)	M_2 (GW)	63.5	63.8	66.3
	S_2 (GW)	17.7	17.3	19.5
	K_1 (GW)	1.2	1.2	1.0
	O_1 (GW)	1.1	1.0	1.1
East China Sea (C4+C5-C2-C3)	M_2 (GW)	57.2	56.5	47.6
	S_2 (GW)	17.7	17.5	15.1
	K_1 (GW)	8.8	8.6	8.5
	O_1 (GW)	14.6	14.5	14.9

is near an order of magnitude larger than S_2 tide, and three orders of magnitude larger than diurnal constituents. The results show that large dissipation of semidiurnal constituents occurs in shallow areas with strong tidal currents, such as the West Korea Bay, eastern coasts of Jiangsu, Zhejiang and Fujian provinces. The bottom friction dissipation in the Okinawa Trough is three orders of magnitude less than that in the Yellow Sea. However, large dissipation of diurnal constituents occurs in the Bohai Strait, the Korea Strait, and the South Yellow Sea. According to equation (6), the estimation of bottom friction dissipation is closely related to the bottom friction coefficient and bottom current velocity cubed. The average values of semidiurnal and diurnal bottom friction dissipation in the Bohai Sea, Yellow Sea and East China Sea are listed in Table 4. The relative differences in bottom friction dissipation between different BFC schemes are depicted in Figures 9E–H. The results show that the variations in bottom friction dissipation are significantly related to the magnitude and spatial distribution of BFCs. The bottom friction dissipation significantly increases in the high-valued BFC areas. Large variations in tidal dissipation mainly appear in the shallow coastal areas and the South Yellow Sea (Figures 9E, F). The bottom friction dissipation of semidiurnal constituents in EXP 1 is 0.208 W/m^2 less than that in EXP 3 in the West Korea Bay, while that in EXP 2 is 0.203 W/m^2 less than EXP 3. For the coastal areas of Jiangsu and Fujian provinces, the bottom friction dissipation of semidiurnal tides in EXP 1 is 0.011 W/m^2 and 0.017 W/m^2 smaller than EXP 3, while those in EXP 2 are 0.011 and 0.019 W/m^2 smaller than EXP 3. The semidiurnal bottom friction dissipation in EXP 1 is 0.0078 W/m^2 larger than EXP 3 in the South Yellow Sea, while those in EXP 2 are 0.0082 smaller than EXP 3. On the other hand, the variations in diurnal tidal dissipation between different BFC schemes mainly appear in the Bohai Sea, the West Korea Bay, the middle of the South Yellow Sea, and the northwestern Ryukyu Island (Figures 9G, H). The average values of diurnal tidal dissipation in EXP 1 and EXP 2 are larger than those in EXP 3 in the Bohai Sea, the West Korea Bay, and the

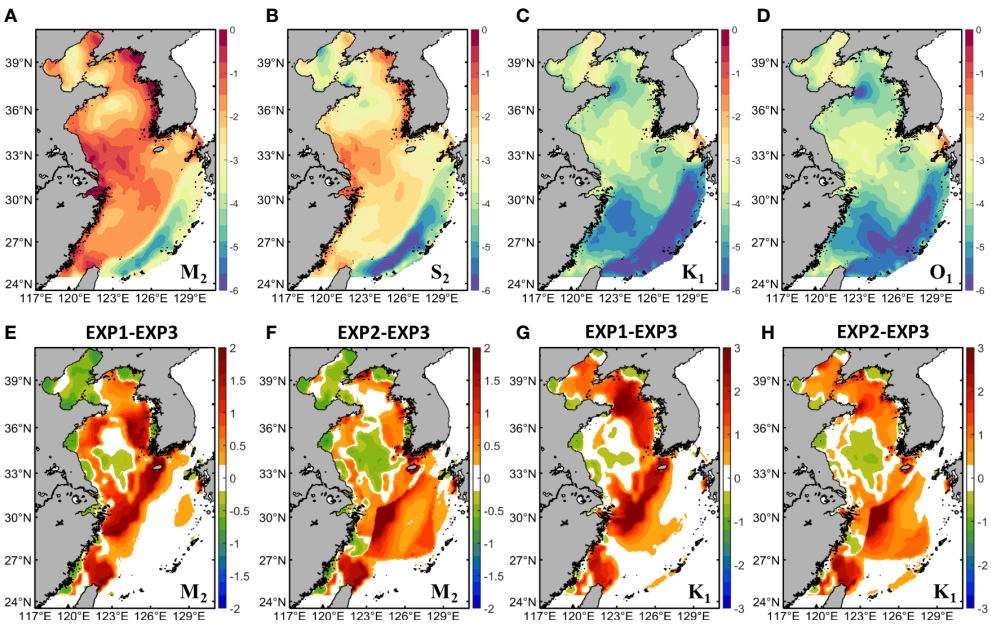


FIGURE 9
(A–D) Estimations of bottom friction dissipation for M_2, S_2, K_1, and O_1 tides in EXP 3 on a log scale, respectively ($log_{10}D$, unit: $log_{10}(W/m^2)$). **(E–H)** Relative differences of semidiurnal and diurnal bottom friction dissipation between EXP 1, EXP 2 to EXP 3, respectively.

TABLE 4 The estimated bottom friction dissipation in the East China Seas.

		EXP 1	EXP 2	EXP 3
Bohai Sea	M_2 (W/m^2)	0.0077	0.0111	0.0218
	K_1 (W/m^2)	0.0011	0.0011	0.0007
Yellow Sea	M_2 (W/m^2)	0.0976	0.0928	0.1689
	K_1 (W/m^2)	0.0002	0.0002	0.0002
East China Sea	M_2 (W/m^2)	0.0496	0.1813	0.0404
	K_1 (W/m^2)	0.0003	0.0003	0.0003

northwestern Okinawa Trough. The average value of diurnal tidal dissipation in EXP 1 is 17.9% larger than that in EXP 3 in the South Yellow Sea, while the average value of diurnal tidal dissipation in EXP 2 is 15.6% smaller than EXP 3.

The variations in bottom friction dissipation are closely related to the magnitude and spatial distribution of BFCs (Figures 9E–H). Signell and Geyer (1991) and Zhong and Li (2006) concluded that the headland-shaped coastal topography combined with strong tidal currents in shallow areas could result in turbulent eddies and therefore strengthened energy dissipation. The tidal sand ridges on the eastern coast of Jiangsu province impede the propagation of tidal currents and result in turbulent eddies and high tidal dissipation. On the other hand, the composition of seabed sediments can influence the magnitude of tidal dissipation as well. The SD-BFCs in the northwestern Ryukyu Island result in high bottom friction dissipation, while the sediment type of this area is sand, and the surrounding sediment environments are composed of clay, volcanic sand gravel, siliceous mud, and calcareous ooze (Dutkiewicz et al., 2015; Pringle et al., 2018). Large grain size and high suspended sediment concentration can also enlarge tidal dissipation, e.g. the eastern coast of Jiangsu province and Hangzhou Bay. The high-valued BFC areas in EXP 3 are closely related to water depth, sediment environment, and coastal topography (Qian et al., 2021), and therefore the bottom friction dissipation increases in the downstream sides of headland topography and those areas with mixed sediment types.

4 Discussions

4.1 Relations between BFC and tidal dynamics versus water depth

The regions that dissipate large amounts of energy are identified from the modelling results. The variations of tidal energy dissipation are greatly similar to the spatial distributions of BFCs, which denotes that the estimation of tidal energy dissipation is mainly influenced by the spatial distributions of BFCs. The specific areas performed with large variations in BFCs and strong tides are the West Korea Bay, the eastern coasts of Jiangsu, Zhejiang and Fujian provinces, and the middle of the South Yellow Sea. The average values of water depth in the West Korea Bay, the eastern coasts of Zhejiang and Fujian provinces, and the South Yellow Sea are 48.8, 34.2, and 51.2 m, respectively. For the West Korea Bay, the increasing BFCs and bottom current velocity relate to larger amounts of tidal dissipation. Based on the tidal propagation in the Yellow Sea, the coastal currents flow along the headland-shaped topography, which will enlarge form drag.

Although the averaged values of tidal dissipation of different BFC schemes in the eastern Jiangsu province are similar, the averaged values of the bottom current velocity of semidiurnal and diurnal tides are ~0.1 m/s and 0.01 m/s, respectively.

The variations in the spatial distributions of BFCs and velocity fields between different BFC schemes are depicted in Figures 2C, D and Figure 10, respectively. Significant variations in the magnitude and spatial distributions of BFCs between EXP 1 and EXP 3 appear in the coastal areas and middle of the South Yellow Sea, while those of EXP 2 in the northwestern Ryukyu Islands are larger than EXP 3 (Figures 2C, D). It can be seen that the variations in the bottom current velocity of semidiurnal and diurnal constituents are negatively related to the spatial distributions of BFCs. In EXP 3, the average BFC values in the West Korea Bay, the eastern coasts of Jiangsu and Fujian provinces are (0.02311, 0.00228, 0.00285), respectively, which are larger than those in EXP 1; however, the average BFC values in the eastern coast of Zhejiang province, the South Yellow Sea, and northwestern Ryukyu Islands are (0.00169, 0.00145, 0.00182), respectively, which are smaller than those in EXP 1. In EXP 2, the average BFC values in the West Korea Bay, the eastern coasts of Jiangsu, Zhejiang and Fujian provinces, and the South Yellow Sea are (0.00143, 0.00197, 0.00151, 0.00252, 0.00124), respectively, which are smaller than those in EXP 3; however, the average BFC value in the northwestern Ryukyu Islands is 0.01553, which are larger than those in EXP 3. Moreover, the variations in BFCs and velocity field of diurnal constituents are similar to semidiurnal constituents, but the values of variations in the velocity field are an order of magnitude less than those from semidiurnal constituents.

Changes in the physical factors that control bottom friction dissipation undoubtedly play a role in these observed changes. Significant effects of bathymetry, bottom friction coefficients and ocean bedforms on tidal energy transport have been investigated (Blakely et al., 2022). The relationships between BFCs, bottom current velocity and bottom friction dissipation versus water depth in the whole East China Seas and those highlighting areas are depicted in Figure 11. The average values of the bottom current velocity of semidiurnal tides are an order of magnitude larger than those of diurnal tides, while the average values of bottom friction dissipation of semidiurnal tides are three orders of magnitude larger than those of diurnal tides. In the East China Seas, the bottom friction dissipation of

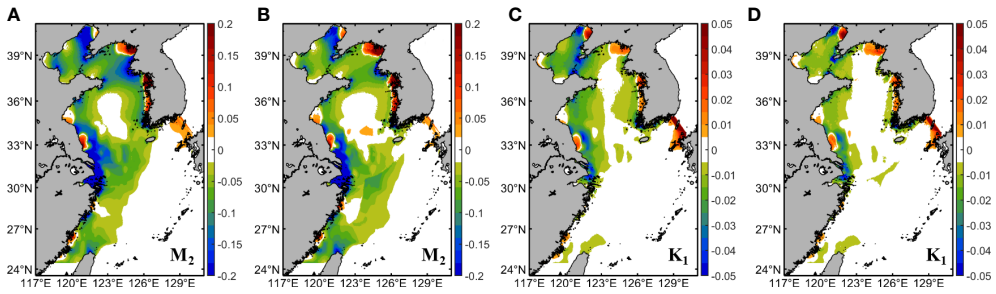


FIGURE 10
(A, B) Differences between the time-averaged semidiurnal bottom current velocity of EXP 1 and EXP 2 to those of EXP 3 (unit: m/s), respectively. **(C, D)** Differences between the time-averaged diurnal bottom current velocity of EXP 1 and EXP 2 to those of EXP 3 (unit: m/s), respectively.

semidiurnal constituents decreases versus the increasing water depth, while that of diurnal constituents has two peaks at 37 m and 97 m (Figure 11A). The average values of BFCs in EXP 2 and EXP 3 first decrease versus the increasing water depth, and then increase until becoming a constant, while the bottom current velocity of semidiurnal and diurnal tides both increase to peaks at 7.5 m and 32.5 m and then decrease. The first peak of diurnal tidal dissipation is proportional to the current velocity, and the second peak is inversely proportional to the current velocity. The variations in bottom friction dissipation of semidiurnal constituents between different BFC schemes are small, but the average values of diurnal tidal dissipation in 37 m of EXP 1 and EXP 2 are 66.7% larger than that in EXP 3. However, the variations in bottom friction dissipation of semidiurnal tides versus water depth in the coastal areas with high-values BFCs are different. For the West Korea Bay, the semidiurnal tidal dissipation first increases versus the increasing water depth and then decreases at 12 m (Figure 11B). The peak values of semidiurnal tidal dissipation in EXP 1 and EXP 2 are 22.0% and 11.6% smaller than EXP 3, while the corresponding values of semidiurnal current velocity in EXP 1 and EXP 2 are 5.0% and 12.6% smaller than EXP 3. At the same time, the average values of BFCs in EXP 2 and EXP 3 decrease when the water depth is less than 67.5 m. For the eastern coasts of Zhejiang and Fujian provinces, the maximum values of semidiurnal tidal dissipation of EXP 1 and EXP 2 are 0.1686 W/m^2

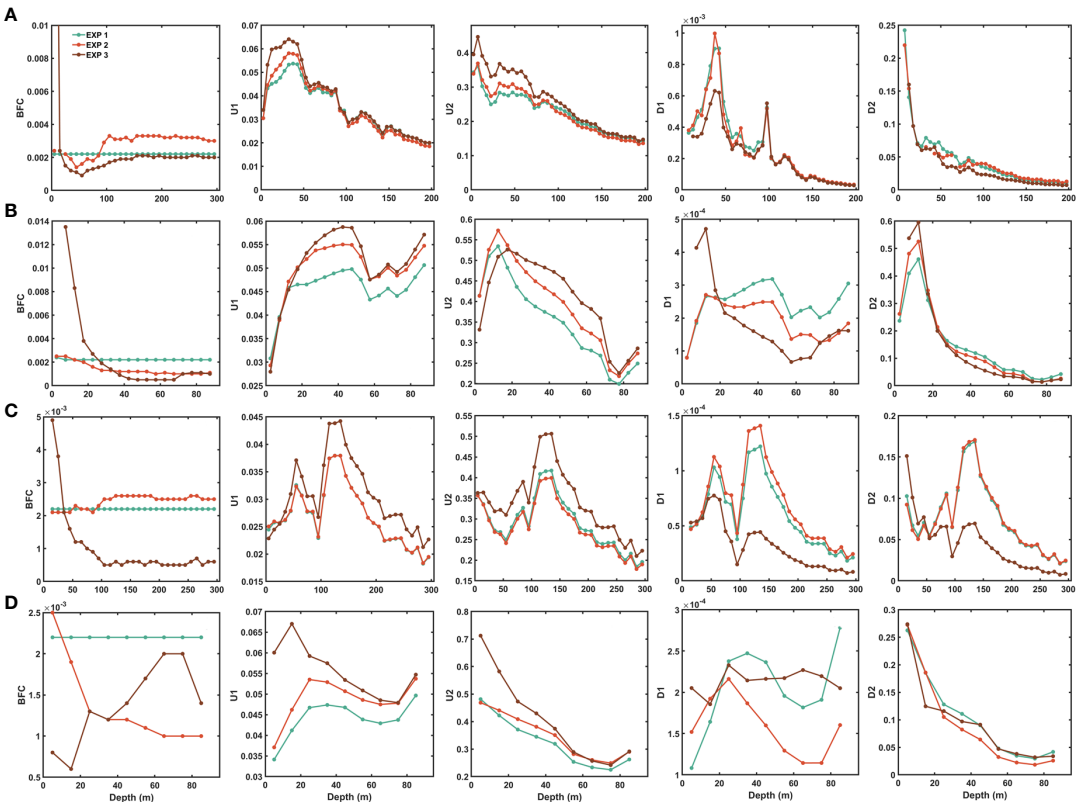


FIGURE 11
Relations of the BFCs, bottom current velocity of diurnal (U1) and semidiurnal (U2) tides, bottom friction dissipation of diurnal (D1) and semidiurnal (D2) tides versus water depth in the East China Seas **(A** row**)**, the West Korea Bay **(B** row**)**, the Zhejiang-Fujian provinces **(C** row**)**, the middle South Yellow Sea **(D** row**)**, respectively."

and 0.1707 W/m^2 in the water depth of around 135 m, while the maximum value of semidiurnal tidal dissipation of EXP 3 is 0.1512 W/m^2 in the water depth around 15 m (Figure 11C). The average values of BFCs in EXP 2 are in the interval of (0.0021, 0.0025), while those in EXP 3 decrease from 0.005 versus the increasing water depth. The corresponding semidiurnal current velocity reaches the maximum value at 135 m, and the maximum velocity of EXP 1 and EXP 2 is 17.6% and 21.1% smaller than EXP 3. On the other hand, the maximum values of diurnal tidal dissipation in EXP 1 and EXP 2 are 1.2×10^{-4} W/m^2 and 1.4×10^{-4} W/m^2 in the water depth of around 135 m, while that in EXP 3 is 0.7×10^{-4} W/m^2 in the water depth around 55 m. Meanwhile, the maximum values of diurnal current velocity in EXP 1 and EXP 2 are 13.6% less than that in EXP 3 in the water depth of around 135 m. For the South Yellow Sea, the semidiurnal tidal dissipation decreases versus the increasing water depth, and then slightly increases when the water depth is larger than 75 m. The relationship between semidiurnal tidal dissipation and water depth is similar to the relationship between the semidiurnal current velocity and water depth. Although the average values of semidiurnal current velocity in EXP 1 and EXP 2 are 14.6% and 6.4% smaller than that in EXP 3, the average value of semidiurnal tidal dissipation in EXP 1 is 22.3% larger than EXP 3, while that in EXP 2 is 11.6% smaller than EXP 3. It can be seen that the variations in bottom friction dissipation are similar to the variations in the tidal current velocity, while the effects of BFCs become more important in the coastal shallow areas.

4.2 Influence on the parameter h/u^3

Tidal mixing is essential in the coastal shallow areas, as it is one of the main mechanisms for the transport of nutrients to the euphotic zones, and also plays an important role in the water mass formation process and thermohaline circulation (Bray, 1988a; Lavin and Organista, 1988; Alvarez-Borrego and Lara-Lara, 1991; Argote et al., 1995). According to the simulated results above, the magnitude and spatial distribution of bottom friction dissipation are closely related to the spatial distribution of BFCs. Simpson and Hunter (1974) suggested a simple model examining the transition between stratified and unstratified regimes controlled by the level of tidal mixing from the observed position of the front. They assumed that, if the area and time of interest were limited, the locus of front could be defined simply by the parameter h/u^3. h/u^3 could be used as the parameter which controlled the formation of a front and used to predict the occurrence of stratification. In general, the Simpson-Hunter number (SH) is frequently presented in the form as:

$$SH = \log_{10}\left(\frac{h}{\left|\overline{u^3}\right|}\right) \tag{7}$$

where h is water depth, and u is the depth-averaged current velocity. Figure 12A shows the spatial distribution of SH numbers in EXP 3. Small values of SH numbers denote that mixed conditions prevail, while large values denote stratified conditions are predicted (Simpson and Hunter, 1974; Pingree and Griffiths, 1978; Argote et al., 1995). The average values of SH numbers in the whole study area, the Bohai Sea, the Yellow Sea, and the East China Sea are (3.98, 3.61, 3.02, 4.57) in EXP 1, (3.96, 3.44, 2.96, 4.59) in EXP 2, (3.83, 3.10, 2.85, 4.50) in EXP 3. Large values of SH numbers mainly appear in the Okinawa Trough, and the middle of the South Yellow Sea, while the coastal areas have small values of *SH* numbers. Figures 12B, C show the variations in the SH numbers between different BFC schemes. Large differences mainly distribute in the shallow water areas with strong tidal currents, i.e. the Bohai Sea, the West Korea Bay, the Hangzhou Bay, and the eastern coasts of Jiangsu and Fujian provinces. The average values of SH numbers in EXP 1 are (0.51, 0.14, 0.22, 0.13, 0.16, 0.39, 0.16) larger than those in EXP 3 in the Bohai Sea, the West Korea Bay, the eastern coasts of Jiangsu, Zhejiang and Fujian provinces, the Hangzhou Bay and the South Yellow Sea, respectively. The average values of SH numbers in EXP 2 are (0.34, 0.05, 0.20, 0.10, 0.21, 0.39, 0.11) larger than those in EXP 3 in the Bohai Sea, the West Korea Bay, the eastern coasts of Jiangsu, Zhejiang and Fujian provinces, the Hangzhou Bay and the South Yellow Sea, respectively. Comparing the cubed depth-averaged current velocity between different BFC schemes (Figures 12D–F), the spatial distributions of SH numbers are closely related to the cubed current velocity. The small valued SH numbers of EXP 3 in the coastal areas and the Bohai Sea may relate with weak stratification effects and strong oceanic mixing. It can be seen that the increasing BFCs in shallow coastal areas, especially in the Bohai Sea, the North Yellow Sea, and the eastern coasts of Jiangsu, Zhejiang and Fujian provinces, can enlarge the values of SH numbers. However, the values of SH numbers in the West Korea Bay and Jianghua Bay with high-valued BFCs are increased by the decreasing current velocity. Simpson and Pingree (1978) found that the partitioning of the seas into stratified and mixed regimes separated by frontal boundaries, while the strong stratification was associated with values of SH numbers larger than 3, and the complete vertical mixing with low values of SH numbers (<1.5). The SH number has been tested by field observations and the databases of temperature and salinity profiles which determined the positions of fronts in the shelf seas (Garrett et al., 1978; Lie, 1989; Glorioso and Flather, 1995; Kobayashi et al., 2006). Du et al. (2022) used a 10-year dataset of satellite-derived suspended sediment concentrations to identify the spatiotemporal variations in suspended sediment fronts on the inner shelf of the East China Seas. They found that the local high-value and low-value SH regions corresponded to the local low-value and high-value frontal probability regions, respectively. However, the critical values of SH numbers for frontal boundaries are varied in different regions. This may be attributable to changes in the main source of heat input, boundary-driven turbulence, and wind force (Simpson and Sharples, 2012).

5 Conclusions

To study the effects of spatial bottom friction parameterization schemes on tidal dynamics, a high-resolution model based on FVCOM was developed and used to simulate the four principal tidal constituents (M_2, S_2, K_1, O_1) in the East China Seas. The four principal tidal constituents in the East China Seas were simulated with different schemes of BFCs: the empirical constant (EC-BFC), sediment-dependent form (SD-BFC) and spatial distribution obtained from the adjoint tidal model with data assimilation

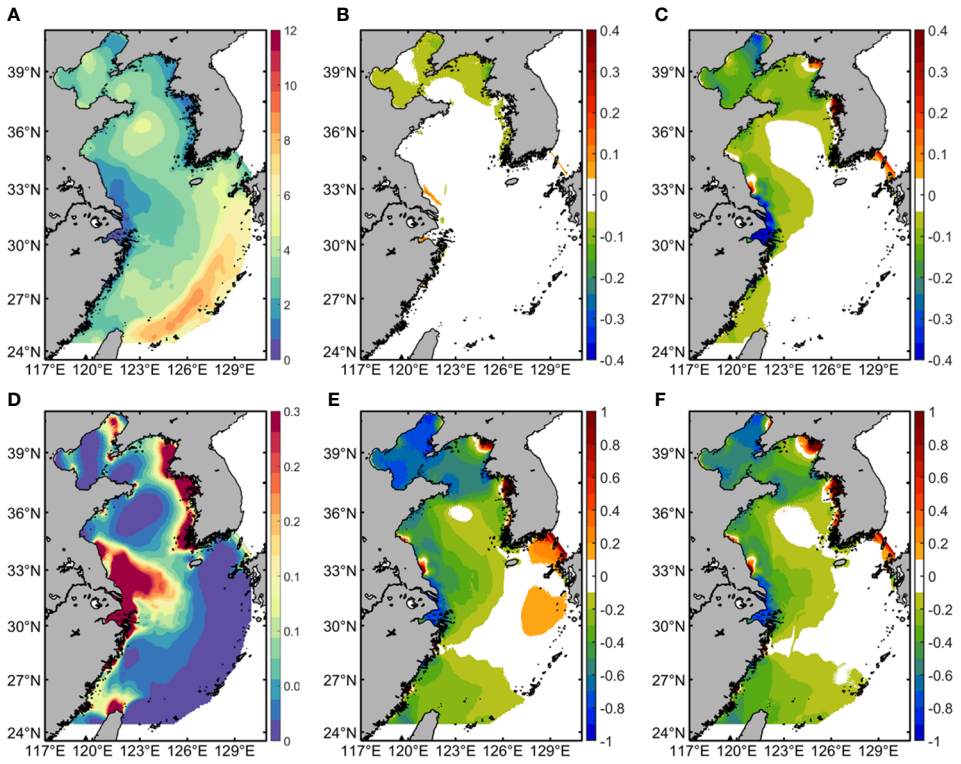


FIGURE 12
(A) Estimations of Simpson-Hunter (SH) number in EXP 3. **(B, C)** Relative differences of SH numbers between EXP 1, EXP 2 to EXP 3, respectively. **(D)** The cubed time-averaged and depth-averaged current velocity ($\overline{u^3}$) in EXP 3. **(E, F)** Relative differences of depth-averaged current velocity between EXP 1, EXP 2 to EXP 3, respectively.

(SV-BFC). The results were evaluated against the observations from satellite altimeters and tidal gauge stations, which demonstrated that the simulated results were reasonable and the simulated results obtained by the SV-BFC fitted observations best. The locations of amphidromic points calculated with EC-BFC and SD-BFC were in the northwest of those from SV-BFC.

The variations in tidal dynamics between different BFC schemes were closely related to the spatial distribution of BFC, especially in the high-valued BFC areas, e.g. the West Korea Bay, the South Yellow Sea, and the eastern coasts of Jiangsu, Zhejiang and Fujian provinces. The average value of the maximum tidal elevation with SV-BFC was approximately 0.26 m larger than those in the EC-BFC and SD-BFC schemes. Meanwhile, the average value of the maximum bottom current velocity with SV-BFC was approximately 0.12 m/s larger than those in the EC-BFC and SD-BFC. The SV-BFC scheme resulted in the increase of tidal energy flux transporting into the Bohai and Yellow Seas, while the tidal energy transporting from the Korea Strait was smaller than those from EC-BFC and SD-BFC. The variations in bottom friction dissipation were closely related to the spatial distribution of BFCs. The high-valued BFC areas of SV-BFC dissipated larger amounts of tidal energy, and the average values of SH numbers were lower than those in EC-BFC and SD-BFC. However, the values of SH numbers in the West Korea Bay and Jianghua Bay with high-valued BFCs were increased because of the decreasing current velocity under the headland-shaped topography. The SD-BFC in the northwestern Ryukyu Island resulted in high bottom friction dissipation, while the sediment type of this area was sand, and the surrounding sediment environments were composed of clay, volcanic sand gravel, siliceous mud, and calcareous ooze.

This study evaluates the effects of bottom friction parameterization schemes on the estimations of tidal energy flux, bottom friction dissipation, and oceanic mixing, and finds that the magnitude of tidal dissipation was closely related to water depth, bottom topography, and sediment types. This study mainly discusses the bottom friction dissipation, while the internal tide dissipation was also needed to be considered especially in deep-sea areas.

Data availability statement

The raw data supporting the conclusions of this article will be made available upon reasonable request and with authors' permission. Requests to access the datasets should be directed to JZ, Jicai_Zhang@163.com.

Author contributions

SQ conducted the data analysis, software, model validation, writing, and visualization. YD and ZW conducted the formal analysis, software. JZ conceived the study, provided the methods and resources, and reviewed and edited the manuscript. JC, DW, and YW conducted the formal analysis and supervision. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the National Key Research and Development Plan of China (Grant No. 2022YFC2808304), the Science and Technology Commission of Shanghai Municipality (STCSM, Grant No. 21JC1402500), and the National Natural Science Foundation of China [Grant No. 41876086].

Acknowledgments

We thank Ms. Ling Tian for her help on the language improvement. We acknowledge all the reviewers for their valuable suggestions.

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Alvarez-Borrego, S., and Lara-Lara, J. R. (1991). "The physical environment and primary productivity of the gulf of California, in the gulf and peninsular provinces of the californias," in *Mere. am. assoc. pet. geol*, vol. 47 . Eds. J. P. Dauphin and B. Simoneit, (Tulsa, USA:American Association of Petroleum Geologists) 555–567.

Argote, M. L., Amador, A., and Lavin, M. F. (1995). Tidal dissipation and stratification in the gulf of California. *J. Geophys. Res.* 100, 16103–16118. doi: 10.1029/95JC01500

Blakely, C. P., Ling, G., Pringle, W. J., Contreras, M. T., Wirasaet, D., Westerink, J. J., et al. (2022). Dissipation and bathymetric sensitivities in an unstructured mesh global tidal model. *J. Geophys. Res.: Oceans* 127, e2021JC018178. doi: 10.1029/2021JC018178

Bo, T., and Ralston, D. K. (2020). Flow separation and increased drag coefficient in estuarine channels with curvature. *J. Geophys. Res.: Oceans.* 125, 1–25. doi: 10.1029/2020JC016267

Bray, N. A. (1988). Thermohaline circulation in the gulf of California. *J. Geophys. Res.* 93, 4993–5020. doi: 10.1029/JC093iC05p04993

Brink, K. H. (2016). Cross-shelf exchange. *Annu. Rev. Mar. Sci.* 8, 59–78. doi: 10.1146/annurev-marine-010814-015717

Chen, Y., and Cheng, P. (2020). Numerical modeling study of tidal energy and dissipation in the East China seas. *J. Xiamen Univ. (Natural Science).* 59 (1), 61–73.

Cheng, R. T., Ling, C.-H., Gartner, J. W., and Wang, P. F. (1999). Estimates of bottom roughness length and bottom shear stress in south San Francisco bay, California. *J. Geophys. Res.: Oceans.* 104, 7715–7728. doi: 10.1029/1998JC900126

Chen, C., Huang, H., Beardsley, R. C., Liu, H., Xu, Q., and Cowles, G. (2007). A finite volume numerical approach for coastal ocean circulation studies: comparisons with finite difference models. *J. Geophys. Res.: Oceans.* 112, 83–87. doi: 10.1029/2006JC003485

Chen, C., Liu, H., and Beardsley, R. C. (2003). An unstructured grid, finite-volume, three-dimensional, primitive equations ocean model: application to coastal ocean and estuaries. *J. Atmos. Ocean. Technol.* 20 (1), 159–186. doi: 10.1175/1520-0426(2003)020<0159:AUGFVT>2.0.CO;2

Chu, D., Niu, H., Wang, Y. P., Cao, A., Li, L., Du, Y., et al. (2021). Numerical study on tidal duration asymmetry and shallow-water tides within multiple islands: An example of the zhoushan archipelago. *Estuar. Coast. Shelf Sci.* 262, 107576. doi: 10.1016/j.ecss.2021.107576

Chu, D., Zhang, J., Wu, Y., Jiao, X., and Qian, S. (2019). Sensitivities of modelling storm surge to bottom friction, wind drag coefficient, and meteorological product in the East China Sea. *Estuar. Coast. Shelf Sci.* 231, 106460. doi: 10.1016/j.ecss.2019.106460

Cowen, R. K., and Sponaugle, S. (2009). Larval dispersal and marine population connectivity. *Annu. Rev. Mar. Sci.* 1, 443–466. doi: 10.1146/annurev.marine.010908.163757

Das, S. K., and Lardner, R. W. (1991). On the estimation of parameters of hydraulic models by assimilation of periodic tidal data. *J. Geophys. Res.* 96, 187–196. doi: 10.1029/91JC01318

Dutkiewicz, A., Müller, R. D., O'Callaghan, S., and Jónasson, H. (2015). Census of seafloor sediments in the world's ocean. *Geology* 43 (9), 795–798. doi: 10.1130/G36883.1

Du, Y., Zhang, J., Wei, Z., Yin, W., Wu, H., Yuan, Y., et al. (2022). Spatio-temporal variability of suspended sediment fronts (SSFs) on the inner shelf of the East China Sea: The contribution of multiple factors. *J. Geophys. Res.: Oceans* 127, 1–25. doi: 10.1029/2021JC018392

Dyer, K. R., and Soulsby, R. L. (1988). Sand transport on the continental shelf. *Annu. Rev. Fluid Mech.* 20, 295–324. doi: 10.1146/annurev.fl.20.010188.001455

Egbert, G. D., and Erofeeva, S. Y. (2002). Efficient inverse modeling of barotropic ocean tides. *J. Atmosp. Ocean. Technol.* 19 (2), 183–204. doi: 10.1175/1520-0426(2002)0192.0.CO;2

Egbert, G., and Ray, R. (2000). Significant dissipation of tidal energy in the deep ocean inferred from satellite altimeter data. *Nature* 405, 775–778. doi: 10.1038/35015531

Egbert, G. D., Ray, R. D., and Bills, B. G. (2004). Numerical modeling of the global semidiurnal tide in the present day and in the last glacial maximum. *J. Geophys. Res. C: Oceans* 109 (3), 1–5. doi: 10.1029/2003JC001973

Fang, G., Kwok, Y. K., Yu, K., and Zhu, Y. (1999). Numerical simulation of principal tidal constituents in the south China Sea, gulf of tonkin and gulf of Thailand. *Continent. Shelf Res.* 19, 845–869. doi: 10.1016/S0278-4343(99)00002-3

Fang, G., Wang, Y., Wei, Z., Choi, B. H., Wang, X., and Wang, J. (2004). Empirical cotidal charts of the bohai, yellow, and East China seas from 10 years of TOPEX/Poseidon altimetry. *J. Geophys. Res.* 109 (C11), 1–13. doi: 10.1029/2004jc002484

Fan, R., Zhao, L., Lu, Y., Nie, H., and Wei, H. (2019). Impacts of currents and waves on bottom drag coefficient in the East China shelf seas. *J. Geophys. Res.: Oceans.* 124, 7344–7354. doi: 10.1029/2019JC015097

Gao, X., Wei, Z., Lv, X., Wang, Y., and Fang, G. (2015). Numerical study of tidal dynamics in the south China Sea with adjoint method. *Ocean Model.* 92, 101–114. doi: 10.1016/j.ocemod.2015.05.010

Garrett, C. (1975). Tides in gulfs. *Deep Sea Res. Oceanogr. Abstr.* 22 (1), 23–35. doi: 10.1016/0011-7471(75)90015-7

Garrett, C., Keeley, J. R., and Greenberg, D. A. (1978). Tidal mixing versus thermal stratification in the gulf of Maine. *Atmos Oceans.* 16 (4), 403–423. doi: 10.1080/07055900.1978.9649046

Glorioso, P. D., and Flather, R. A. (1995). A barotropic model of the currents off SE south America. *J. Geophys. Res.* 100 (C7), 13427–13440. doi: 10.1029/95JC00942

Guo, X., and Yanagi, T. (1998). Three-dimensional structure of tidal current in the East China Sea and the yellow Sea. *J. Oceanogr.* 54, 651–668. doi: 10.1007/BF02823285

Huang, X., Zhang, R., Ma, Q., Jiang, B., and Sun, J. (2017). Numerical simulation of tidal waves in bohai Sea and yellow Sea based on FVCOM. *J. Dalian Ocean Univ.* 32 (5), 617–624.

Huettel, M., Berg, P., and Kostka, J. E. (2014). Benthic exchange and biogeochemical cycling in permeable sediments. *Annu. Rev. Mar. Sci.* 6, 23–51. doi: 10.1146/annurev-marine-051413-012706

Jeffreys, H. (1921). "Tidal friction in shallow seas," in *Philosophical transactions of the royal society of London. Series a, containing papers of a mathematical or physical character*. 221, 239–264.

Kang, S. K., Lee, S., and Lie, H. (1998). Fine grid tidal modeling of the yellow and East China seas. *Continent. Shelf Res.* 18 (7), 739–772. doi: 10.1016/S0278-4343(98)00014-4

Kobayashi, S., Simpson, J. H., Fujiwara, T., and Horsburgh, K. J. (2006). Tidal stirring and its impact on water column stability and property distributions in a semi-enclosed shelf sea (Seto inland Sea, Japan). *Continent. Shelf Res.* 26 (11), 1295–1306. doi: 10.1016/j.csr.2006.04.006

Lavin, M. F., and Organista, S. (1988). Surface heat flux in the northern gulf of California. *J. Geophys. Res.* 93, 14033–14038. doi: 10.1029/JC093iC11p14033

Lee, J. C., and Jung, K. T. (1999). Application of eddy viscosity closure models for the M_2 tide and tidal currents in the yellow Sea and the East China Sea. *Continent. Shelf Res.* 19 (4), 445–475. doi: 10.1016/S0278-4343(98)00087-9

Lie, H. J. (1989). Tidal fronts in the south-eastern hwanghae (Yellow Sea). *Continent. Shelf Res.* 9 (6), 527–546. doi: 10.1016/0278-4343(89)90019-8

Li, L., Guan, W., Hu, J., Cheng, P., and Wang, X. H. (2018). Responses of water environment to tidal flat reduction in xiangshan bay: Part I hydrodynamics. *Estuar. Coast. Shelf Sci.* 206, 14–26. doi: 10.1016/j.ecss.2017.11.003

Li, P., Li, L., Zuo, J., Chen, M., and Zhao, W. (2005). Tidal energy fluxes and dissipation in the bohai Sea, the yellow Sea and the East China Sea. *Period. Ocean Univ. China* 35 (5), 713–718.

Lu, X., and Zhang, J. (2006). Numerical study on spatially varying bottom friction coefficient of a 2D tidal model with adjoint method. *Continent. Shelf Res.* 26, 1905–1923. doi: 10.1016/j.csr.2006.06.007

McWilliams, J. C. (2006). *Fundamentals of geophysical fluid dynamics* (New York: Cambridge Univ. Press).

Mellor, G. L., and Yamada, T. (1982). Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys.* 20 (4), 851–875. doi: 10.1029/RG020i004p00851

Mofjeld, H. O. (1988). Depth dependence of bottom stress and quadratic drag coefficient for barotropic pressure-driven currents. *J. Phys. Oceanogr.* 18, 1658–1669. doi: 10.1175/1520-0485(1988)018<1658:DDOBSA>2.0.CO;2

Munk, W. (1997). Once again: Once again–tidal friction. *Prog. Oceanogr.* 40 (1), 7–35. doi: 10.1016/S0079-6611(97)00021-9

Munk, W., and Wunsch, C. (1998). Abyssal recipes II: energetics of tidal and wind mixing. *Deep Sea Res. Part I: Oceanogr. Res. Pap* 45, 1977–2010. doi: 10.1016/S0967-0637(98)00070-3

Pawlowicz, R., Beardsley, B., and Lentz, S. (2002). Classical tidal harmonic analysis including error estimates in MATLAB using T_TIDE. *Comput. Geosci.* 28 (8), 929–937. doi: 10.1016/S0098-3004(02)00013-4

Pingree, R. D., and Griffiths, D. K. (1978). Tidal fronts on the shelf seas around the British isles. *J. Geophys. Res.: Oceans* 83 (C9), 1–8. doi: 10.1029/JC083iC09p04615

Pringle, W. J., Wirasaet, D., Suhardjo, A., Meixner, J., Westerink, J. J., Kennedy, A. B., et al. (2018). Finite-element barotropic model for the Indian and Western pacific oceans: Tidal model-data comparisons and sensitivities. *Ocean Model.* 129, 13–38. doi: 10.1016/j.ocemod.2018.07.003

Qian, S., Wang, D., Zhang, J., and Li, C. (2021). Adjoint estimation and interpretation of spatially varying bottom friction coefficients of the M_2 tide for a tidal model in the bohai, yellow and East China seas with multi-mission satellite observations. *Ocean Model.* 161, 101783. doi: 10.1016/j.ocemod.2021.101783

Signell, R., and Geyer, W. (1991). Transient eddy formation around headlands. *J. Geophys. Res.: Oceans* 96 (C2), 2561–2575. doi: 10.1029/90JC02029

Simpson, J. H., and Hunter, J. R. (1974). Fronts in the Irish Sea. *Nature* 250 (5465), 404–406. doi: 10.1038/250404a0

Simpson, J. H., and Pingree, R. D. (1978). "Shallow sea fronts produced by tidal stirring," in *Oceanic fronts in coastal processes*. Eds. M. J. Bowman and W. E. Esaias (New York: Springer-Verlag), 29–42.

Simpson, J. H., and Sharples, J. (2012). *Introduction to the physical and biological oceanography of shelf seas* (New York: Cambridge University Press), 466.

Taylor, G. I. (1920). "I. tidal friction in the Irish Sea," in *Philosophical transactions of the royal society of London. Series a, containing papers of a mathematical orphysical character*. 220, 1–33. doi: 10.1098/rsta.1920.0001

Trowbridge, J. H., and Lentz, S. J. (2018). The bottom boundary layer. *Annu. Rev. Mar. Sci.* 10, 397–420. doi: 10.1146/annurev-marine-121916-063351

Wang, X., Chao, Y., Shum, C., Yi, Y., and Fok, H. S. (2012). Comparison of two methods to assess ocean tide models. *J. Atmosp. Ocean. Technol.* 29 (8), 1159–1167. doi: 10.1175/JTECH-D-11-00166.1

Wang, D., Liu, Q., and Lv, X. (2014). A study on bottom friction coefficient in the bohai, yellow, and East China Sea. *Math. Problems Eng.* 2014, 1–7. doi: 10.1155/2014/432529

Wang, D., Zhang, J., and Wang, Y. P. (2021). Estimation of bottom friction coefficient in multi-constituents tidal models using the adjoint method: temporal variations and spatial distributions. *J. Geophys. Res.: Oceans* 126, 1–20. doi: 10.1029/2020JC016949.

Wu, R., Jiang, Z., and Li, C. (2018). Revisiting the tidal dynamics in the complex zhoushan archipelago waters: a numerical experiment. *Ocean Model.* 132, 139–156. doi: 10.1016/j.ocemod.2018.10.001

Zhang, X., Chu, D., and Zhang, J. (2021). Effects of nonlinear terms and topography in a storm surge model along the southeastern coast of China: a case study of typhoon chan–hom. *Natural Hazards.* 107, 551–574. doi: 10.1007/s11069-021-04595-y

Zhang, J., Lu, X., Wang, P., and Wang, Y. P. (2011). Study on linear and nonlinear bottom friction parameterizations for regional tidal models using data assimilation. *Continent. Shelf Res.* 31, 555–573. doi: 10.1016/j.csr.2010.12.011

Zhao, B., Fang, G., and Cao, D. (1993). Numerical modeling on the tides and tidal currents in the eastern China seas. *Yellow Sea Res.* 5, 41–61.

Zhong, L., and Li, M. (2006). Tidal energy fluxes and dissipation in the Chesapeake bay. *Continent. Shelf Res.* 26 (6), 752–770. doi: 10.1016/j.csr.2006.02.006

Zhu, X., and Liu, G. (2012). Numerical study on the tidal currents, tidal energy fluxes and dissipation in the China seas. *Oceanol. ET Limnol. Sinica.* 43 (3), 669–677.

Zhu, X., Song, D., Bao, X., and Liu, G. (2014). Tidal energy flux and dissipation in the Northwest pacific. *J. Trop. Oceanogr.* 33 (1), 1–9.

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 01 February 2023
DOI 10.3389/fmars.2022.1046964

Check for updates

OPEN ACCESS

EDITED BY
Shiqiu Peng
South China Sea Institute of Oceanology, Chinese Academy of Sciences, China

REVIEWED BY
Mingwei Li,
Harbin Engineering University, China
Sin-Chi Kuok,
University of Macau, China
Zitong Chen,
Guangzhou Institute of Tropical and Marine Meteorology (GITMM), China

*CORRESPONDENCE
Xiang Wang
xiangwangcn@nudt.edu.cn

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

RECEIVED 17 September 2022
ACCEPTED 10 November 2022
PUBLISHED 01 February 2023

CITATION
Pan X, Wang X, Zhao C, Wu J, Wang H, Wang S and Chen S (2023) USFP: An unbalanced severe typhoon formation prediction framework based on transfer learning.
Front. Mar. Sci. 9:1046964.
doi: 10.3389/fmars.2022.1046964

COPYRIGHT
© 2023 Pan, Wang, Zhao, Wu, Wang, Wang and Chen. 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(s) 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.

USFP: An unbalanced severe typhoon formation prediction framework based on transfer learning

Xiaotian Pan[1], Xiang Wang[1*], Chengwu Zhao[1], Jianping Wu[1], Huizan Wang[1], Senzhang Wang[2] and Sihao Chen[3]

[1]College of Meteorology and Oceanography, National University of Defense Technology, Changsha, China, [2]School of Computer Science and Engineering, Central South University, Changsha, China, [3]College of Computer Science and Technology, National University of Defense Technology, Changsha, China

Introduction: Severe typhoons, as extreme weather events, can cause a large number of casualties and property damage in coastal areas. There are mainly three kinds of methods for the prediction of severe typhoon formation, which are the numerical-based methods, the statistical-based methods, and the machine learning-based methods. However, existing methods do not consider the unbalance between the number of ordinary typhoon samples and severe typhoon samples, which makes the accuracies of existing methods in the prediction of severe typhoons much lower than that of ordinary typhoons.

Methods: In this paper, we propose an unbalanced severe typhoon formation prediction (USFP) framework based on transfer learning. We first propose a severe typhoon pre-learning model which is used to learn prior knowledge from a constructed balanced dataset. Then, we propose an unbalanced severe typhoon re-learning model which utilizes the prior knowledge learning from the pre-learning model. Our USFP framework fuses three different variables, which are atmospheric variables, sea surface variables, and ocean hydrographic variables.

Results: Extensive experiments based on datasets of three different regions show that our USFP framework outperforms the numerical model IFS of ECMWF and existing machine learning methods.

KEYWORDS

severe typhoon, unbalanced data, transfer learning, deep learning, spatio-temporal

1 Introduction

A tropical cyclone (TC) is a powerful and profound tropical weather system. As extreme weather, severe typhoons not only cause great economic damage to coastal areas but also greatly endanger people's lives and property. According to statistics from the National Meteorological Administration of China, in 2018, about 32 million people were affected by typhoons, and the direct economic loss reached 69.73 billion RMB (China Meteorological Administration, 2020). Severe typhoons, which are more powerful than ordinary typhoons, bring more serious disasters.

In recent years, machine learning-based methods have been widely used in meteorology and oceanography. (Mecikalski et al., 2021; Wikner et al., 2021; Fei et al., 2022). Tropical cyclones (including severe typhoons) are a high-impact and disastrous weather phenomenon in meteorology and oceanography. There are mainly three kinds of methods used to predict the formation of tropical cyclones: the numerical-based methods, the statistical-based methods, and the machine learning-based methods. First, the typical numerical-based methods are the hurricane weather research forecast model (HWRF), the Global Forecast System (GFS), and the Integrated Forecasting System (IFS), which are used to forecast future weather by numerically solving a set of hydrodynamic and thermodynamic equations and the main forecasting method used by many official organizations in the world (John, 2019). Second, the typical statistical-based methods of TC prediction are the Statistical Typhoon Intensity Prediction Scheme (STIPS) and the Statistical Hurricane Intensity Prediction Scheme (SHIPS), which are based on the numerical-based methods to build forecasting models considering future changes in the atmospheric and oceanic conditions. The third is the machine learning-based methods, which are data-driven methods that do not consider physical mechanisms. Meteorologists apply machine learning models such as AdaBoost Algorithm (Zhang et al., 2019), SVM (Richman et al., 2017; Kim et al., 2019), CNN (Matsuoka et al., 2018), and Fuzzy Neural Network (Yip and Yau, 2011) to tropical cyclone prediction. However, existing studies do not consider the unbalanced characteristics of the typhoon samples. The number of samples between ordinary typhoons and severe typhoons is quite different as shown in Table 1. This makes the prediction accuracy of severe typhoons will be decreased.

TABLE 1 The number of positive and negative samples in different regions counted from the World Meteorological Organization (WMO) version of the International Best Trajectory Archive for Climate Management (IBTrACS) Global Tropical Cyclone Best Trajectory Dataset.

	Western Pacific	Eastern Pacific	North Atlantic
Positive Samples	237	97	105
Negative samples	2482	833	1452

In this paper, we propose an unbalanced severe typhoon formation prediction (USFP) framework, which is based on the transfer learning method (Yosinski et al., 2014; Long et al., 2015). Our USFP framework can effectively minimize the impact of unbalanced data on severe typhoon formation prediction. It first learns prior knowledge from a balanced dataset extracted from the unbalanced dataset with the severe typhoon pre-learning model shown in Section 4.2. Then, the prior knowledge is transferred to the unbalanced severe typhoon re-learning model for training as described in Section 4.3. Our USFP framework fuses three variables: atmospheric variables, sea surface variables, and ocean hydrographic variables. Extensive experiments in three different regions show that our USFP framework outperforms the numerical model IFS of ECMWF and existing machine learning methods.

Contributions of this paper include:

1. We propose an unbalanced severe typhoon formation prediction (USFP) framework based on transfer learning. The USFP framework is trained by transferring prior knowledge obtained from a balanced dataset to the unbalanced severe typhoon re-learning model. To the best of our knowledge, we are the first to use transfer learning to improve the prediction accuracies of severe typhoons.
2. We design a customized loss function to optimize our USFP framework, which assigns different weights to different categories of ordinary typhoon samples and severe typhoon samples.
3. We fuse multiple data from atmospheric variables, sea surface variables, and ocean hydrographic variables to predict severe typhoon formation. Extensive experiments performed on data from three regions show that our USFP framework can effectively improve the forecasting effect of unbalanced severe typhoon formation.

The rest of this paper is organized as follows: Section 2 reviews the current studies in the field; Section 3 defines the unbalanced severe typhoon problem in this paper; Section 4 describes the structure of the USFP framework; Section 5 demonstrates the effectiveness of the USFP framework through experiments; Section 6 summarizes the current work in this paper and gives an outlook on future work.

2 Related work

At present, there is a lack of research on the formation of unbalanced severe typhoons. Therefore, this section reviews

research in the field of tropical cyclone forecasting and data unbalance, which is used to explore methods to solve the problem of unbalanced severe typhoon formation prediction.

2.1 Tropical cyclone forecasting

2.1.1 Numerical-based methods

The numerical-based methods forecast the environmental field through the meteorological marine environmental conditions and then extrapolate the forecast of typhoon elements, which is done through model initialization and physical process parameterization. At present, the numerical-based methods have an accuracy rate of 70-80% (Halperin et al., 2013) and are used in many official organizations like the hurricane weather research forecast model (HWRF), the Global Forecast System (GFS), and the Integrated Forecasting System (IFS) of the European Centre for Medium-Range Weather Forecast (ECMWF). For example, Elsberry et al. (2021) improved the Pacific typhoon intensity prediction technology based on ECMWF and successfully predicted the rapid intensification events after the formation of tropical cyclones. Na et al. (2018) evaluated the intensity forecast error of tropical cyclones and the analysis found that the official is more mature in predicting the weakening problem than the intensification problem. The numerical-based methods have had a lot of fruitful research work in typhoon physical law perception and forecasting. However, the numerical-based methods have shortcomings such as incomplete expression of complex physical processes, low accuracy of typhoon intensity prediction, and high computational cost.

2.1.2 Statistical-based methods

The statistical-based methods are based on numerical weather forecasts and take into account changes in the future atmospheric environment and ocean conditions to build forecast models. The method achieves the prediction of TC by analyzing the regularity of a large amount of data and representing it with a functional relationship equation. Some of the typical methods are the Statistical Typhoon Intensity Prediction Scheme (STIPS) (Demaria and Kaplan, 1994) and the Statistical Hurricane Intensity Prediction Scheme (SHIPS) (Fritsch and Chappell, 1980; DeMaria and Kaplan, 1999; Knaff et al., 2005). The STIPS is a multiple linear regression model based on a statistical-dynamical framework, which is constructed by using a large amount of environmental information obtained from the Navy Operational Global Analysis and Prediction System (NOGAPS). The SHIPS is commonly used in the Atlantic and East Pacific regions, which has good results in tropical cyclone forecasting. The regression coefficients of the SHIPS are updated each year after the hurricane season with the latest samples and improved operational forecasting. The statistical-based methods have been relatively well applied in TC intensity prediction. However, because the statistical-based methods are often helpless in the face of massive data, it is difficult to extract critical and effective forecast information.

2.1.3 Machine learning-based methods

In recent years, machine learning has achieved good results in numerous fields. Therefore, many scholars have applied it to the field of tropical cyclones forecasting. These methods compensate for the shortcomings of the statistical-based methods and the numerical-based methods to some extent. For example, Wijnands et al. (2016) of the University of Melbourne, Australia, used the Peter Clark algorithm to select the best predictors for short-term formation forecasts of tropical cyclones. Zhang et al. (2015) of Brookhaven National Laboratory in the United States tried to use nonlinear ensemble machine learning classifiers to determine whether a mesoscale convective system would evolve into a tropical cyclone in different prediction periods. Ahijevych et al. (2016) of the National Center for Atmospheric Research in the United States used the random forest algorithm to predict the possibility of a mesoscale convective system developing into a tropical cyclone within 2 hours. Based on WindSat satellite ocean surface wind and precipitation data, Park et al. (2016) from the Busan Institute of Ocean Science and Technology in South Korea used decision trees to analyze the intensity of tropical cyclones. Higa et al. (2021) successfully estimated typhoon intensity with high accuracy by using the VGG-16 model to process a single satellite image and combining the knowledge of the meteorological domain. The machine learning-based methods as data-driven methods can ignore the imprecise physical mechanisms of typhoon formation and have significant advantages in capturing the nonlinear relationship between forecast factors and forecast targets (Reichstein et al., 2019).

2.2 Unbalanced data problems

2.2.1 Data sampling-based methods

The data sampling-based methods are to manually balance the unbalanced datasets by over-sampling or under-sampling, which are widely used to solve unbalanced problems in the field of machine learning. Han et al. (2005) designed an improved oversampling algorithm based on the SMOTE algorithm, which only uses the minority class samples on the boundary to synthesize new samples, thereby improving the class distribution of the samples. Yan and Cao (2019) proposed two feature-based oversampling methods to rebalance binary and multi-class time series datasets with good results in terms of statistical significance. In contrast, Liu et al. (2009) used an ensemble learning mechanism to optimize the traditional undersampling method with a better training effect. an ensemble learning-based undersampling technique. However, the data sampling-based methods such as oversampling and

undersampling cannot meet the requirement of the actual forecast of severe typhoons and are not suitable for solving the problems faced in this paper.

2.2.2 Algorithm improvement-based methods

The core idea of the algorithm improvement-based methods is to make the model more focused on the sample size less class as a way to improve the prediction accuracy of small sample data. For example, Mo et al. (2019) proposed an unbalanced sample classification algorithm based on the deep residual network, which has a recognition rate close to 100% for small sample unbalanced datasets. Hu et al. (2018) proposed a multi-task learning framework using an attribute attention mechanism to capture key information and improved the accuracy of crime charge prediction for small samples. Cui et al. (2019) designed a weight adjustment scheme to rebalance the loss using the effective number of samples per class, which resulted in a class-balanced loss that can achieve significant performance gains on long-tailed datasets. Geng and Luo, (2019) modified the standard CNN to a cost-sensitive network (CS-CNN), which can use the category-dependent cost matrix to penalize misclassified samples. Zhan (2020) used the loss function of Focal loss (Lin et al., 2017) to establish a network model based on the DNN-LSTM, which had a better result for the problem of unbalanced data.

2.2.3 Transfer learning-based methods

The transfer learning-based approach is to use the pre-trained model parameters for the training of new models, which helps to improve the accuracy of samples with fewer data in the dataset. Taherkhani et al. (2020) proposed the AdaBoost-CNN model based on transfer learning with 16.98% higher accuracy compared to the classical Adaboost. Al-Stouhi and Reddy (2016) proposed a Rare-Transfer transfer learning algorithm with a label-dependent update mechanism that can effectively handle rare class classification problems. Troncoso et al. (2018) transformed the problem of predicting extreme monsoons into an unbalanced binary classification problem and performed transfer learning on a series of related technical models. Singh et al. (2020) used VGG-19 as the base model and supplemented it with several techniques to achieve better results than existing frameworks. Lee et al. (2016) used a CNN classifier model based on transfer learning to normalize the data by thresholding the large-class data and obtained better results on small classes of data.

Transfer learning is not only applied to the problem of imbalanced data, some researchers have also applied transfer learning to the prediction of tropical cyclones. Deo et al. (2017) assessed the relationship between different types of cyclones by using transfer learning and traditional neural network methods to achieve more stable intensity predictions for tropical cyclones. Pang et al. (2021) combined a deep convolutional generative adversarial network (DCGAN) and the YOLOv3 model to propose a New Detection Framework of Tropical Cyclones (NDFTC) with good stability and accuracy. Combinido et al. (2018). used a Visual GeometrCombinidoy Group 19-1ayer CNN (VGG19) model to estimate TC intensity on TC grayscale infrared images obtained from various geostationary satellites, which achieved lower RMSE. The transfer learning-based methods can improve the focus on the minority class while maintaining the classification accuracy of the majority class. When forecasting the formation of severe typhoons, it is necessary to focus on the accuracy of forecasting the formation of severe typhoons but the forecast accuracy of non-severe typhoons cannot be ignored. Therefore, we use the idea of transfer learning to build the USFP framework to solve the problem of unbalanced severe typhoon formation.

3 Problem definitions

The formation of typhoons requires a combination of both atmospheric and oceanic factors. To simulate the atmospheric and oceanic factors of typhoon formation, we convert the variables into multidimensional tensors. Since the atmospheric variables and ocean hydrographic variables are in a 3D space, the atmospheric environment field corresponding to a typhoon can be represented as L×W×H×A three-dimensional grid data, where L and W represent longitude and latitude, H represents the height of the atmosphere and A represents atmospheric environment variables. Similarly, the ocean hydrographic environment field corresponding to a typhoon can be represented as L×W×D×R three-dimensional grid data, where D represents the depth of the ocean and R represents ocean hydrographic variables. And since the sea surface variables are in a 2D space. The sea surface environment field corresponding to a typhoon can be represented as a L×W×O two-dimensional grid data, where L and W represent longitude and latitude and O represents sea surface variables. The prediction variables including the atmosphere, sea surface and ocean hydrography can be represented by $X=[\ X^P,X^S,X^O]$. Since the above spatial environment variables are time dependent, our problem can be regarded as a spatio-temporal prediction problem to predict whether a strong typhoon will form or not. Given typhoon spatio-temporal data: $X=x_{t-6b}$ (b=0,1,⋯,t/6), the predicted typhoon state can be defined as:

$$Y_{t+6k} = f(x_t, x_{t-6}, \cdots, x_{t-6b}), k = 1, 2, 3 \cdots, b = 0, 1, 2 \cdots \quad (1)$$

where t is the prediction moment, X_t represents the atmospheric, sea surface, and ocean hydrographic variables at the prediction moment, Y_{t+6k} represents the predicted typhoon state, k represents the timestep of the prediction and b represents the lookback step before the prediction moment. We consider typhoons with wind speed reaching 84 kt as strong typhoons, which are positive samples. The rest of the typhoons are as non-

strong typhoons, which are negative samples. When $Y_{t+6k}=1$, it means that a severe typhoon is formed, which is a positive sample. When $Y_{t+6k}=0$, it means that a severe typhoon is not formed, which is a negative sample. However, according to the defined standard, the number of positive and negative samples obtained is very unbalanced as shown in Table 1. Therefore, the severe typhoon formation prediction problem can be regarded as an unbalanced spatio-temporal series binary classification problem.

4 Methods

This section presents the details of the unbalanced severe typhoon formation prediction (USFP) framework. The first subsection introduces the overall architecture of the framework and the training method of the framework. The second subsection introduces the main structure of the severe typhoon pre-learning model. The last subsection details the structure of the unbalanced severe typhoon re-learning model and the loss function designed for it.

4.1 Architecture

The overall framework is shown in Figure 1. As we can see, the framework consists of two parts: the first part is the severe typhoon pre-learning model and the second part is the unbalanced severe typhoon re-learning model.

The first part is to learn from the balanced dataset. The structure and parameters of the feature extraction part learned from the balanced dataset are what we call prior knowledge in this paper. The severe typhoon pre-learning model uses 2D Convolutional Neural Networks (2DCNN) and 3D Convolutional Neural Networks (3DCNN) to extract features from high-dimensional data. The model weights are adjusted by adding classifiers. We train the severe typhoon pre-learning model by constructing a balanced dataset. The trained model architecture and the model weight parameters of the feature extraction part are saved as prior knowledge for transfer learning.

The second part is to apply the prior knowledge to train the unbalanced dataset. First, we transfer the obtained prior knowledge to the unbalanced severe typhoon re-learning model. Then, the unbalanced severe typhoon re-learning

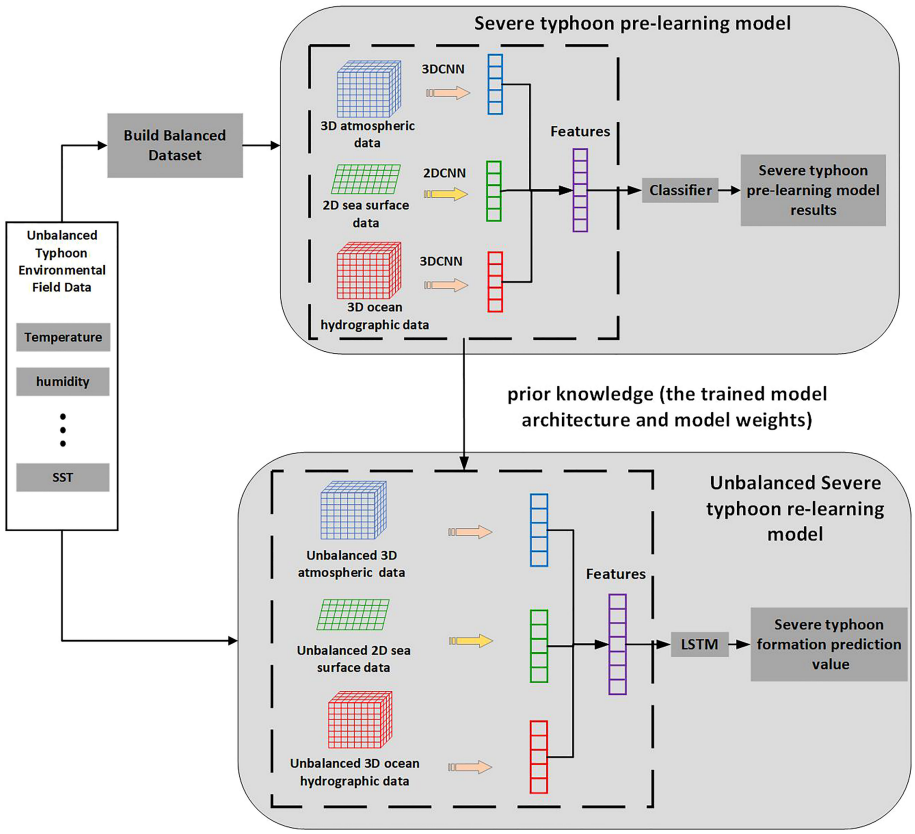


FIGURE 1
The overall structure of the USFP framework.

model with prior knowledge is trained using an unbalanced dataset. In addition, we design the unbalanced severe typhoon (UST) loss function to optimize the model by assigning different weights to the ordinary typhoon and the severe typhoon samples. Finally, the most accurate classification results are obtained by adjusting the parameters of the model and used as the USFP framework results of severe typhoon formation prediction.

4.2 Severe typhoon pre-learning model

To obtain prior knowledge, we construct a balanced dataset for training the severe typhoon pre-learning model according to the definition of positive and negative samples in Section 3. We fuse atmospheric variables, sea surface variables, and ocean hydrographic variables as prediction variables to form the typhoon environmental field data. We take temperature (t), relative humidity (rh), geopotential height (z), u-component of wind (u) and v-component of wind (v) as the basic atmospheric variables associated with typhoons. The sea surface temperature (sst) is used as the basic sea surface variable associated with typhoons. The seawater temperature (st), eastward seawater velocity (water_u) and northward seawater velocity (water_v) are used as the basic ocean hydrographic variables associated with typhoons. We classify and label these environmental variable datasets according to the typhoon historical best track dataset. The labeled data are sampled to construct a balanced dataset for training a severe typhoon pre-learning model.

We use different networks to extract the features of different environmental variables. Sea surface variables, as 2D environmental field information, need to be extracted features with a 2DCNN network. Atmospheric variables and ocean hydrographic variables, as 3D environmental field information, need to be extracted features with the 3DCNN network. By using the 2D convolution kernel and the 3D convolution kernel respectively, the corresponding 2D feature map and 3D feature map are obtained. The feature maps are passed through the flatten layer and the fully connected layer to obtain the feature vectors. The model structure and model weights of this part are used as prior knowledge extracted from the balanced dataset. The model architecture that is used as prior knowledge is shown in Figure 2.

According to the previous research (Chen et al., 2019), the structure of this feature extraction part is generalized into the formula as follows:

$$Features = Flatten\left(3DCNN\left(X^{P}_{t-6b}\right)\right) + Flatten\left(2DCNN\left(X^{S}_{t-6b}\right)\right) + Flatten\left(3DCNN\left(X^{O}_{t-6b}\right)\right) \quad (2)$$

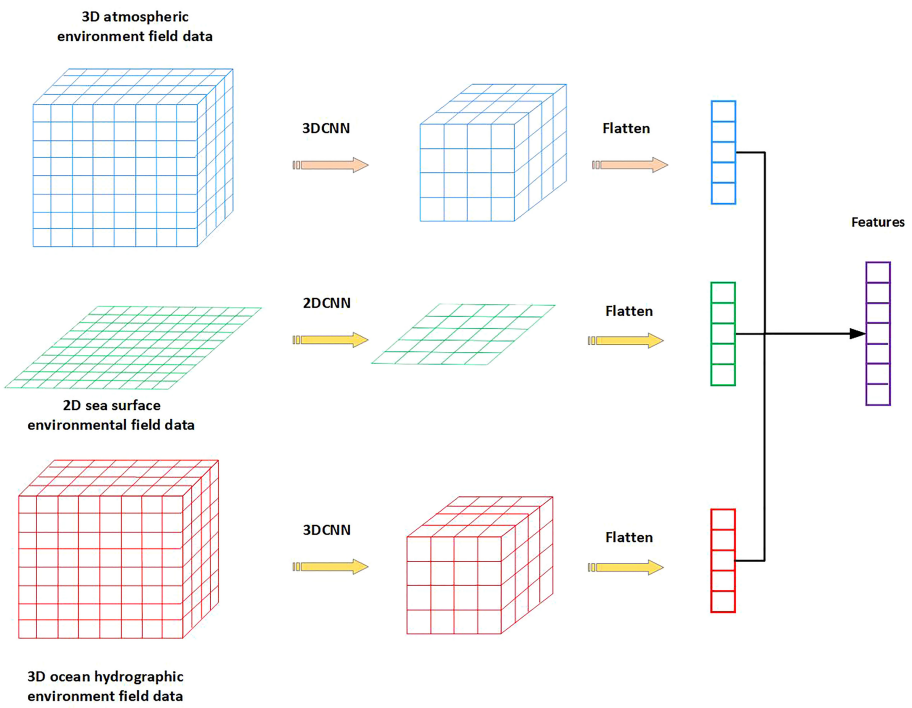


FIGURE 2
The model architecture used as prior knowledge in the strong typhoon pre-learning model.

The features learned by 3DCNN and 2DCNN are trained with the classifier. For the severe typhoon pre-learning model, we choose the commonly used binary cross-entropy loss function. The formula is as follows:

$$Loss = -Y_{t+6k} log Y'_{t+6k} - (1 - Y_{t+6k}) log\left(1 - Y'_{t+6k}\right) = \begin{cases} -log Y'_{t+6k}, & Y_{t+6k} = 1 \\ -log\left(1 - Y'_{t+6k}\right), & Y_{t+6k} = 0 \end{cases} \tag{3}$$

Where Y'_{t+6k} is the model prediction result and Y_{t+6k} is the real label value. $Y_{t+6k} = 1$ for positive samples and $Y_{t+6k} = 0$ for negative samples.

The model architecture and model weights of the trained feature extraction part are saved as prior knowledge extracted from the balanced dataset for the next transfer learning step.

4.3 Unbalanced severe typhoon re-learning model

The unbalanced severe typhoon re-learning model is composed of the transferred feature extraction component (prior knowledge) and the LSTM model. The input to the unbalanced strong typhoon relearning model is the unbalanced dataset. The unbalanced dataset is trained with prior knowledge to obtain feature vectors, which are used as the input to the LSTM. The LSTM implements the prediction of strong typhoon formation.

The LSTM model is an improvement of the recurrent neural network model, which can keep the error at a constant level and enhance the robustness. Figure 3 shows the operations performed by the LSTM unit at time t, where X_t refers to the input at the current time, C_{t-6} refers to the cell state 6 hours before time t, h_{t-6} refers to the hidden state 6 hours before time t, and C_t refers to the information that can be stored in the LSTM cell. Through the three control units of input gate, output gate and forget gate in LSTM, it is determined which of the input information will be forgotten and which will be retained. Finally, the cell state C_t and the hidden state h_t corresponding to time t are obtained. LSTM can be represented as:

$$h_t, C_t = LSTM([h_{t-6}, X_t], C_{t-6}) \tag{4}$$

The LSTM model obtains the predicted value Y_{t+6k} after 6k hours at time t. Thereby, the unbalanced severe typhoon re-learning model can be expressed simply as:

$$Y_{t+6k} = LSTM\left(Features\left(X^P_{t-6b}, X^S_{t-6b}, X^O_{t-6b}\right)\right) \tag{5}$$

We expect the unbalanced severe typhoon re-learning model to pay more attention to the severe typhoon samples in the classification process of unbalanced datasets. Therefore, the unbalanced severe typhoon loss function designed in this paper assigns different weights to the strong typhoon samples, which is calculated as follows:

$$UST\ Loss = \begin{cases} -\alpha\left(1 - Y'_{t+6k}\right)^{\gamma} log Y'_{t+6k}, & Y_{t+6k} = 1 \\ -(1-\alpha) Y'^{\gamma}_{t+6k} log\left(1 - Y'_{t+6k}\right), & Y_{t+6k} = 0 \end{cases} \tag{6}$$

where γ is the weight factor used to focus on difficult and misclassified samples and α is the balance factor used to balance the unbalanced proportion of positive and negative samples. The higher the value of γ is set, the more it focuses on difficult samples. The value of α is appropriately weighted according to

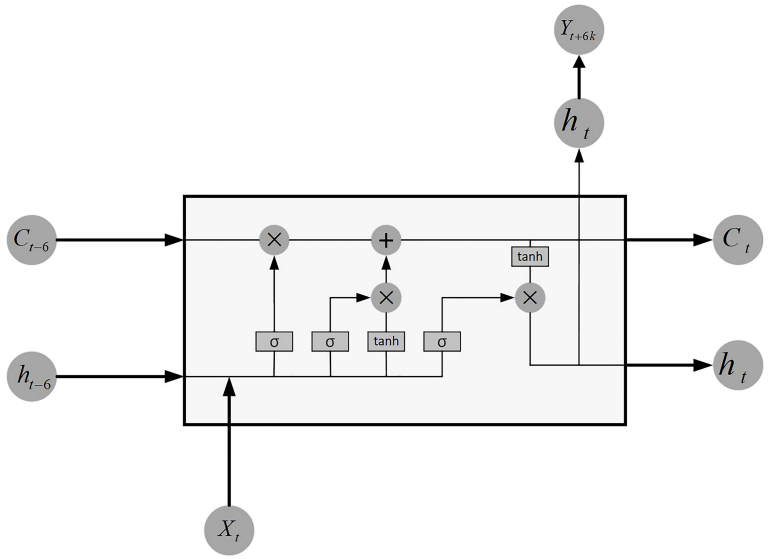


FIGURE 3
The classifier LSTM unit.

the sample ratio setting in the experiments. In the experiment, $\gamma = 2$ and $\alpha = 0.9$. The analysis of γ is shown in Section 5.4.2.4.

After building the unbalanced strong typhoon relearning model, it is necessary to freeze and retrain the prior knowledge part of the model. Adjustments to the number of frozen and retrained layers are called fine-tuning operations in transfer learning. By freezing some layers in the prior knowledge, the number of training parameters of the model can be adjusted. In experiments, we try to find the best combination between the number of frozen layers and the number of retraining layers to achieve better predictions. The analysis of fine-tuning experimental results is presented in Section 5.4.2.1.

5 Experiments

This section details the experiments performed with the USFP framework. It mainly includes the experimental dataset, the evaluation metrics of the framework, the implementation of the experiment, and the analysis of the experiment results.

5.1 Datasets

The typhoon track dataset used in this paper is the World Meteorological Organization (WMO) version of the International Best Trajectory Archive for Climate Management (IBTrACS) Global Tropical Cyclone Best Trajectory Dataset. The atmospheric variable and sea surface variable dataset used in this paper are the ERA-Interim Reanalysis Dataset. The ocean hydrographic variable dataset used in this paper is the Hybrid Coordinate Ocean Model (HYCOM) dataset. The data for each moment of each typhoon in the three datasets correspond to each other. The first two datasets were recorded from 1979 to 2016. The ocean hydrographic datasets were recorded from 1994 to 2015. The wind speed is recorded every 6 hours from the time of tropical cyclone formation. Since typhoons are high-impact weather, the surrounding environmental fields change drastically. Therefore, high-resolution and multi-level data were selected during the experiment. According to previous research on high-dimensional data (Wang et al., 2021), this paper uses a neural network to reduce the dimension of high-dimensional data. In addition, compared with traditional deep learning datasets, typhoon forecasts need to ensure timeliness. The longer the timeliness of the general typhoon forecast business forecast, the more valuable it is for reference. But too long timeliness will also lead to a reduction in the number of training samples. Therefore, the prediction time of 24 hours was chosen to ensure a balance between the sample size and the timeliness of the model.

In the experiment, this paper selects typhoons in the Western Pacific (WP), Eastern Pacific (EP), and North Atlantic (NA) regions as samples. According to previous research (Camargo et al., 2007; Chen et al., 2019), temperature (t), relative humidity (rh), geopotential height (z), u component of wind (u), and v component of wind (v) are selected as atmospheric variables. The atmospheric pressure level of 1000/975/925/850/800/700/600/500/400/300/200/100hpa is chosen. Sea surface temperature (sst) is selected as the sea surface variable. According to previous research (Shay et al., 2000; Wu et al., 2007; Goni et al., 2009; Lin et al., 2009; Vissa et al., 2013), ocean heat content (UOHC), eddy currents and other ocean features play an important role in the intensification of tropical cyclones. Therefore, seawater temperature (st), eastward sea water velocity (water_u), and northward water velocity (water_v) are selected as ocean hydrographic variables. The ocean depth of 100/90/80/70/60/50/45/40/35/30/25/20/15/12/10/8/6/4/2m is chosen.

This part is about constructing input datasets and output datasets. For the 3DCNN part, the dimension of the input dataset is N×T×L×W×H×A or N×T×L×W×D×R. For the 2DCNN part, the dimension of the input dataset is N×T×L×W×O. For the LSTM classifier part, the dimension of the input dataset is N×T×K. Among them, N represents the number of samples. L and W range in 33 ~ 161. T represents the time step and $T = 5$. $H = 12$. $D = 19$. A = 5. $R = 3$. $O = 1$. K represents the length of the eigenvector. The samples at the moment when the maximum wind speed near the typhoon center value reaches the severe typhoon standard are marked as 1 and the rest of ordinary typhoons samples are marked as 0. In the experiment without ocean hydrographic variables, the data set is constructed by random sampling of the atmospheric environmental field data and the sea surface environmental field data. In the ocean hydrographic variables experiment, we intercept atmospheric and sea surface variables recorded from 1994 to 2015, which are recorded at the same time as the ocean hydrographic variables. After screening the original dataset, we constructed a balanced dataset with the same number of positive and negative samples and an unbalanced dataset with a positive and negative sample ratio of about 1:9 in different regions. The

TABLE 2 The number of positive and negative samples in the experiment.

Experiment Type	Sample	WP	EP	NA
No ocean hydrographic variables experiment	Positive samples	100	90	100
	Negative samples	900	810	900
Ocean hydrographic variables experiment	Positive samples	100	50	77
	Negative samples	900	430	693

specific sample size is shown in Table 2. Furthermore, the dataset is split into a 70% training set and a 30% test set.

5.2 Evaluation metrics

To measure the performance of the model, the evaluation metric for traditional binary classification models is generally accuracy (Acc). The specific formula is as follows:

$$Acc = \frac{TP+TN}{TP+FN+FP+TN} \tag{7}$$

However, it can be seen from the formula that when the dataset is unbalanced, the model training process will be more biased towards negative samples to make TN much larger than TP to obtain a larger Acc value. Therefore, according to the evaluation metrics selected when solving unbalanced data in the official Tensorflow document (Abadi et al., 2015), we select the values of ROC_AUC, PR_AUC, and F1 as evaluation metrics to measure the actual performance of the model.

ROC_AUC refers to the area under the ROC curve. The abscissa of the ROC curve is the false positive rate ($FPR = \frac{FP}{FP+TN}$) and the ordinate is the true positive rate ($TPR = \frac{TP}{TP+FN}$). The larger the ROC_AUC value, the more likely the current classification model will place positive samples in front of negative samples, which can better classify these samples.

$$ROC_AUC = \int ROC \tag{8}$$

PR_AUC refers to the area under the PR curve. The abscissa of the PR curve is the recall rate ($recall = \frac{TP}{TP+FN}$) and the ordinate is the precision rate ($precision = \frac{TP}{TP+FP}$). When the PR_AUC value is larger, the positive sample classification effect is better.

$$PR_AUC = \int PR \tag{9}$$

F1 is the harmonic mean of precision and recall. Since the precision rate and the recall rate are contradictory to a certain extent, the F1 value is used to evaluate the precision rate and the recall rate as a whole.

$$F1 = \frac{2\times precision\times recall}{precision+recall} \tag{10}$$

5.3 Implementation

The USFP framework is implemented in tensorflow2.4 using Keras. The experimental model was trained using a Tesla V100 GPU card. Since there are some missing data in the typhoon environmental field dataset, the missing data are filled in and the whole data are normalized before the experiment. The constructed balanced dataset is used to train the severe typhoon pre-learning model.

In the severe typhoon pre-learning model 3DCNN module, three 3D convolutional layers, a maximum pooling layer, a flat layer, and a fully connected layer are used. For learning the atmospheric environmental field data features, the size of each layer of the convolutional layer is 5×5×1, the stride is 2×2×1, and the number of filters is 64, 64, and 128. The size of the max pooling layer is 5×5×1 and the stride is 2×2×1. For learning the characteristics of the ocean hydrographic environmental field data, the size of each layer of the convolutional layer is 3×3×1, the stride is 2×2×1, and the number of filters is 64, 64, and 128. The size of the max pooling layer is 3×3×1 and the stride is 2×2×1. The high-dimensional data is dimensionally reduced using a flatten layer. Both variables output 100 feature vectors through the fully connected layer. In the severe typhoon pre-learning model 2DCNN module, three 2D convolutional layers, one max pooling layer, one flatten layer, and one fully connected layer are used. The size of each layer of the convolutional layer is 5×5, the stride is 2×2, and the number of filters is 32, 64, and 128. The size of the max pooling layer is 5×5 and the stride is 2×2. The high-dimensional data is dimensionally reduced using a flatten layer. Then, the data outputs 100 feature vectors through the fully connected layer. Besides, the feature extraction part of this model needs to be encapsulated using the TimeDistributed layer wrapper to assign the same weights to the data in the time latitude. The activation function of all layers is 'relu'.

After training the severe typhoon pre-learning model, the obtained prior knowledge is transferred to the unbalanced severe typhoon re-learning model and combined with LSTM for training. The L2 regularization is added to the fully connected layer to prevent overfitting of the model during training. After the framework structure is constructed, the hyperparameters are tuned to obtain the best prediction results.

5.4 Analysis

5.4.1 Result analysis

In order to prove that the USFP framework proposed in this paper can effectively improve the unbalanced data problem, experiments compare the framework with traditional machine learning-based methods. Since the environmental data used in this paper has a spatiotemporal dimension, in order to ensure the consistency of the comparative experiments, the machine learning model involved in the comparison needs to be able to consider the spatiotemporal relationship in the dataset. Therefore, this paper selects the ConvLSTM model (Shi et al., 2015), which can learn spatiotemporal features of data and the hybrid CNN_LSTM model (Chen et al., 2019), which has a good effect on typhoon formation and intensity prediction, as the comparison objects. As a spatiotemporal sequence prediction model, the ConvLSTM model can well capture the spatial information of the data based on the LSTM model. The typhoon spatiotemporal depth mixed prediction model [75]

has a good performance in the formation of typhoon and the prediction of typhoon intensity, and the model has a high generalization ability. In addition, this paper also compares the proposed framework with other methods combining CNN and LSTM, such as 2DCNN+LSTM, 3DCNN+LSTM. Furthermore, since these models do not take into account the imbalance of data set samples, some traditional methods of dealing with imbalanced data are added to the original model for comparison. These machine learning model methods selected in the experiments all use the best parameters provided by the original author's paper, and use the same training and test sets for training and testing. The specific experimental results are shown in Table 3.

As shown in Table 3, the USFP framework in the WP region can achieve the best ROC_AUC value of 0.735, the best PR_AUC value of 0.3, and the best F1 value of 0.311. In comparison, the best ROC_AUC of traditional machine learning-based methods is only 0.666, the best PR_AUC is only 0.262, and the best F1 is only 0.233. This is because traditional machine learning-based methods do not consider the unbalance of the data and the model cannot obtain good results based on the actual unbalanced dataset. In addition, the traditional methods of handling unbalanced data assign different weights to different samples based on the ratio between samples, which cannot reflect the actual distribution of data features. Therefore, these methods also cannot lead to better classification performance of the model. To ensure the generalization ability of the framework on different datasets and the robustness of the framework, experiments are also conducted on the EP and NA regions. In the EP region, the best ROC_AUC of traditional machine learning-based methods is 0.562, the best PR_AUC is 0.203, and the best F1 is 0.200. The USFP framework can achieve the best ROC_AUC value of 0.625, the best PR_AUC value of 0.205, and the best F1 value of 0.252. In the NA region, the best ROC_AUC of traditional machine learning-based methods is 0.698, the best PR_AUC is 0.270, and the best F1 is 0.270. The USFP framework can achieve the best ROC_AUC value of 0.746, the best PR_AUC value of 0.365, and the best F1 value of 0.328.

5.4.2 Parametric analysis

5.4.2.1 Fine tuning

The USFP framework is fine-tuned in the experiments according to the fine-tuning method in Section 4.3. The experimental results are shown in Figure 4. As the number of frozen layers increases, the three types of evaluation metrics fluctuate to varying degrees in different regions. Through experimental verification, it can be concluded that in the WP area, when there is no frozen layer, the PR_AUC value reaches the best value; and when the number of frozen layers is two, the F1 and ROC_AUC values reach the best value. In the NA area, when the number of frozen layers is 15, the three types of evaluation metrics all reach the best values. In the EP area, when the number of frozen layers is 8, the ROC_AUC value reaches the optimal value; when the number of frozen layers is 2, the F1 value and PR_AUC value reach the optimal value.

5.4.2.2 Severe typhoon pre-learning model LR

The gradient descent algorithm is a commonly used optimization algorithm for deep learning, which can calculate the gradient through partial derivatives. The learning rate (LR) is one of the important parameters to control the update rate of parameters. According to previous studies (Kornblith et al., 2019), the learning effect of the severe typhoon pre-learning model has a great influence on the effect of the USFP framework. Therefore, the LR of the severe typhoon pre-learning model is selected as the adjustment parameter for analysis. The specific results are shown in Figure 5.

As shown in Figure 5, as the LR of the severe typhoon pre-learning model decreases, different metrics in different regions have certain fluctuations. In the WP region, ROC_AUC, PR_AUC, and F1 reach their maximum values when the severe typhoon pre-learning model LR is 0.001, 0.0003, and 0.0003 respectively. In the NA area, each metric increases and then decreases as the LR of the severe typhoon pre-learning model decreases. ROC_AUC, PR_AUC, and F1 reached the maximum

TABLE 3 Comparison of experimental results of different models.

Model Type	WP			EP			NA		
	ROC_AUC	PR_AUC	F1	ROC_AUC	PR_AUC	F1	ROC_AUC	PR_AUC	F1
ConvLSTM	0.596	0.129	0.204	0.519	0.151	0.181	0.520	0.104	0.185
2DCNN-LSTM	0.571	0.109	0.208	0.518	0.180	0.148	0.521	0.163	0.130
3DCNN-LSTM	0.601	0.262	0.233	0.524	0.141	0.181	0.698	0.259	0.270
Hybrid CNN-LSTM	0.638	0.224	0.229	0.562	0.203	0.179	0.635	0.264	0.204
Hybrid CNN-LSTM +focal loss	0.582	0.225	0.202	0.525	0.197	0.163	0.637	0.270	0.205
Hybrid CNN-LSTM + focal loss + class weight	0.666	0.221	0.194	0.512	0.102	0.181	0.578	0.170	0.181
USFP framework	**0.735**	**0.300**	**0.311**	**0.625**	**0.205**	**0.252**	**0.746**	**0.365**	**0.328**

WP, Western Pacific; EP, Eastern Pacific; NA, North Atlantic. The bold values in the table represent the largest value in each column.

value when the LR of the severe typhoon pre-learning model is 0.0003, 0.0004, and 0.0002 respectively. In the EP region, ROC_AUC, PR_AUC, and F1 all reach the maximum value when the severe typhoon pre-learning model LR is 0.0003. It can be concluded that when the LR value of the severe typhoon pre-learning model is selected as 0.0003, the unbalanced strong typhoon formation prediction framework has a good prediction effect in each region. Therefore, the LR value of the severe typhoon pre-learning model should be chosen to be 0.0003, so that the framework has better generalization ability.

5.4.2.3 Epochs

In this paper, we analyzed the epochs parameter to determine the period to obtain the best performance of the model. As shown in Figure 6, in both the WP region and the NA region, the ROC_AUC values tend to stabilize after 30 epochs of training. The ROC_AUC value in the EP region fluctuates relatively wildly, but after 10 rounds of training, the value is relatively stable within a certain range. Therefore, in our experiments, a relatively suitable training period is around 30 epochs.

5.4.2.4 Weight factor

This section conducts a parametric analysis for the unbalanced severe typhoon loss function in the unbalanced severe typhoon re-learning model. Since the value of the balance factor α is determined by the sampling ratio in the experiment, the α is set to 0.9 in the experiments. We performed a parametric analysis of the weighting factor γ. As can be seen from the Table 4, the ROC_AUC value in the WP region increases and then decreases as γ increases. There are some fluctuations in the F1 and PR_AUC values in the WP region. When the γ is 2, the best effect of the model is achieved on the WP region. Therefore, the γ is set to 2 in this paper.

5.4.3 Variables analysis

To study the influence of ocean hydrographic variables on the formation of severe typhoons, we selected seawater temperature

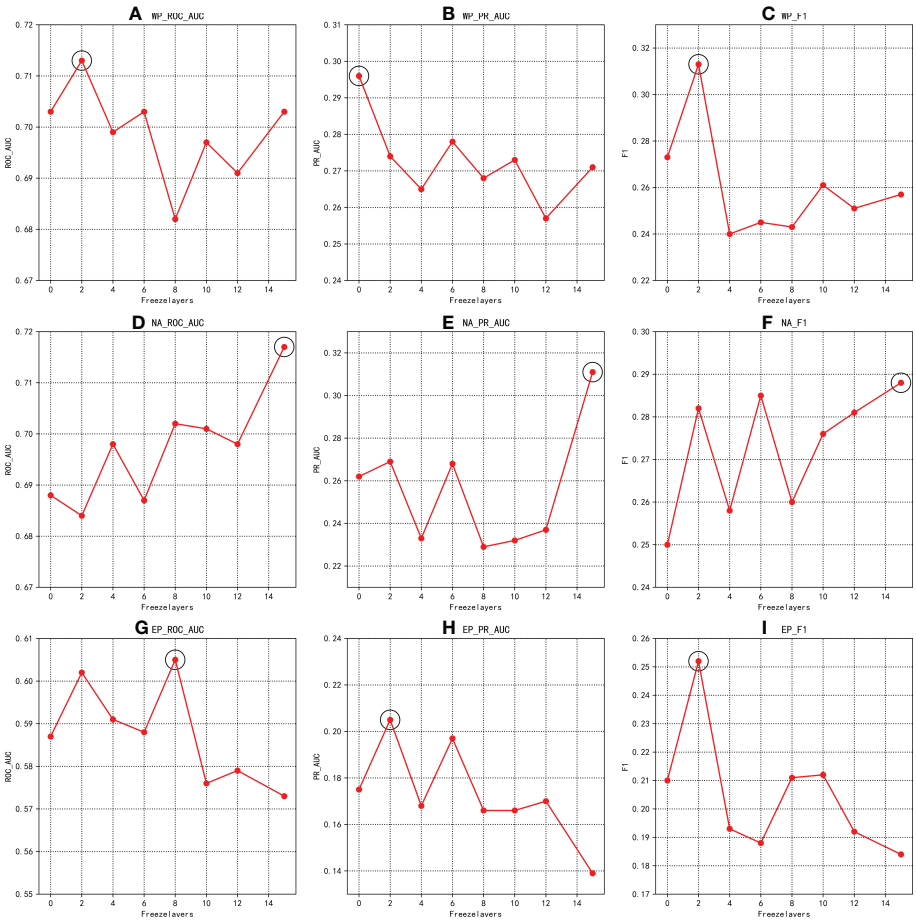


FIGURE 4
In different regions, the three types of evaluation metrics change with the number of frozen layers. **(A–C)** represent the results of the three indicators for the WP region. **(D–F)** represent the results of the three indicators for the NA region. **(G–I)** represent the results of the three indicators for the EP region.

TABLE 4 Parametric analysis for weight factor γ.

Weight factor γ	WP		
	ROC_AUC	PR_AUC	F1
0.5	0.683	0.228	0.268
1	0.704	0.267	0.310
2	**0.735**	**0.300**	**0.311**
5	0.698	0.224	0.253
10	0.659	0.284	0.259

WP, Western Pacific; EP, Eastern Pacific; NA, North Atlantic. The bold values in the table represent the largest value in each column.

(st), eastward seawater velocity (water_u), and northward current velocity (water_v) as ocean hydrographic variables. Atmospheric environmental field data, sea surface environmental field data, and ocean hydrographic environmental field data are used together as input data for the experiments. The experimental results are shown in Table 5.

In the WP region, the USFP framework can achieve the best ROC_AUC value of 0.754, the best PR_AUC value of 0.286, and the best F1 value of 0.382. In comparison, the best ROC_AUC of traditional machine learning-based methods is only 0.712, the best PR_AUC is only 0.243, and the best F1 is only 0.205. In the EP region, the best ROC_AUC of the traditional machine

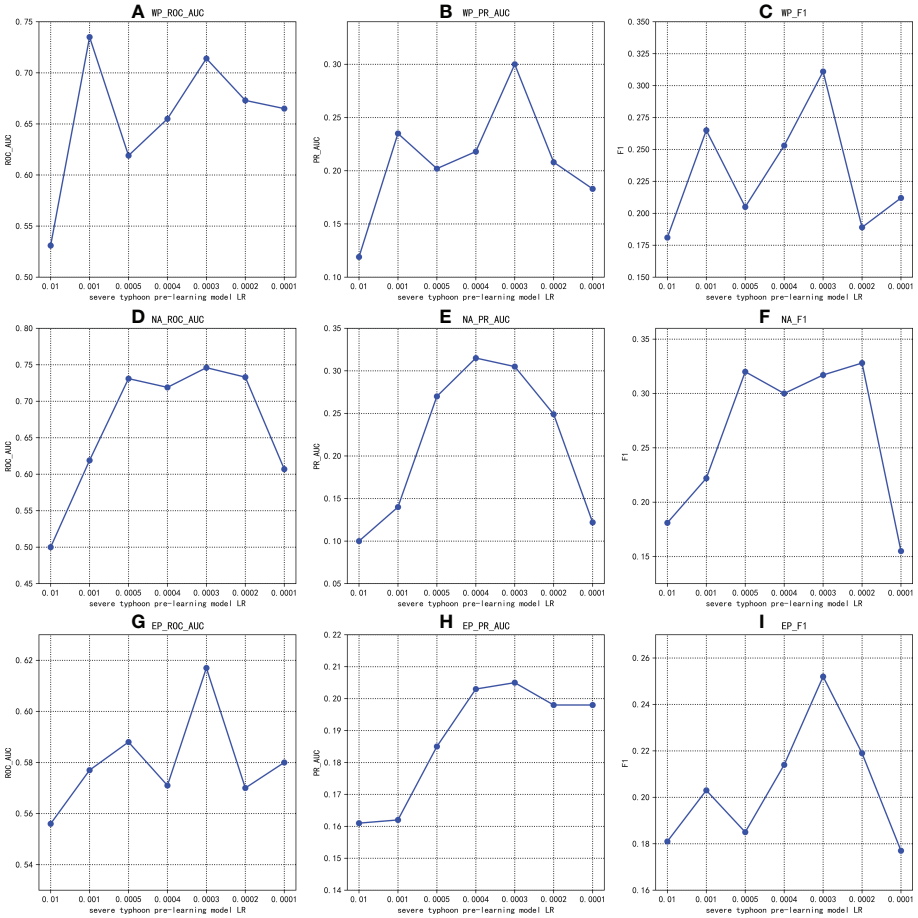


FIGURE 5
In different regions, the three types of evaluation metrics change with the severe typhoon pre-learning model LR. **(A–C)** represent the results of the three indicators for the WP region. **(D–F)** represent the results of the three indicators for the NA region. **(G–I)** represent the results of the three indicators for the EP region.

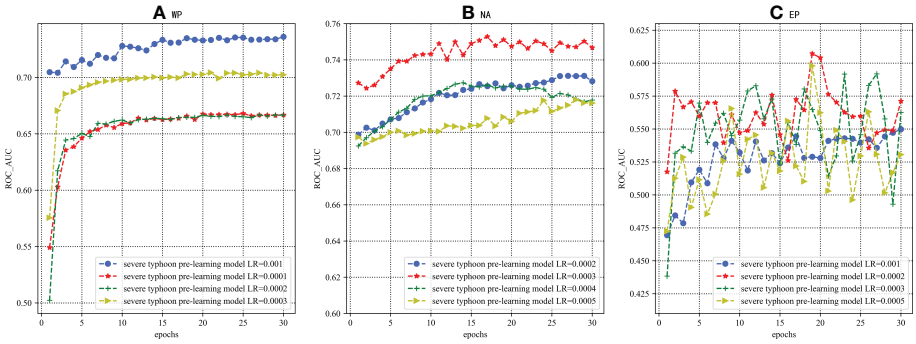


FIGURE 6
The ROC_AUC results with the change of epochs.

learning-based method is 0.574, the best PR_AUC is 0.227, and the best F1 is 0.217. The USFP framework can achieve the best ROC_AUC value of 0.671, the best PR_AUC value of 0.252, and the best F1 value of 0.263. In the NA region, the best ROC_AUC of the traditional machine learning-based methods is 0.697, the best PR_AUC is 0.341, and the best F1 is 0.222. The USFP framework can achieve the best ROC_AUC value of 0.783, the best PR_AUC value of 0.397, and the best F1 value of 0.365.

Besides, we compare the experimental results with the ocean hydrographic variables to the experimental results without the ocean hydrographic variables. The comparison results are shown in Table 6. After adding the ocean hydrographic data, the evaluation metrics for all three regions (except the PR_AUC metric in the WP region) have improved. This proves that ocean hydrographic information has some influence on the formation of severe typhoons and helps to improve the accuracy of severe typhoon formation prediction.

5.4.4 Compare with the numerical model IFS

The numerical model prediction results used in this paper were obtained from the THORPEX Interactive Grand Global Ensemble (TIGGE) Model Tropical Cyclone Track Dataset. The dataset contains historical forecasting results for multiple official organizational models. We chose the historical forecasts of the European Centre for Medium-Range Weather Forecasts (ECMWF) ensemble forecast system (IFS), which has the longest record year in this dataset, as the object of comparison for the experiment. The historical forecast results document the predictions of tropical cyclone tracks from 2006 to 2016. Firstly, the temporal and central latitudinal and longitudinal positions provided by the World Meteorological Organization (WMO) version of the International Best Track Archive for Climate Management (IBTrACS) global tropical cyclone best track dataset are matched one-to-one with the historical prediction results of the IFS model to obtain the tropical cyclone samples. Secondly, the intensity values of tropical cyclone samples filtered from the global tropical cyclone best track dataset were used as labels to count the number of positive and negative samples for the three regions. The results are shown in Table 7.

Finally, the 24-hour step forecasts of the tropical cyclone samples filtered from the historical forecast results of the IFS model were used as model results and evaluated using the same evaluation metrics. The evaluation results are compared with the experimental results of the USFP framework. The comparison results are shown in Table 8.

The results of the USFP framework without ocean hydrographic variables and the USFP framework with ocean hydrographic variables outperform the prediction results of the

TABLE 5 Comparison of experimental results of different models after adding ocean hydrographic variables.

Model Type	WP			EP			NA		
	ROC_AUC	PR_AUC	F1	ROC_AUC	PR_AUC	F1	ROC_AUC	PR_AUC	F1
3DCNN-LSTM	0.712	0.201	0.062	0.574	0.193	0.197	0.653	0.206	0.182
Hybrid CNN-LSTM	0.681	0.170	0.181	0.491	0.112	0.077	0.697	0.341	0.222
Hybrid CNN-LSTM +focal loss	0.642	0.172	0.197	0.559	0.227	0.217	0.650	0.181	0.204
Hybrid CNN-LSTM + focal loss + class weight	0.705	0.243	0.205	0.551	0.175	0.162	0.684	0.204	0.181
USFP framework	**0.754**	**0.286**	**0.382**	**0.671**	**0.252**	**0.263**	**0.783**	**0.397**	**0.365**

WP, Western Pacific; EP, Eastern Pacific; NA, North Atlantic. The bold values in the table represent the largest value in each column.

TABLE 6 Comparison of results between the USFP frameworks containing different environmental field variables.

Model Type	WP			EP			NA		
	ROC_AUC	PR_AUC	F1	ROC_AUC	PR_AUC	F1	ROC_AUC	PR_AUC	F1
USFP framework (No ocean hydrographic variables)	0.735	**0.300**	0.311	0.625	0.205	0.252	0.746	0.365	0.328
USFP framework (ocean hydrographic variables)	**0.754**	0.286	**0.382**	**0.671**	**0.252**	**0.263**	**0.783**	**0.397**	**0.365**

WP, Western Pacific; EP, Eastern Pacific; NA, North Atlantic. The bold values in the table represent the largest value in each column.

TABLE 7 The number of positive and negative samples screened by official data statistics.

	WP	EP	NA
Positive sample	81	48	34
Negative sample	422	262	314

TABLE 8 Comparison of results between the USFP framework and the numerical model IFS of ECMWF.

Model Type	WP			EP			NA		
	ROC_AUC	PR_AUC	F1	ROC_AUC	PR_AUC	F1	ROC_AUC	PR_AUC	F1
the numerical model IFS of ECMWF	0.546	**0.320**	0.184	0.490	0.172	0.041	0.539	0.339	0.150
USFP framework (No ocean hydrographic variables)	0.735	0.300	0.311	0.625	0.205	0.252	0.746	0.365	0.328
USFP framework (ocean hydrographic variables)	**0.754**	0.286	**0.382**	**0.671**	**0.252**	**0.263**	**0.783**	**0.397**	**0.365**

WP, Western Pacific; EP, Eastern Pacific; NA, North Atlantic. The bold values in the table represent the largest value in each column.

IFS model for all metrics except the PR_AUC metric in the WP region. This proves that the USFP framework outperforms the numerical model IFS of ECMWF in the field of unbalanced severe typhoon formation prediction.

5.4.5 Example prediction

We performed a USFP framework forecast using the example of strong typhoon Roke in 2011. The generation of strong typhoon was successfully predicted by using USFP framework. The value of F1 is 1 because there is only one positive sample for a single sample and the calculated value of ROC_AUC is 0.734 and the calculated value of PR_AUC is 0.2. By checking the historical prediction results of the numerical model IFS of ECMWF, we can conclude that the numerical model did not accurately predict the generation of severe typhoons.

6 Conclusions and future work

In this paper, we define the severe typhoon formation prediction problem as a classification problem of spatio-temporal series prediction and propose an unbalanced severe typhoon formation prediction (USFP) framework. The framework fuses atmospheric, sea surface, and ocean hydrographic variables and uses a severe typhoon pre-learning model to obtain prior knowledge from the constructed balanced dataset. Then, we transfer the prior knowledge to the severe typhoon re-learning model to predict the formation of severe typhoons. Extensive experiments show that the USFP framework proposed in this paper is more accurate than the numerical model IFS of ECMWF and existing machine learning models.

Additional analysis and experiments on the parameters of the framework can lead to better results of the framework. We analyzed three parameters in our experiments: the number of frozen layers, the LR of the severe typhoon pre-learning model, and the number of epochs. The number of freezing layers is 2 or 15 layers to obtain better results. The optimal LR of the severe typhoon pre-learning model is 0.0003. The optimal epochs to adjust the iterations is about 30 rounds. In addition, we conducted a comparative experiment on the choice of environmental field variables. The experimental results show that adding the ocean hydrographic environmental field variables can help improve the prediction effect of the framework on the formation of severe typhoons.

In the future, we will further refine the details of the parameters in the USFP framework and use various data such as satellite image data to improve the framework functionality. In addition, since the framework is purely data-driven, it suffers from the problem of uninterpretability. Therefore, in further developmental work, we will try to integrate this framework with traditional physical models to improve the application ability of the model in practical systems.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: https://www.ncdc.noaa.gov/ibtracs/, https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era-interim, https://www.hycom.org/dataserver/gofs-3pt1/reanalysis, https://rda.ucar.edu/datasets/ds330.3/index.html#!description.

Author contributions

XP and XW developed the idea. XP designed the research, analyzed the data, and wrote the manuscript. XW guided the whole study and revised the manuscript. CZ provides support and assistance on dataset acquisition. XW, HW, and SW helped to revise the manuscript. XW, HW, SW, and JW gave comments on the manuscript. SC helped with the English translation of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This research is partially supported by the National Natural Science Foundation of China (42276205) and the National Natural Science Foundation of China (41875121).

Acknowledgments

Special thanks to the Institute of Numerical Meteorology and Oceanography, College of Meteorology and Oceanography, National University of Defense Technology for the equipment support.

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., et al. (2015). TensorFlow: Large-scale machine learning on heterogeneous systems. In *12th USENIX symposium on operating systems design and implementation*. (OSDI 16) (pp. 265–283).

Ahijevych, D., Pinto, J. O., Williams, J. K., and Steiner, M. (2016). *Probabilistic forecasts of mesoscale convective system initiation using the random forest data mining technique*. (Boston MA, USA: Weather & Forecasting) 31 (2), 581–599.

Al-Stouhi, S., and Reddy, C. K. (2016). Transfer learning for class imbalance problems with inadequate data. *Knowledge Inf. Syst.* 48 (1), 201–228. doi: 10.1007/s10115-015-0870-3

Camargo, S. J., Emanuel, K. A., and Sobel, A. H. (2007). Use of a genesis potential index to diagnose enso effects on tropical cyclone genesis. *J. Climate* 20 (19), 4819–4834. doi: 10.1175/JCLI4282.1

Chen, R., Wang, X., Zhang, W., Zhu, X., Li, A., and Yang, C. (2019). A hybrid CNN-LSTM model for typhoon formation forecasting. *GeoInformatica* 23 (3), 375–396. doi: 10.1007/s10707-019-00355-0

China Meteorological Administration (2020). *China Meteorological disaster yearbook 2019* (Beijing, China: Meteorological Press).

Combinido, J. S., Mendoza, J. R., and Aborot, J. (2018). "A convolutional neural network approach for estimating tropical cyclone intensity using satellite-based infrared images," in *International Conference on Pattern Recognition*. (Chinese Acad Sci, Inst Automat, Beijing, PEOPLES R CHINA: IEEE), 1474–1480.

Cui, Y., Jia, M., Lin, T. Y., Song, Y., and Belongie, S. (2019). "Class-balanced loss based on effective number of samples," in *IEEE Conference on Computer Vision and Pattern Recognition*. (Long Beach, CA: IEEE Comp Soc).

Demaria, M., and Kaplan, J. (1994). A statistical hurricane intensity prediction scheme (ships) for the atlantic basin. *Wea Forecasting* 9 (2), 209–220. doi: 10.1175/1520-0434(1994)009<0209:ASHIPS>2.0.CO;2

DeMaria, M., and Kaplan, J. (1999). *An updated statistical hurricane intensity prediction scheme (SHIPS) for the Atlantic and Eastern north pacific basins* Vol. 14 (Boston MA, USA: Weather & Forecasting), 326–337.

Deo, R. V., Chandra, R., and Sharma, A. (2017). Stacked transfer learning for tropical cyclone intensity prediction. *arXiv preprint* arXiv:1708.06539. doi: 10.48550/arXiv.1708.06539

Elsberry, R. L., Tsai, H. C., Chin, W. C., and Marchok, T. P. (2021). Predicting rapid intensification events following tropical cyclone formation in the western north pacific based on ecmwf ensemble warm core evolutions. *Atmosphere* 12. doi: 10.3390/atmos12070847

Fei, T., Huang, B., Wang, X., Zhu, J., Chen, Y., Wang, H., et al. (2022). A hybrid deep learning model for the bias correction of SST numerical forecast products using satellite data. *Remote Sens.* 14 (6), 1339. doi: 10.3390/rs14061339

Fritsch, J. M., and Chappell, C. F. (1980). Numerical prediction of convectively driven mesoscale pressure systems. part II. mesoscale model. *J. Atmospheric Sci.* 37 (8), 1734–1762. doi: 10.1175/1520-0469(1980)037<1734:NPOCDM>2.0.CO;2

Geng, Y., and Luo, X. (2019). Cost-sensitive convolutional neural networks for imbalanced time series classification. *Intelligent Data Analysis* 23 (2), 357–370.

Goni, G., DeMaria, M., Knaff, J., Sampson, C., Ginis, I., Bringas, F., et al. (2009). Applications of satellite-derived ocean measurements to tropical cyclone intensity forecasting. *Oceanography* 22 (3), 190–197. doi: 10.5670/oceanog.2009.78

Halperin, D. J., Fuelberg, H. E., Hart, R. E., Cossuth, J. H., Sura, P., and Pasch, R. J. (2013). An evaluation of tropical cyclone genesis forecasts from global numerical models. *Weather Forecasting* 28 (6), 1423–1445.

Han, H., Wang, W. Y., and Mao, B. H. (2005). *Borderline-smote: a new over-sampling method in imbalanced data sets learning* (Beijing, China: Lecture Notes in Computer Science).

Higa, M., Tanahara, S., Adachi, Y., Ishiki, N., Nakama, S., Yamada, H., et al. (2021). Domain knowledge integration into deep learning for typhoon intensity classification. *Sci. Rep.* 11 (1), 1–10. doi: 10.1038/s41598-021-92286-w

Hu, Z., Li, X., Tu, C., Liu, Z., and Sun, M.. (2018). "Few-shot charge prediction with discriminative legal attributes[C]," in *Proceedings of the 27th International Conference on Computational Linguistics*. (Santa Fe, New Mexico, USA: Association for Computational Linguistics), 487–498.

John, P. C. (2019). *National hurricane center forecast verification report, (2019)* (Miami, FL, USA: National Hurricane Center).

Kim, M., Park, M. S., Im, J., Park, S., and Lee, M. I. (2019). Machine learning approaches for detecting tropical cyclone formation using satellite data. *Remote Sens.* 11 (10), 1195. doi: 10.3390/rs11101195

Knaff, J. A., Sampson, C. R., and Demaria, M. (2005). *An operational statistical typhoon intensity prediction scheme for the western north pacific* Vol. 20 (Boston MA, USA: Weather & Forecasting), 688–699.

Kornblith, S., Shlens, J., and Le, Q. V. (2019). "Do better imagenet models transfer better?," in *IEEE Conference on Computer Vision and Pattern Recognition*. (Long Beach, CA: IEEE Comp Soc), 2661–2671.

Lee, H., Park, M., and Kim, J. (2016). "Plankton classification on imbalanced large scale database *via* convolutional neural networks with transfer learning," in *IEEE International Conference on Image Processing ICIP*. (Phoenix, AZ: IEEE), 3713–3717.

Lin, I. I., Chen, C. H., Pun, I. F., Liu, W. T., and Wu, C. C. (2009). Warm ocean anomaly, air sea fluxes, and the rapid intensification of tropical cyclone nargis. *Geophys Res. Lett.* 36, L03817. doi: 10.1029/2008GL035815

Lin, T. Y., Goyal, P., Girshick, R., He, K., and Dollár, P. (2017). Focal loss for dense object detection. *IEEE Trans. Pattern Anal. Mach. Intell.* 99, 2999–3007. doi: 10.1109/iccv.2017.324

Liu, X. Y., Wu, J., and Zhou, Z. H. (2009). Exploratory undersampling for class-imbalance learning. *IEEE Trans. Syst. Man Cybernetics Part B* 39 (2), 539–550. doi: 10.1109/tsmcb.2008.2007853

Long, M., Cao, Y., Wang, J., and Jordan, M. (2015). "Learning transferable features with deep adaptation networks," in *International conference on machine learning* ((Lille, FRANCE: PMLR), 97–105.

Matsuoka, D., Nakano, M., Sugiyama, D., and Uchida, S. (2018). *Deep learning approach for detecting tropical cyclones and their precursors in the simulation by a cloud-resolving global nonhydrostatic atmospheric model* Vol. 5 (Japan: Progress in Earth and Planetary Science).

Mecikalski, J. R., Sandmæl, T. N., Murillo, E. M., Homeyer, C. R., Bedka, K. M., Apke, J. M., et al. (2021). A random-forest model to assess predictor importance and nowcast severe storms using high-resolution radar–GOES satellite–lightning observations. *Monthly Weather Review* 149 (6), 1725–1746. doi: 10.1175/MWR-D-19-0274.1

Mo, X., Jiao, H., Liu, Y., and Zhang, J. (2019). "A classification algorithm of unbalanced data samples based on ResNet. RCAE 2019," in *RCAE 2019: Proceedings of the 2019 The 2nd International Conference on Robotics, Control and Automation Engineering*. (New York, NY, United States: Association for Computing Machinery).

Na, W., McBride, J. L., Zhang, X.-H., and Duan, Y. H.. (2018). *Understanding biases in tropical cyclone intensity forecast error*. (Boston MA, USA: Weather & Forecasting) 33 (1), 129–138.

Pang, S., Xie, P., Xu, D., Meng, F., Tao, X., Li, B., et al. (2021). NDFTC: A new detection framework of tropical cyclones from meteorological satellite images with deep transfer learning. *Remote Sens.* 13 (9), 1860. doi: 10.3390/rs13091860

Park, M. S., Kim, M., Lee, M. I., Im, J., and Park, S. (2016). Detection of tropical cyclone genesis *via* quantitative satellite ocean surface wind pattern and intensity analyses using decision trees. *Remote Sens. Environ.* 183, 205–214. doi: 10.1016/j.rse.2016.06.006

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., et al. (2019). Deep learning and process understanding for data-driven earth system science. *Nature* 566 (7743), 195. doi: 10.1038/s41586-019-0912-1

Richman, M. B., Leslie, L. M., Ramsay, H. A., and Klotzbach, P. J. (2017). Reducing tropical cyclone prediction errors using machine learning approaches. *Proc. Comput. Sci.* 114, 314–323. doi: 10.1016/j.procs.2017.09.048

Shay, L. K., Goni, G. J., and Black, P. G. (2000). Effects of a warm oceanic feature on hurricane opal. *Mon Weather Rev.* 128, 1366–1383. doi: 10.1175/1520-0493(2000)128<1366:EOAWOF>2.0.CO;2

Shi, X., Chen, Z., Wang, H., Yeung, D. Y., Wong, W. K., and Woo, W. C. (2015). *Convolutional LSTM network: A machine learning approach for precipitation nowcasting* Vol. 28 (Montreal, CANADA: Advances in neural information processing systems).

Singh, R., Ahmed, T., Kumar, A., Singh, A. K., Pandey, A. K., and Singh, S. K. (2020). Imbalanced breast cancer classification using transfer learning. *IEEE/ACM Trans. Comput. Biol. Bioinf.* 18 (1), 83–93. doi: 10.1109/TCBB.2020.2980831

Taherkhani, A., Cosma, G., and McGinnity, T. M. (2020). AdaBoost-CNN: An adaptive boosting algorithm for convolutional neural networks to classify multi-class imbalanced datasets using transfer learning. *Neurocomputing* 404, 351–366. doi: 10.1016/j.neucom.2020.03.064

Troncoso, A., Ribera, P., Asencio-Cortés, G., Vega, I., and Gallego, D. (2018). Imbalanced classification techniques for monsoon forecasting based on a new climatic time series. *Environ. Model. software* 106, 48–56. doi: 10.1016/j.envsoft.2017.11.024

Vissa, N. K., Satyanarayana, A. N. V., and Kumar, B. P. (2013). Intensity of tropical cyclones during pre-and post-monsoon seasons in relation to accumulated tropical cyclone heat potential over bay of Bengal. *Nat. Hazard* 68 (2), 351–371. doi: 10.1007/s11069-013-0625-y

Wang, X., Li, X., Zhu, J., Xu, Z., Ren, K., Zhang, W., et al. (2021). "A local similarity-preserving framework for nonlinear dimensionality reduction with neural networks," in *International conference on database systems for advanced applications* (Cham: Springer), 376–391.

Wijnands, J. S., Qian, G., and Kuleshov, Y. (2016). Variable selection for tropical cyclogenesis predictive modeling. *Monthly Weather Rev.* 144 (12), 4605–4619. doi: 10.1175/mwr-d-16-0166.1

Wikner, A., Pathak, J., Hunt, B. R., Szunyogh, I., Girvan, M., and Ott, E. (2021). Using data assimilation to train a hybrid forecast system that combines machine-learning and knowledge-based components. *Chaos (Woodbury, N.Y.)* 31 (5), 053114. doi: 10.1063/5.0048050

Wu, C. C., Lee, C. Y., and Lin, I. I. (2007). The effect of the ocean eddy on tropical cyclone intensity. *J. Atmos Sci.* 64 (10), 3562–3578. doi: 10.1175/JAS4051.1

Yan, Q., and Cao, Y. (2019). *Optimizing shapelets quality measure for imbalanced time series classification* Vol. 4) (Netherlands: Applied Intelligence).

Yip, Z. K., and Yau, M. K. (2011). Application of artificial neural networks on north atlantic tropical cyclogenesis potential index in climate change. *J. Atmospheric Oceanic Technol.* 29 (9), 1202–1220. doi: 10.1175/jtech-d-11-00178.1

Yosinski, J., Clune, J., Bengio, Y., and Lipson, H. (2014). *How transferable are features in deep neural networks?* (Montreal, CANADA: MIT Press).

Zhan, F. (2020). "Research on bank fraud transaction detection based on LSTM-focalloss," in *ACAI 2020: 2020 3rd International Conference on Algorithms, Computing and Artificial Intelligence*. (New York, NY, United States: Association for Computing Machinery).

Zhang, W., Fu, B., Peng, M. S., and Li, T.. (2015). *Discriminating developing versus nondeveloping tropical disturbances in the western north pacific through decision tree analysis* (Boston MA, USA: Weather & Forecasting) 30 (2), 446–454.

Zhang, T., Lin, W., Lin, Y., Zhang, M., and Xue, W. (2019). *Prediction of tropical cyclone genesis from mesoscale convective systems using machine learning* Vol. 34 (Boston MA, USA: Weather and Forecasting).

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 02 February 2023
DOI 10.3389/fmars.2023.1112336

Check for updates

OPEN ACCESS

EDITED BY
Shiqiu Peng,
State Key Laboratory of Tropical Oceanography (CAS), China

REVIEWED BY
Javier Zavala-Garay,
Department of Marine and Coastal Sciences, Rutgers, The State University of New Jersey, United States
Chuanyu Liu,
Institute of Oceanology (CAS), China

*CORRESPONDENCE
Guijun Han
guijun_han@tju.edu.cn
Zenghua Ji
jzh1999@tju.edu.cn

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

RECEIVED 30 November 2022
ACCEPTED 20 January 2023
PUBLISHED 02 February 2023

CITATION
Wu X, Han G, Li W, Ji Z, Cao L and Dong W (2023) A hybrid deep learning model for predicting the Kuroshio path south of Japan. *Front. Mar. Sci.* 10:1112336. doi: 10.3389/fmars.2023.1112336

COPYRIGHT
© 2023 Wu, Han, Li, Ji, Cao and Dong. 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(s) 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.

A hybrid deep learning model for predicting the Kuroshio path south of Japan

Xiaobo Wu[1], Guijun Han[1*], Wei Li[1,2], Zenghua Ji[1*], Lige Cao[1] and Wanqiu Dong[1]

[1]School of Marine Science and Technology, Tianjin University, Tianjin, China, [2]Tianjin Key Laboratory for Oceanic Meteorology, Tianjin Meteorological Service, Tianjin, China

At present, many prediction models based on deep learning methods have been widely used in ocean prediction with satisfactory results. However, few deep learning models are used to predict the Kuroshio path south of Japan. In this study, a hybrid deep learning prediction model is constructed based on the long short-term memory (LSTM) neural network, combined with the complex empirical orthogonal function (CEOF) and bivariate empirical mode decomposition (BEMD), called CEOF-BEMD-LSTM. We train the model by using a 50-year (1958-2007) long time series of daily mean positions of the Kuroshio path south of Japan extracted from a regional ocean reanalysis dataset. During the test period of 15 years (2008-2022) by using daily altimetry dataset, our model shows a good performance for the Kuroshio path prediction with the lead time of 120 days, with 0.44° root-mean-square error (RMSE) and 0.75 anomaly correlation coefficient (ACC). This model also has good prediction skill score (SS). Moreover, the CEOF-BEMD-LSTM model successfully hindcasts the formation of the latest Kuroshio large meander since the summer of 2017. Predictions of the Kuroshio path for the coming 120 days (from January1 to April 30, 2023) indicate that the Kuroshio will continue to remain in the state of the large meander. Besides, predictor(s) of the Kuroshio path south of Japan need to be sought and added in future research.

KEYWORDS

Kuroshio path prediction, south of Japan, ocean reanalysis, complex empirical orthogonal function, empirical mode decomposition, long short-term memory, altimetry data

1 Introduction

The Kuroshio is the western boundary current of the North Pacific Subtropical Gyre. It originates from the bifurcated North Equatorial Current on the eastern side of the Philippines, flows into the East China Sea *via* the east of Taiwan Island, and then veers eastward through the Tokara Strait into the sea south of Japan (Usui et al., 2006; Usui, 2019; Qiu et al., 2021). As the second warmest current globally, the Kuroshio brings a large amount of heat to the southern coast of Japan (Tsujino et al., 2006). Due to topographic constraints, the Kuroshio path south of Japan exhibits three typical paths (Kawabe, 1995): the typical large

meander (tLM) path, offshore non-large meander (oNLM) path, and nearshore non-large meander (nNLM) path (Figure 1). The variation of the Kuroshio path has large effects on climate, fisheries, and ship navigation (Nakata et al., 2000; Tsujino et al., 2006; Sugimoto et al., 2020; Sugimoto et al., 2021). Therefore, it is of great significance to conduct prediction of the Kuroshio path south of Japan.

Numerical prediction plays a dominant role in predicting the variation of the Kuroshio path south of Japan. Several studies used experiments to predict the variability of the Kuroshio path using data assimilation models, and the results showed that the predictive limit for the Kuroshio path south of Japan is about a couple of months (Komori et al., 2003; Kamachi et al., 2004; Miyazawa et al., 2005; Usui et al., 2006). The Japan Coastal Ocean Predictability Experiment (JCOPE), which is aimed to describe the Kuroshio path, Kuroshio Extensions, and Oyashio variability (Miyazawa et al., 2008; Miyazawa et al., 2009); It provides the Kuroshio prediction two months ahead of time.

As one of the most popular and consequential technologies, deep learning methods have been widely used for ocean prediction (Reichstein et al., 2019). Among them, the recurrent neural network (RNN) and its variants are known to work well in processing time series data to find the time-varying principles hidden in the time series data (Song et al., 2020). The long short-term memory (LSTM) neural network, as one of the essential variants of the RNN, can detect even minor changes from the time series and avoids the problem of vanishing gradient and exploding gradient (Hochreiter and Schmidhuber, 1997). In recent years, the LSTM neural network had a good performance in the time series prediction of ocean variables (Liu et al., 2018; Xiao et al., 2019; Shao et al., 2021b). However, few deep learning models are used to predict the Kuroshio path south of Japan.

In this study, we present a hybrid deep learning prediction model, combining the complex empirical orthogonal function (CEOF) analysis, bivariate empirical mode decomposition (BEMD) analysis, and LSTM neural network, named CEOF-BEMD-LSTM model, to predict the Kuroshio path south of Japan. The rest of this paper is organized as follows. In section 2, we introduce the data and methods. In section 3, we describe the prediction experiments and results of the Kuroshio path. Summary and discussion are given in section 4.

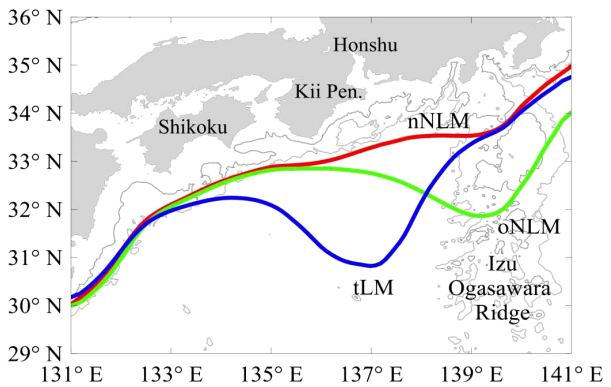


FIGURE 1
Three typical Kuroshio paths south of Japan: nNLM (red line), oNLM (green line), and tLM (blue line) derived from a regional ocean reanalysis described in section 2.1. Thin black contours are isobaths of 1 000 and 2 000 m.

2 Data and methods

2.1 Data

We use the sea-surface height (SSH) data from an ocean reanalysis dataset, which is produced by a Northwest Pacific regional ocean reanalysis system, called China Ocean ReAnalysis (CORA, http://www.cmoc-china.cn; see Han et al., 2013). The CORA system uses a sequential three-dimensional variational (3D-Var) scheme implemented within a multigrid framework (Li et al., 2008), and assimilates satellite remote sensing sea-surface temperature (SST), altimetry SSH anomaly (SSHA), and *in-situ* temperature/salinity profiles into the parallelized Princeton Ocean Model with generalized coordinate system (POMgcs; Mellor et al., 2002; Ezer and Mellor, 2004). The daily reanalysis dataset from January 1958 to December 2007 is used for this study.

We also use the daily absolute dynamic topography (ADT) data of the Ssalto/Duacs altimeter products from January 2008 to December 2022 from the Copernicus Marine and Environment Monitoring Service (CMEMS) (https://marine.copernicus.eu) to conduct prediction experiments. The study domain is from 131°E to 141°E and from 29°N to 36°N (see Figure 1). The spatial resolution of both datasets in the study region is 0.25° × 0.25°.

In this study, the Kuroshio path south of Japan is defined by the 70-cm SSH isoline and 110-cm ADT isoline, respectively. The discrepancy between the definitions with these two datasets results from different reference mean sea surfaces used (Yang and Liang, 2019), but both definitions can capture the Kuroshio axis position well (Wu et al., 2022). The time series of the Kuroshio path (hereafter the Kuroshio path data), which are the latitude value in degrees (°) corresponding to each longitude of the study area, are selected as the truth to conduct subsequent prediction experiments. One part of the Kuroshio path data ranging from 1958 to 2007 serves as the training dataset to train the prediction model, and the other part from 2008 to 2022 is used as the testing dataset to test the prediction model.

2.2 Methods

2.2.1 EOF and CEOF analyses

The empirical orthogonal function (EOF) analysis is widely used in dimensionality reduction and pattern extraction in atmospheric and oceanic sciences (Hannachi et al., 2007). However, the EOF analysis cannot deal with propagating features. Therefore, the CEOF analysis is introduced to solve such problem.

In this study, the Kuroshio path data can be expressed as matrix **X**:

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1n} & \cdots & x_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{m1} & \cdots & x_{mn} & \cdots & x_{mN} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{M1} & \cdots & x_{Mn} & \cdots & x_{Mn} \end{pmatrix} \quad (1)$$

where the dimensions are $M \times N$, with M representing the spatial dimension and N representing the temporal dimension.

The matrix X is first normalized, expressed as X':

$$X' = \sigma^{-1}(X - \bar{X}) \tag{2}$$

where σ is the standard deviation matrix and $\bar{X}$is the climatology.

In the EOF analysis, the spatial modes (EOFs) and associated temporal coefficients (PCs) are obtained by performing a Jacobi decomposition on the covariance matrix of X'.

In the CEOF analysis, a Hermite matrix (U) is constructed by applying the Hilbert transform to the matrix X'. It can be further expanded as:

$$U = BP \tag{3}$$

where P is composed of the complex EOFs (aka, spatial modes, hereafter CEOFs), while B is composed of the corresponding complex PCs (aka, temporal coefficients, hereafter CPCs). In this study, the temporal coefficients (PCs and CPCs) will be taken as the raw data for the input of the deep learning prediction model. Detailed information about the use of the CEOF analysis in this study is given in sections 3.1 and 3.2.

2.2.2 BEMD analysis

The empirical mode decomposition (EMD) analysis is an efficient method for data denoising (Huang et al., 1998). Rilling et al. (2007) purposed the BEMD to handle bivariate (complex) time series. This method considers the complex signal as a superposition of fast and slow oscillation components. First, the poles of the projection vectors of the complex signal in different directions and their envelopes are obtained; then, the mean of the envelope is defined as the slow oscillation signal, and the fast oscillation signal is obtained by separating it from the original signal, which is called the complex intrinsic mode function (CIMF). The BEMD analysis achieves the direct decomposition of a complex signal, and avoids the inconsistency between real and imaginary decompositions (Ma et al., 2015).

$$y(t) = \sum_{i=1}^{m} \mathrm{CIMF}_i(t) \tag{4}$$

After the BEMD analysis, the original signal $y(t)$ can be decomposed into m CIMFs, and expressed as:

In this study, the BEMD analysis is applied to the CPCs. Details about the use of the BEMD analysis in this study are provided in section 3.2.

2.2.3 LSTM neural network

The LSTM neural network can tackle the long-term dependence of sequence data well, and is regarded as a state-of-the-art method for time series prediction. As a variant of the RNN, it solves the problem of gradient vanishing and gradient explosion that exist in the traditional RNN (Hochreiter and Schmidhuber, 1997). Figure 2 shows the structure of an LSTM cell. The LSTM cell is made up of forget gate, input gate, and output gate. Specifically, the forget gate mainly selectively forgets the previous cell state quantity; the input gate mainly selectively memorizes the new input information; and the output gate selectively outputs the updated cell state quantity. The main calculation is defined by a series of equations as follows:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{5}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{6}$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_C) \tag{7}$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{8}$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{9}$$

$$h_t = o_t * \tanh(C_t) \tag{10}$$

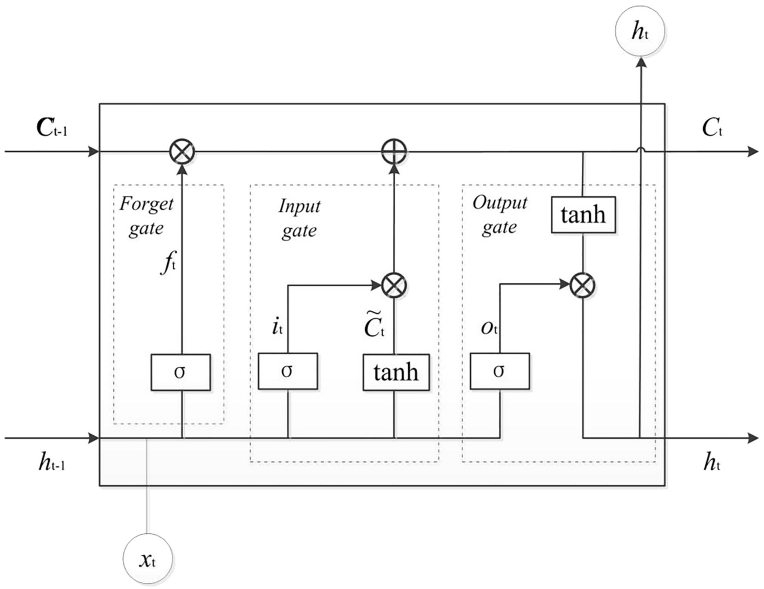


FIGURE 2
Structure of an LSTM cell.

where f_t, i_t, and o_t represent the outputs of forget gate, input gate, and output gate, respectively. C_t is the cell state vector, and σ is the sigmoid function. W_f, W_i, Wo, and W_C are the corresponding weights; b_f, b_i, bo, and b_c are the corresponding biases. h_t is the output, $\tilde{C}_t$ is the new memory vector, and x_t is the input.

In this study, we build a 4-layer deep neural network model to conduct 120-day Kuroshio path prediction experiments based on the LSTM neural network. By trial and error, the size of the time window used to predict the Kuroshio path is set to 30, which means that we use the preceding 30-day Kuroshio path data for prediction. Besides, the adaptive moment estimation (Adam) is taken as the gradient optimization algorithm, which provides an optimized method for solving sparse gradients and noise problems (Song et al., 2020). The rectified linear unit (ReLU) function is used as the activation function. This function avoids the gradient vanishing problem of the sigmoid function and tanh function, and it has a high calculation efficiency.

2.2.4 Evaluation criteria

To evaluate the performance of the prediction models, we employ root-mean-square error (RMSE), anomaly correlation coefficient (ACC), and prediction skill score (SS) as the evaluation criteria. These calculation formulas are defined as follows:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{j=1}^{m} x_{i,j}^{P} - x_{i,j}^{T}}{m}} \tag{11}$$

$$ACC = \frac{\sum_{i=1}^{n}(x_{i,j}^{P} - \bar{x}_j^{P})(x_{i,j}^{T} - \bar{x}_j^{T})}{\sqrt{\sum_{i=1}^{n}(x_{i,j}^{P} - \bar{x}_j^{P})^2 (x_{i,j}^{T} - \bar{x}_j^{T})^2}} \tag{12}$$

$$\mathrm{SS} = 1 - \frac{\mathrm{MSE}(p,o)}{\mathrm{MSE}(c,o)} \tag{13}$$

where $x_{i,j}^{P}$ and $x_{i,j}^{T}$ are prediction and true values, respectively, of the Kuroshio path of the i^{th} grid point on the j^{th} day; . $\bar{x}_j^{P} = \sum_{i=1}^{n} x_{i,j}^{P}$. and $\bar{x}_j^{T} = \sum_{i=1}^{n} x_{i,j}^{T}$ are prediction and true mean values of the Kuroshio path on the j^{th} day, respectively;m is the number of days of testing data; n is the number of spatial grid points representing the Kuroshio path; RMSE is the root-mean-square error of the i^{th} grid point; and ACC is the spatial anomaly correlation coefficient of the j^{th} day. $\mathrm{MSE}(p,o) = \frac{1}{m}\sum_{i=1}^{m}(x_{i,j}^{P} - x_{i,j}^{T})^2$ denotes the mean square error between prediction and observations; $\mathrm{MSE}(c,o) = \frac{1}{m}\sum_{i=1}^{m}(x_i^{C} - x_{i,j}^{T})^2$denotes the mean square error between climatology and observations, in which x_i^{C} is climatological value of the Kuroshio path of the i^{th} grid point.

3 Prediction experments

3.1 Comparison of EOF-LSTM and CEOF-LSTM models

First, we construct the CEOF-LSTM (EOF-LSTM) prediction model, based on the CEOF (EOF) analysis and LSTM neural network only (see Figure 6). To compare the performance from the CEOF analysis with the EOF analysis for the Kuroshio path prediction, we conduct 120-day Kuroshio path prediction experiments. After the CEOF (EOF) analysis, the Kuroshio path data are decomposed into several modes and separated into CEOFs (EOFs) and CPCs (PCs). To reduce the cost time, we use the first 16 CEOFs (18 EOFs) and their corresponding CPCs (PCs), accounting for 99% of the total variance, as input parameters to predict the CPC (PC) time series for different lead times. All the modes used pass the North significance test (North et al., 1982); and these CEOFs (EOFs) and CPCs (PCs) are able to reconstruct the main characteristics of the Kuroshio path. After training the model, the CPC (PC) time series of the significant Kuroshio path are predicted, and the Kuroshio paths are reconstructed by using these predicted CPCs (PCs) and the CEOFs (EOFs) from the CEOF (EOF) analysis. The residual from the unused higher-order modes of the CEOF (EOF) analysis (accounting for 1% of the total variance) at the start time serves as a correction to obtain the final prediction in a form of persistence. Such a correction can improve the prediction skill in the first three days of the lead time. Compared with the prediction experiment without the correction, the RMSEs of the CEOF-LSTM (EOF-LSTM) model prediction results are reduced by 12.2% (6.2%), 6.1% (2.7%), and 1.8% (0.7%), respectively; and the ACC values are increased by 0.0038 (0.0032), 0.0022 (0.0014), and 0.0007 (0.0007), respectively.

Figure 3 shows the averaged RMSE and ACC of the predictions for different lead times using CEOF-LSTM (red line) and EOF-LSTM (green line) models. The dashed black line indicates the climatological standard deviation of the 50-year (1958-2007) Kuroshio path. It can be seen that the RMSE of the CEOF-LSTM model is significantly smaller than that of the EOF-LSTM model for each lead time. The ACC of the CEOF-LSTM model is below 0.6 (the black dashed line in Figure 3B, which has the spatial ACC of 0.6, a thumb rule for measuring "usefulness" of predictions; Pendlebury et al., 2003) at the lead time of 100 days, while the ACC of the EOF-LSTM is below 0.6 as early as 35 days. When the lead time is 120 days, the RMSE of the CEOF-LSTM model is reduced by 12%, and the ACC is improved by 0.23 compared to the EOF-LSTM model.

Figures 4A, C depict temporal-spatial distributions of RMSE for the 1-120 days Kuroshio path predictions using the EOF-LSTM and CEOF-LSTM models. The solid black contour indicates the prediction range where the RMSE at each location of the Kuroshio path reaches its climatological standard deviation. Both RMSEs exhibit similar spatial distributions, but the RMSE of the CEOF-LSTM model is smaller. To compare the RMSEs of these two models better, we also calculate the temporal-spatial distributions of difference between the RMSE of the CEOF-LSTM model and that of the EOF-LSTM model, as shown in Figure 5A. The RMSE of the CEOF-LSTM model is significantly smaller than that of the EOF-LSTM model in all regions, especially in the region of 137°-140°E. It can be reduced by as much as 0.3°.

In summary, the CEOF analysis is significantly better than the EOF analysis for predicting the Kuroshio path south of Japan. It may be due to these following reasons: The CEOF analysis can resolve propagating wave signals (Bouzinac et al., 1998), which are closely related to the variation of the Kuroshio path, while the EOF analysis cannot reveal such signal characteristics. The LSTM neural network can capture and learn these signal features during the training process, and thus the prediction of the CEOF-LSTM model is better. Considering the comparison results and explanations above,

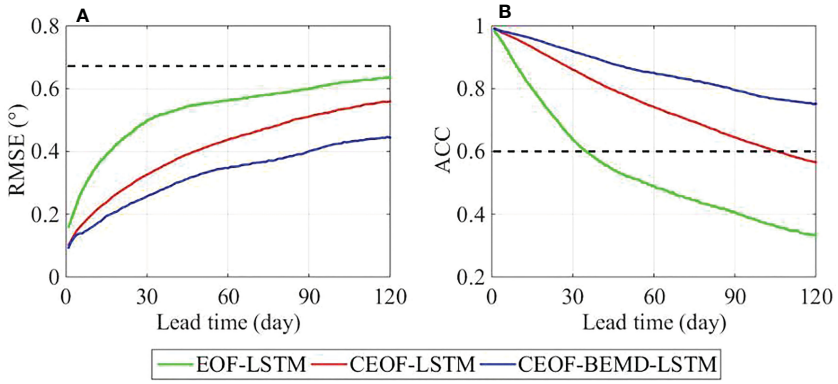


FIGURE 3
(A) Space-averaged RMSE (°) of the prediction using the EOF-LSTM, CEOF-LSTM, and CEOF-BEMD-LSTM models. The dashed black line indicates the climatological standard deviation of the Kuroshio path. **(B)** Averaged ACC values of the predictions using the EOF-LSTM, CEOF-LSTM, and CEOF-BEMD-LSTM models. The dashed black line indicates spatial ACC of 0.6, a thumb rule for measuring "usefulness" of predictions.

we conduct further predictions based on the CEOF-LSTM model in the following-up experiments.

3.2 Prediction experiments using CEOF-BEMD-LSTM model

To improve the performance of the CEOF-LSTM model, we add the BEMD analysis to the prediction model (Rilling et al., 2007; Shao et al., 2021a), called the CEOF-BEMD-LSTM model. Figure 6 shows the framework of the CEOF-BEMD-LSTM model, which can be broken down into three parts: (A) data preprocessing, (B) LSTM prediction, and (C) correction. During the first part of data preprocessing, the Kuroshio path data are first divided into training dataset and testing dataset. Then, the CEOF analysis decomposes the training dataset into CEOFs and CPCs. The CPCs of the testing dataset are obtained by projecting the testing dataset onto the CEOFs. Next, the BEMD analysis is conducted on the first 16 CPCs (accounting for 99% of the total variance) to extract the CIMFs. Each CPC is decomposed into 16 CIMFs; and all the CIMFs serve as the inputs for the LSTM neural network. In the second part of LSTM prediction, the LSTM neural network is used to predict the CIMFs. After training and predicting, the predictions of the CPCs are obtained by using the predictions of the CIMFs, which are the outputs of the LSTM neural network. Based on the predictions of the CPCs and the CEOFs obtained from the CEOF analysis, the predictions of the Kuroshio path are reconstructed. In the last part of the correction, the final prediction is obtained by adding the residual, which consists of unused higher-order modes of the CEOF analysis (accounting for 1% of the total variance) at the start time, as a correction in a form of persistence. Similar to the results presented in section 3.1, the prediction skill in the first three days of the lead time is better than that of the prediction experiment without the correction. Specifically, the RMSEs of the prediction results are reduced by 14.2%, 6.4%, and 1.9%, respectively; and the ACC values are increased by 0.0038, 0.0023, and 0.0006, respectively.

In this section, we compare the predictions of the CEOF-BEMD-LSTM model with those of the CEOF-LSTM model to evaluate the performance of the CEOF-BEMD-LSTM model. Figure 3 shows the averaged RMSE and ACC values of the predictions for different lead times using the CEOF-LSTM (red line) and CEOF-BEMD-LSTM (blue line) models. The dashed black line indicates the climatological standard deviation of the 50-year (1958-2007) Kuroshio path. The RMSE of the CEOF-LSTM model is significantly larger than that of the CEOF-BEMD-LSTM model when the lead time is 120 days, and exceeding the climatological standard deviation of the Kuroshio path. Compared with the CEOF-LSTM model (red line), the performance of the CEOF-BEMD-LSTM model (blue line) exhibits better prediction results, with smaller RMSE and larger ACC (blue line). Even to 120 days, the RMSE of the CEOF-BEMD-LSTM model is much smaller than the climatological standard deviation of the Kuroshio path, and the ACC of the CEOF-BEMD-LSTM model exceeds 0.7. Compared with the CEOF-LSTM model, the RMSE of the CEOF-BEMD-LSTM model in the prediction results of day 120 is reduced by 26%, and the ACC is improved by 0.19.

Figures 4C, E show temporal-spatial distributions of RMSE for the 1-120 days Kuroshio path prediction results using the CEOF-LSTM and CEOF-BEMD-LSTM models. The solid black contour indicates the prediction range, where the RMSE of each location of the Kuroshio path achieves its climatological standard deviation. The RMSEs of both models expand progressively with increasing lead time; and the RMSEs of both models in the Kuroshio large meander region gradually converge downstream and attain their maxima in the Izu-Ogasawara Ridge (IOR) region with the same lead time. This is probably because the Kuroshio path in this region changes frequently, leading to lower signal-to-noise ratios and larger errors for the predictions. More importantly, the prediction range of the CEOF-BEMD-LSTM model exceeds 120 days (Figure 4E), while the prediction ranges of the CEOF-LSTM model are under 120 days in the upper Kuroshio (131°-135°E) and IOR (Figures 4C) and their RMSEs are larger. We also depict the temporal-spatial distributions of the difference between the RMSE of the CEOF-BEMD-LSTM model and that of the CEOF-LSTM model in the 1-120 days predictions. Figure 5B clearly shows that the RMSE of the CEOF-BEMD-LSTM model is smaller than that of the CEOF-LSTM model in all locations.

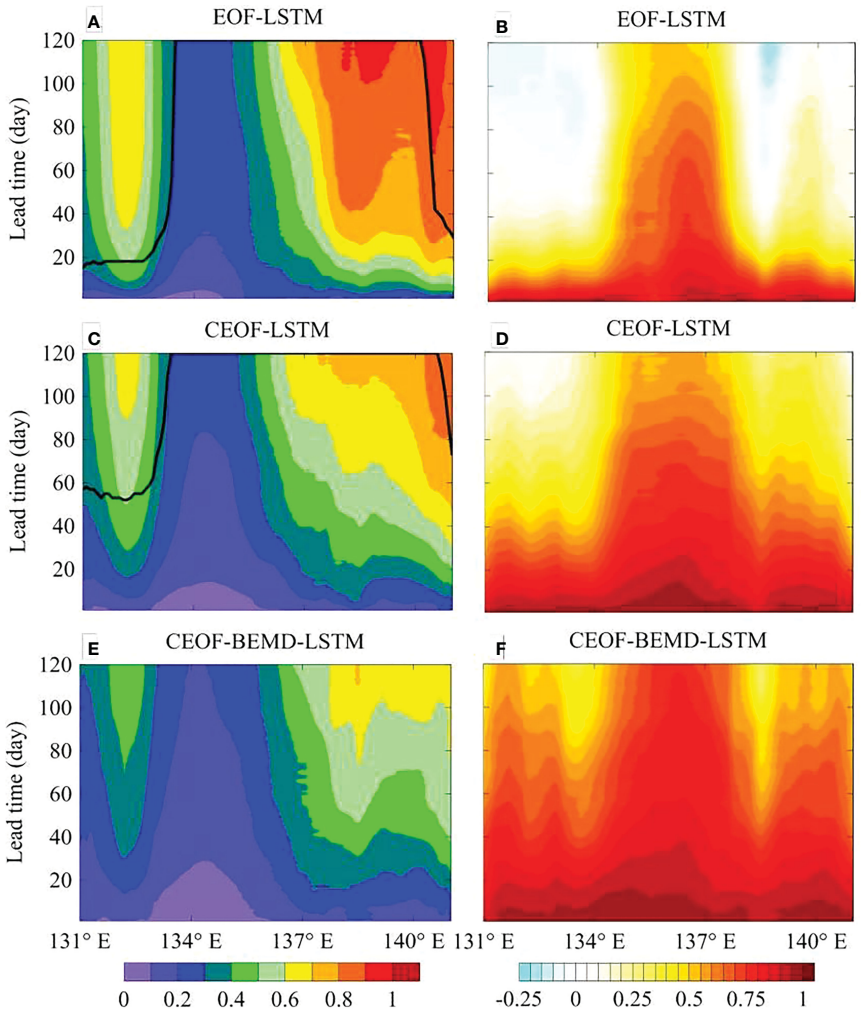


FIGURE 4
Left panels **(A, C, E)**: Temporal-spatial distributions of RMSE (°) of the 1-120 days Kuroshio path predictions using EOF-LSTM, CEOF-LSTM, and CEOF-BEMD-LSTM models. The solid black contour indicates the prediction range where the RMSE at each location of the Kuroshio path reaches its climatological standard deviation. Right panels **(B, D, F)**: Same as the left panels, except for temporal-spatial distributions of prediction skill score.

We also calculate the prediction skill score (SS) with each model to further evaluate the predictions. The SS is positive (negative) when the accuracy of the prediction is greater (less) than the accuracy of the climatology (Murphy, 1988). Meanwhile, the closer the SS approaches toward 1, the better the prediction. The temporal-spatial distributions of SS for the 1-120 days Kuroshio path predictions using the EOF-LSTM, CEOF-LSTM, and CEOF-BEMD-LSTM models are described in the right panels of Figure 4. The SSs of all models show similar

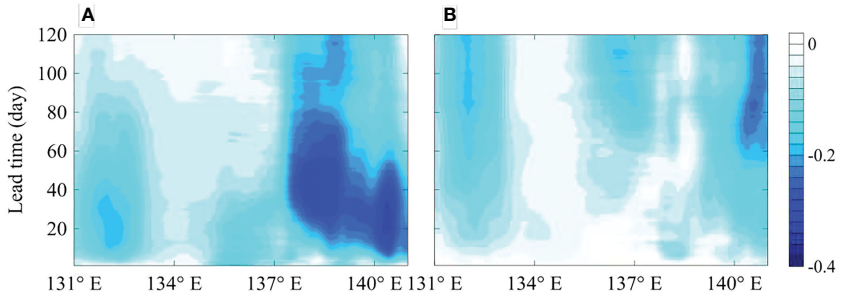


FIGURE 5
(A) Temporal-spatial distributions of difference between the RMSE of the CEOF-LSTM model and that of the EOF-LSTM model in the 1-120 days predictions. Negatives value means that the RMSE of the CEOF-LSTM model is smaller than that of the EOF-LSTM model. **(B)** Same as **(A)**, except for the difference between the RMSE of CEOF-BEMD-LSTM model and that of the CEOF-LSTM model. Negatives value means that the RMSE of the CEOF-BEMD-LSTM model is smaller than that of the CEOF-LSTM model.

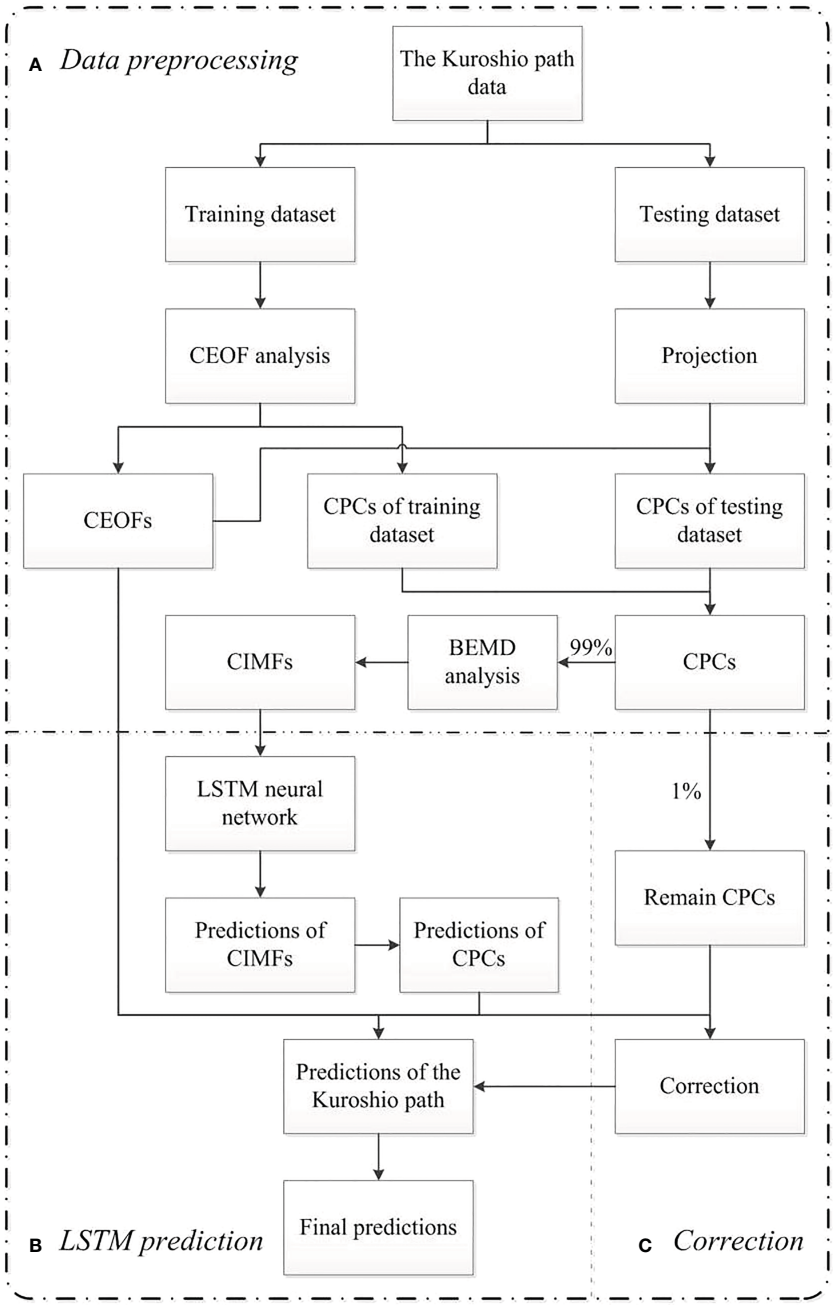


FIGURE 6
Framework of CEOF-BEMD-LSTM model.

distributions. Specially, the SSs decrease gradually as the lead time increases. They are larger in the 135°-138°E region and gradually decrease to the east and west (Figures 4D, F). Moreover, the CEOF-BEMD-LSTM model demonstrates the best SS. When the lead time is 120 days, the SS of this model is still maintained above 0.3, being larger than other models'. In the meanwhile, the SS remains relatively high in the IOR despite the large RMSE in the region. To summarize, the CEOF-BEMD-LSTM model exhibits the best prediction skill in the 120-day Kuroshio path prediction experiments.

3.3 The latest Kuroshio large meander prediction

The latest Kuroshio large meander occurred in August 2017, and is the second Kuroshio large meander in this century. As a unique phenomenon, the Kuroshio large meander has a significant impact on climate change along the southern coast of Japan (Sugimoto et al., 2020; Sugimoto et al., 2021). Therefore, Kuroshio large meander prediction is one of the important goals to conduct the experiments

for predicting the Kuroshio path south of Japan. Considering the best performance of the CEOF-BEMD-LSTM model presented in section 3.2, we use the CEOF-BEMD-LSTM model to predict the latest Kuroshio large meander next.

Figure 7 depicts the predictions of the Kuroshio path with the lead time of 120 days from July 1 (1-day) to October 28 (120-day), 2017. In general, the Kuroshio large meander transforms from the nearshore non-large meander. However, this one switched from the offshore non-large meandering path (Figure 7A), accompanied by a smaller meander in the IOR. Then, the smaller meander continued to be advected downstream with decreasing amplitude and eventually disappeared (Figures 7B, C). Meanwhile, a trigger meander from upstream was advected to the southern sea of Honshu with increasing amplitude, eventually forming a stable Kuroshio large meander path (Figures 7D–I). Overall, the predicted Kuroshio path captures this process, namely, which implies that the Kuroshio path prediction with the CEOF-BEMD-LSTM model can predict the latest Kuroshio large meander formation process. Noted that the prediction errors exist within the region of Kuroshio large meanders and the prediction magnitude is smaller than the actual Kuroshio path (Figures 7C–I).

The latest Kuroshio large meander has lasted for five years and remains so. In the final part of this section, we use the CEOF-BEMD-LSTM model to predict the Kuroshio path south of Japan for 120 days from January 1 (1-day) to April 30 (120-day), 2023. The prediction results indicate that the Kuroshio will remain in the state of the large meander (Figure 8). The position of the large meander will gradually shift westward from 137.4°E on January 1, 2023 (dashed red line) to 136.5°E on April 30, 2023 (dashed brown line).

4 Summary and discussion

In this study, a hybrid deep learning prediction model, called CEOF-BEMD-LSTM model, is developed for predicting the Kuroshio path south of Japan based on the CEOF analysis, BEMD analysis, and LSTM neural network. To evaluate the performance of this model, we use the Kuroshio path data obtained from the CORA reanalysis dataset from 1958 to 2007 (50 years) as a training dataset, and its counterpart from the altimetry data from 2008 to 2022 as a testing dataset, to conduct 120-day Kuroshio path prediction experiments. Prediction results show that the CEOF-BEMD-LSTM model has good performance in the 120-day prediction range evaluated by using two common deterministic skill metrics, the ACC and RMSE. Even when the lead time is 120 days, the RMSE is about 0.44°, which is less than the climatological standard deviation, and the ACC can still reach 0.75, which is greater than 0.6 (a widely used measure for forecast verification; Pendlebury et al., 2003). This model also exhibits a good prediction skill score (SS). Besides, the model successfully hindcasts the formation of the latest Kuroshio large meander since the summer of 2017. Finally, we predict the Kuroshio path from January 1 to April 30, 2023, and the predictions indicate that the Kuroshio will continue to be a large meander.

Comparatively speaking, the prediction range of the traditional numerical prediction is usually 60 days (Komori et al., 2003; Miyazawa et al., 2005; Usui et al., 2006). At present, the JCOPE operational system provides the prediction of the Kuroshio path south of Japan with a two-month lead time (https://fra-roms.fra.go.jp/fra-roms/). However, since its predictions are given in figures with no statistics of the prediction results available, we cannot compare our results with theirs

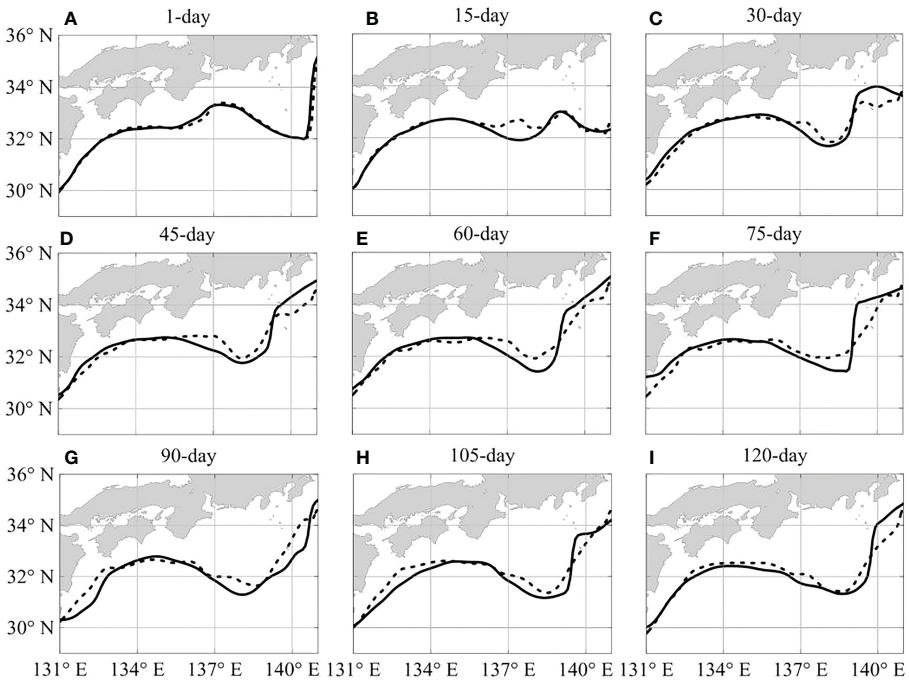


FIGURE 7
Prediction results of the Kuroshio path with the lead time of 120 days from July 1 (1-day) to October 28 (120-day), 2017 **(A–I)**. The solid curve represents the true path, and the dashed one represents the prediction.

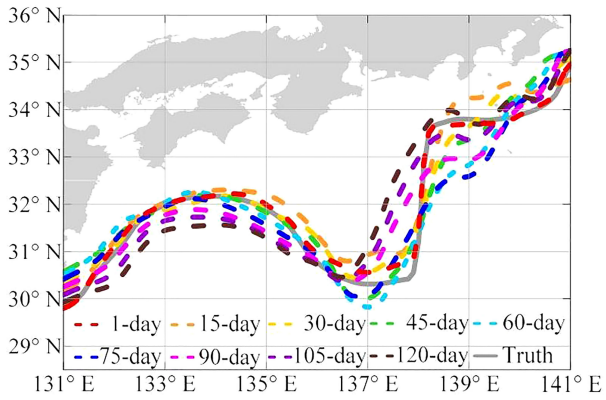


FIGURE 8
Prediction results of the Kuroshio path with the lead time of 120 days from January 1 (1-day) to April 30 (120-day), 2023. The solid gray curve represents the truth at lead-time of one day.

quantitatively. Currently, its two-month predictions show that the Kuroshio will continue to be a large meander, as our model does.

Some recent studies showed that the inclusion of appropriate predictors can improve the prediction range of deep learning models (e.g., Oh and Suh, 2018; Liang et al., 2021). In future research, great effort is required to seek predictor(s) of the Kuroshio path south of Japan, which should be encoded in a reasonable way.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: http://www.cmoc-china.cn/pages/productService.html https://resources.marine.copernicus.eu/product-detail/SEALEVEL_GLO_PHY_L4_NRT_OBSERVATIONS_008_046/DATA-ACCESS https://resources.marine.copernicus.eu/product-detail/SEALEVEL_GLO_PHY_L4_MY_008_047/DATA-ACCESS.

Author contributions

XW and ZJ constructed the prediction model and conducted the prediction experiments. XW wrote the initial draft and revised the manuscript. GH and WL proposed the main ideas and revised the manuscript. LC and WD provided high-performance data processing and revised the manuscript. All authors contributed to the article and approved the revised version.

Funding

This research is sponsored by the National Natural Science Foundation of China (grant 41876014).

Acknowledgments

The authors thank the following data and tool providers: the National Marine Data and Information Service (http://www.cmoc-china.cn) for providing the daily CORA dataset, the Copernicus Marine and Environment Monitoring Service for providing the altimetry dataset (https://resources.marine.copernicus.eu), Google for providing the open source machine learning framework TensorFlow (https://www.tensorflow.org/), Keras (https://keras.io/) for conducting deep learning prediction experiments, and Scikit-learn (http://scikit-learn.org/stable/) for providing the machine learning library based on which the experiments are conducted. The authors sincerely appreciate the reviewers for their valuable comments and suggestions on improving the quality of this paper.

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Bouzinac, C., Vazquez, J., and Font, J. (1998). Complex empirical orthogonal functions analysis of ERS-1 and TOPEX/POSEIDON combined altimetric data in the region of the Algerian current. *J. Geophys. Res.-Oceans* 103 (C4), 8059–8071. doi: 10.1029/97JC02909

Ezer, T., and Mellor, G. L. (2004). A generalized coordinate ocean model and a comparison of the bottom boundary layer dynamics in terrain-following and in z-level grids. *Ocean Model.* 6 (3-4), 379–403. doi: 10.1016/s1463-5003(03)00026-x

Han, G., Li, W., Zhang, X., Wang, X., Wu, X., Fu, H., et al. (2013). A new version of regional ocean reanalysis for coastal waters of China and adjacent seas. *Adv. Atmos. Sci.* 30 (004), 974–982. doi: 10.1007/s00376-012-2195-4

Hannachi, A., Jolliffe, I. T., and Stephenson, D. B. (2007). Empirical orthogonal functions and related techniques in atmospheric science: a review. *Ini. J. Climatol.* 27, 1119–1152. doi: 10.1002/joc.1499

Hochreiter, S., and Schmidhuber, J. (1997). Long short-term memory. *Neural Computation.* 9 (8), 1735–1780. doi: 10.1162/neco.1997.9.8.1735

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., et al. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. London Ser. A-Mathemat. Phys. Sci.* 454, 903–995. doi: 10.1098/rspa.1998.0193

Kamachi, M., Kuragano, T., Sugimoto, S., Yoshita, K., Sakurai, T., Nakano, T., et al. (2004). Short-range prediction experiments with operational data assimilation system for the Kuroshio south of Japan. *J. Oceanogr.* 60, 269–282. doi: 10.1023/B:JOCE.0000038333.97882.51

Kawabe, M. (1995). Variations of current path, velocity, and volume transport of the Kuroshio in relation with the large meander. *J. Phys. Oceanogr.* 25, 3103–3117. doi: 10.1175/1520-0485(1995)025,3103:VOCPVA.2.0.CO;2

Komori, N., Awaji, T., Ishikawa, Y., and Kuragano, T. (2003).). short-range forecast experiments of the Kuroshio path variabilities south of Japan using TOPEX/Poseidon altimetric data. *J. Geophys. Res.-Oceans* 108 (C1), 3010. doi: 10.1029/2001JC001282

Liang, X. S., Xu, F., Rong, Y., Zhang, R., Tang, X., and Zhang, F. (2021). El Niño modoki can be mostly predicted more than 10 years ahead of time. *Sci. Rep.* 11 (1), 17860. doi: 10.1038/s41598-021-97111-y

Liu, J., Zhang, T., Han, G., and Gou, Y. (2018). TD-LSTM: Temporal dependence-based LSTM networks for marine temperature prediction. *Sensors* 18 (11), 3797. doi: 10.3390/s18113797

Li, W., Xie, Y., He, Z., Han, G., Liu, K., Ma, J., et al. (2008). Application of the multigrid data assimilation scheme to the China seas' temperature forecast. *J. Atmos. Oceanic Technol.* 25 (11), 2106–2116. doi: 10.1175/2008JTECHO510.1

Ma, Y., Li, G., Wang, Y., Zhou, H., and Zhang, B. (2015). Random noise attenuation by f-x spatial projection-based complex empirical mode decomposition predictive filtering. *Appl. Geophys.* 12, 47–54. doi: 10.1007/s11770-015-0467-3

Mellor, G. L., Häkkinen, S. M., Ezer, T., and Patchen, R. C. (2002). "A generalization of a sigma coordinate ocean model and an intercomparison of model vertical grids," in *Ocean forecasting*. Eds. N. Pinardi and J. Woods (Berlin, Heidelberg: Springer), 55–72. doi: 10.1007/978-3-662-22648-3_4

Miyazawa, Y., Kagimoto, T., Guo, X., and Sakuma, H. (2008). The Kuroshio large meander formation in 2004 analyzed by an eddy-resolving ocean forecast system. *J. Geophys. Res.-Oceans.* 113, C10015. doi: 10.1029/2007JC004226

Miyazawa, Y., Yamane, S., Guo, X., and Yamagata, T. (2005). Ensemble forecast of the Kuroshio meandering. *J. Geophys. Res.-Oceans.* 110, C10026. doi: 10.1029/2004JC002426

Miyazawa, Y., Zhang, R., Guo, X., Tamura, H., Ambe, D., and Lee, J. S. (2009). Water mass variability in the western north pacific detected in a 15-year eddy resolving ocean reanalysis. *J. Oceanogr.* 65 (6), 737–756. doi: 10.1007/s10872-009-0063-3

Murphy, A. H. (1988). Skill scores based on the mean square error and their relationships to the correlation coefficient. *Mon. Wea. Rev.* 116 (12), 2417–2424. doi: 10.1175/1520-0493(1988)116<2417:SSBOTM>2.0.CO;2

Nakata, H., Funakoshi, S., and Nakamura, M. (2000). Alternating dominance of postlarval sardine and anchovy caught by coastal fishery in relation to the Kuroshio meander in the enshu-nada Sea. *Fish. Oceanogr.* 9 (3), 248–258. doi: 10.1046/j.1365-2419.2000.00140.x

North, G. R., Bell, T. L., Cahalan, R. F., and Moeng, F. J. (1982). Sampling errors in the estimation of empirical orthogonal functions. *Mon. Wea. Rev.* 110 (7), 699–706. doi: 10.1175/1520-0493(1982)110<0699:SEITEO>2.0.CO;2

Oh, J., and Suh, K. D. (2018). Real-time forecasting of wave heights using EOF-wavelet-neural network hybrid model. *Ocean engineering.* 150, 48–59. doi: 10.1016/j.oceaneng.2017.12.044

Pendlebury, S. F., Adams, N. D., Hart, T. L., and Turner, J. (2003). Numerical weather prediction model performance over high southern latitudes. *Mon. Wea. Rev.* 131 (2), 335–353. doi: 10.1175/1520-0493(2003)131<0335:NWPMPO>2.0.CO;2

Qiu, B., Chen, S., Schneider, N., Oka, E., and Sugimoto, S. (2021). On reset of the wind-forced decadal Kuroshio extension variability in late 2017. *J. Climate.* 33 (24), 10813–10827. doi: 10.1175/JCLI-D-20-0237.1

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., et al. (2019). Deep learning and process understanding for data-driven earth system science. *Nature* 566, 195–204. doi: 10.1038/s41586-019-0912-1

Rilling, G., Flandrin, P., Goncalves, P., and Lilly, J. M. (2007). Bivariate empirical mode decomposition. *IEEE Signal Process. Lett.* 14 (12), 936–939. doi: 10.1109/LSP.2007.904710

Shao, Q., Hou, G., Li, W., Han, G., Liang, K., and Bai, Y. (2021a). Ocean reanalysis data-driven deep learning forecast for sea surface multivariate in the South China Sea. *Earth Space Sci.* 8, e2020EA001558. doi: 10.1029/2020EA001558

Shao, Q., Li, W., Han, G., Hou, G., Liu, S., Gong, Y., et al. (2021b). A deep learning model for forecasting sea surface height anomalies and temperatures in the South China Sea. *J. Geophys. Res.-Oceans.* 126, e2021JC17515. doi: 10.1029/2021JC017515

Song, T., Jiang, J., Li, W., and Xu, D. (2020). A deep learning method with merged LSTM neural networks for SSHA prediction. *IEEE J-STARS.* 13, 2853–2860. doi: 10.1109/JSTARS.2020.2998461

Sugimoto, S., Qiu, B., and Kojima, A. (2020). Marked coastal warming off Tokai attributable to Kuroshio large meander. *J. Oceanogr.* 76 (2), 141–154. doi: 10.1007/s10872-019-00531-8

Sugimoto, S., Qiu, B., and Schneider, N. (2021). Local atmospheric response to the Kuroshio large meander path in summer and its remote influence on the climate of Japan. *J. Climate.* 34 (9), 3571–3589. doi: 10.1175/JCLI-D-20-0387.1

Tsujino, H., Usui, N., and Nakano, H. (2006). Dynamics of Kuroshio path variations in a high-resolution general circulation model. *J. Geophys. Res.-Oceans.* 111, C11001. doi: 10.1029/2005JC003118

Usui, N., Nagai, T., Saito, H., Suzuki, K., and Takahashi, M. (2019). *Progress of studies on Kuroshio path variations south of Japan in the past decade, in the Kuroshio current: physical, biogeochemical, and ecosystem dynamic*, Vol. 9. Eds. T. Nagai, H. Saito, K. Suzuki, et al, American Geophysical Union (AGU). 147–161. doi: 10.1002/9781119428428.ch9

Usui, N., Tsujino, H., Fujii, Y., and Kamachi, M. (2006). Short-range prediction experiments of the Kuroshio path variabilities south of Japan. *Ocean Dynam.* 56, 607–623. doi: 10.1007/s10236-006-0084-z

Wu, X., Zhao, Y., Han, G., Li, W., Shao, Q., Cao, L., et al. (2022). Temporal-spatial oceanic variation in relation with the three typical Kuroshio paths South of Japan. *Acta Oceanol. Sin.* 41 (2), 15–25. doi: 10.1007/s13131-021-1941-9

Xiao, C., Chen, N., Hu, C., Wang, K., Gong, J., and Chen, Z. (2019). Short and mid-term sea surface temperature prediction using time-series satellite data and LSTM-AdaBoost combination approach. *Remote Sens. Environ.* 233, 111358. doi: 10.1016/j.rse.2019.111358

Yang, Y., and Liang, X. S. (2019). New perspectives on the generation and maintenance of the Kuroshio large meander. *J. Phys. Oceanogr.* 49 (8), 2095–2113. doi: 10.1175/JPO-D-18-0276.1

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 03 February 2023
DOI 10.3389/fmars.2023.1097209

Check for updates

OPEN ACCESS

EDITED BY
Shiqiu Peng,
State Key Laboratory of Tropical Oceanography (CAS), China

REVIEWED BY
YoungHo Kim,
Pukyong National University, Republic of Korea
Zheqi Shen,
Hohai University, China

*CORRESPONDENCE
Wansuo Duan
duanws@lasg.iap.ac.cn

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

RECEIVED 13 November 2022
ACCEPTED 20 January 2023
PUBLISHED 03 February 2023

CITATION
Jiang L, Duan W, Wang H, Liu H and Tao L (2023) Evaluation of the sensitivity on mesoscale eddy associated with the sea surface height anomaly forecasting in the Kuroshio Extension.
Front. Mar. Sci. 10:1097209.
doi: 10.3389/fmars.2023.1097209

COPYRIGHT
© 2023 Jiang, Duan, Wang, Liu and Tao. 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(s) 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.

Evaluation of the sensitivity on mesoscale eddy associated with the sea surface height anomaly forecasting in the Kuroshio Extension

Lin Jiang[1,2], Wansuo Duan[2,3]*, Hui Wang[1], Hailong Liu[2,3] and Lingjiang Tao[4]

[1]Institute of Marine Science and Technology, Shandong University, Qingdao, China, [2]State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics (LASG), Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China, [3]College of Marine Science, University of Chinese Academy of Sciences, Qingdao, China, [4]Department of Atmospheric and Oceanic Sciences and Institute of Atmospheric Sciences, Fudan University, Shanghai, China

The sensitivity of the sea surface height anomaly (SSHA) forecasting on the accuracy of mesoscale eddies over the Kuroshio Extension region, which was determined by the conditional non-linear optimal perturbation (CNOP) method using a two-layer quasigeostrophic model, is evaluated by adopting multiply realistic marine datasets through an advanced particle filter assimilation method. It is shown that, if additional observations are preferentially assimilated to the sensitive area of mesoscale eddies identified by the CNOP, where the eddies present a clear high- to low-velocity gradient along the eddy rotation, the forecasting skill of the SSHA can be more significantly improved. It is also demonstrated that the forecasts of the SSHA in the region where the large-scale mean flow possesses much stronger barotropic and/or baroclinic instability tend to exhibit stronger sensitivity to the accuracy of the initial field in the sensitive area of mesoscale eddies. Therefore, more attention should be preferentially paid to the assimilation of the additional observations of the mesoscale eddies for the SSHA forecast in the region with a strong velocity shear of ocean circulation. The present study verifies the sensitivity on mesoscale eddies of SSHA forecasts derived by the two-layer quasigeostrophic model using multiply sets of realistic oceanic data, especially including observation and reanalysis data, which further additionally demonstrates the importance of targeted observations of mesoscale eddies to the SSHA forecast in the regions of strong velocity shear of ocean circulation and provides a more credible scientific basis for the field campaign of the targeted observations for mesoscale eddies associated with the SSHA forecasting.

KEYWORDS

mesoscale eddy, sensitivity, data assimilation, sea surface height anomaly, forecasting

1 Introduction

Sea surface height anomaly (SSHA) is considered one of the features for the surface and subsurface dynamics of the ocean and directly or indirectly reflects information on the main dynamic processes, including mesoscale eddies, waves, currents, and tides (Tanajura et al., 2016; Song et al., 2021). The SSHA has provided a wealth of information about ocean circulation and atmosphere–ocean interactions (Tandeo et al., 2014). The forecasting of SSHA is crucial for predicting future extreme or very hazardous phenomena such as extremely high waves, hurricanes, and other phenomena (Tanajura et al., 2015). Highly accurate SSHA forecasts will also be allowed to provide a sufficient basis for ship navigation, marine engineering, and industrial development, as well as fishery resources forecast (Solanki et al., 2015; Lumban-Gaol et al., 2017). Currently, the study on the SSHA assimilation and, thus, highly accurate SSHA prediction has been a hot topic in physical oceanography and meteorological sciences (Tanajura et al., 2015; Yavuzdoğan and Tanır Kayıkçı, 2020).

The prediction of SSHA has Is been a challenge (Song et al., 2021). For decades, numerical models based on dynamical/physical equations have played a dominant role in ocean predictions and are also often used to predict SSHA. The data assimilation method can provide more accurate initial conditions for numerical models by combining limited observations and model output, thus increasing the medium- and short-term prediction skills of the model. Many studies have been devoted to the studies on the assimilation of SSHA observations, and great progress has been made in improving the SSHA prediction, but there are still considerable uncertainties (Fang, 2006; Agarwal et al., 2022). Moreover, numerical models involve a lot of complex physical processes, and their corresponding integration requires a lot of computational resources. Then, if rich SSHA data provided by satellite altimeters are assimilated, it will need relatively high computational cost and even lead to an overfitting situation (Li et al., 2010; Song et al., 2021), which, together with the effect of the model errors, causes this assimilation to not necessarily provide positive effects on predictions. To address these embarrassments, it is essential to thin the data and devise an appropriate assimilating strategy to highly and effectively initialize numerical models associated with SSHA forecasting (Li et al., 2010; Zanna et al., 2018; Fraser et al., 2019).

Weiss and Grooms (2017) demonstrated that assimilating the observations on mesoscale eddies can achieve a more accurate ocean state than doing it over the whole model field; especially, they found that when fewer sea surface height (SSH) observations on the mesoscale eddies are assimilated, it improves the accuracy of the initial field more effectively and reduces more errors of the SSH predictions made by a two-layer quasigeostrophic (QG) model. Therefore, appropriate initialization of mesoscale eddies can lead to a much greater improvement in the prediction of the future SSH. In their work, the assimilation strategy of mesoscale eddies was to assimilate the observations on evenly distributed regular grids over eddies. However, mesoscale eddies are usually irregular in shape and asymmetric in the flow field, which reduces the stability of the vortex structure and presents a highly non-linear nature (Tang et al., 2020). Considering this point, Jiang et al. (2022) inferred that there should exist an area where the data assimilation should be preferentially implemented for the initialization of irregular eddies, rather than the evenly distributed regular grids on the eddies suggested by Weiss and Grooms (2017). Furthermore, they adopted an advanced approach of conditional non-linear optimal perturbation (CNOP; Mu et al., 2003) and revealed such area by using the two-layer QG model as adopted in Weiss and Grooms (2017). Exactly, this area is located on the eddies and presents a clear high- to low-velocity gradient along the eddy rotation. In this area, Jiang et al. (2022) provided a more effective assimilation strategy to mesoscale eddies associated with the improvement of the SSH anomaly (SSHA) forecasting skill. This useful area may represent the sensitive area of the initial field for SSHA forecasts, and the relevant assimilation strategy could provide an idea to design an observational array on mesoscale eddies for greatly improving the SSHA forecasting skill (Jiang et al., 2022). Such thought is related to the target observation, a new observational strategy for numerical weather forecasting and climate predictions (Snyder, 1996). Note that the CNOP approach has been successfully applied to the identification of the sensitive area for target observations of the forecasts for high-impact air–sea environmental events, such as the El Niño-Southern Oscillation, Indian Ocean Dipole, Kuroshio large meander, and Tropical Cyclone [see the review of Duan et al. (2022)], and the sensitivity on mesoscale eddy of the SSHA forecasting revealed by Jiang et al. (2022) is its another new attempt, which is still limited within the frame of the conceptual two-layer QG model.

To further verify the sensitivity on mesoscale eddy of the SSHA forecasting provided by Jiang et al. (2022), this study would examine it in realistic circumstances, where three sets of more realistic marine data are adopted and an advanced particle filter assimilation method is used. In addition, it is known that the Kuroshio Extension (KE) region, as a continuation of Kuroshio, has been observed as having the highest mesoscale eddy kinetic energy (EKE) in the global ocean (Wyrtki et al., 1976; Ferrari and Wunsch, 2009), and increasing attention has been paid to the potential role of eddies there in affecting the relevant ocean and overlying atmosphere (Qiu and Chen, 2005; Waterman et al., 2011; Nakamura et al., 2015; Yang and Liang, 2018). In the present study, the investigation will focus on mesoscale eddies in the KE region.

The rest of this paper is organized as follows. The data and algorithms adopted in the present study are introduced in Section 2, and the experimental design is described in Section 3. Section 4 evaluates the sensitivity on the accuracy on mesoscale eddies of SSHA forecasting, and an interpretation of the results is also presented there. Finally, a summary and discussion are provided in Section 5.

2 Data and algorithms

In this section, we will introduce more realistic oceanic data for the evaluation of the sensitivity on mesoscale eddies and the algorithms that are used to identify mesoscale eddies and assimilate ocean data. The details are as follows.

2.1 Data

The daily ocean grid data of the SSHA and surface current velocity components are used in the present study. The three sets of

data covering the KE region (32°N–38°N, 140°E–180°E) are, respectively, extracted from the time series of an ocean circulation model data, reanalysis data, and observation data from 2008 to 2017. The model data are from the output of the LICOM3, a global ocean general circulation model developed by the State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics (LASG)/Institute of Atmospheric Physics (IAP) of the Chinese Academy of Sciences. The LICOM3 has a free sea surface and Arakawa B grid and uses primitive equations with Boussinesq and hydrostatic approximations. The eddy-resolving simulation of LICOM3 is forced by a surface–atmospheric dataset for driving ocean models based on a Japanese 55-year atmospheric reanalysis (Tsujino et al., 2018), with the initial condition being from the Mercator Ocean analysis (Lellouche et al., 2018; Li et al., 2020). The grid space of the data is 10 km, but in the present study, we transfer the data having a resolution of 20 km and make it agree with the resolutions of the other two sets of data as follows. The reanalysis data come from the Global Ocean Reanalysis and Simulations (GLORYS), which was obtained from a global marine data assimilation and reanalysis system implemented under the framework of the MyOcean project composed of a consortium of 60 partners across Europe and structured around a core team of Marine Core Service operators, aiming to carry out simulations of the global ocean using a higher resolution grid under the constraints of integrating assimilated data (Ferry et al., 2010). At present, the system has been upgraded to GLORYS2 (Parent et al., 2011), and four versions of the global marine reanalysis data products were released, among which the latest version GLORYS2V4 contributes to the reanalysis data used in this study; these data were made by the Nucleus for European Models of the Ocean model (NEMO) on the horizontal resolution 0.25° × 0.25° (corresponding to the grid space of the latitudinal approximately 15 km and longitudinal approximately 25 km over the KE region) with 75 vertical levels through a surface forcing of the European Centre for Medium-Range Weather Forecasts reanalysis (ERA) interim and the assimilation of SSHA, sea surface temperature (SST), sea ice concentrations (SIC), and *in-situ* temperature and salinity (T/S) profiles. For the observation data used in the present study, they are from the Archiving, Validation, and Interpolating of Satellite Oceanographic (AVISO) altimeter data distributed by the Copernicus Marine and Environment Monitoring Service (CMEMS) and also on a 0.25° × 0.25° grid; these data were processed through optimal interpolation from all the delayed-time merging of multiple altimeter satellites (such as ERS-1/2, Topes/Poseidon, ENVISAT, and Jason-1/Jason-2).

2.2 Algorithms

2.2.1 Vortex identification algorithm

The present study adopts the SSHA-based eddy identification algorithm suggested by Chelton et al. (2011) to determine the vortex position. Specifically, the contours are first extracted with a given interval according to the SSHA maps, and an anticyclonic (cyclonic) eddy is defined as having one local maximum (minimum) of the SSHA enclosed by closed SSHA contours, with the eddy edge being the outermost closed contour.

2.2.2 Particle filtering assimilation

The particle filter (PF) is a sequential Monte Carlo procedure, which is often used to derive the probability distributions of state variables through a large number of independent random samples (i.e., particles). Such particles are directly sampled from the state space and are properly located, weighted, and propagated sequentially by the application of the Bayesian rule through assimilating the information contained in observations (Moradkhani et al., 2005). To facilitate the readers, we describe it as follows.

Suppose that the evolution of the state vector X_k is controlled by Eq. (1).

$$X_{k+1} = M(X_k, \zeta_k) \tag{1}$$

where X_k is the state vector at time t_k, M is the model propagation operator, and ζ_k is the white noise sequence with the mean value of 0. A group of Monte Carlo samples (i.e., particles) of weights are generated to approximate the prior probability density functions (PDFs) as in Eq. (2).

$$p_N(X_k) = \sum_{i=1}^{N} w_k^i \delta(X_k - X_k^i),\ i = 1, 2, \ldots, N \tag{2}$$

where X_k^i and w_k^i denote the ith particle of the model state and its weight, $\delta(\bullet)$ is the Dirac delta function, whose value is zero anywhere except at zero and whose integral over the entire real line is equal to one. The initial state of each particle is obtained by uniform sampling ($w_0^i = 1/N$)from the initial probability density distribution $p(X_0)$ of the state vector (Xie and Verbraeck, 2018).

Then, prediction and filtering are iterated. A very important concept in the PF method is the sequential importance sampling (SIS) used for selecting the particle weights through the information from a number of discrete observations. When an observation Y_k at t_k becomes available, the weight will change at each point in the domain according to the Bayes' theorem, thus yielding the new weight

$$w_k^i = \frac{p(Y_k|X_k^i)}{p(Y_k)} w_{k-1}^i \tag{3}$$

where $p(Y_k|X_k^i)$ is the PDF of the observations given the model state X_k^i, and $p(Y_k)$ is the PDF of the observation. The latter can be considered as a normalization factor, which ensures that the sum of the weights of all particles is equal to one (Kramer and Dijkstra, 2013). Assuming that the error distribution of a measurement H is a multivariate normal distribution and Σ denotes the error covariance matrix of the observations, then for a Gaussian distributed prior, $p(Y_k|X_k^i)$ can be expressed as in Eq. (4)

$$p(Y_k|X_k^i) \sim \exp[-\frac{1}{2}(Y_k - \mathrm{H}(X_k^i))^T \sum{}^{-1} (Y_k - \mathrm{H}(X_k^i))] \tag{4}$$

where H is the projection of the model state X_k^i into the observation space Y_k. With Eq. (4), the weight w_k^i can be calculated. It is noted that, if several observations at different grids are simultaneously assimilated, the weight w_k^i is updated according to Eq. (5).

$$w_k^i \sim \exp[-\frac{1}{2}\sum_{j=1}^{m}(Y_k - \mathrm{H}(X_k^i))^T \sum{}^{-1} (Y_k - \mathrm{H}(X_k^i))] \tag{5}$$

The core of the PF method is to change the weight of each ensemble member according to the observation information. Thus,

this assimilation method can be applied not only to model forward integrations but also to offline model ensemble prediction datasets. In fact, Kramer and Dijkstra (2013) have applied the PF method in offline ensemble data to determine the optimal observation location for the predictions of El Nino-Southern Oscillation. Duan and Feng (2018) also used this offline method to investigate an optimal observational array for improving two favors of El Niño predictions in the whole Pacific [also see Hou et al. (2022)]. In this paper, we would also use the PF method to evaluate the sensitivity of the SSHA forecast of mesoscale eddies by using the above realistic oceanic data.

3 Experimental design

The sensitivity on the accuracy of the mesoscale eddy associated with SSHA forecasting, as mentioned in Section 1, was determined by the CNOP method using the two-layer QG model (Jiang et al., 2022), mainly concerned with the lead time of 7 days of the SSHA forecasts. In this paper, the sensitivity will be quantitatively assessed by comparing the effect on the 7-day SSHA forecasting of different assimilation strategies implemented on the mesoscale eddies over the KE region, where the 7-day SSHA forecasting, for comparison, inherits the work of Jiang et al. (2022) and also of Weiss and Grooms (2017). The PF method, as described in Section 2, was employed to conduct the assimilation experiments. In order to generate the particles of the PF, the vortex identification algorithm is carried out four times each month every 7 days, specifically on days 1, 8, 15, and 22, for the three sets of data from 2008 to 2017, respectively. Particularly, we select the eddies that are identified on June 1 and December 1 of each year and possess a clear high- to low-velocity gradient along the eddy rotation featured by the CNOP-type perturbations in Jiang et al. (2022). Then, the flow field confined in a rectangle of a certain radius centered on the selected vortex center and its 7-day development are regarded as the "Truth Run" in the LICOM3 model data and the GLORYS2V4 reanalysis data; as a result, 55 "Truth Runs" for the LICOM3 data and 41 for the GLORYS2V4 data are obtained. For these two sets of data, synthetic "Observation" is then produced by adding a normally distributed stochastic noise N (0,0.25) (i.e., observational errors) on the "Truth Run." However, for the AVISO observation data, we also assume that its observational errors are stochastic noises satisfying a normal distribution. Then, we also yield stochastic noise with a normally distributed noise of N (0,0.25) and superimpose them on the "Observation" data to offset the observational errors, finally constructing 55 synthetic "Truth Runs" for the "Observation" in AVISO data. Thus, the Truth Run and the corresponding Observation in the three sets of data are determined for the assimilation experiments. Note that the standard deviation o 0.25 here is not proportional to that of the realistic observations (such as the AVISO) due to the limitation of the use of offline data and associated PF assimilation, but it is experimentally obtained to satisfy the need of evaluating the sensitivity on mesoscale eddy.

The identified eddies in each year and their corresponding 7-day developments can be regarded as the samples (i.e., the particles of the PF) of the predictions to the Truth Runs. By statistics, there are 4,366 particles in the LICOM data, 5,521 in the GLORYS2V4 data, and 5,619 in the AVISO data for each cyclonic vortex and 3,475, 5,526, and 7,221 particles for each anticyclonic vortex, respectively. The samples in different datasets make up an ensemble of equally weighted particles, and the ensemble mean can be regarded as the "Control Run" of each "Truth Run," respectively. When the observation information is introduced to the Control Run by the PF assimilation, the weight of each particle will change and then the corresponding ensemble mean is updated. This updated ensemble mean is hereafter referred to as the "Assimilation Run." To facilitate understanding, the logic of the Truth Run, Control Run, and Assimilation Run is shown in Figure 1. For the PF assimilation, four strategies are designed as shown in Figure 2, which are respectively referred to as PF1, PF2, PF3, and PF4. The PF1 assimilates the observations located in the sensitive area identified by Jiang et al. (2022), where the eddies present a clear high- to low-velocity gradient along the eddy rotation, while the PF2/3/4 assimilate the observations in three non-sensitive areas, which are respectively obtained by rotating 90°/180°/270° along the vortex rotation direction starting from the sensitive area. These four areas have a common area size and do not overlap each other, consequently covering the whole eddy. For PF1/2/3/3/4, they each assimilate 10 groups of observations, whose

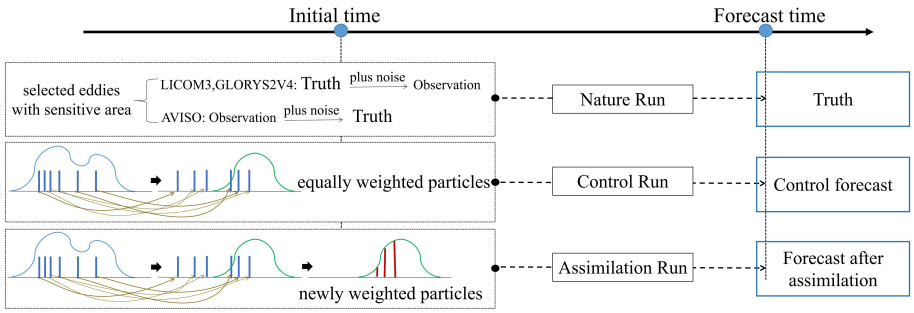


FIGURE 1
A diagram showing the validation scheme. The prior PDF (blue thin curve) of a system is sampled by a number of particles at the initial time, which are indicated by the blue vertical bars. These particles are all propagated forward in time, indicated by the brown lines. In the Control Run, the equally weighted particles make up an ensemble mean forecast, i.e., the control forecast in the figure; in the Assimilation Run, a group of newly weighted particles (red vertical bars) is obtained through the PF assimilation method using the observation information (green thin curve), and the ensemble mean of the newly weighted particles constitutes an updated forecast, i.e., the forecast from the assimilation. The comparison between the improvement of the Assimilation Run and that of the Control Run against the Nature Run would reveal the usefulness of assimilated observations, where the Nature Runs are obtained by taking the model runs for the LICOM3 and GLORYS2V4 reanalysis data and superimposing the noise to the observation for the AVISO.

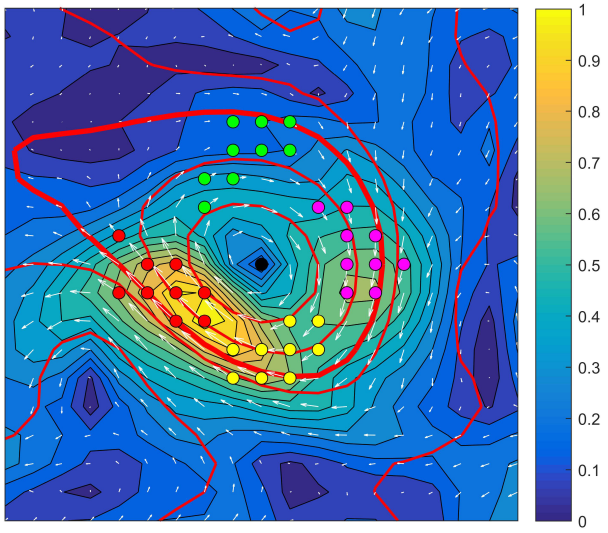


FIGURE 2
Four kinds of assimilation strategies. The colored bold points denote the (synthetic) observations and the red, green, purple, and yellow dots correspond to the (synthetic) observations assimilated by PF1, PF2, PF3, and PF4, respectively. The PF1 assimilates nine observations in the sensitive area, while the PF2, PF3, and PF4 assimilate nine observations in the non-sensitive area obtained by rotating 90°, 180°, and 270° along the vortex rotation direction starting from the PF1 assimilation area, respectively. Note that the distribution of the nine observations in each area shows an example of randomly selected 10 groups of observational locations.

locations are randomly selected from a corresponding area, so as to make the results more reliable in statistics.

The extent of the error reduction from assimilation is evaluated by Eq. (6).

$$b = \frac{\mathrm{d}F_1 - \mathrm{d}F_2}{\mathrm{d}F_1} \times 100\,\% \tag{6}$$

where $\mathrm{d}F_1$ is the forecast error of the Control Run with respect to the Truth Run and $\mathrm{d}F_2$ is that of the Assimilation Run with respect to the Truth Run. It is noted that these forecast errors are both measured by the root mean square error (RMSE) with the formula as in Eq. (7).

$$\mathrm{RMSE} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(P_i - T_i)^2}, \tag{7}$$

where m represents the total number of grid points in the concerned forecast area, and T_i P_i are the truth and its prediction on the ith grid point, respectively.

4 Evaluation of the sensitivity on the mesoscale eddy

In this section, we evaluate the sensitivity on mesoscale eddies of SSHA forecasting by using the realistic oceanic data provided in Section 2; particularly, we focus on the KE region and separate the circulation fields of strong and weak dynamical instabilities to do it.

4.1 The validity of the sensitivity on mesoscale eddy in promoting the SSHA forecasting skill

The assimilation strategies shown in Figure 2 are implemented to the Control Run of each Truth Run, and the extent of the error reduction from assimilation, i.e., b in Eq. (6), is calculated. Figure 3 shows the box plots of b for the three sets of SSHA data over the KE region, where the values of b are relative to the selected eddies and the randomly selected 10 groups of observations in each area on the eddies (see Section 3). It is easily seen that, for all the three datasets, the values of b in PF1 are always obviously larger than those in PF2, PF3, and PF4. Furthermore, when we examine respectively the three datasets to count the number of eddies that exhibit the largest value of b among the PF1, PF2, PF3, and PF4 strategies in terms of the ensemble mean after assimilating 10 groups of observations, we find that the number of eddies using the PF1 assimilation strategy is the highest (see Table 1). This indicates that the PF1 strategy, for the collected eddies, has a larger probability to significantly enhance the corresponding SSHA forecasting skill, as compared with the PF2/3/4 strategies. This result, combined with the sensitivity on the accuracy of the mesoscale eddy revealed by Jiang et al. (2022), shows that PF1 could be the optimal assimilation strategy for SSHA forecasting. This implies that additional observations should be preferentially implemented in the areas with a clear high- to low-velocity gradient along the rotation direction on mesoscale eddies. Consequently, the sensitive area of the mesoscale eddy associated with SSHA forecasting determined by CNOP is effective even when using more realistic marine data including model data, reanalysis, and observations. This also sheds light on that the sensitive area identified by the conceptual QG model in Jiang et al. (2022) could be robust, which therefore could provide reliable scientific guidance for implementing additional observations of actual mesoscale eddies in realistic field campaigns for improving SSHA forecasting skill.

4.2 Modulating effect of ocean circulation instability on the sensitivity on mesoscale eddies

We have verified that the PF1 assimilation strategy of the sensitive areas of mesoscale eddies is superior to the PF2/3/4 strategies of non-sensitive areas for improving the SSHA forecasting skill. It is noted that we are concerned about the mesoscale eddies over the KE region. Meanwhile, it is known that there exist baroclinic (BC) and/or barotropic (BT) instabilities of ocean circulation in the KE region in the presence of strong shear of the eastward-flowing jet; furthermore, notable differences in instability strengths exist between the upstream and downstream KE regions with the former having much stronger instability (Spall, 2000; Williams et al., 2007; Stammer et al., 2012; Bishop, 2013). Then how will these instabilities affect the sensitivity on mesoscale eddies? To address this question, we will separate the upstream and downstream regions of the KE and further analyze the sensitivity on mesoscale eddies.

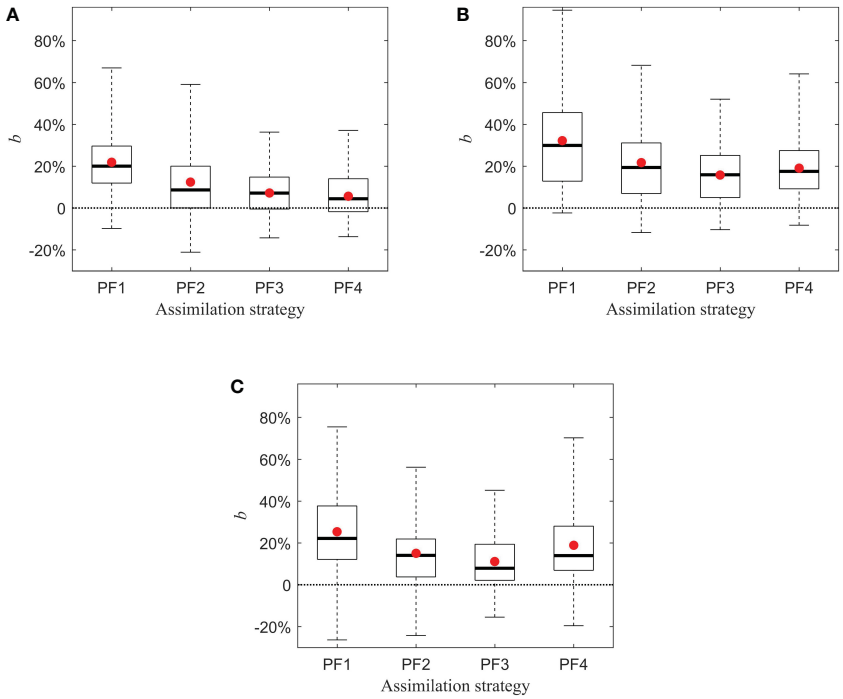


FIGURE 3
The box plots of *b* for the PF1, PF2, PF3, and PF4 assimilation strategies using **(A)** LICOM3 data, **(B)** GLORYS2V4 data, and **(C)** AVISO data, with respect to selected eddies and 10 groups of observations for each area.

Referring to Yang and Liang (2016), we recognize the regions west of 154°E as the upstream region and the east of 154°E as the downstream region. Then, there are 40/15, 20/21, and 30/25 eddies in the upstream/downstream KE regions from LICOM3 data, GLORYS2V4 data, and AVISO data, respectively. For these eddies, we investigate the sensitivity on mesoscale eddies of the SSHA forecasting using the assimilation strategies as in Section 3 and identify the differences between the upstream and downstream regions.

The results are plotted in Figure 4 and Table 2. Obviously, all of the three datasets demonstrate that the value of *b* is still the largest when using the PF1 strategy in either upstream or downstream regions, and the number of eddies with PF1 being the most effective assimilation strategy is also the largest (see Table 2). This indicates that the advantages of the PF1 strategy are still valid over both upstream and downstream regions. When we further compare the degrees of improvements of the SSHA forecasting skill due to assimilation between upstream and downstream, it seems that PF1 provides an improvement in the upstream KE region with almost the same degree as in the downstream region, according to the arithmetic mean and median of improvements; however, when we count the number of eddies that possess the largest value of *b* in PF1, PF2, PF3, and PF4, it is found that the percentage of the number of eddies with PF1 being the best assimilation strategy to the total number of eddies in the upstream region is obviously larger than that in the downstream region (see Table 2). Obviously, this indicates that assimilating additional observations located in sensitive areas of mesoscale eddies in the upstream region, compared with doing it in the downstream region, possesses a greater probability to improve the corresponding SSHA forecasting skill, although the improvements in these two regions are of less different amplitudes.

To sum up, it is particularly noteworthy that, for more realistic ocean data of LICOM3 data, GLORYS2V4 data, and AVISO data investigated here, the assimilation implemented in the sensitive area of mesoscale eddies in the upstream region exhibits more advantages than that conducted in the downstream region to improve the SSHA forecasting skill. Therefore, the sensitivity on mesoscale eddies of the SSHA forecasting is more prominent in the upstream KE region than in the downstream region, and it is more effective for improving the SSHA forecasting level to assimilate additional observations located in the sensitive areas of mesoscale eddies in the upstream KE region.

TABLE 1 The numbers of eddies with the largest value of *b* occurring in PF1, PF2, PF3, and PF4, respectively.

	PF1	PF2	PF3	PF4
LICOM3 data	33	9	7	6
GLORYS2V4 data	18	13	3	7
AVISO data	33	6	7	9

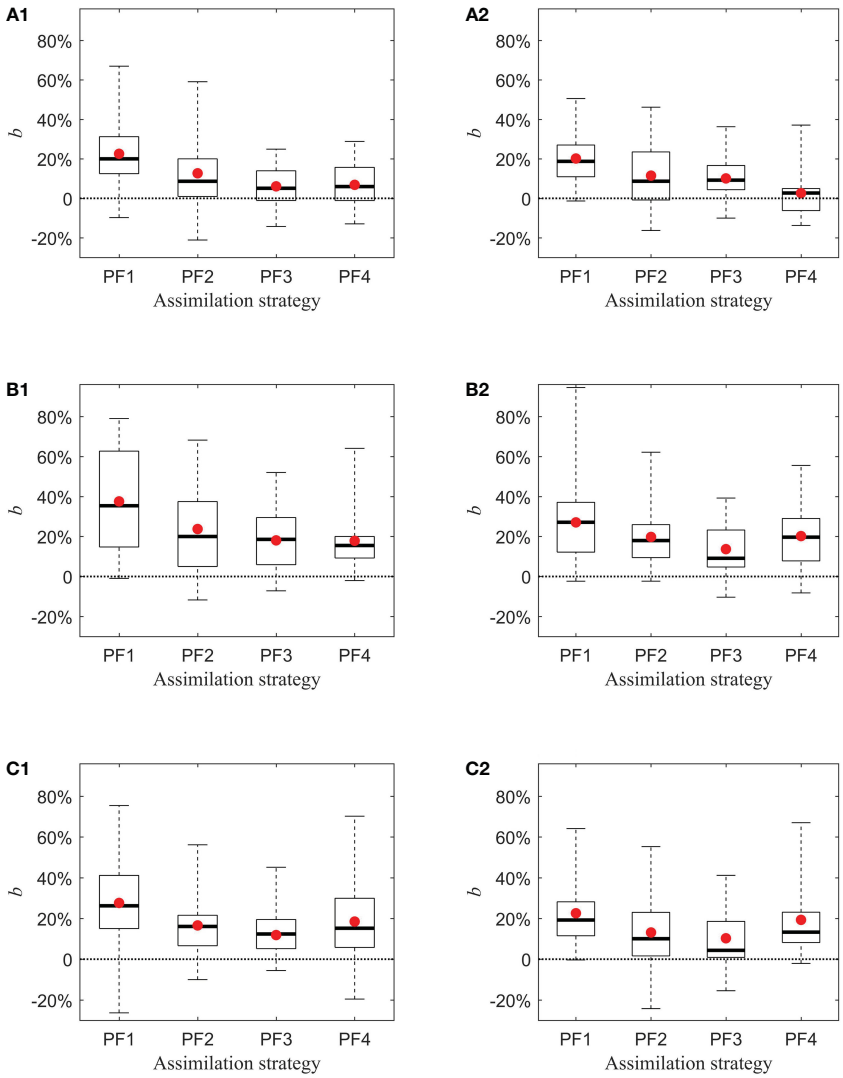


FIGURE 4
The box plots of *b* of PF1, PF2, PF3, and PF4 in the upstream (1) and downstream (2) of the KE, with respect to selected eddies and 10 groups of observations for each area. **(A–C)** For the LICOM model data, GLORYS2V4 data, and AVISO data, respectively.

4.3 Interpretation

In this section, we would interpret why the SSHA forecasting skill is more prominently improved in the upstream KE region than in the downstream KE region by preferentially implementing the additional assimilation in the sensitive area of mesoscale eddies. In fact, this conclusion involves the energy conversion of different spatial scales. Results in previous studies have shown that mesoscale eddies tend to extract energy from mean flow (ocean circulation) in the upstream KE region along the stream direction through an eddy–wave interaction, but they are inclined to transmit energy to mean flow in the downstream KE region (Hall, 1991; Yang and Liang 2018). Therefore, the stronger BT or BC instability in the upstream KE region would induce a stronger eddy–wave interaction and make the circulation mean flow transmit more energies to the mesoscale eddies; as such, the mesoscale eddies in the upstream KE region possess more energies. Then, on mesoscale eddies, the initial perturbations located in the sensitive area, where a clear high- to low-velocity gradient along the eddy rotation is presented, would stimulate a much larger positive BT conversion rate according to the equation $\mathrm{BT} = -\frac{V'^2}{(\bar{V}+V')} \cdot \frac{\partial \bar{V}}{\partial t}$ with

TABLE 2 The number of eddies with the largest value of *b* in PF1/PF2/PF3/PF4 and the percentages of the number of eddies with the largest value of *b* in the PF1 to the total eddy number in the upstream and downstream KE regions.

PF1/PF2/PF3/PF4	Upstream KE	Downstream KE
LICOM data	26/5/3/6 (65.00%)	7/4/4/0 (46.67%)
GLORYS2V4 data	10/6/2/2 (50.00%)	8/7/1/5 (38.10%)
AVISO data	19/3/5/3 (63.33%)	14/3/2/6 (56.00%)

the negative velocity tendency $\frac{\partial \bar{V}}{\partial t} < 0$ due to a high- to low-velocity gradient in the sensitive area, where V' is the velocity component of the initial perturbations and $\bar{V}$ represents the flow velocity of the mesoscale eddy in a natural coordinate system [the details are referred to as in Jiang et al. (2022)]. Consequently, the energies transmitted from the mean flow to the mesoscale eddies in the upstream KE region would be provided much more to the perturbation and would enhance its much quicker growth, finally yielding a much greater impact on the SSHA forecasting in the upstream KE region. That is to say, the SSHA forecasting in the upstream region is much more sensitive to the accuracy of the mesoscale eddies there, especially sensitive to the accuracy of the flow field in the sensitive area of mesoscale eddies. Then, if we give priority to implementing additional observations in the sensitive area of mesoscale eddies in the upstream KE region and assimilate them to the Control Run for improving its initial field, the SSHA forecasting skill would have more probabilities to achieve much greater improvement. On the contrary, for the downstream KE region with much weaker BT or BC instability, the energies transmitted from the ocean circulation to the mesoscale eddies are much less, and thus, there are not enough energies provided to promote the growth of the initial perturbations even in the sensitive areas of mesoscale eddies, eventually exerting a weaker impact on the SSHA forecasting there. Therefore, the sensitivity on mesoscale eddies in the downstream region of weaker instability is not as strong as that in the upstream region of stronger instability, and assimilating additional observations in the sensitive areas in the downstream region to the Control Run is certainly less effective than doing it in the upstream region for improving the SSHA forecasting skill.

Combining the above numerical results and theoretical reasoning, it is concluded that the stronger the dynamical instability of mean flow (or ocean circulation), the stronger the sensitivity on mesoscale eddies of SSHA forecasting. Therefore, in the regions with a strong instability of ocean circulation, such as in upstream KE, more attention should be preferentially paid to assimilating additional observations (i.e., the targeted observations) on mesoscale eddies, so as to efficiently improve the accuracy of mesoscale eddies and, thus, greatly increase the SSHA forecasting skill.

5 Conclusion and discussion

In this paper, the sensitivity on mesoscale eddies of SSHA forecasting, determined by CNOP through a conceptual QG model (Jiang et al., 2022), is further evaluated using three sets of realistic marine data, particularly including the LICOM3 model data, GLORYS2V4 reanalysis data, and AVISO altimeter observation data. An advanced PF assimilation method is implemented over the KE region to improve the initial accuracy of mesoscale eddies there and then increase the corresponding SSHA forecasting. Four assimilation strategies are tested, which are relevant with the assimilation in the sensitive area, where a clear high- to low-velocity gradient along the eddy rotation is presented, and those in the other three non-sensitive areas of mesoscale eddies. The results demonstrate that the assimilation implemented in the sensitive area of mesoscale eddies is most effective for promoting the SSHA forecasting skill. This sheds light on the fact that the sensitivity on mesoscale eddy of SSHA forecasting determined by the QG model together with the CNOP approach in Jiang et al. (2022) is reasonable even in realistic marine data. It is therefore concluded that the sensitivity on mesoscale eddies obtained by the QG model is reliable for providing scientific guidance for targeting observation of actual mesoscale eddies associated with SSHA forecasting in the KE region. That is to say, additional observations in the sensitive areas of mesoscale eddies should be preferentially implemented and/or assimilated in order to greatly improve the forecasting skills of SSHA in the KE region.

The above sensitivity on mesoscale eddies is also tested by separating the upstream and downstream regions of KE. The upstream region presents the oceanic circulation with much stronger dynamical (BT and/or BC) instability, while the downstream region provides much weaker instability. It is shown that the assimilation implemented in the sensitive areas of mesoscale eddies in the upstream region has more probabilities than that in the downstream region for improving the SSHA forecasting skill. Theoretically, the stronger eddy–wave interaction induced by the stronger instability in the upstream KE region tends to make the ocean circulation transmit more energies to the mesoscale eddies there, thus being favorable for more energies further provided to the disturbances on the eddies through the mechanism of the BT instability [see Jiang et al. (2022)] and finally yielding a much greater impact on the SSHA forecasting in the upstream KE region due to the growth of disturbances, and the strongest instability in the sensitive area of mesoscale eddies would enhance most the growth of the disturbances of the SSHA. It is therefore certain that, if additional observations are preferentially implemented in the sensitive area in the upstream KE region and assimilated to the model fields, the growth of initial errors there would be greatly suppressed and the corresponding SSHA forecasting skill would be much more significantly improved. It is suggested that more attention should be preferentially paid to the assimilation of the targeted observations on mesoscale eddies located in the area where the background ocean circulation presents stronger instability, such as the upstream KE region, in order to efficiently improve the SSHA forecasting skill.

The sensitivity on mesoscale eddies of SSHA prediction was revealed in Jiang et al. (2022) and in the present study, and it is further practically evaluated through three sets of realistic ocean data; moreover, more concerns are additionally suggested to the sensitivity on mesoscale eddies over the upstream region for improving the SSHA forecasting skill there. However, due to the limitation of the offline data adopted in the present study, we have to deduce the possible dynamical mechanism responsible for the relationship between the sensitivity on mesoscale eddy and the dynamical instability of the background mean flow; therefore, a quantitative evaluation is expected to verify the dynamics of the modulation role of the mean flow to the sensitivity on mesoscale eddy of the SSHA forecasts by using specific ocean models, such as the Regional Ocean Modeling System. Also, the PF assimilation method used in this paper has the advantages of easy operation and offline implementation, whereas the phenomenon of particle degeneracy sometimes occurs and severely influences the quality of assimilating results; as such, other data assimilation methods, such as the ensemble Kalman filter and four-dimensional variational methods, are anticipated to be

applied in specific models to examine the sensitivity on mesoscale eddies of the SSHA predictions. Doing such would also help investigate the sensitivity on mesoscale eddies from the three-dimensional structure on mesoscale eddy, rather than from the frame of two-dimensional motion of mesoscale eddy in the present study. The interactions among eddies would induce anomalies of ocean state and play a significant effect on the underlying atmosphere, such as winds, clouds, precipitation, and typhoons (Chelton, 2013; Renault et al., 2019) by the air–sea interaction, and thus, relevant studies are also expected. The present study focuses on the accuracy of mesoscale eddies but is related to the forecast of SSHA. In fact, the predictions of mesoscale eddy and its moving track and intensity are essential for describing further ocean state and its underlying atmosphere, and therefore, the corresponding predictability study should be carried out comprehensively although it is much more challenging. It is expected that the present study and the associated study of Jiang et al. (2022) can provide useful ideas to address the predictability of mesoscale eddies themselves.

Data availability statement

The raw data supporting the conclusions of this article will be made available upon reasonable request and with authors'permission. Open source datasets include the GLORYS2V4 data and AVISO data (both available online at https://resources.marine.copernicus.eu.).

Author contributions

LJ and WD conceived the research, designed the experiments, performed the simulations, and analyzed the results. All authors contributed to the article and approved the submitted version.

Funding

The study was supported by the National Natural Science Foundation of China (Grant No. 41930971).

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Agarwal, N., Sharma, R., and Kumar, R. (2022). Impact of along-track altimeter sea surface height anomaly assimilation on surface and sub-surface currents in the bay of Bengal. *Ocean Model.* 169 (9),101931. doi: 10.1016/j.ocemod.2021.101931

Bishop, S. P. (2013). Divergent eddy heat fluxes in the kuroshio extension at 144°–148° E. part II: Spatiotemporal variability. *J. Phys. Oceanogr.* 43 (11), 2416–2431. doi: 10.1175/jpo-d-13-061.1

Chelton, D. (2013). Mesoscale eddy effects. *Nat. Geosci.* 6 (8), 594–595. doi: 10.1038/ngeo1906

Chelton, D. B., Schlax, M. G., and Samelson, R. M. (2011). Global observations of nonlinear mesoscale eddies. *Prog. Oceanogr.* 91 (2), 167–216. doi: 10.1016/j.pocean.2011.01.002

Duan, W. S., and Feng, F. (2018). Application of particle filter assimilation in the target observation for El niño–southern oscillation. *Chin. J. Atmos. Sci. (in Chinese)* 42 (3), 677–695. doi: 10.3878/j.issn.1006-9895.1711.17264

Duan, W. S., Yang, L. C., Mu, M., Wang, B., Shen, X. S., Meng, Z. Y., et al. (2022). Advances in predictability study on weather and climate in China. *Adv. Atmos. Sci.* in reviewing.

Fang, M. (2006). Study of assimilating the satellite Sea surface temperature and Sea surface height anomaly into an East China Sea model - a progress report. 2nd Dragon Symposium, Santorini, Greece. European Space Agency Publications Division.

Ferrari, R., and Wunsch, C. (2009). Ocean circulation kinetic energy: Reservoirs, sources, and sinks. *Annu. Rev. Fluid Mech.* 41 (1), 253–282. doi: 10.1146/annurev.fluid.40.111406.102139

Ferry, N., Parent, L., Garric, G., Barnier, B., and Jourdain, N. (2010). Mercator Global eddy permitting, ocean reanalysis GLORYS1V1: Description and results. *Mercator-Ocean Q. Newslett.* 36, 15–27.

Fraser, R., Palmer, M., Roberts, C., Wilson, C., Copsey, D., and Zanna, L. (2019). Investigating the predictability of north Atlantic sea surface height. *Climate Dyn.* 53 (3-4), 2175–2195. doi: 10.1007/s00382-019-04814-0

Hall, M. M. (1991). Energetics ot the kuroshio extension at 35°N,152°E. *J. Phys. Oceanogr.* 21, 958–975. doi: 10.1175/1520-0485(1991)0210958:EOTKEA>2.0.CO;2

Hou, M., Tang, Y., Duan, W., and Shen, Z. (2022). Toward an optimal observational array for improving two flavors of El niño predictions in the whole pacific. *Climate Dyn.* doi: 10.1007/s00382-022-06342-w

Jiang, L., Duan, W., and Liu, H. (2022). The most sensitive initial error of Sea surface height anomaly forecasts and its implication for target observations of mesoscale eddies. *J. Phys. Oceanogr.* 52, 723–740. doi: 0.1175/jpo-d-21-0200.1

Kramer, W., and Dijkstra, H. A. (2013). Optimal localized observations for advancing beyond the ENSO predictability barrier. *Nonlinear Process. Geophys.* 20 (2), 221–230. doi: 10.5194/npg-20-221-2013

Lellouche, J., Greiner, E., Galloudec, O., Garric, G., Regnier, C., Drevillon, M., et al. (2018). Recent updates to the Copernicus marine service global ocean monitoring and forecasting real-time 1/12° high-resolution system. *Ocean Sci.* 14, 1093–1126. doi: 10.5194/os-14-1093-2018

Li, Y., Liu, H., Ding, M., Lin, P., Yu, Z., Yu, Y., et al. (2020). Eddy-resolving simulation of CAS-LICOM3 for phase 2 of the ocean model intercomparison project. *Adv. Atmos. Sci.* 37 (10), 1067–1080. doi: 10.1007/s00376-020-0057-z

Li, X., Zhu, J., Xiao, Y., and Wang, R. (2010). A model-based observation-thinning scheme for the assimilation of high-resolution SST in the shelf and coastal seas around China. *J. Atmos. Ocean. Technol.* 27, 1044–1058. doi: 10.1175/2010jtecho709.1

Lumban-Gaol, J., Leben, R. R., Vignudelli, S., Mahapatra, K., Okada, Y., Nababan, B., et al. (2017). Variability of satellite-derived sea surface height anomaly, and its relationship with bigeye tuna (Thunnus obesus) catch in the Eastern Indian ocean. *Eur. J. Remote Sens.* 48 (1), 465–477. doi: 10.5721/EuJRS20154826

Moradkhani, H., Hsu, K.-L., Gupta, H., and Sorooshian, S. (2005). Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter. *Water Resour. Res.* 41 (5), W05012. doi: 10.1029/2004wr003604

Mu, M., Duan, W. S., and Wang, B. (2003). Conditional nonlinear optimal perturbation and its applications. *Nonlinear Process. Geophys.* 10 (6), 493–501. doi: 10.5194/npg-10-493-2003

Nakamura, H., Isobe, A., Minobe, S., Mitsudera, H., Nonaka, M., and Suga, T. (2015). "Hot spots" in the climate system–new developments in the extratropical ocean–atmosphere interaction research: a short review and an introduction. *J. Oceanogr.* 71 (5), 463–467. doi: 10.1007/s10872-015-0321-5

Parent, L., Ferry, N., Garric, G., Legalloudec, O., Testut, C.-e., Barnier, B., et al. (2011). GLORYS2: A global ocean reanalysis simulation of the period 1992-present. *13. Abstract retrieved from Abstracts in Geophysical Research Abstracts.*

Qiu, B., and Chen, S. (2005). Variability of the kuroshio extension jet, recirculation gyre, and mesoscale eddies on decadal time scales. *J. Phys. Oceanogr.* 35 (11), 2090–2103. doi: 10.1175/jpo2807.1

Renault, L., Masson, S., Oerder, V., Jullien, S., and Colas, F. (2019). Disentangling the mesoscale ocean-atmosphere interactions. *J. Geophys. Res.: Oceans* 124 (3), 2164–2178. doi: 10.1029/2018jc014628

Snyder, C. (1996). Summary of an informal workshop on adaptive observations and FASTEX. *Bull. Am. Meteorol. Soc.* 77, 953–961. doi: 10.1175/1520-0477-77.5.953

Solanki, H. U., Bhatpuria, D., and Chauhan, P. (2015). Signature analysis of satellite derived SSHa, SST and chlorophyll concentration and their linkage with marine fishery resources. *J. Mar. Syst.* 150, 12–21. doi: 10.1016/j.jmarsys.2015.05.004

Song, T., Han, N., Zhu, Y., Li, Z., Li, Y., Li, S., et al. (2021). Application of deep learning technique to the sea surface height prediction in the south China Sea. *Acta Oceanol. Sin.* 40 (7), 68–76. doi: 10.1007/s13131-021-1735-0

Spall, M. A. (2000). Generation of strong mesoscale eddies by weak ocean gyres. *J. Mar. Res.* 58 (1), 97–116. doi: 10.1357/002224000321511214

Stammer, D., Marotzke, J., Maier-Reimer, E., Hernández-Deckers, D., Haak, H., Fast, I., et al. (2012). An estimate of the Lorenz energy cycle for the world ocean based on the STORM/NCEP simulation. *J. Phys. Oceanogr.* 42 (12), 2185–2205. doi: 10.1175/jpo-d-12-079.1

Tanajura, C. A. S., Lima, L. N., and Belyaev, K. P. (2015). Assimilation of satellite surface-height anomalies data into a hybrid coordinate ocean model (HYCOM) over the Atlantic ocean. *Oceanology* 55 (5), 667–678. doi: 10.1134/s0001437015050161

Tanajura, C. A. S., Lima, L. N., and Belyaev, K. (2016). Impact on oceanic dynamics from assimilation of satellite surface height anomaly data into the hybrid coordinate ocean model ocean model (HYCOM) over the Atlantic ocean. *Oceanology* 56 (4), 509–514. doi: 10.1134/s000143701603022x

Tandeo, P., Chapron, B., Ba, S., Autret, E., and Fablet, R. (2014). Segmentation of mesoscale ocean surface dynamics using satellite SST and SSH observations. *IEEE Trans. Geosci. Remote Sens.* 52 (7), 4227–4235. doi: 10.1109/tgrs.2013.2280494

Tang, Q., Gulick, S. P. S., Sun, J., Sun, L., and Jing, Z. (2020). Submesoscale features and turbulent mixing of an oblique anticyclonic eddy in the gulf of Alaska investigated by marine seismic survey data. *J. Geophys. Res.: Oceans* 125 (1). doi: 10.1029/2019jc015393

Tsujino, H., Urakawa, S., Nakano, H., Small, R., Kim, W., Yeager, S., et al. (2018). JRA-55 based surface dataset for driving ocean-sea-ice models (JRA55-do). *Ocean Model.* 130, 79–139. doi: 10.1016/j.ocemod.2018.07.002

Waterman, S., Hogg, N. G., and Jayne, S. R. (2011). Eddy–mean flow interaction in the kuroshio extension region. *J. Phys. Oceanogr.* 41 (6), 1182–1208. doi: 10.1175/2010jpo4564.1

Weiss, J. B., and Grooms, I. (2017). Assimilation of ocean sea-surface height observations of mesoscale eddies. *Chaos* 27 (12), 126803. doi: 10.1063/1.4986088

Williams, R. G., Wilson, C., and Hughes, C. W. (2007). Ocean and atmosphere storm tracks: The role of eddy vorticity forcing. *J. Phys. Oceanogr.* 37 (9), 2267–2289. doi: 10.1175/jpo3120.1

Wyrtki, K., Magaard, L., and Hager, J. (1976). Eddy energy in the oceans. *J. Geophys. Res.* 81 (15), 2641–2646. doi: 10.1029/JC081i015p02641

Xie, X., and Verbraeck, A. (2018). A particle filter-based data assimilation framework for discrete event simulations. *Simulation* 95 (11), 1027–1053. doi: 10.1177/0037549718798466

Yang, Y., and Liang, X. S. (2016). The instabilities and multiscale energetics underlying the mean–Interannual–Eddy interactions in the kuroshio extension region. *J. Phys. Oceanogr.* 46 (5), 1477–1494. doi: 10.1175/jpo-d-15-0226.1

Yang, Y., and Liang, X. S. (2018). On the seasonal eddy variability in the kuroshio extension. *J. Phys. Oceanogr.* 48 (8), 1675–1689. doi: 10.1175/jpo-d-18-0058.1

Yavuzdoğan, A., and Tanır Kayıkçı, E. (2020). A copula approach for sea level anomaly prediction: a case study for the black Sea. *Surv. Rev.* 53 (380), 436–446. doi: 10.1080/00396265.2020.1816314

Zanna, L., Brankart, J. M., Huber, M., Leroux, S., Penduff, T., and Williams, P. D. (2018). Uncertainty and scale interactions in ocean ensembles: From seasonal forecasts to multidecadal climate predictions. *Q. J. R. Meteorol. Soc.* 145 (S1), 160–175. doi: 10.1002/qj.3397

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 06 February 2023
DOI 10.3389/fmars.2023.1112065

Check for updates

OPEN ACCESS

EDITED BY
Shiqiu Peng,
State Key Laboratory of Tropical Oceanography, South China Sea Institute of Oceanology, Chinese Academy of Sciences

REVIEWED BY
Zheqi Shen,
Hohai University, China
Huizan Wang,
National University of Defense Technology, China

*CORRESPONDENCE
Xiaomeng Huang
hxm@mail.tsinghua.edu.cn

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

RECEIVED 30 November 2022
ACCEPTED 23 January 2023
PUBLISHED 06 February 2023

CITATION
Wang M, Wang D, Xiang Y, Liang Y, Xia R, Yang J, Xu F and Huang X (2023) Fusion of ocean data from multiple sources using deep learning: Utilizing sea temperature as an example. *Front. Mar. Sci.* 10:1112065. doi: 10.3389/fmars.2023.1112065

COPYRIGHT
© 2023 Wang, Wang, Xiang, Liang, Xia, Yang, Xu and Huang. 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(s) 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.

Fusion of ocean data from multiple sources using deep learning: Utilizing sea temperature as an example

Mingqing Wang[1,2], Danni Wang[2], Yanfei Xiang[1], Yishuang Liang[2], Ruixue Xia[1], Jinkun Yang[3], Fanghua Xu[1] and Xiaomeng Huang[1]*

[1]Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing, China, [2]Intelligent Forecasting Division, Ninecosmos Science and Technology Ltd., Wuxi, China, [3]National Marine Data and Information Service, Ministry of Natural Resources, Tianjin, China

For investigating ocean activities and comprehending the role of the oceans in global climate change, it is essential to gather high-quality ocean data. However, existing ocean observation data have deficiencies such as inconsistent spatial and temporal distribution, severe fragmentation, and restricted observation depth layers. Data assimilation is computationally intensive, and other conventional data fusion techniques offer poor fusion precision. This research proposes a novel multi-source ocean data fusion network (ODF-Net) based on deep learning as a solution for these issues. The ODF-Net comprises a number of one-dimensional residual blocks that can rapidly fuse conventional observations, satellite observations, and three-dimensional model output and reanalysis data. The model utilizes vertical ocean profile data as target constraints, integrating physics-based prior knowledge to improve the precision of the fusion. The network structure contains channel and spatial attention mechanisms that guide the network model's attention to the most crucial features, hence enhancing model performance and interpretability. Comparing multiple global sea temperature datasets reveals that the ODF-Net achieves the highest accuracy and correlation with observations. To evaluate the feasibility of the proposed method, a global monthly three-dimensional sea temperature dataset with a spatial resolution of 0.25°×0.25° is produced by fusing ocean data from multiple sources from 1994 to 2017. The rationality tests on the fusion dataset show that ODF-Net is reliable for integrating ocean data from various sources.

KEYWORDS

data fusion, three-dimensional ocean datasets, deep learning, attention mechanisms, physics-based prior knowledge

1 Introduction

Ocean science research has acquired international attention in recent years due to the ocean's importance as a regulator of the Earth's system and its importance in controlling and preventing global climate change (Cheng et al., 2020). To understand and predict climate change and the evolution of the marine environment, researchers collected high-quality ocean data to conduct scientific investigations and numerical simulations. Currently, *in situ* observations are used to collect the vast majority of ocean data. Oceanographic float observations can provide more precise data on the interior of the ocean, but their sparsity, uneven distribution, and low resolution make them challenging to employ directly in ocean research and numerical model simulations (Su et al., 2018; Su et al., 2021). Satellite remote sensing monitoring of the ocean has advanced rapidly in recent years, allowing for continuous observations of the ocean over a wide area and for long periods of time. However, ocean satellites are unable to observe subsurface and deeper ocean structures and processes due to their limited observation depth (Chapman and Charantonis, 2017). Combining the benefits of multi-source observation data to build 3D gridded ocean datasets is therefore an important and challenging problem.

A large number of researches have been conducted on multi-sensor sea surface satellite data fusion, including optimum interpolation methods, Bayesian methods, and variational methods (Zhu et al., 2018; Xiao et al., 2021). For example, NCEP developed RTG SST, a satellite-based SST analysis dataset for real-time global SST monitoring, and OISST, an SST analysis dataset for optimum interpolation (Thiébaux et al., 2003; Chelton and Wentz, 2005). To acquire high spatial and temporal resolution SST from the merging of coastal multi-satellite SST and *in situ* observation data, Chao et al. (2009) utilized the two-dimensional variational (2DVAR) data assimilation method. To achieve the fusion of multi-sensor SST data, Zhu et al. (2018) employed the Spatiotemporal Hierarchical Bayesian Model. Successive correction analysis (SCA), optimum interpolation (OI), variational methods (3DVAR and 4DVAR), and Kalman filter (KF) are the primary assimilation techniques utilized in ocean research (Cressman, 1959; Danard et al., 1968; Lorenc, 1981; Courtier et al., 1994; Evensen, 1994). Many objective analysis datasets, e.g., the EN4 analysis dataset (Good et al., 2013), the global gridded Argo dataset (Zhang et al., 2022), and reanalysis datasets, e.g., the Simple Ocean Data Assimilation (SODA) reanalysis (Carton and Giese, 2008), the Estimating the Circulation and Climate of the Ocean (ECCO) reanalysis, and the Hybrid Coordinate Ocean Model (HYCOM) reanalysis, have been developed using data assimilation methods. However, as the volume, velocity, variety, and veracity of ocean observation data continue to grow, conventional data assimilation and fusion systems are facing increasingly complicated issues (Bauer et al., 2015; Stammer et al., 2016). Existing methods for fusing ocean data always rely heavily on a prior knowledge of linear principles, normal distributions, and appropriate error covariances. This limits their suitability in realistic nonlinear ocean systems, and the resulting fusion accuracy still needs improving. In 3D ocean data assimilation, typical observation profiles are assimilated at each grid point, layer by layer. The generation of spurious high-frequency signals in the vertical direction is one problem, while the huge increase in observation data volume and the large computational cost of the assimilation approach are others. Therefore, scientists are focusing on new methodologies, particularly artificial intelligence (AI), to rapidly fuse different ocean observation datasets.

AI techniques have been extremely successful in the fields of audio, picture, video, and natural language processing because of their ability to fit nonlinear systems and capture high-dimensional features (Hinton and Salakhutdinov, 2006; Kahou et al., 2016; Yu and Deng, 2016; Jiao and Zhao, 2019; Strubell et al., 2019). Scientific data and techniques made possible by advances in AI have aided researchers in the atmospheric and oceanic sciences (Overpeck et al., 2011; Reichstein et al., 2019). There have been many significant advances in ocean research, including wave forecasting (Bento et al., 2021), sea ice forecasting (Andersson et al., 2021), mesoscale eddy identification (Vafaei et al., 2022), subsurface temperature reconstruction (Su et al., 2021), and ENSO prediction (Ham et al., 2019). Multi-sensor sea surface satellite data has been fused using AI techniques in the field of ocean data fusion. To implement wind speed inversion over the ocean, Chu et al. (2020) used a multimodal deep learning approach to combine disparate GNSS-R data. Xiao et al. (2021) presented a genetic algorithm-aided deep neural network model to enhance the SST field's resolution and accuracy. Although the AI model has acquired sufficient accuracy in merging surface satellite data, experts are sometimes suspicious of its results because it is uncertain which factors influence the model's decisions. To the best of our knowledge, few academic institutions have used AI methods to generate the reanalysis data set. Therefore, it is crucial to ensure interpretability in data fusion methods.

In order to overcome the shortcomings of current data fusion and assimilation methods, this paper proposes a novel multi-source ocean data fusion method based on deep learning to achieve intelligent fusion of *in situ* observations, sea surface satellite data, numerical model data, objective analysis data, and reanalysis data. For the objective of integrating sources into common "multidimensional grids", the ODF-Net combines several spatial-temporal scales by applying appropriate transforms to disparate ocean data (Salcedo-Sanz et al., 2020). By using physics-based prior knowledge, vertical profile observations, and gradient information as objective constraints, the model is able to reduce high-frequency spurious signals in the vertical direction of ocean data. The addition of global attention mechanisms (GAM), comprising channel and spatial attention mechanisms, improves both the model's fusion performance and interpretability. Finally, the ODF-Net is utilized to fuse multi-source ocean data from 1994 to 2017 to create a global 0.25°×0.25° monthly 3D sea temperature fusion dataset named ODF-ST dataset.

The remainder of the article is organized as follows. Section 2 introduces all data used in the study, as well as data processing and sample production methods. Section 3 introduces the ODF-Net, including the network structure, attention mechanisms, and objective function design. Section 4 validates the performance and interpretability of the ODF-Net, and evaluates the ODF-ST dataset to verify the practicality of the model. The conclusions and a discussion of future work are provided in Section 5.

2 Data

To undertake intelligent data fusion, we collected ocean data from a range of sources, including *in situ* observations, ocean satellite observations, and 3D gridded data (e.g., numerical model data, objective analysis data, and reanalysis). This study covers the majority of the entire marine domain (180°W–180°E, 60°S–65°N). The proposed strategy was discussed over time (every month from 1994 to 2017) and depth (from the sea surface to 1000 m). The horizontal resolution of the target grid is 0.25°, while the vertical resolution is 23 standard levels (0, 4, 8, 12, 20, 30, 40, 50, 70, 90, 125, 150, 200, 250, 300, 350, 400, 500, 600, 700, 800, 900, and 1000 m). Because different ocean data had diverse spatial and temporal resolutions and distributions, we employed interpolation to ensure consistency in both space and time.

2.1 1D observation profile

In this work, high-precision *in situ* observation profiles acquired from the UK Met Office Hadley Centre's EN4 temperature and salinity profiles dataset version 4.2.1 (subsequently referred to as EN4-profiles) were used as model training labels. The World Ocean Database (WOD), the Arctic Synoptic Basin-wide Observation (ASBO), the Global Temperature and Salinity Profile Program (GTSPP), and the Argo Global Data Assembly Centers (GDACs) provide the fundamental observations in EN4-profiles (Good et al., 2013). EN4-profiles are widely used to evaluate model simulations as " ground truth" (Kumar et al., 2017). We selected high-quality temperature profiles through quality flags.

EN4-profiles have a discontinuous and irregular spatial and temporal distribution and need to be interpolated into the previously mentioned target grid. First, we utilized linear interpolation to interpolate EN4-profiles to 23 standard levels. The processed profiles were then interpolated level by level onto the previously described 0.25° horizontal grid. Because the observation profiles are extremely sparse, a spatial-temporal weighted interpolation method (Zeng and Levy, 1995) was used to increase the number of samples and improve the interpolated data accuracy. For each horizontal objective grid to be interpolated, a spatial-temporal domain with a spatial radius R_s and a temporal radius R_t are specified as the interpolation neighborhood centered on the target grid. The objective grid will be null if there are no observation profiles in the interpolation neighborhood. The monthly temperature $T_{obj,i}$ of the level i of the objective grid is computed as

$$T_{obj,i} = \frac{\sum_{k=1}^{N} w_k T_k}{\sum_{k=1}^{N} w_k} \quad (1)$$

where N represents the total number of observations in the neighborhood, T_k represents the k-th temperature observation in the neighborhood, and w_k represents the interpolation weight of T_k. The w_k is calculated as

$$w_k = \frac{2 - \left[\frac{(x_k - x_0)^2 + (y_k - y_0)^2}{R_s^2} + \frac{(t_k - t_0)^2}{R_t^2}\right]}{2 + \left[\frac{(x_k - x_0)^2 + (y_k - y_0)^2}{R_s^2} + \frac{(t_k - t_0)^2}{R_t^2}\right]} \quad (2)$$

Where x_k, y_k, t_k represent the longitude, latitude, and time corresponding to T_k, x_0, y_0 and t_0 represent the longitude, latitude, and time corresponding to $T_{obj,i}$, respectively. In this work, R_s is 0.48°, R_t is 15 days, and t_0 is the 16th day of each month.

2.2 2D sea surface datasets and 3D gridded datasets

To yield sea surface information for ocean data fusion, multi-source ocean satellite observation and analysis data were used as model training input. The surface variables include sea surface temperature (SST), sea level anomaly (SLA), and sea surface wind (SSW). Three satellite SST analysis datasets were collected, including NOAA's Optimum Interpolation Sea Surface Temperature (OISST, version 2) (Reynolds et al., 2007), the Extended Reconstructed Sea Surface Temperature (ERSST, version 5) (Huang et al., 2017), and the Hadley Centre Global Sea Ice and Sea Surface Temperature (HadISST) (Rayner et al., 2003). The OISST data is daily with a horizontal resolution of 0.25°, the ERSST data is monthly with a horizontal resolution of 2°, and the HadiSST data is monthly with a horizontal resolution of 1°. The satellite SLA is a daily Aviso-SLA (version 4.0) dataset from Copernicus Marine Environment Monitoring Service (CMEMS) with a horizontal resolution of 0.25°. NASA's monthly Cross-Calibrated Multi-Platform (CCMP, Version 2) wind data with a horizontal resolution of 0.25° is provided by the satellite SSW (Atlas et al., 2011). All sea surface data were collected between 1994 and 2017.

To generate subsurface information for ocean data fusion, numerical model data, objective analysis data, and reanalysis data were used. The addition of numerical model data and reanalysis data increased the physical rationality of the 3D ODF-ST dataset. Monthly historical simulation data (r1i1p1f1) from NCAR's CESM2 Earth system model (Danabasoglu et al., 2020) and monthly historical simulation data (r1i1p1f1) from CMA's BCC-CSM2-HR climate system model (Wu et al., 2021) were included in numerical model data. The Hadley Center's EN4 monthly objective analysis data (version 4.2.1, subsequently referred to as EN4-analysis) is used for the objective analysis, which has a horizontal resolution of 1°. (Good et al., 2013). SODA (version 3) monthly ocean reanalysis data from the University of Maryland (Carton et al., 2018), ECCO (version 4) monthly reanalysis dataset with a horizontal resolution of 0.5° from NASA (Forget et al., 2015), and HYCOM daily reanalysis data with a horizontal resolution of 1/12° from the US Naval Research Laboratory are the reanalysis datasets used (Chassignet et al., 2007). All of these 3D gridded datasets were collected between 1994 and 2017.

To produce the monthly average dataset for sea surface data, daily OISST and AVSIO SLA analysis data were averaged. Using bilinear interpolation, all sea surface data were uniformly interpolated to the previously described 0.25° horizontal grid. The daily HYCOM reanalysis data for 3D gridded data were averaged to generate a monthly average dataset. All 3D gridded data were linearly interpolated to vertical standard levels before being uniformly interpolated to the previously defined horizontal grid with 0.25° resolution using a bilinear interpolation method.

3 Methodology

We propose a data-level fusion architecture, depicted in Figure 1, to accommodate the multi-dimensionality and heterogeneity inherent in ocean data gathered from various sources. We first transform, align, and organize heterogeneous data such as multi-source ocean data and spatiotemporal information into regular samples, and then build a deep learning-based data fusion model to automatically extract multi-level features from the samples and finally generate fused data. One benefit of a data-level fusion architecture is that it allows for the fusion of data from multiple sources while utilizing a single network. Data-level fusion is superior to feature-level and decision-level fusion methods in terms of reducing the number of model parameters(Kopuklu et al., 2018). Furthermore, as model fusion takes place at the data level, the correspondence between different datasets can be automatically extracted.

In order to organize data from multiple sources into samples for model training, we produced samples based on the observation profiles collected in section 2.1, where one observation profile is equivalent to one sample. To better incorporate prior knowledge, such as the vertical structure of sea temperature, into the model and to suppress the spurious high-frequency signal of fusion data in the vertical direction, we chose to use the observation profile interpolated to standard layers as the label, as shown in Figure 1. Therefore, a label has a dimension of 1×D, where D is the number of vertical standard levels (23 in this work). The deepest effective sea temperatures were used to fill the missing data produced by seafloor terrain, and the filled data were not taken into account in the loss function. We reserved 10% of all samples as the test set for validation of model performance.

The dimension of sample features is C×D, where C=M+1 is the number of channels in the input layer of the fusion model, M is the number of 3D gridded datasets, and M=6 in the current study. As shown in Figure 1, the first six channels are vertical profiles of EN4-analysis, HYCOM, SODA, ECCO, CESM2, and BCC-CSM2-HR that correspond to the sample label in the temporal and spatial dimensions. The last channel provides the spatial data (latitude and longitude), the temporal data (year and month), and the sea surface data (ERSST, OISST, HadiSST, AVSIO SLA, and CCMP) corresponding to the sample label, with the remaining locations filled with zeros.

3.1 Model structures of the ODF-Net

The ODF-Net, shown in Figure 2A, is a variant of the 1D-ResNet model. There are three distinct sections to this design. To accomplish the task of extracting shallow features from multi-source ocean data, block A contains a 1D convolutional layer that uses a 3×1 kernel in addition to two GAMs. Blocks B and C, depicted in Figures 2B, C, respectively, are composed mostly of 1D convolutional layers with a 3×1 kernel, dropout layers (dropout probabilities of 0.5), and skip-connection structures to harvest the deep features. Block D is a decoding block that fully integrates sea temperature data from many sources through the use of a combination of shallow and deep features.

In the ODF-Net, we choose to use the more advanced Adaptively Parametric Rectifier Linear Unit (APReLU) activation function rather than the more common ReLU activation function used in the 1D-ResNet. When the original features are less than zero, APReLU runs each sample through a small fully connected network to produce matching weights, which are then used as coefficients of the original features to provide a more flexible method of nonlinear transformation (Zhao et al., 2020).

3.2 GAMs in the ODF-Net

Improving the model's interpretability assists with both understanding the deep learning model's complex decision-making foundation and guaranteeing the model's reliability (Xu et al., 2021). Giving neural networks an attention mechanism improves their ability to learn by focusing on the relevant important feature while discarding the rest. In order to emphasize the interaction of multiple sources of information at different depth levels and to enable the model to capture important features in both dimensions, we modify the GAM (Liu et al., 2021) and redesign the sub-modules. Figure 3A

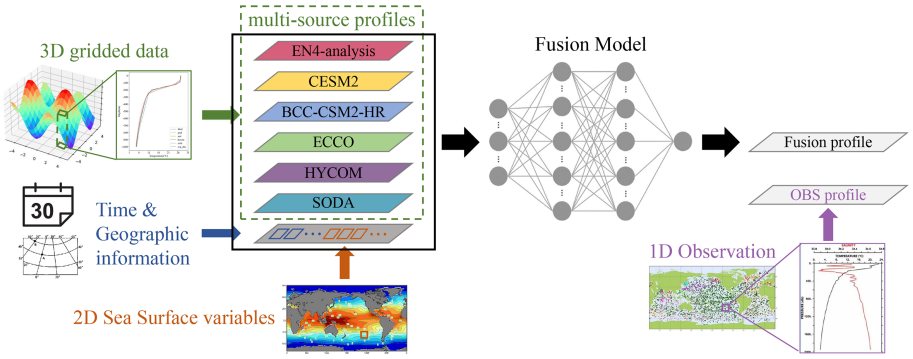


FIGURE 1
Framework for the intelligent fusion of multi-source heterogeneous ocean data.

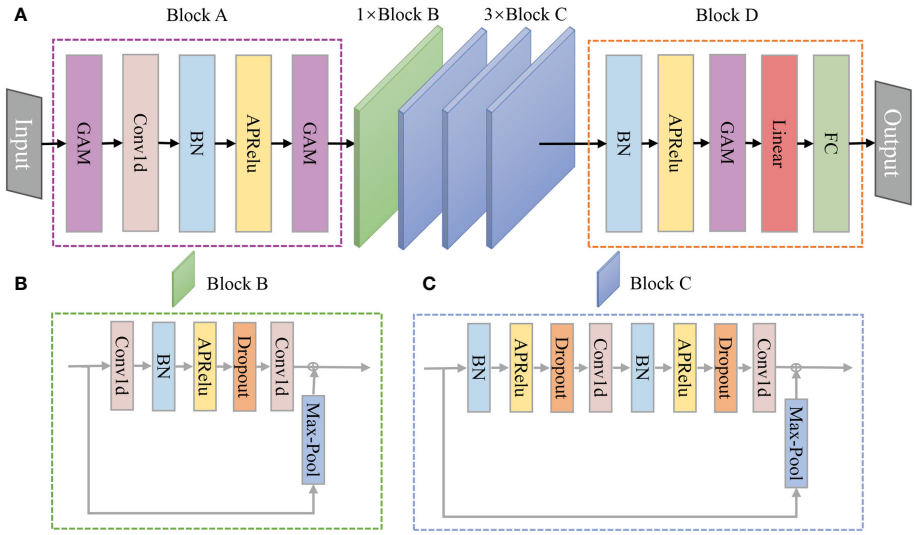


FIGURE 2
Structures of **(A)** the ODF-Net, **(B)** Block B, and **(C)** Block C.

depicts the GAM's overall attention mechanism process, which sequentially combines channel attention and spatial attention. Given input features F_1 the intermediate state F_2 and output F_3 are defined as

$$\mathbf{F}_2 = \mathbf{M}_\mathbf{C}(\mathbf{F}_1) \otimes \mathbf{F}_1 \tag{3}$$

$$\mathbf{F}_3 = \mathbf{M}_\mathbf{S}(\mathbf{F}_2) \otimes \mathbf{F}_2 \tag{4}$$

where M_c and M_s are the channel attention and spatial attention weights, respectively, indicating the model's degree of attention to distinct channels and depth levels of input features. ⊗ denotes the multiplication operation by element. For the first GAM, in particular, channel attention reflects the importance of different sources, while spatial attention reflects the model's concentration on diverse physical depth levels.

A major part of the channel attention submodule is shown in Figure 3B. When extracting information in two dimensions, 2D permutation is employed, and then a two-layer neuron network is used to magnify the dependence between channels and depth levels across dimensions. Each channel's weights are then calculated using the sigmoid function. Figure 3C depicts the spatial attention submodule, which uses two convolutional layers to aggregate information from different depths and a sigmoid function to determine the weights of each depth.

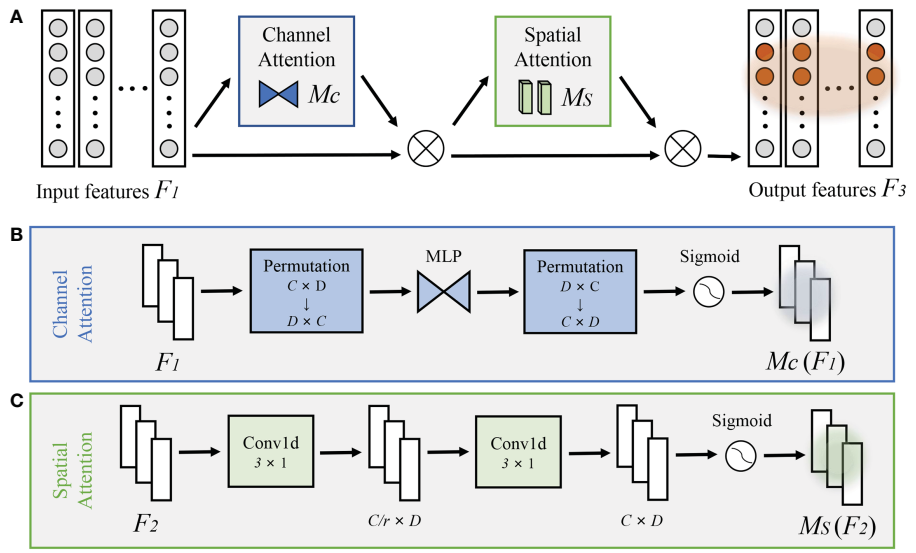


FIGURE 3
(A) The overview of redesigned GAM. The structures of **(B)** channel attention submodule, and **(C)** spatial attention submodule.

3.3 Design of loss function

In order to integrate prior knowledge such as the vertical structure of sea temperature into the model and reduce the spurious high-frequency signal of fusion data in the vertical direction, the vertical gradient and integral of the sea temperature profile were added to the loss function, which is calculated as

$$\mathbf{Loss}_1 = \mathbf{RMSE} + \alpha \cdot \mathbf{Loss}_{\mathbf{ST_Grad}} + \beta \cdot \mathbf{Loss}_{\mathbf{Cumsum}} \quad (5)$$

where $\mathrm{Loss}_{\mathrm{ST_Grad}}$ is the sea temperature gradient constraint, $\mathrm{Loss}_{\mathrm{Cumsum}}$ is the sea temperature profile integral constraint. α and β are hyperparameters, which were determined as 0.004 and 0.02 respectively through comparison experiments. The $\mathrm{Loss}_{\mathrm{ST_Grad}}$ and $\mathrm{Loss}_{\mathrm{Cumsum}}$ are defined as

$$\mathbf{Loss}_{\mathbf{ST_Grad}} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\mathbf{ST_Grad}_{\mathbf{label},\,i} - \mathbf{ST_Grad}_{\mathbf{pred},\,i})^2}, \quad (6)$$

$$\mathbf{ST_Grad}_k = \frac{T_{k+1} - T_k}{Z_{k+1} - Z_k}$$

$$\mathbf{Loss}_{\mathbf{Cumsum}} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\mathbf{Cumsum}_{\mathbf{label},\,i} - \mathbf{Cumsum}_{\mathbf{pred},\,i})^2}, \quad (7)$$

$$\mathbf{Cumsum} = \sum_{k=1}^{M} T_k$$

where ST_Grad and Cumsum represent the vertical gradient and the vertical integral of the sea temperature profile, respectively. T_i denotes the sea temperature value of the i-th level and Z_i denotes the vertical depth of the i-th level.

Meanwhile, to determine whether the loss function with gradient and integral constraints may improve the model's fusion performance, a comparative experiment was run with the loss function Loss_2=RMSE, and the experimental results are provided in Table 1. R^2 is the coefficient of determination, a value ranging from 0 to 1 that indicates how effectively a statistical model predicts an outcome. The closer a model's R^2 is to 1, the better it is at making predictions. Equation 8 gives the calculation of R^2.

$$R^2 = \frac{\sum_{i=1}^{N}(f_i - \bar{y})^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \quad (8)$$

where f_i and y_i represent the i-th prediction and label, respectively. $\bar{y}$ denotes the averaged value of all labels. The model with Loss_1 has nearly the same root mean square error (RMSE) as the model with Loss_2 but the sea temperature profile gradient error (STPGE) is reduced by 4%, indicating that the addition of physics-based prior knowledge constraints in the loss function has an enhancement effect on the vertical structure of the fusion sea temperature.

3.4 Ablation studies

Four sets of experiments (Exp0–Exp3) were designed to validate the favorable impacts of GAM and APReLU on the fusion model. Exp0 is the baseline, which does not include GAM or APReLU. Exp1 includes GAM, and Exp2 includes APReLU. Exp3 includes both GAM and APReLU, i.e., the ODF-Net. In addition, we designed Exp4 to do the same fusion task using Transformer (Vaswani et al., 2017), a state-of-the-art model for sequence-to-sequence learning, in order to validate the ODF-Net's performance in comparison to other models. The transformer's hyperparameters were tuned, and the encode dimension, number of attention heads, number of identical layers, query vector length, and key vector length were all set to 128, 8, 6, 16, and 16, respectively.

Table 2 shows a comparison of test results from the five sets of experiments. Comparing Exp1 and Exp0, GAM reduces RMSE by 2.87% and STPGE by 4.28%; comparing Exp2 and Exp0, APReLU reduces RMSE by 4.33% and STPGE by 5.32%; comparing Exp3 and Exp0, GAM and APReLU reduce RMSE by 6.98% and STPGE by 6.86%. As a result, GAM and APReLU considerably increase model performance. The comparison of Exp4 and Exp3 shows that the ODF-Net outperforms the Transformer in all metrics, including RMSE, STPGE, and R^2. The quantities of trainable parameters are listed in the third column of Table 2 with M representing 10^6, the higher the value, the more complicated the model. The amount of trainable parameters in the ODF-Net is only about 60% of the Transformer, demonstrating that the ODF-Net we developed is lightweight and high-performance.

4 Results and discussion

This section begins with an evaluation of the ODF-Net's performance using *in situ* observations from the reserved test set. The model's interpretability was then examined by collecting the attention weights of the first GAM in the ODF-Net. Finally, the global sea temperature dataset developed by the ODF-Net was examined to verify the method's practicality.

4.1 Performance of the ODF-Net

To verify the accuracy of the fusion model, we compared the RMSE, STPGE, and R^2 of all eight data sources, including ODF-Net

TABLE 1 Comparison of metrics on test set using different loss functions.

Loss function	RMSE(°C)	STPGE(°C/m)	R^2
Loss_1	0.74	0.0418	0.992
Loss_2	0.73	0.0435	0.988

TABLE 2 Comparison of metrics on test set using different model structures.

	Model structure	Parameters	RMSE(°C)	STPGE(°C/m)	R^2
Exp0	Baseline	0.43M	0.79	0.0436	0.9911
Exp1	Baseline+GAM	0.47M	0.77	0.0429	0.9915
Exp2	Baseline+APReLU	0.95M	0.76	0.0425	0.9918
Exp3	**Baseline+GAM+APReLU**	**0.98M**	**0.74**	**0.0418**	**0.9922**
Exp4	Transformer	1.59M	0.77	0.0438	0.9912

The bold values are generated by our model (ODF-Net).

predictions, six sets of fused data, and their ensemble average data (EAD), with *in situ* observations. Table 3 summarizes the accuracy of different datasets over the test set using the metrics mentioned above. As Table 3 suggests, ODF-Net predictions are optimal on all evaluation metrics. ODF-Net predictions have an average RMSE of 0.74°C, while other data sources have RMSEs greater than 0.9°C. ODF-Net predictions have an average STPGE of 0.042°C/m, while other data sources have average STPGEs greater than 0.046°C/m. The ODF-Net improves both the accuracy and the vertical structure of the fusion sea temperature.

The spatial distribution of RMSE is depicted in Figures 4A, B. ODF-Net predictions are more accurate than EAD's in almost all regions, especially in areas with large gradients such as the Gulf Stream, the Kuroshio Extension, and the West Wind Drift, where the improvement is noticeable. Figure 4C shows the spatial distribution of the percentage improvement in R^2 of ODF-Net predictions with observation profiles versus R^2 of EAD predictions with observation profiles. Statistical examination of the test set reveals that R^2 of ODF-Net predictions is greater than EAD for 78.51% of the profiles. The improvement is more significant in areas with large gradients, such as the boundary current regions, which are similar to the spatial distribution of RMSE, suggesting that the deep learning model learns more correct information from multiple datasets.

The distribution and variation of sea temperatures in different ocean areas and depth levels exhibit distinct characteristics due to the effects of several factors, such as solar radiation, land-sea distribution, ocean currents, and monsoons (Chen et al., 2002; Li et al., 2020). To examine the fusion effect, we evaluated the accuracy of ODF-Net predictions at different depth levels by region. The global ocean was divided into five oceans, which are the Pacific Ocean, the Atlantic Ocean, the Indian Ocean, the Arctic Ocean, and the Southern Ocean. ODF-Net performance in the five oceans and global regions is then discussed.

Here, we examined the time average RMSEs of eight data sources at different depth levels in the five oceans and global regions, including CESM2 (green line), BCC-CSM2-HR (red line), ECCO (purple line), SODA (pink line), EN4-analysis (orange line), HYCOM (brown line), EAD (gray line), and ODF-Net predictions (blue line) (Figure 5). The time average RMSEs in the vertical direction of eight data sources varied in each of the five oceans, but the RMSE of ODF-Net predictions at different depth levels is significantly lower than that of other data sets in each ocean as well as the global region. This indicates that the proposed ODF-Net performs better on a global scale. The accuracy improvement in ODF-Net predictions above the thermocline is greater than in other deeper layers, particularly at the thermocline with the largest vertical gradient. This might be attributed to the large number of thermocline observations, which enables the ODF-Net to learn the bias between other data sources and observations.

The Taylor diagram incorporates numerous assessment measures that are commonly used to evaluate model performance, including the correlation coefficient (COEF), root mean square error (RMSE), and standard deviation (STD) (Taylor, 2001). Taylor diagrams were utilized to more thoroughly and objectively analyze the statistical connections between various data points and observations in this work. Taylor diagrams of sea temperatures in the five oceans and global regions based on nine data sources, including projections (black dots) and observations (red dots) from ODF-Net are shown in Figure 6. Figure 6 shows that the results from several datasets vary widely, while the predictions generated by ODF-Net consistently and

TABLE 3 Comparison of metrics on test set of different data sources.

	RMSE(°C)	STPGE(°C/m)	R^2
EN4-analysis	0.90	0.046	0.989
CESM2	2.17	0.058	0.933
BCC-CSM2-HR	2.48	0.057	0.913
ECCO	2.33	0.060	0.923
HYCOM	1.19	0.057	0.980
SODA	1.01	0.048	0.986
EAD	1.21	0.049	0.979
ODF-Net predictions	**0.74**	**0.042**	**0.992**

The bold values are generated by our model (ODF-Net).

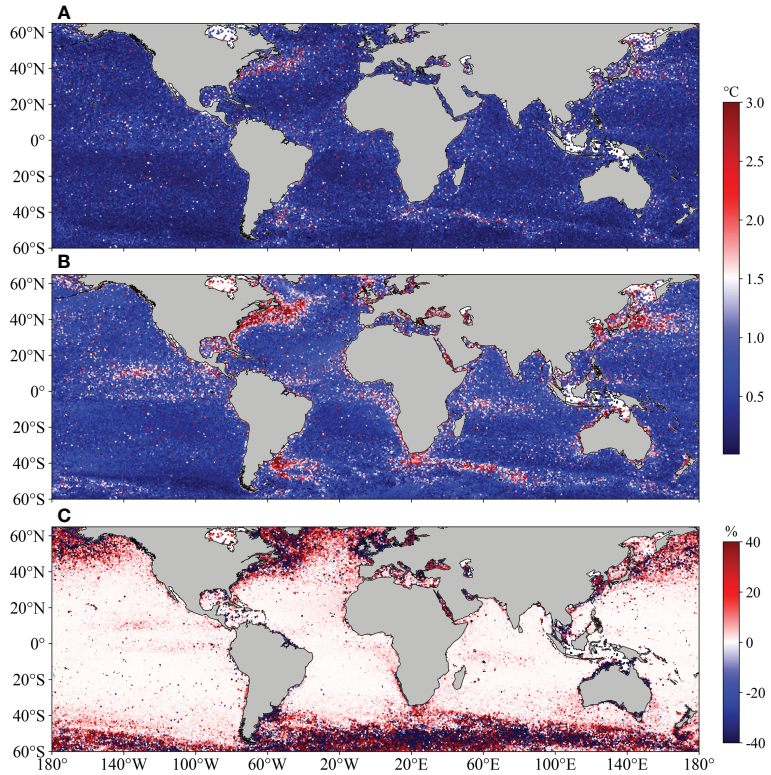


FIGURE 4
Spatial distribution of RMSE between **(A)** the ODF-Net, **(B)** EAD and EN4-profiles observations; **(C)** Spatial distribution of increase percentage in R^2 of the ODF-Net relative to EAD.

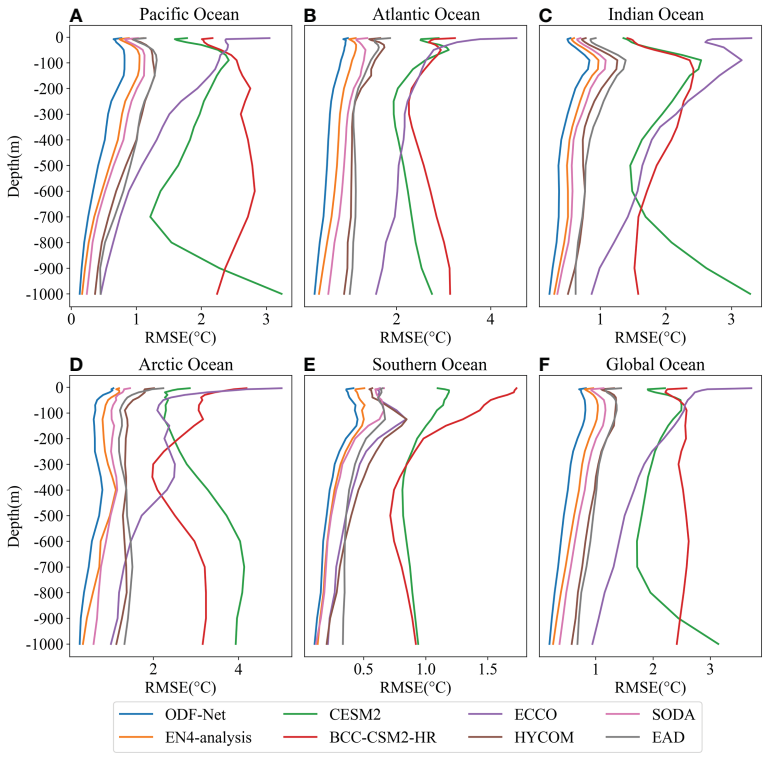


FIGURE 5
Vertical distribution of time-averaged RMSEs in **(A)** Pacific Ocean, **(B)** Atlantic Ocean, **(C)** Indian Ocean, **(D)** Arctic Ocean, **(E)** Southern Ocean, and **(F)** Global Ocean.

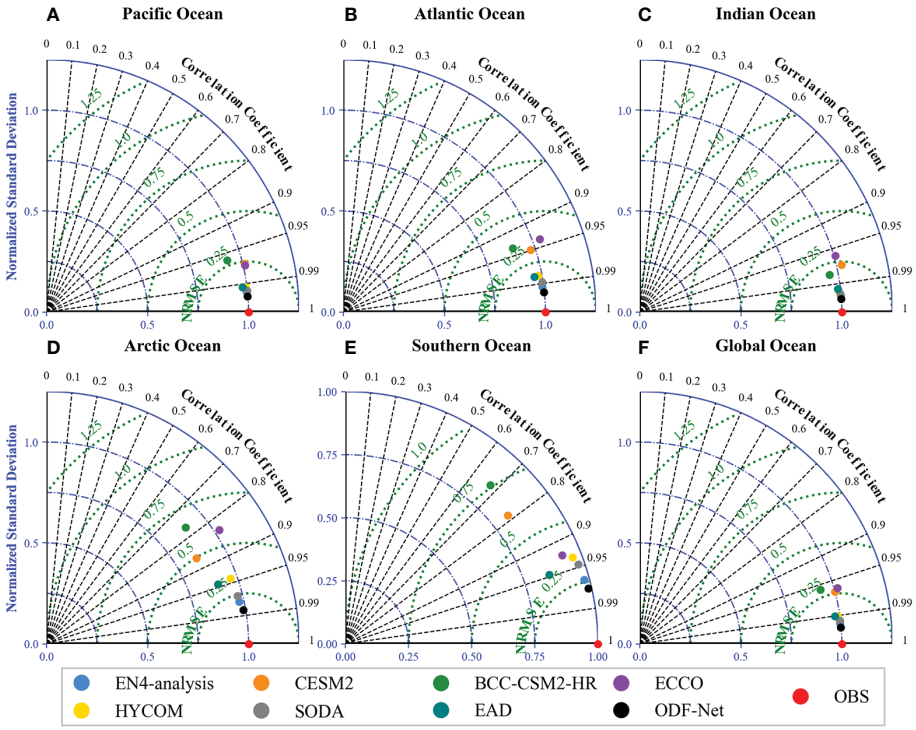


FIGURE 6
Time-averaged Taylor diagram in **(A)** Pacific Ocean, **(B)** Atlantic Ocean, **(C)** Indian Ocean, **(D)** Arctic Ocean, **(E)** Southern Ocean, and **(F)** Global Ocean.

comprehensively outperform those generated by any other dataset. ODF-Net predictions are unbiased with a high level of correlation. The ODF-Net also has different outcomes depending on location. The performance of the ODF-Net in the Pacific, Atlantic, and Indian Oceans was comparable to that of the global region (Figures 6A–C). Figures 6D, E illustrate that the performance of the ODF-Net model in the Northern and Southern Oceans still needs to be improved. The main reason is that the performance of ODF-Net is highly dependent on high-quality observations, which are much less available in the Arctic and Southern Oceans than in other regions.

4.2 Interpretability of the ODF-Net

Attention weights, as an intermediate output of the network model, can be used as a convenient tool to explain model decisions. Many studies have discussed the model interpretation ability of the attention weight distribution for neural network models based on attention mechanisms (Pruthi et al., 2019; Serrano and Smith, 2019; Wiegreffe and Pinter, 2019). In this study, we utilized attention weight distribution to analyze the contribution of multi-source ocean data as well as spatiotemporal information to the ODF-Net fusion process. Since multi-source data were sent directly to the first GAM in the ODF-Net, the attention weights could represent the contributions of the original ocean data as well as of the spatiotemporal information. We obtained the channel attention weight M_c and the spatial attention weight M_s of the first GAM, and since the GAM used a combination of channel and spatial attention serially, the global attention weight M_{global} is defined as:

$$\mathbf{M_{global}} = \mathbf{M_C} \otimes \mathbf{M_S} \tag{9}$$

Figure 7 shows a heat map of the global attention weight M_{global}. The top six rows illustrate the contribution of each of the 3D gridded datasets to the final fusion task, while each column represents the contribution of ocean data at various depth levels. As shown in Figure 7, the top three 3D gridded data contributors, in order, are EN4-analysis, SODA, and HYCOM, whereas ECCO, CESM2, and BCC-CMS2-HR contribute relatively little. The distribution of attention weights is rather reasonable, as the error between EN4-analysis and observed EN4-profiles is the smallest on the test set, followed by SODA and HYCOM, whereas the average errors are larger for ECCO, CESM2, and BCC-CMS2-HR. This implies that the ODF-Net has given more attention to high-precision data. In the spatial dimension, the attention weights of the shallow levels above 100 m are greater than those of the deeper levels below 100 m for EN4-analysis and SODA, the two datasets that received the most attention, and for other datasets, the weights of the deep levels are greater than those of the shallow levels. This is due to the fact that sea temperatures are more stable at deeper levels, and 3D gridded datasets at deeper levels are more accurate than those at shallow levels. Due to the extremely high correlation of sea temperature variations with latitude, latitude has the greatest weight in the last channel. In contrast, the attention weights of temporal information and sea surface data are not significantly different. The analysis of attention weights shows that the ODF-Net pays more attention to the data sources that are more accurate. At the same time, information from different depth levels of the same data source makes different contributions to the fusion process.

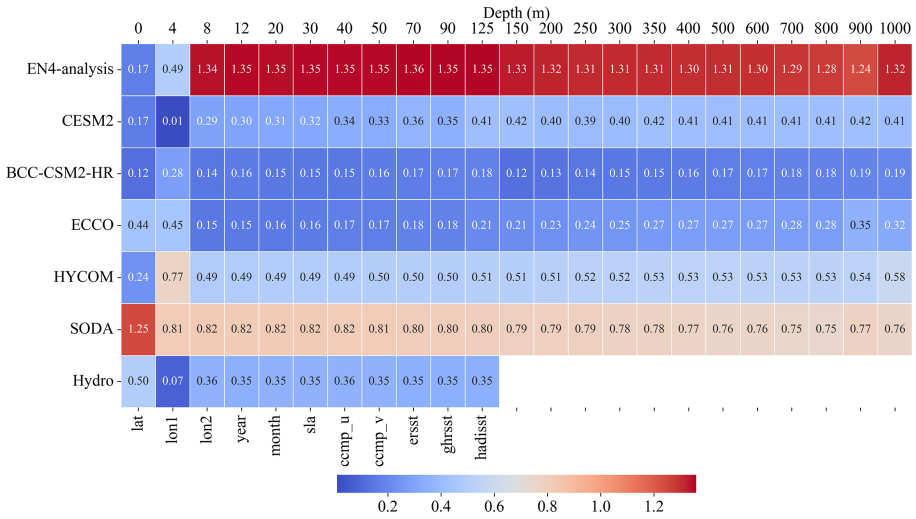


FIGURE 7
Heat map of GAM attention weights.

4.3 Evaluation of ODF-ST dataset

Since the purpose of this study is to accomplish accuracy, high resolution, and spatiotemporal continuous intelligent fusion of multi-dimensional, multi-source heterogeneous ocean data, we used the ODF-Net to generate a 3D sea temperature fusion dataset (ODF-ST dataset). ODF-ST dataset spans the years 1994 to 2017, with a spatial extent of the global ocean (180°W–180°E, 60°S–65°N), a monthly temporal resolution, and a spatial resolution of 0.25°. To assess the spatial rationality of the ODF-ST dataset, we compared the fusion SST with that of OISST, the spatial distribution of fusion sea temperature profiles, and sea temperatures at different levels with WOA18 (Boyer et al., 2018). To assess the temporal rationality of the ODF-ST dataset, we compared the fusion sea temperature profiles with Tropical Atmosphere Ocean Array (TAO) monthly observation profiles. Finally, we evaluated the ENSO index time series calculated with fusion sea temperature.

The global climatic SSTs (averaged from 1994 to 2017) of the ODF-ST dataset and the OISST are shown in Figure 8, where both SSTs have a "low-high-low" distribution from north to south, with a notable high-value area in the mid-western Pacific Ocean and a center SST of nearly 29°C. The transition of 25°C isotherms in the east-central Pacific Ocean (red frame area) is similar. The differences in SSTs are in the range of 0.5°C in most regions, indicating that the distribution of ODF-ST SST is acceptable and trustworthy. ODF-ST SST is significantly higher than OISST in Hudson Bay, the Mediterranean Sea, the southwestern coast of Africa, and the western Okhotsk Sea, but markedly lower than OISST in the northwestern and southwestern Atlantic Ocean and the southeastern Okhotsk Sea, possibly due to more complex changes in nearshore currents and sparse observations.

The distribution of global climatic sea temperatures (averaged from 1995 to 2017) between the ODF-ST dataset and WOA18 at 500 m (Figures 9A–C) and 900 m (Figures 9D–F) shows that the ODF-ST sea temperature is very similar to WOA18. From north to south, both 500 m sea temperatures exhibit a "low-high-low-high-low" distribution, with noticeable high-value areas in the northwest Atlantic Ocean, the Mediterranean Sea, the southwest Indian Ocean, and the northwest Pacific Ocean, with the center sea temperature of the high-value area in the northwest Atlantic Ocean being around 17°C. Both 900 m sea temperatures have obvious high-value areas in the mid-eastern Atlantic Ocean, the Mediterranean Sea, and the Gulf of Aden, with the center sea temperature of the high-value area in the Mediterranean Sea being around 15°C. The isotherm trends are also strikingly similar, with homologous transitions of 500 m sea temperatures in the northern Atlantic Ocean and the southern Indian Ocean(red frame area). The differences between 500 m and 900 m sea temperatures of the ODF-ST dataset and WOA18 in most regions are lower than 0.25°C. Differences in 500 m sea temperatures have noticeable positive and negative oscillations in the Gulf Stream, the Kuroshio Extension, the North Pacific Current, and the West Wind Drift. Differences in 900 m sea temperatures have noticeable positive and negative oscillations in the Nansha Islands and the West Wind Drift.

To evaluate the rationality of ODF-ST sea temperature in the vertical direction, we averaged sea temperatures from the ODF-ST dataset and WOA18 in the longitudinal and latitudinal directions and then compared their sea temperature profiles along the latitudinal (Figures 10A–C) and longitudinal (Figures 10D–F) directions. The profile along the latitudinal direction demonstrates that the ODF-ST dataset and WOA18 both have a high sea temperature value of 28°C in-depth levels over 100m near 5°N. The isotherms exhibit evident grooves in the northern and southern hemispheres' mid-latitudes. The difference in sea temperature is mostly less than 0.25°C. ODF-ST sea temperature is about 0.5°C higher than WOA18 in depth over 80 m in the Northern Hemisphere.

The profile along the longitudinal direction shows that the ODF-ST dataset's isotherm change is essentially consistent with that of WOA18. Both have three regions with strong gradient variations located at 60°W–80°W, 0°–50°E, and 110°E–150°E, respectively. The differences in sea temperature are also higher in these regions, while

10.3389/fmars.2023.1112065

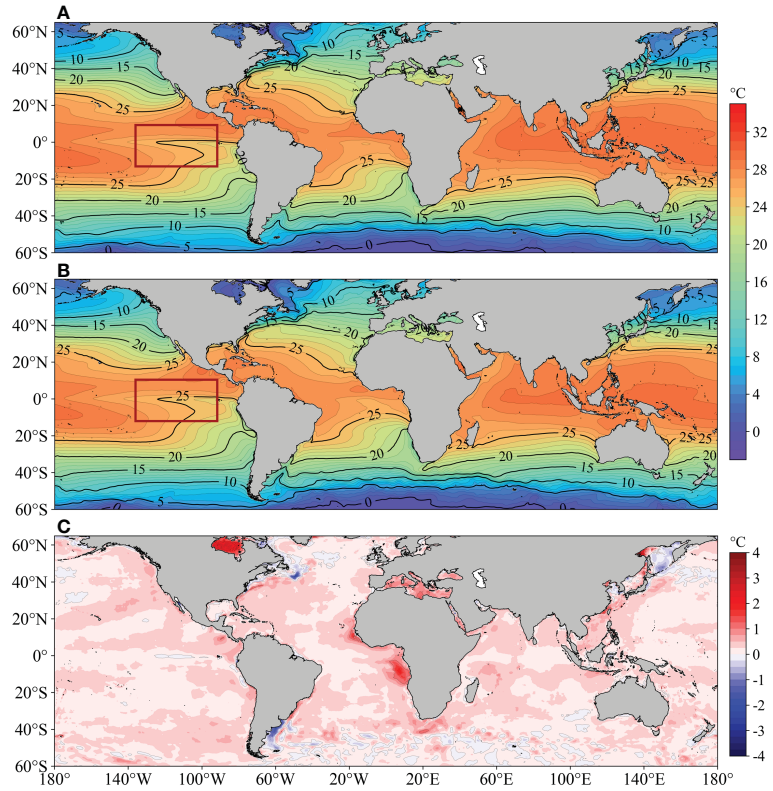


FIGURE 8
Global climatic (1994~2017) SST distribution of **(A)** ODF-ST dataset, **(B)** OISST, and **(C)** their difference.

differences in other regions are less than 0.25°C. The highly consistent ODF-ST sea temperature with WOA18 in the vertical direction implies that the spatial distribution of ODF-ST sea temperature is reasonable.

We used monthly TAO observation profiles from 1994 to 2017 to conduct a comparative analysis of temporal correlation in order to assess the temporal rationality of the ODF-ST sea temperature. Figure 11 depicts, for each observation site, the spatial distribution

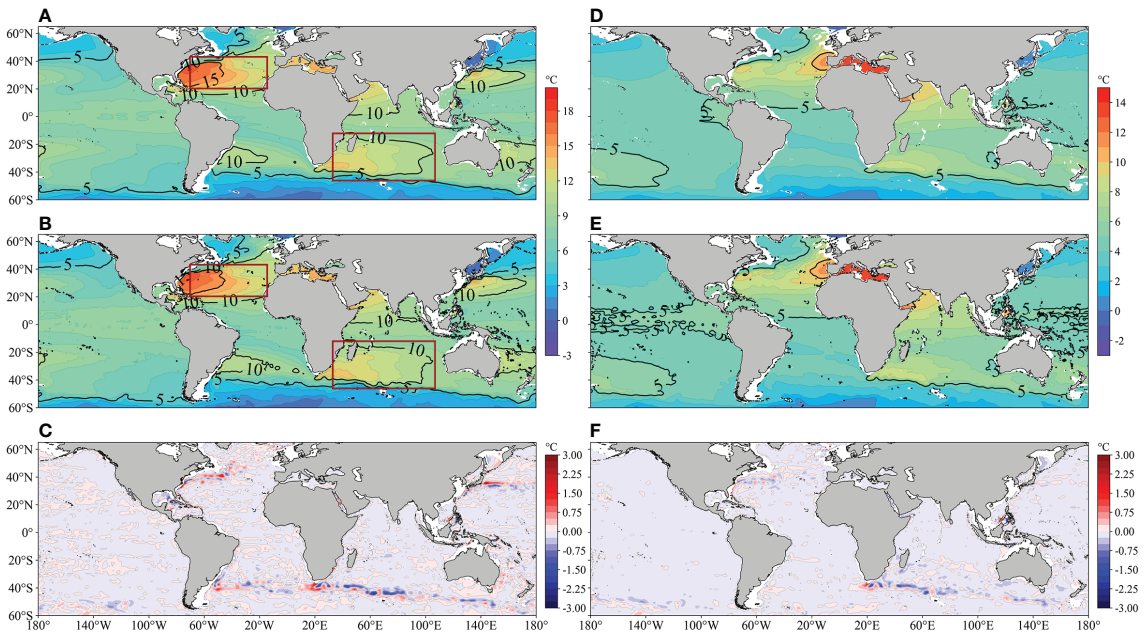


FIGURE 9
Global climatic (1995~2017) 500m sea temperature distribution of **(A)** ODF-ST dataset, **(B)** WOA18, and **(C)** their difference; Global climatic (1995~2017) 900m sea temperature distribution of **(D)** ODF-ST dataset, **(E)** WOA18, and **(F)** their difference.

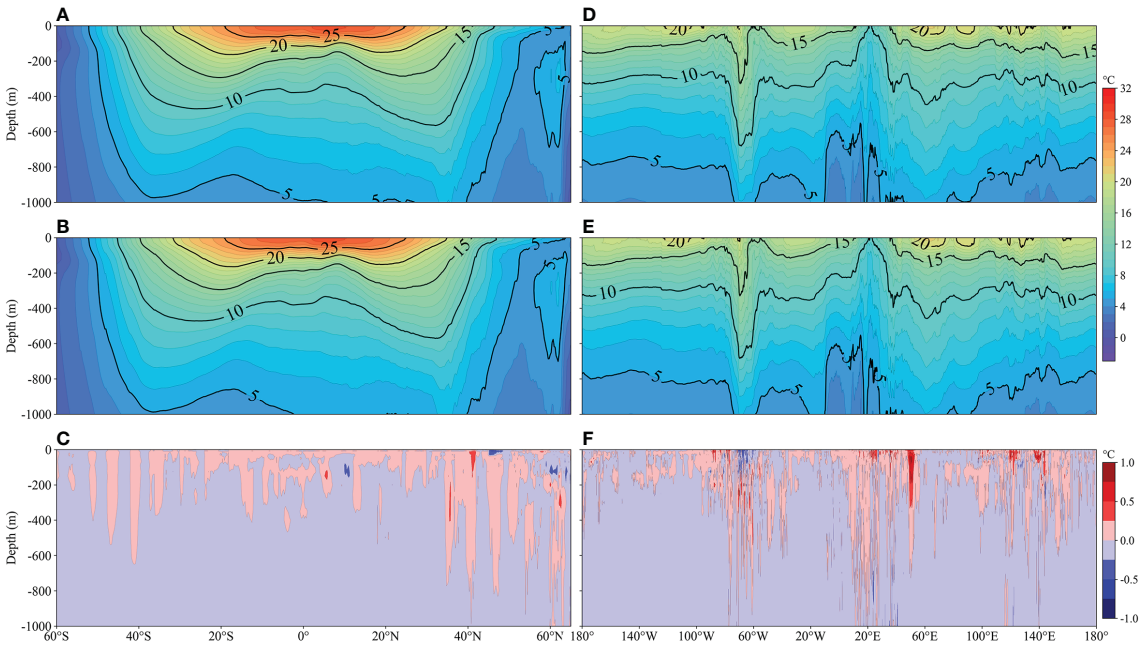


FIGURE 10
Longitudinal averaged climatic (1995~2017) sea temperature profiles along latitudinal direction of **(A)** ODF-ST dataset, **(B)** WOA18, **(C)** their difference and longitudinal direction of **(D)** ODF-ST dataset, **(E)** WOA18, and **(F)** their difference.

of the temporal correlation between the ODF-ST sea temperature and TAO observation profiles.The Pearson correlation coefficient has a range of 0.9755 to 0.9995 and a mean value of 0.996.Therefore, in the sea area where the TAO array is deployed, the temporal distribution of ODF-ST sea temperature is reasonable.

To verify the ODF-ST dataset's ability to capture the ENSO signal, Figure 12 illustrates the ENSO index time series in the Nino3.4 region (5°S–5°N, 170°W–120°W) of multiple sources from January 1994 to December 2017. Variations of the ENSO index in the ODF-ST dataset are generally consistent with those of ERSST and HadiSST and can reflect the significant El Niño years (1995, 1998, 2003, 2007, 2010, 2015) and La Niña years (1999, 2000, 2008, 2011, 2012, 2017).

We compared the Pearson correlation coefficient and RMSE of the anomalies in the Nino3.4 region of average sea temperatures above 100 m from different sources with the ENSO index provided by the U.S. Climate Prediction Center (NOAA/CPC) and discovered that the Pearson correlation coefficients of the ODF-ST dataset, HYCOM, ECCO, and SODA were all above 0.9 or higher, and the RMSEs of them were all below 0.5°C, indicating that the ODF-ST dataset's sea temperature field could fairly reflect the ENSO signal.

5 Conclusion and future work

This paper presents an ODF-Net model for fusing ocean data from multiple sources, including 1D observation profiles, 2D sea surface datasets, and 3D gridded datasets. This approach is distinguished by its precision, speed, and interpretability. Instead of performing level-by-level, single-point ocean data assimilation, the vertical profile of the ocean is employed as the objective constraint.

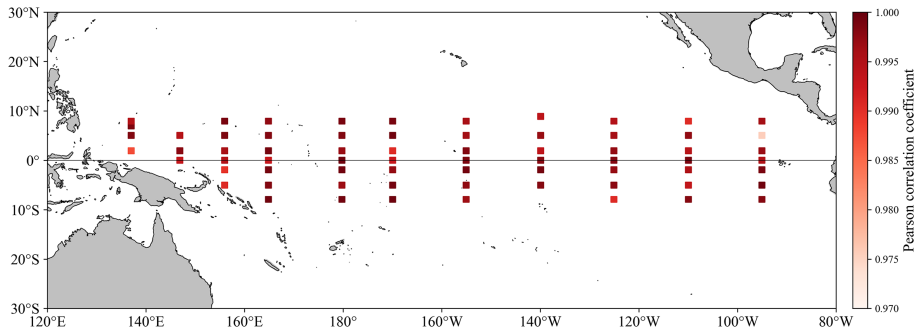


FIGURE 11
Spatial distribution of the temporal Pearson correlation coefficient between the ODF-ST sea temperature and TAO observation profiles.

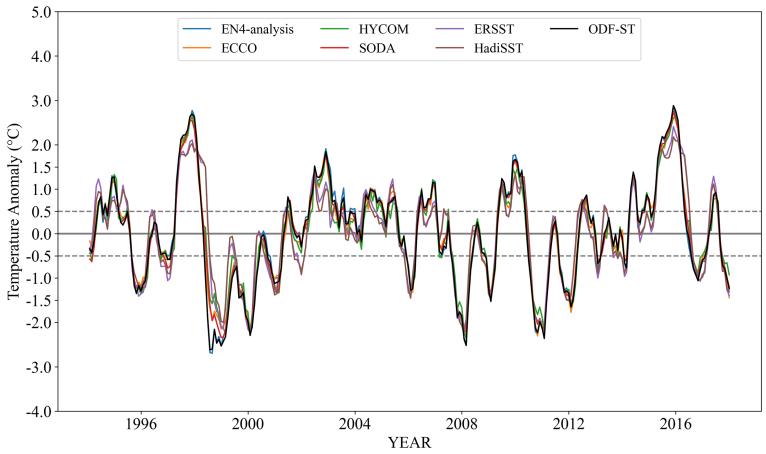


FIGURE 12
ENSO index time series in the Nino3.4 region of various sources.

This enables us to incorporate physics-based previous knowledge and eliminate vertical high-frequency spurious signals. Global attention mechanisms are intended to guide ODF-net to crucial features from a diverse number of data sources and depth levels. The ODF-Net fusion sea temperature has a lower RMSE (0.74°C), a lower STPGE (0.042°C/m), and a higher R^2 than the fused data sources and EAD (0.99). A heat map of global attention weights was utilized to demonstrate the interpretability of the model. The ODF-Net assigned different weights to various characteristics of the datasets.

The most significant outcome of this study is a novel approach and paradigm for solving the age-old problem of integrating data from various, divergent ocean sources into a single whole. Nevertheless, the existing ODF-Net has only combined and investigated sea temperature; we will expand to include more factors of the marine environment. By adding new factors and examining their influence on the fusion outcomes, it is possible to further improve the fusion performance of the model. Moreover, the single-moment and single-profile data could be substituted with time-series ocean element fields in a realistic geographical region as fusion factors.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

MW and XH conceived and designed the research methodology. YX collected and pre-processed data. DW and MW conducted software, model validation and visualization. MW, DW and XH wrote and revised the manuscript. YL, RX, FX and JY gave important advice in the research and thesis writing process. All authors contributed to the article and approved the submitted version.

Funding

This work is supported by the National Key Research and Development Program of China (2021YFC3101600,2020YFA0607900, 2020YFA0608000) and the National Natural Science Foundation of China (42125503, 42075137).

Conflict of interest

Authors MW, DW and LY are employed by Ninecosmos Science and Technology Ltd.

The remaining 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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Andersson, T. R., Hosking, J. S., Pérez-Ortiz, M., Paige, B., Elliott, A., Russell, C., et al. (2021). Seasonal Arctic sea ice forecasting with probabilistic deep learning. *Nat. Commun.* 12 (1), 1–12. doi: 10.1038/s41467-021-25257-4

Atlas, R., Hoffman, R. N., Ardizzone, J., Leidner, S. M., Jusem, J. C., Smith, D. K., et al. (2011). A cross-calibrated, multiplatform ocean surface wind velocity product for meteorological and oceanographic applications. *B. Am. Meteorol. Soc* 92 (2), 157–174. doi: 10.1175/2010BAMS2946.1

Bauer, P., Thorpe, A., and Brunet, G. (2015). The quiet revolution of numerical weather prediction. *Nature* 525 (7567), 47–55. doi: 10.1038/nature14956

Bento, P. M. R., Pombo, J. A. N., Mendes, R. P. G., Calado, M. R. A., and Mariano, S. J. P. S. (2021). Ocean wave energy forecasting using optimised deep learning neural networks. *Ocean Eng.* 219, 108372. doi: 10.1016/j.oceaneng.2020.108372

Boyer, T. P., Baranova, O. K., Coleman, C., Garcia, H. E., Grodsky, A., Locarnini, R. A., et al. (2018). World ocean atlas 2018. *NOAA Atlas NESDIS* 87, 1–207.

Carton, J. A., Chepurin, G. A., and Chen, L. (2018). SODA3: A new ocean climate reanalysis. *J. Climate* 31 (17), 6967–6983. doi: 10.1175/JCLI-D-18-0149.1

Carton, J. A., and Giese, B. S. (2008). A reanalysis of ocean climate using simple ocean data assimilation (SODA). *Mon. Weather Rev.* 136 (8), 2999–3017. doi: 10.1175/2007MWR1978.1

Chao, Y., Li, Z., Farrara, J. D., and Hung, P. (2009). Blending sea surface temperatures from multiple satellites and *in situ* observations for coastal oceans. *J. Atmos. Ocean. Tech.* 26 (7), 1415–1426. doi: 10.1175/2009JTECHO592.1

Chapman, C., and Charantonis, A. A. (2017). Reconstruction of subsurface velocities from satellite observations using iterative self-organizing maps. *IEEE Geosci. Remote S.* 14 (5), 617–620. doi: 10.1109/LGRS.2017.2665603

Chassignet, E. P., Hurlburt, H. E., Smedstad, O. M., Halliwell, G. R., Hogan, P. J., Wallcraft, A. J., et al. (2007). The HYCOM (hybrid coordinate ocean model) data assimilative system. *J. Mar. Syst.* 65 (1-4), 60–83. doi: 10.1016/j.jmarsys.2005.09.016

Chelton, D. B., and Wentz, F. J. (2005). Global microwave satellite observations of sea surface temperature for numerical weather prediction and climate research. *B. Am. Meteorol. Soc* 86 (8), 1097–1116. doi: 10.1175/BAMS-86-8-1097

Chen, G., Chapron, B., Ezraty, R., and Vandemark, D. (2002). A global view of swell and wind sea climate in the ocean by satellite altimeter and scatterometer. *J. Atmos. Ocean. Tech.* 19 (11), 1849–1859. doi: 10.1175/1520-0426(2002)019<1849:AGVOSA>2.0.CO;2

Cheng, L., Abraham, J., Zhu, J., Trenberth, K. E., Fasullo, J., Boyer, T., et al. (2020). Record-setting ocean warmth continued in 2019. *Adv. Atmos. Sci.* 37, 137–142. doi: 10.1007/s00376-020-9283-7

Chu, X., He, J., Song, H., Qi, Y., Sun, Y., Bai, W., et al. (2020). Multimodal deep learning for heterogeneous GNSS-r data fusion and ocean wind speed retrieval. *IEEE J. STARS.* 13, 5971–5981. doi: 10.1109/JSTARS.2020.3010879

Courtier, P., Thépaut, J. N., and Hollingsworth, A. (1994). A strategy for operational implementation of 4D-var, using an incremental approach. *Q. J. R. Meteor. Soc* 120 (519), 1367–1387. doi: 10.1002/qj.49712051912

Cressman, G. P. (1959). An operational objective analysis system. *Mon. Weather Rev.* 87 (10), 367–374. doi: 10.1175/1520-0493(1959)087<0367:AOOAS>2.0.CO;2

Danabasoglu, G., Lamarque, J. F., Bacmeister, J., Bailey, D. A., DuVivier, A. K., Edwards, J., et al. (2020). The community earth system model version 2 (CESM2). *J. Adv. Model. Earth Sy.* 12 (2), e2019MS001916. doi: 10.1029/2019MS001916

Danard, M. B., Holl, M. M., and Clark, J. R. (1968). Fields by correlation assembly–a numerical analysis technique. *Mon. Weather Rev.* 96 (3), 141–149. doi: 10.1175/1520-0493(1968)096<0141:FBCAAN>2.0.CO;2

Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res. Oceans* 99 (C5), 10143–10162. doi: 10.1029/94JC00572

Forget, G., Campin, J. M., Heimbach, P., Hill, C. N., Ponte, R. M., and Wunsch, C. (2015). ECCO version 4: An integrated framework for non-linear inverse modeling and global ocean state estimation. *Geosci. Model. Dev.* 8 (10), 3071–3104. doi: 10.5194/gmd-8-3071-2015

Good, S. A., Martin, M. J., and Rayner, N. A. (2013). EN4: Quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. *J. Geophys. Res. Oceans* 118 (12), 6704–6716. doi: 10.1002/2013JC009067

Ham, Y. G., Kim, J. H., and Luo, J. J. (2019). Deep learning for multi-year ENSO forecasts. *Nature* 573 (7775), 568–572. doi: 10.1038/s41586-019-1559-7

Hinton, G. E., and Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *science* 313 (5786), 504–507. doi: 10.1126/science.1127647

Huang, B., Thorne, P. W., Banzon, V. F., Boyer, T., Chepurin, G., Lawrimore, J. H., et al. (2017). Extended reconstructed sea surface temperature, version 5 (ERSSTv5): upgrades, validations, and intercomparisons. *J. Climate* 30 (20), 8179–8205. doi: 10.1175/JCLI-D-16-0836.1

Jiao, L., and Zhao, J. (2019). A survey on the new generation of deep learning in image processing. *IEEE Access* 7, 172231–172263. doi: 10.1109/ACCESS.2019.2956508

Kahou, S. E., Bouthillier, X., Lamblin, P., Gulcehre, C., Michalski, V., Konda, K., et al. (2016). Emonets: Multimodal deep learning approaches for emotion recognition in video. *J. Multimodal User In.* 10 (2), 99–111. doi: 10.1007/s12193-015-0195-2

Kopuklu, O., Kose, N., and Rigoll, G. (2018). Motion fused frames: Data level fusion strategy for hand gesture recognition. in. *Proc. IEEE Conf. CVPR Workshops*, 2103–2111. doi: 10.1109/CVPRW.2018.00284

Kumar, A., Wen, C., Xue, Y., and Wang, H. (2017). Sensitivity of subsurface ocean temperature variability to specification of surface observations in the context of ENSO. *Mon. Weather Rev.* 145 (4), 1437–1446. doi: 10.1175/MWR-D-16-0432.1

Li, Y. Y., Dong, Q., Ren, Y. Z., Kong, F. P., and Yin, Z. (2020). Spatiotemporal characteristics of sea surface salinity of Indian and pacific oceans. *J. Remote Sens. (Chinese)* 24 (10), 1193–1205. doi: 10.11834/jrs.20209068

Liu, Y., Shao, Z., and Hoffmann, N. (2021). Global attention mechanism: Retain information to enhance channel-spatial interactions. *arXiv preprint arXiv*. doi: 10.48550/arXiv.2112.05561

Lorenc, A. C. (1981). A global three-dimensional multivariate statistical interpolation scheme. *Mon. Weather Rev.* 109 (4), 701–721. doi: 10.1175/1520-0493(1981)109<0701:AGTDMS>2.0.CO;2

Overpeck, J. T., Meehl, G. A., Bony, S., and Easterling, D. R. (2011). Climate data challenges in the 21st century. *science* 331 (6018), 700–702. doi: 10.1126/science.1197869

Pruthi, D., Gupta, M., Dhingra, B., Neubig, G., and Lipton, Z. C. (2019). Learning to deceive with attention-based explanations. *arXiv preprint arXiv*. doi: 10.48550/arXiv.1909.07913

Rayner, N. A. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., et al. (2003). Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res. Atmos* 108 (D14), 4407. doi: 10.1029/2002JD002670

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., and Carvalhais, N. (2019). Deep learning and process understanding for data-driven earth system science. *Nature* 566 (7743), 195–204. doi: 10.1038/s41586-019-0912-1

Reynolds, R. W., Smith, T. M., Liu, C., Chelton, D. B., Casey, K. S., and Schlax, M. G. (2007). Daily high-resolution-blended analyses for sea surface temperature. *J. Climate* 20 (22), 5473–5496. doi: 10.1175/2007JCLI1824.1

Salcedo-Sanz, S., Ghamisi, P., Piles, M., Werner, M., Cuadra, L., Moreno-Martínez, A., et al. (2020). Machine learning information fusion in earth observation: A comprehensive review of methods, applications and data sources. *Inform. Fusion* 63, 256–272. doi: 10.1016/j.inffus.2020.07.004

Serrano, S., and Smith, N. A. (2019). Is attention interpretable? *arXiv preprint arXiv*. doi: 10.48550/arXiv.1906.03731

Stammer, D., Balmaseda, M., Heimbach, P., Köhl, A., and Weaver, A. (2016). Ocean data assimilation in support of climate applications: status and perspectives. *Annu. Rev. Mar. Sci.* 8, 491–518. doi: 10.1146/annurev-marine-122414-034113

Strubell, E., Ganesh, A., and McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. *arXiv preprint arXiv*. doi: 10.48550/arXiv.1906.02243

Su, H., Huang, L., Li, W., Yang, X., and Yan, X. H. (2018). Retrieving ocean subsurface temperature using a satellite-based geographically weighted regression model. *J. Geophys. Res. Oceans* 123 (8), 5180–5193. doi: 10.1029/2018JC014246

Su, H., Wang, A., Zhang, T., Qin, T., Du, X., and Yan, X. H. (2021). Super-resolution of subsurface temperature field from remote sensing observations based on machine learning. *Int. J. Appl. Earth Obs.* 102, 102440. doi: 10.1016/j.jag.2021.102440

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res. Atmos* 106 (D7), 7183–7192. doi: 10.1029/2000JD900719

Thiébaux, J., Rogers, E., Wang, W., and Katz, B. (2003). A new high-resolution blended real-time global sea surface temperature analysis. *B. Am. Meteorol. Soc* 84 (5), 645–656. doi: 10.1175/BAMS-84-5-645

Vafaei, B., Ezam, M., Saghaei, A., and Bidokhti, A. A. (2022). Automatic identification and tracking of meso-scale eddies in the Persian gulf using the pattern mining approach. *Int. J. Environ. Sci. Technol.* 19 (7), 6011–6022. doi: 10.1007/s13762-021-03779-0

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). Attention is all you need. *NIPS* 30, 5998-6008. doi: 10.48550/arXiv.1706.03762

Wiegreffe, S., and Pinter, Y. (2019). Attention is not not explanation. *arXiv preprint arXiv*. doi: 10.48550/arXiv.1908.04626

Wu, T., Yu, R., Lu, Y., Jie, W., Fang, Y., Zhang, J., et al. (2021). BCC-CSM2-HR: a high-resolution version of the Beijing climate center climate system model. *Geosci. Model. Dev.* 14 (5), 2977–3006. doi: 10.5194/gmd-14-2977-2021

Xiao, C., Hu, C., Chen, N., Zhang, X., Chen, Z., and Tong, X. (2021). A genetic algorithm-assisted deep neural network model for merging microwave and infrared daily Sea surface temperature products. *Front. Environ. Sci.* 421. doi: 10.3389/fenvs.2021.748913

Xu, J., Yang, J., Xiong, X., Li, H., Huang, J., Ting, K. C., et al. (2021). Towards interpreting multi-temporal deep learning models in crop mapping. *Remote Sens. Environ.* 264, 112599. doi: 10.1016/j.rse.2021.112599

Yu, D., and Deng, L. (2016). *Automatic speech recognition* Vol. 1 (Berlin: Springer).

Zeng, L., and Levy, G. (1995). Space and time aliasing structure in monthly mean polar-orbiting satellite data. *J. Geophys. Res. Atmos.* 100 (D3), 5133–5142. doi: 10.1029/94JD03252

Zhang, C., Wang, D., Liu, Z., Lu, S., Sun, C., Wei, Y., et al. (2022). Global gridded argo dataset based on gradient-dependent optimal interpolation. *J. Mar. Sci. Eng.* 10 (5), 650. doi: 10.3390/jmse10050650

Zhao, M., Zhong, S., Fu, X., Tang, B., Dong, S., and Pecht, M. (2020). Deep residual networks with adaptively parametric rectifier linear units for fault diagnosis. *IEEE T. Ind. Electron.* 68 (3), 2587–2597. doi: 10.1109/TIE.2020.2972458

Zhu, Y., Bo, Y., Zhang, J., and Wang, Y. (2018). Fusion of multisensor SSTs based on the spatiotemporal hierarchical Bayesian model. *J. Atmos. Ocean. Tech.* 35 (1), 91–109. doi: 10.1175/JTECH-D-17-0116.1

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 09 February 2023
DOI 10.3389/fmars.2022.1084784

Check for updates

OPEN ACCESS

EDITED BY
Yasumasa Miyazawa,
Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Japan

REVIEWED BY
Shuhei Masuda,
Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Japan
Tengfei Xu,
Ministry of Natural Resources, China
Clemente Augusto Souza Tanajura,
Federal University of Bahia, Brazil

*CORRESPONDENCE
David E. Gwyther
david.gwyther@gmail.com
Moninya Roughan
mroughan@unsw.edu.au

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

RECEIVED 31 October 2022
ACCEPTED 20 December 2022
PUBLISHED 09 February 2023

CITATION
Gwyther DE, Roughan M, Kerry C and Keating SR (2023) Impact of assimilating repeated subsurface temperature transects on state estimates of a western boundary current.
Front. Mar. Sci. 9:1084784.
doi: 10.3389/fmars.2022.1084784

COPYRIGHT
© 2023 Gwyther, Roughan, Kerry and Keating. 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(s) 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.

Impact of assimilating repeated subsurface temperature transects on state estimates of a western boundary current

David E. Gwyther[1]*, Moninya Roughan[1]*, Colette Kerry[1] and Shane R. Keating[2]

[1]Coastal and Regional Oceanography Lab, School of Biological, Earth and Environmental Sciences, UNSW Sydney, Sydney, NSW, Australia, [2]School of Mathematics and Statistics, UNSW Sydney, Sydney, NSW, Australia

Western Boundary Currents and the eddies they shed are high priorities for numerical estimation and forecasting due to their economic, ecological and dynamical importance. However, the rapid evolution, complex dynamics and baroclinic structure that is typical of eddies and the relatively sparse sampling in western boundary currents leads to significant challenges in understanding the 3-dimensional structure of these boundary currents and mesoscale eddies. Here, we use Observing System Simulation Experiments (OSSEs) to explore the impact of assimilating synthetic subsurface temperature observations at a range of temporal resolutions, to emulate expendable bathythermograph transects with different repeat frequencies (weekly to quarterly). We explore the improvement in the representation of mesoscale eddies and subsurface conditions in a dynamic western boundary current system, the East Australian Current, with a data-assimilating regional ocean model. A characterisation of the spatial and temporal ocean variability spectrum demonstrates the potential for undersampling and aliasing by a lower sampling frequency. We find that assimilating subsurface temperature data with at least a weekly repeat time best improves subsurface representation of this dynamic, eddy-rich region. However, systemic biases introduced by the data assimilation system hinder the ability of the model to produce more accurate subsurface representation with fortnightly or monthly sampling. Removal of this bias may improve subsurface representation in eddy-rich regions with fortnightly or even less frequent observations. These results highlight the value of both increased subsurface observation density in regions of dynamic oceanography as well as continued development of data assimilation techniques in order to optimise the impact of existing observations.

KEYWORDS

Western Boundary Current (WBC), East Australian current, expendable bathythermograph (XBT), observing system simulation experiment (OSSE), data assimilation (DA)

1 Introduction

Subtropical Western Boundary Currents (WBCs) are narrow, rapidly flowing warm water currents that are important for ecosystems, climate, weather and cross-shelf exchange (e.g. Lambaerts et al., 2013; Malan et al., 2020; Oliver et al., 2021; Li et al., 2022a). As fast-flowing WBC jets become unstable, they shed *O*(100) km-wide mesoscale eddies. The formation, structure and evolution of these eddies and associated structures are important due to the impact they have on the transport of heat and salt (Abernathey and Haller, 2018), weather (Frenger et al., 2013), mixing (Klocker and Abernathey, 2014), and the delivery of nutrients (Everett et al., 2012).

Given their location adjacent to populous coastlines, WBCs have a pivotal role in coastal fisheries and other blue economies (e.g. Young et al., 2011; Brieva et al., 2015), weather and climate, and search-and-rescue and navigation. An impediment to manyof these end-users is the limitation in model predictability resulting from the dynamically changing eddy field. Mesoscale dynamics are inherently sensitive, where divergent evolution results from small differences in initial conditions. This leads to a timescale limit on predictability for techniques such as search and rescue, navigation and other methods that require accurate forecasts of ocean weather.

Accurate estimates and predictions of WBCs and eddy-rich regions are generally sought through data-assimilating models (Oke et al., 2013). The technique of data assimilation combines observations with a model forecast, often in an iterative process, to produce an analysis or optimal estimate of the ocean state. Hence, the evolution of eddies can be continually updated within model forecasts, providing a best estimate of eddy structure, timing and location (Oke et al., 2013).

One of the key ways in which operational forecasting systems differ from non-operational, research-focussed assimilation systems is the types of observations that can be assimilated. For example, operational forecast systems typically assimilate sea surface height (SSH), sea surface temperature (SST), a smaller number of subsurface observations, such as vertical temperature profiles from expendable bathythermograph (XBT) probes and Argo floats, as well as occasionally wind stress and sea surface salinity observations (e.g. Brassington et al., 2007).

In contrast, hindcast reanalyses with a research focus (e.g. Kerry et al., 2016; Siripatana et al., 2020), can augment these traditional observation types with more novel, often delayed-mode observations, e.g. high frequency radar-inferred surface currents, hydrography from autonomous gliders, and measurements from subsurface moorings. The constraint limiting operational forecast systems from assimilating the full gamut of non-traditional and subsurface observations, is the ability to have data prepared and available in near-real time for the next operational window — which can often not be met when quality control or other data preparation must be conducted with human supervision, or obviously if there is a delay in data recovery. As a result, operational systems will often be limited to just surface observations combined with a small number of near-real time subsurface measurements.

The East Australian Current (EAC; see Figure 1 for region) is one such WBC with routine subsurface sampling. The EAC flows southwards from 27.5°S as a coherent jet, before beginning to meander and lose coherency between 31°–33°S and then feeding an 'evolving' field of cyclonic and anticyclonic eddies in the Tasman sea. Like many other WBCs, an important source of subsurface real-time measurements in the EAC are repeated XBT transects and Argo floats. A long-term program of XBT deployments from ships of opportunity has been operated along repeat transects in the southern Pacific Ocean since the late 1980s (and 1991 for the two transects through the EAC region, named 'PX30' and 'PX34') by Scripps Institution of Oceanography, the Australian Commonwealth Scientific and Industrial Research Organisation and the New Zealand National Institute for Water and Atmospheric Research. The original use of this data was for estimating boundary current heat budgets (e.g. Roemmich and Cornuelle, 1990; Morris et al., 1996; Roemmich et al., 2005). However the near-real time data delivery and consistent transects lends itself well to data assimilation into ocean models. As the XBT data is delivered to the global telecommunications system it is readily available in near realtime for ingestion into operational modelling systems and reanalyses (e.g. Carton et al., 2000).

Argo floats are a second source of subsurface observations that have also been used in operational forecasting. Argo floats return vast amounts of deep (to 2000 m) vertical temperature and salinity profiles over a much broader area of the ocean and have revolutionised understanding of the ocean (Wijffels et al., 2016; Wong et al., 2020). However, they still have relatively low spatial distribution and are not measuring systematically e.g. along repeat transects at routine time and space scales. Thismeans the data cannot easily be used for closed box heat budgets, their Lagrangian paths could lead to sample aliasing, nor can we systematically assess observation impact. For these reasons we do not consider Argo data in this analysis.

It has been shown with data assimilation experiments that weekly subsurface temperature (XBT-like) observations have a significant impact on representation of the EAC: improving mean surface and subsurface circulation patterns, upper ocean heat content estimates (Gwyther et al., 2022), as well as baroclinic mode structure and eddy representation (Gwyther et al., 2023). However, the actual EAC XBT observing system only employs an approximately quarterly transect repeat time (or less). Hence, there is strong motivation to assess how impactful the existing XBT system is on representation of the EAC and its eddy field. Further it is useful to explore how representation of the EAC System is improved by increasing the observation sampling frequency.

This assessment is conducted with Observing System Simulation Experiments (OSSEs), which are a method of

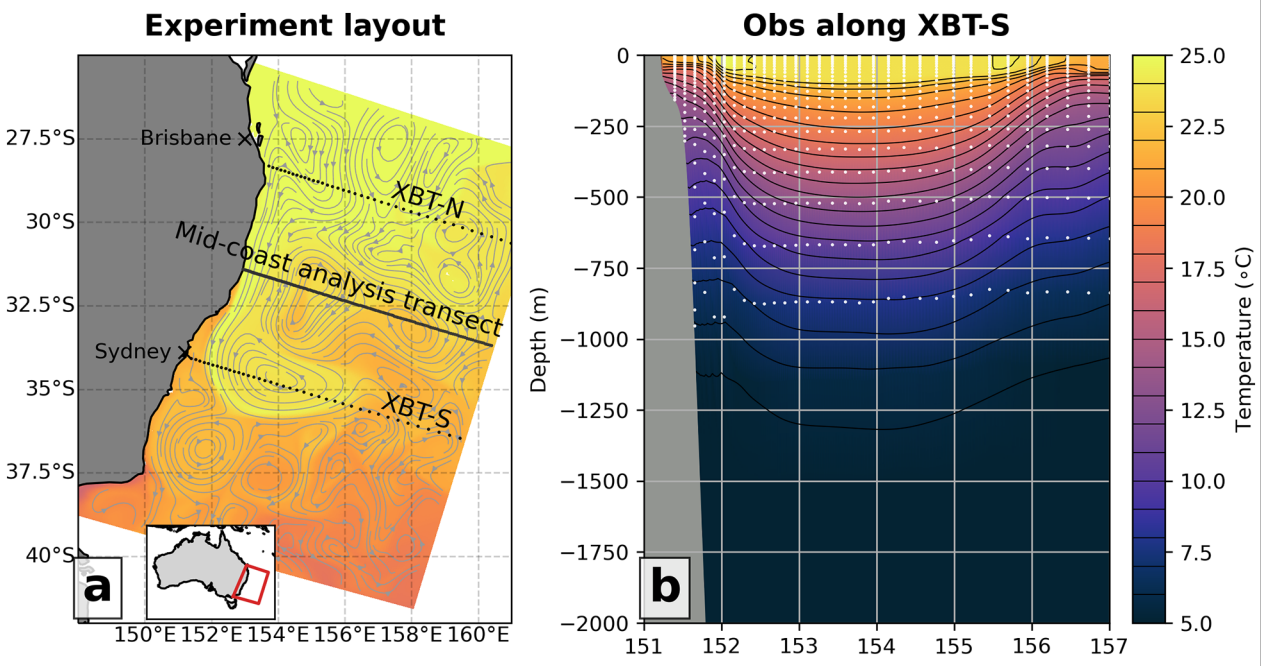


FIGURE 1
(A) A snapshot of model SST at 11-March 2012 is shown for the East coast of Australia. Dots mark deployment locations of XBTs along the XBT-N (beginning at ~28.5°S) and XBT-S (beginning at ~34°S) transects and the line shows the analysis transect beginning at ~31°S. Major east coast cites Brisbane and Sydney are marked, and grey vectors show model surface currents at the same time. **(B)** Temperature measurement locations are marked as grey dots in this vertical slice of temperature (at the same time as in panel **A**) along the XBT-S deployment line (see **A**). Inset in **(A)** shows model domain.

assessing observation impact, where a free-running simulation is taken as the true ocean or reference state, from which synthetic observations are extracted (e.g. Halliwell et al., 2014; Halliwell et al., 2015; Gasparin et al., 2019; Moore et al., 2020). These observations are then assimilated into a simulation that had an initial perturbation applied, and the resulting estimate can be compared to the ref state (see Figure 2 of Gwyther et al., 2022 for a schematic of the full OSSE procedure). We are thus able to assess the impact of the observing strategy on the representation of key ocean properties. In this study, we compare how several temporal sampling frequencies impact representation of subsurface temperature and eddy kinetic energy. This approach has the advantage of being able to assess the impact of a range of sampling configurations on ocean state estimates, without the time or cost of obtaining the ocean observations.

We present a series of model experiments that assimilate synthetic XBT observations with increased temporal resolution approximately matching the existing XBT transect network in the EAC System. In particular, we focus on assessing the impact of different sampling frequency on subsurface temperature fields and eddy kinetic energy, which have previously been shown as challenging to represent accurately (Gwyther et al., 2022; Gwyther et al., 2023). We characterise the ocean variability spectrum in order to demonstrate the potential for undersampling and aliasing by low frequency observations. Lastly, we explicitly separate the systemic error that is introduced by the data assimilation system from the endemic error resulting from inadequate representation ofthe mesoscale dynamics.

2 Methods

Numerical simulations are conducted with the Regional Ocean Modeling System (ROMS; Shchepetkin and McWilliams, 2005) which is a finite-difference primitive equation model with a terrain-following *s*-coordinate. The model domain extends from 27°S to 38°S and over 700 km offshore (with meridional grid resolution of 2.5 km linearly increasing to 6 km towards the east), and with constant meridional resolution of 5 km). There are 30 model layers in the vertical, with the model *s*-coordinate configured for more resolution in the surface boundary layer. This discretisation leads to cell thicknesses in the EAC of 1–3 m immediately below the surface, ~50–100 m thick cells in several hundred metres of water and ~300 m thick cells in the deep ocean below 3000 m. The model grid is rotated by 20° clockwise so as to approximately align the model coordinates with the along-shore and across-shelf directions (Model grid shown in Figure 1A. The bathymetry is sourced from the Geoscience Australia 50 m multibeam survey (Whiteway, 2009). The model domain is identical to that used in several other studies that explore the

EAC velocity variability (e.g. Kerry and Roughan, 2020), biogeochemistry (e.g. Rocha et al., 2019), eddy dynamics (e.g. Li et al., 2021; Li et al., 2022b) and observation impact (e.g. Kerry et al., 2018; Gwyther et al., 2022; Gwyther et al., 2023).

We use two model configurations: a free-running and a data assimilating configuration. The free-running model uses lateral boundary forcing (currents, temperature and salinity conditions) from BRAN2020 (Chamberlain et al., 2021) and surface forcing conditions from the Bureau of Meteorology Atmospheric high-resolution Regional Reanalysis for Australia (BARRA-R; Su et al., 2019). More details are given in Gwyther et al. (2022) and Gwyther et al. (2023).

The data assimilating configuration used for the OSSEs is based on the model setup developed by Kerry et al. (2016), and uses an Incremental Strong Constraint 4-Dimensional Variational scheme (IS4D-VAR; e.g. Moore et al., 2011). This scheme calculates the differences between a free-running 'forecast' and observations over a chosen assimilation window (5 days in our case), where model and observations have associated error fields. The data assimilation scheme then generates new initial and boundary conditions such that a new 'analysis' simulation running with these adjusted initial, boundary and surface forcing conditions has minimised differences (in a least-squares sense) from the observations. The cycle then increments forward, using the previous analysis as the initial conditions for the new forecast. This data assimilating configuration has also been used in previous studies (e.g. Kerry et al., 2016; Kerry et al., 2018; Kerry et al., 2020; Siripatana et al., 2020) including for OSSEs (Gwyther et al., 2022; Gwyther et al., 2023). The lateral forcing conditions are from BRAN2020, while atmospheric conditions are sourced from the Australian Bureau of Meteorology's ACCESS reanalysis (Puri et al., 2013). Both free-running and data assimilating configurations use a bulk flux parameterisation (Fairall et al., 1996) for calculating surface fluxes. This difference in surface forcing conditions between configurations is a necessary requirement for 'fraternal twin'-type OSSEs. As summarised by Halliwell et al. (2014), a balance must be sought between slightly different configurations (forcing conditions, mixing parameters or parameterisations) that introduce error and realistically test the data assimilation system, and a long-term bias that the assimilation cannot correct for. An analysis of long-term mean bias resulting from the different forcing conditions showed that bias is minimal and constrained to the surface ocean, where it can be readily corrected by assimilating SST (see Gwyther et al., 2023).

We use the free-running configuration as the reference state (referred to as the 'ref state'), to which a series of data-assimilating simulations (the OSSEs) are compared. Values are extracted from the ref state and a representative level of error is added as a normally distributed perturbation with standard deviation equal to the observation error. The observation error is set at 0.04 m for SSH, 0.5°C for SST, and with a depth-dependent profile for XBT observations, ranging from 0.6°C to 0.12°C (more information is given in Gwyther et al., 2022). These values are then taken as the synthetic observations, and are assimilated into a perturbed data-assimilating simulation (the OSSE). Following the procedure of (Gwyther et al., 2022), we generate a perturbed simulation by initialising the free-running simulation with an 8-day offset. This perturbation is enough to cause a free-running simulation to diverge significantly from the ref state, and is thusan effective test of the performance of the data assimilation system in assimilating observations into a realistically-diverged background state. Several different perturbations were tested, including a 1-day offset, 1-month, 1-year and a climatological state, but all were found to eventually generate similarly high levels of error (not shown). We thus chose a 8-day offset as it relatively quickly diverged, but it still had mesoscale features in the approximately correct locations, which is analogous to initialisation error in a true data assimilation system. For a more detailed description of the OSSE method, Gwyther et al. (2022) gives further information and includes a schematic of the procedure (their Figure 2).

Previous work has showed that the free-running model produces accurate seasonal and interannual representations of the EAC, including the eddy field, the separation latitude, eddy kinetic energy and volume transport (e.g. Kerry et al., 2016; Kerry and Roughan, 2020; Li et al., 2021). This demonstrates that the free-running ref state is representative of the true ocean. Hence, by analogy, the impact of assimilating the synthetic observations should translate to the true ocean.

This study assesses the performance of four different hypothetical XBT observing strategies through OSSEs, by comparing these against the ref state. All OSSEs assimilate the same synthetic surface observations representing SST and SSH by extracting observations from the ref state simulation with the appropriate location and timing, as per Gwyther et al. (2022; 2023). Synthetic SSH observations are based on the spatial and temporal coverage of the global ocean along-track multi-mission sea level altimeter data. Synthetic SST observations are based on the spatial and temporal coverage of the near-real time Himawari-8 satellite product. Each OSSE also assimilates subsurface XBT-like temperature observations, also extracted from the ref state, along two transects: at ~28.5°S and ~34°S, with a horizontal observation spacing of ~12.5 km at the continental shelf break. However for each experiment the temporal repeat time of the XBT transects is different, ranging from weekly (the 'W-12.5' OSSE), fortnightly (the '2W-12.5' OSSE), monthly (the 'M-12.5' OSSE) to quarterly (the 'Q-12.5' OSSE). Experiment names, and temporal and spatial resolutions are shown in Table 1.

Subsurface observations consist of vertical temperature profiles down to 900 m, with 10 m vertical resolution (or one observation per model layer for model vertical layer spacing greater than 10 m). These observations are taken along two transects, one in the north (referred to as XBT-N; ~28.5°S) and

TABLE 1 The configurations of the free-running ref state and the four experiments are described, including the XBT horizontal and vertical spacing and XBT transect repeat time.

Exp. name	Model configuration details	XBT horizontal and vertical spacing	Transect repeats
ref state	Free-running simulation covering period of Nov 2011–Jan 2013.	Not applicable.	Not applicable
W-12.5	4DVar simulations assimilate synthetic along-track SSH and SST, as well as XBT observations that are extracted from the ref state along two transects.	Transects at ~28.5°S (XBT-N) and ~34°S (XBT-S). Horizontal spacing is 12.5-30 km with full transect observed in ~5 days. Observations extend from 5 m to 900 m, with spacing of at least 10 m.	Weekly (7days)
2W-12.5			Fortnightly (14 days)
M-12.5			Monthly (30 days)
Q-12.5			Quarterly (90 days)

one in the south (referred to as XBT-S; ~34°S) of the domain. The two XBT lines are chosen to represent the approximate location and vertical sampling rates of the PX30 and PX34 lines (Figure 1A shows transect locations and Figure 1B shows vertical distribution of observations along XBT-S). For ease of implementation, in our experiments, the XBT lines are oriented along grid coordinates (i.e. normal to the shore), whereas in reality the PX30 and PX34 XBT lines are along shipping routes from Brisbane to Nouméa, New Caledonia and Sydney to Wellington, New Zealand and thus at an angle to our model grid. The unique niche of the XBT dataset is the relatively fine spacing (10–100 km) between vertical profiles along defined repeat transects, allowing the resolution of a broad spectrum of processes, from eddies to basin-scale circulation (Smith et al., 1999).

All OSSEs assimilate surface observations (SSH and SST; see above) and subsurface temperature observations along both the northern transect and the southern transects, with horizontal XBT spacing of 12.5 km to 30 km (every 5 model cells) and 5-days to sample the transects. Each OSSE has different XBT transect repeat times: The W-12.5 OSSE has a transect repeat every week, the 2W-12.5 OSSE has transect repeats every two weeks, the M-12.5 has transect repeats every month (30 days) and the Q-12.5 has transect repeats every quarter year (90 days). The Q-12.5 OSSE represents the true Scripps XBT lines most closely in spatial resolution and temporal repeat time. The other OSSEs (W-12.5, 2W-12.5 and M-12.5) represent how a higher-frequency sampling scheme will impact simulated representation of the ocean.

3 Results

3.1 Modes of variability in the surface and subsurface EAC

We firstly use the ref state to explore the important spatial and temporal modes of variability of the EAC over the one-year simulation. The goal of this is to gain an understanding of the key frequencies and scales of variability, so that we can better interpret the effectiveness of observing strategies with different sampling times and lengths.

The time evolution of several surface and subsurface quantities in the ref state simulation at ~34°S are shown in Figure 2. The surface fields clearly display the seasonal cycle (transition of high to low SST anomaly from 2012-02 to 2012-08; Figure 2A) and the passage of anticyclonic and cyclonic eddies towards the south-west (anticyclonic at 2012-03 and cyclonic at 2012-11; Figure 2B). The subsurface fields at 500 m are less influenced by seasonal processes, instead being dominated by the passage of eddies (Kerry et al., 2018), and are below the EAC core depth (Gwyther et al., 2022). The most noticeable feature in the Hovmöller diagram of temperatureat 500 m is the temperature increases associated with the passing of warm core, anticyclonic eddies (and vice versa for some cyclonic eddies, e.g. mid-August 2012; Figure 2C). Eddy kinetic energy (EKE; defined as the squared anomaly in velocities from the 2012–2013 mean) at 500 m increases as the largest eddies cross the transect (Figure 2D). The consistent slopes of EKE features in the Hovmöller diagram capture the south-westwards trajectory of the largest (i.e. high EKE) eddies in this region.

The variability in the EKE at 500 m can be explored further with a frequency-wavenumber spectrum analysis (Figure 3A). The spectrum is calculated by taking the 2-dimensional fourier transform of a longitude-time field, which in our case is the EKE anomaly at 500 m (Figure 2A) and the x and y axes are scaled to show periodicity and wavelength. The $\log_{10}$ wavelength-period power spectrum shows several features: higher power at long wavelengths and monthly–seasonal timescales; and, a distinct band of increased power that runs from approximately fortnightly–monthly and very long wavelengths, through decreasing period and wavelengths to approximately daily periods and 30–40 km wavelengths. This band of increased variability is likely associated with the Rossby-mesoscale-submesoscale energy pathway and dynamics, as identified by, for example, by Torres et al. (2018). On either side of this feature (short wavelength and long period, or long wavelength and short period) there is comparatively low power. The two diagonal grey lines in Figure 3A represent a non-dispersive relationship, $\omega=ck$,

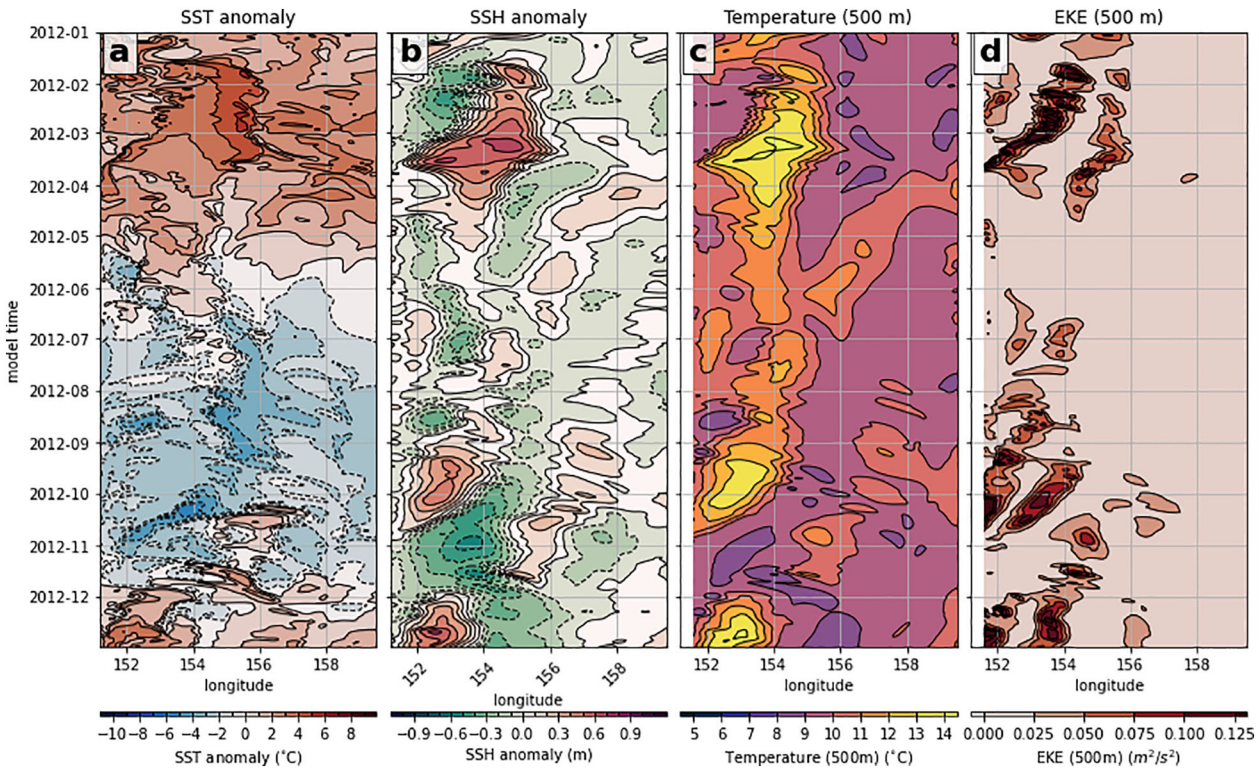


FIGURE 2
Hovmöller diagrams show the longitude–time variability in the ref state along the XBT-S transect (see Figure 1) for two surface quantities **(A)** SST anomaly, **(B)** SSH anomaly, and two subsurface quantities **(C)** temperature at 500 m and **(D)** EKE at 500 m. All anomalies are calculated by subtracting the mean value at each longitude over the time period Jan-2012–Jan-2013. Contour intervals are indicated with marks in the colourbar.

for different values of the phase speed c. These represent the frequency-wavenumber relationships for mesoscale anomalies propagating with phase speeds of c = 8 km/day and 20 km/day. The bracketing of the high-power band by these relationships further confirms that the source of this band of power is mesoscale interactions.

The frequency-wavenumber power spectrum can be further analysed at several important periods, as shown in Figure 3B). Here, we show the spectral variance which we calculate by scaling the power spectrum by the frequency and wavenumber, leading to a more even weighting of signals across all frequencies and wavenumbers. The power contained at each chosen time-scale is shown to decrease with increasing period (e.g. compare 7 days to 90 days; Figure 3B). This shows that sampling at lower frequencies will truncate a significant portion of the EKE variability spectrum, and could alias high-frequency EKE energy into lower frequency modes.

Singular spectrum analysis (SSA; see Elsner and Tsonis, 1996) is used to decompose the linearly detrended time series of EKE at 500 m in the ref state at each spatial location. Different time bands are chosen to perform SSA and the variance explained by the reconstructed components are plotted for each spatial location. Benefits of SSA include that it is data-driven and works only in the time-dimension. As a result, it is less affected by the choice of basis vectors and boundary effects (e.g. EOF analyses). The percentage of total variance in EKE at 500 m that is captured over four chosen time bands are shown in Figure 4. The percentage of variance explained by processes with weekly or greater periodicity is high (spatial mean explained variance is 88%; Figure 4A), which shows that in this simulation almost all variability in the EKE field is occurring with a weekly or longer period. For fortnightly and monthly periods or longer, the percentageof explained variance drops to spatial mean values of 73% and 49%, respectively (Figures 4B, C). For quarterly periods or longer, the percentage of variance captured in that time range is small, 19% on average over the model domain (Figure 4D). Together, this illustrates that the proportion of the ocean variability that is captured by subsurface sampling will decrease with decreasing sampling frequency.

3.2 Mean conditions under more frequent sampling

Given that EKE appears to vary more in its frequency spectrum rather than wavenumber spectrum (Figure 3), and

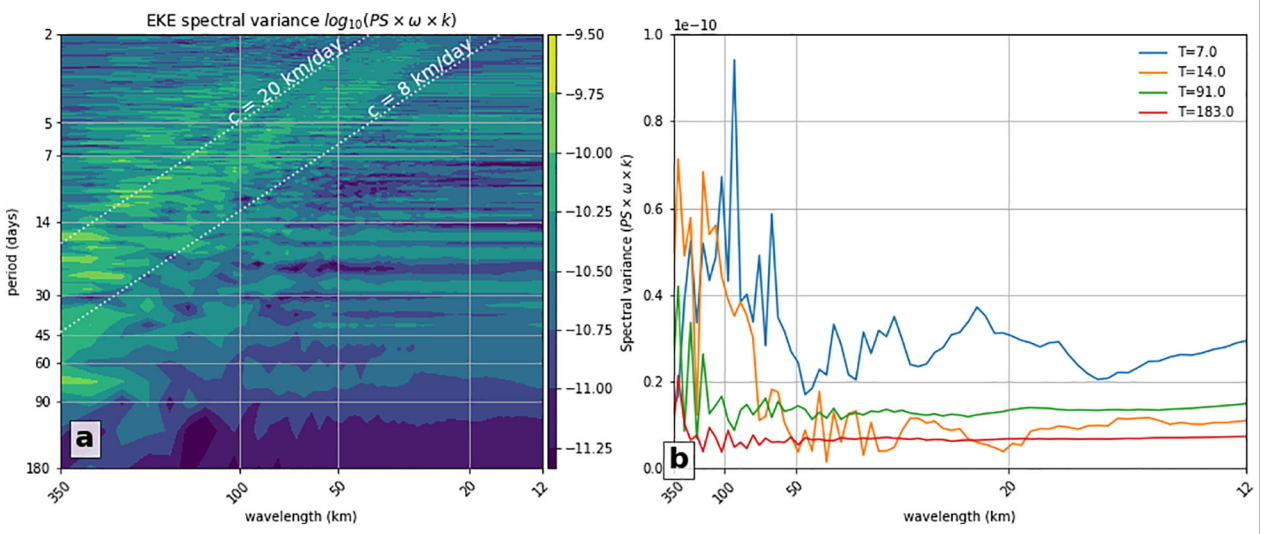


FIGURE 3
The variability of the EKE anomaly at 500 m along the XBT-S transect is explored for the ref state over the period Jan 2012–Jan 2013, with the **(A)** $\log_{10}$ wavelength–period power spectrum of the EKE at 500 m and **(B)** power spectrum of the EKE at 500 m at four selected periodicities. In both panels, the power spectrum is the variance-preserving spectral density, calculated as the power scaled by the frequency and wavenumber, which more evenly weights signalsat higher frequencies and smaller length-scales. The chosen periodicities in **(B)** are 7 days, 14 days, ~90 days and ~180 days. The EKE anomaly is calculated by subtracting the EKE from the time-mean EKE at each longitude. The grey dashed lines in **(A)** show the non-dispersive relationship, $\omega=ck$, between frequency ω and wavenumber k. The slope of the line, c, corresponds to the eddy phase speed, which is shown for two values: 8 km/day and 20 km/day.

the increasing percentage of total variance explained over longer sampling windows, there is justification for a suite of OSSEs that test the impact on EAC and eddy representation from different XBT repeat frequencies. We can assess the improvement in mean ocean conditions by comparing OSSEs with different XBT repeat frequency to the ref state.

The transect mean RMS error in temperature along three representative transects (the XBT-N, mid-coast and XBT-S lines; see Figure 1 for locations) show the improvement in representation with more frequent sampling (Figure 5). For all transects shown, the improvement in temperature RMS by increasing from quarterly to monthly sampling is minimal (Figures 5A–C). In contrast, increasing sampling to weekly decreases RMS, especially in the more dynamic region south of the separation zone (Figure 5C). However, there is not consistent improvement in time series of RMS for the most rapid sampling scheme.

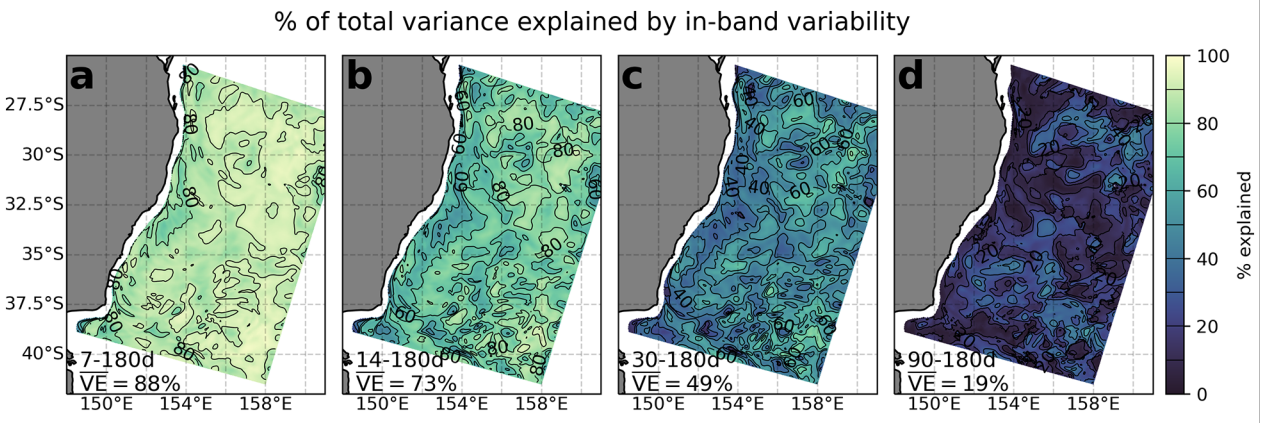


FIGURE 4
Singular spectrum analysis is used to quantify the percentage of the total variance in EKE at 500 m explained in four selected time bands. The variance with a period greater than **(A)** 7 days, **(B)** 14 days, **(C)** 30 days and **(D)** 90 days, and less than 180 days, is expressed as a fraction of the total variance. A high value indicates that a large amount of the full spectrum of variability at that location has a period within the indicated time band. Annotations in each plot show the band over which variance is being explained, and the spatial mean of the in-band variance explained. The contour interval is 10%.

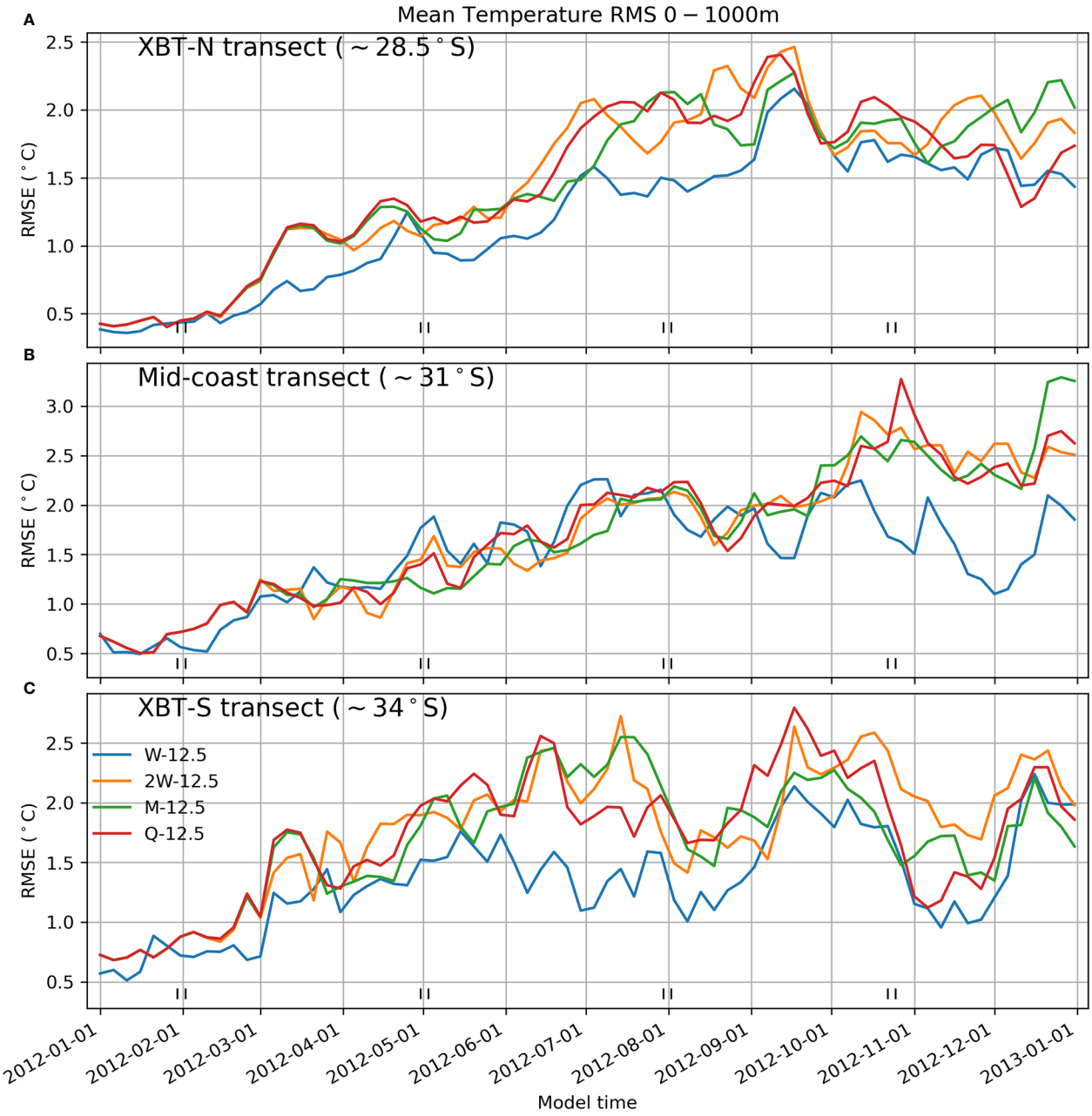


FIGURE 5
The transect-mean RMS error in temperature for **(A)** the XBT-N transect at ~28.5°S, **(B)** the mid-coast analysis transect at ~31°S and **(C)** the XBT-S transect at ~34°S for the four different OSSEs. RMS error compares the spatial-mean difference in temperature between each OSSE and the ref state. Black ticks indicate timings of quarterly XBT data assimilated into the Q-12.5 OSSE.

The vertical structure of temperature RMS, along the transects, shows an improvement when increasing the XBT repeat frequency (Figure 6). This improvement occurs both in the top 1000 m, where there are observations, and across the full water column. For example, at ~34°S, the RMS error decreases from 1.0°C to 0.94°C when increasing sampling from quarterly to monthly (see Figures 6L, K), which is otherwise not highlighted in the spatial mean RMS (Figure 5). However, there is greater improvement in RMS when increasing from monthly to weekly sampling, for example, decreasing RMS in the top 1000 m (where RMS error is highest) from 1.84°C to 1.63°C (cf. Figures E, G) at 31°S, and a larger improvement again at 34°S from 2-weekly to weekly Figures 6I, J.

The spatial structure of temperature RMS clearly shows how the error is decreased by increasing XBT repeat times (Figure 7), particularly in the eddy region (indicated by the blue box, in

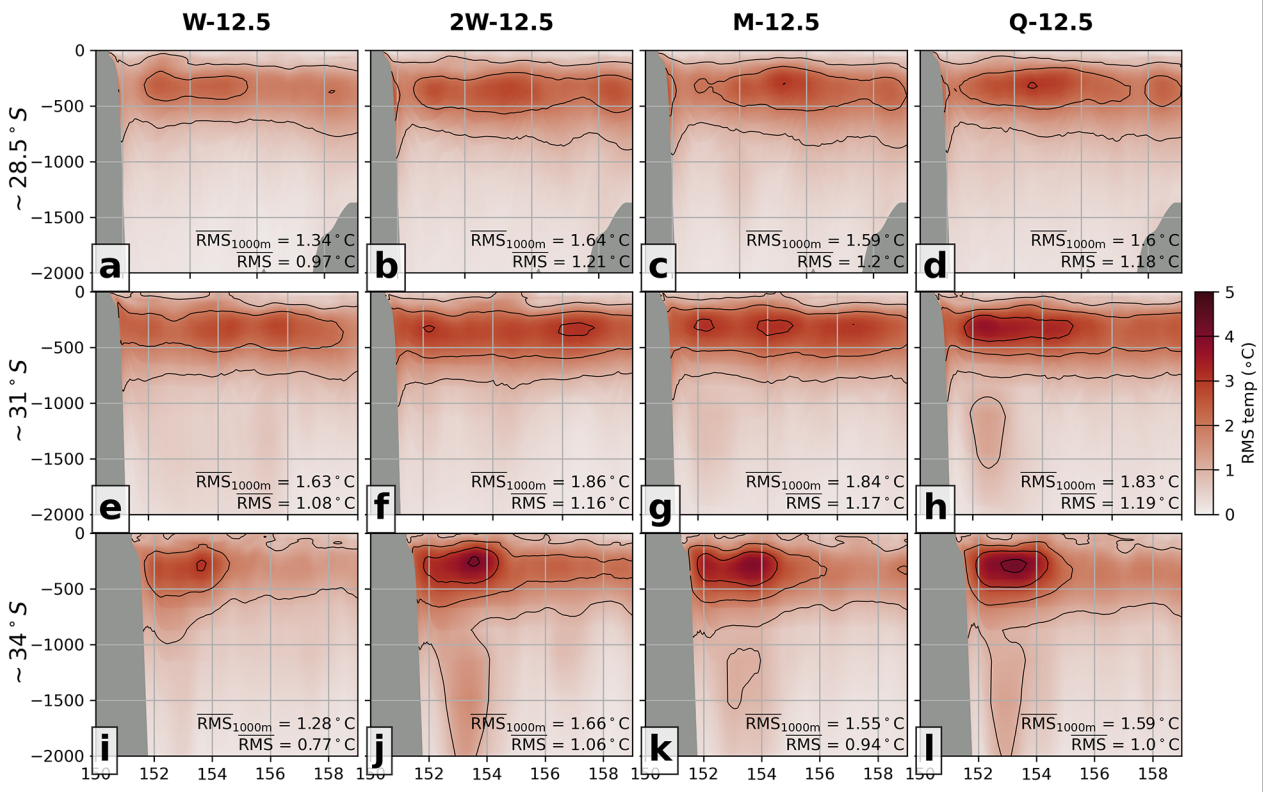


FIGURE 6
The RMS error in temperature, calculated for the time-mean, at the **(A–D)** XBT-N transect, **(E–H)** the mid-coast analysis transect, and **(I–L)** XBT-S transect, for the four different OSSEs: **(A, E, I)** Weekly 12.5 km, **(B, F, J)** Two-weekly 12.5 km, **(C, G, K)** Monthly 12.5 km and **(D, H, L)** Quarterly 12.5. RMS error is relative to the ref state, and mean values are shown in each panel for full-depth and the top 1000 m. The contour interval of 1°C is indicated with marks in the colourbar.

Figure 7A). At 250 m, error is highest in the eddy region (~152° E, 36°S) but decreases as the repeat time is decreased, from a mean value of 3.1°C to 2.4°C for the eddy region (Figures 7A–D). The pattern is similar at 500 m with the W-12.5 OSSE having the lowest RMS compared to the M-12.5 and Q-12.5 OSSEs (cf. Figures 7E, G–H). At 1000 m, RMS is relatively low for all OSSEs, which reflects the lower natural variability at this depth (Figures 7I–L).

Representation of the surface and subsurface salinity show no improvement with more frequent XBT observation repeats (Figure S1). This suggests that assimilation of SSH, SST and subsurface temperature is not enough to improve representation of subsurface salinity, despite any covariance of salinity and temperature properties. In our 4DVar configuration, we estimate the background error covariance matrix by factorisation (Weaver and Courtier, 2001) and prescribe univariate covariance. Covariances are assumed to be static with isotropic horizontal and vertical length scales, with flow-dependence being introduced by updating the tangent-linear and adjoint models with the time evolution of the background. Consequently, salinity is not set to covary with temperature, which may contribute to explaining this result. Compare this to Balmaseda et al. (2013), who show improvement in salinity through assimilation of temperature observations *via* the balance operators.

3.3 Partitioning the source of error

The RMS error in each OSSE field is the combined effect of error introduced from the data assimilation system and the error resulting from dynamic features. A bias is calculated as the time-mean OSSE field minus the time-mean — this represents the systemic error introduced through data assimilation. This bias is then subtracted from the OSSE field, and a bias-corrected RMS error is calculated — this error field now represents the differences in representation resulting from representation of dynamic ocean features.

The bias in each OSSE, calculated as the time mean of the OSSE field minus the time mean of the same field in the ref state, shows the time-averaged difference between the OSSE and the ref state (Figure 8). The bias in the SST is close to zero for all different XBT repeat times (Figures 8A–D). However, there is a cold bias at 250 m of between -1.4°C to -1.8°C for the whole domain, which is 10—

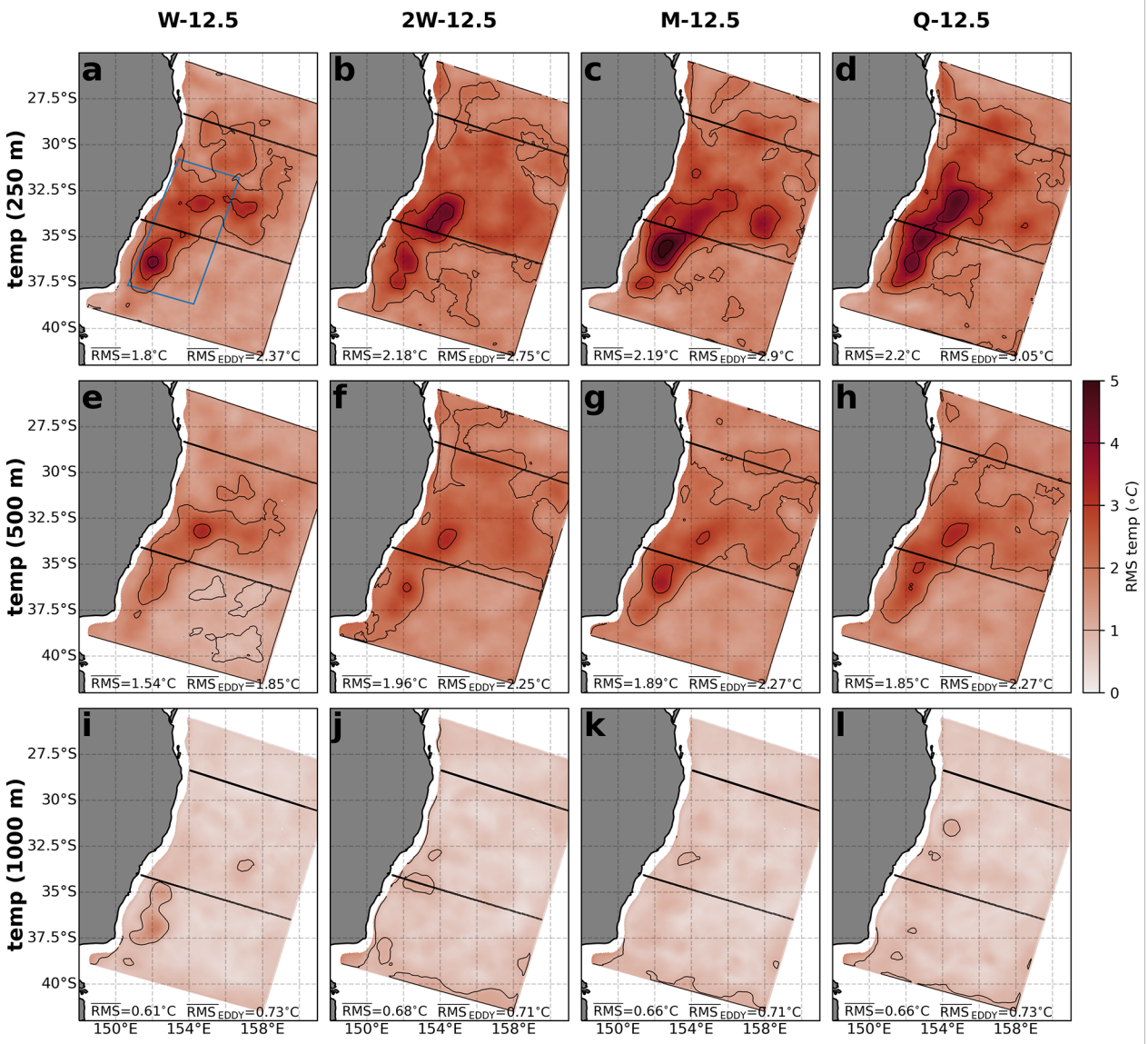


FIGURE 7
The RMS error in temperature, calculated for the time-mean, at depths of **(A–D)** 250 m, **(E–H)** 500 m, and **(I–L)** 1000 m, for the four different OSSEs: **(A, E, I)** Weekly 12.5 km, (**B, F, J**) Two-weekly 12.5 km, (**C, G, K**) Monthly 12.5 km and **(D, H, L)** Quarterly 12.5 km. RMS error is relative to the ref state, and mean values are shown in each panel for the full domain and the eddy-rich region designated by the blue box in panel **(A)**. The contour interval of 1°C is indicated with marks in the colourbar.

13% of the mean temperature at 250 m (Figures 8E–H). At 500 m (Figures 8I–L), this bias is smaller, between -1.2°C to -1.6°C – though relative to the mean temperature at 500 m it is 13—17%; higher than at 250 m). By 1000 m, this has switched to a warm bias, particularly in the eddy region (Figures 8M–P). In almost all OSSEs, particularly in the eddy region (blue box; Figure 8A), the bias is reduced by more frequent repeat times for XBT observations (cf. -1.8°C and -2.5°C; Figures 8E, H).

The bias represents the time-mean difference in conditions, which we hypothesise is imposed by the data assimilation process. It causes some oceanic features to be too cold in the subsurface–500 m range, and to be too warm at depth. As a result, subtracting this bias from OSSE temperature fields before calculating the bias-corrected RMS will give the error resulting from improper representation of the dynamics. The change in this bias-corrected RMS between the different OSSEs more clearly represents how increasingly frequent XBT repeat times better capture ocean dynamics. The bias-corrected RMS in temperature fields for each OSSE are shown in Figure 9. There is a consistent improvement in bias-corrected error at 250 m for the whole domain and for the eddy region as XBT repeat frequency is increased (Figures 9A–D). At least fortnightly

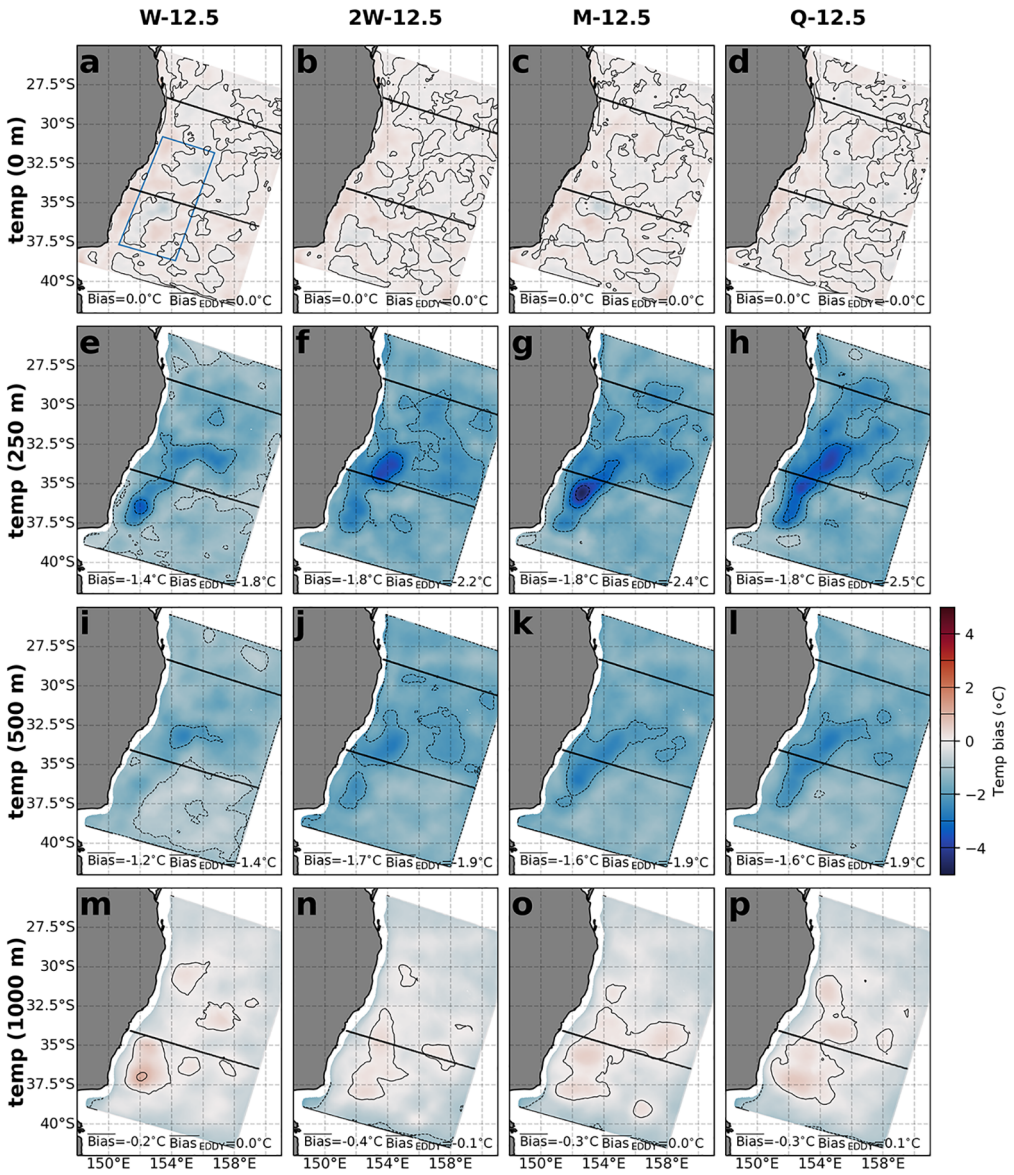


FIGURE 8
The bias in temperature at depths of **(A–D)** 0 m, **(E–H)** 250 m, **(I–L)** 500 m and **(M–P)** 1000 m, **(E–H)** 250 m, **(I–L)** 500 m and **(M–P)** 1000 m, for the four different OSSEs: **(A, E, I, M)** Weekly 12.5 km, **(B, F, J, N)** Two-weekly 12.5 km, **(C, G, K, O)** Monthly 12.5 km and **(D, H, L, P)** Quarterly 12.5 km. Bias is calculated as the time-mean difference between the OSSE and the ref state temperature fields. A positive value indicates the OSSE is warmer than the ref state, and vice versa. The mean values are shown in each panel for the full domain and the eddy-rich region designated by the blue box in panel **(A)**. The contour interval of 1°C is indicated with marks in the colourbar.

XBT observations are required to reduce RMS in the eddy region below 30°S. The improvement at 500 m and 1000 m is again greatest for weekly sampling (Figures 9E, I), though the improvement over fortnightly and slower sampling is minimal.

4 Discussion

There are different modes of variability present in the ocean, from fast (e.g. tidal) to slow (e.g. climate modes) timescales and from small (e.g. eddy cascade) to large (e.g. gyre circulation) spatial scales. When designing an observation strategy, a choice must be made as to what portion of this period-wavelength phase spectrum should be sampled. Processes that are outside of the sampled portion of the spectrum are not measured enough (in time or space) to discern the process and may be aliased into thesampled portion of the spectrum. These processes either require more rapid or finer spaced sampling (i.e. for discerning small or fast scale processes), or, longer or broader sampling (i.e. for capturing large scale or slow processes).

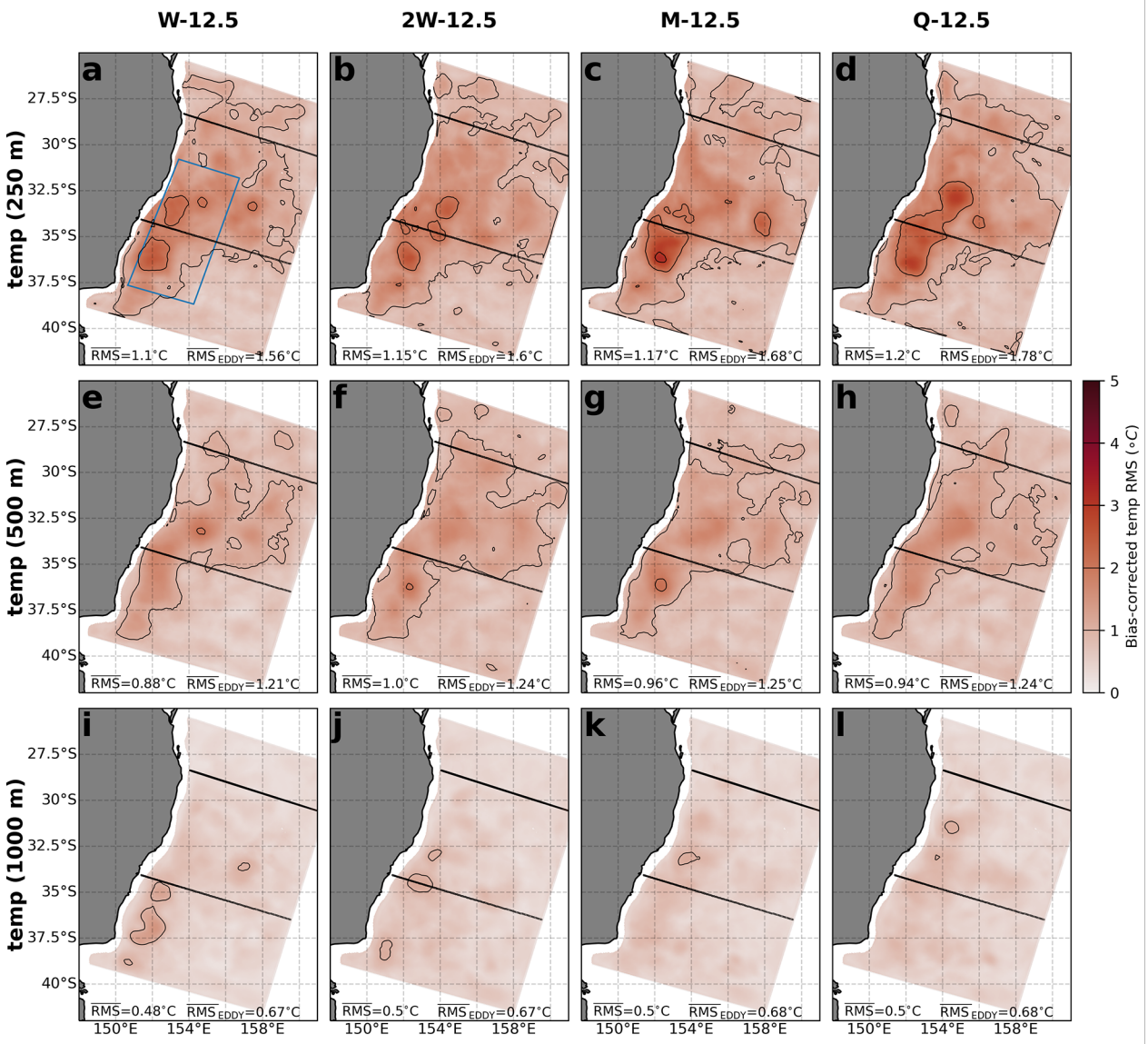


FIGURE 9
The bias-corrected RMS in temperature at depths of **(A–D)** 250 m, **(E–H)** 500 m, **(I–L)** 1000 m, for the four different OSSEs: **(A, E, I)** Weekly 12.5 km, **(B, F, J)** Two-weekly 12.5 km, **(C, G, K)** Monthly 12.5 km and **(D, H, L)** Quarterly 12.5 km.

Likewise, model resolution must be fine enough that any small scale processes that are observed can actually be resolved by the model simulation – though there could be inherent limits to predictability at very fine, submesoscale resolutions (Sandery and Sakov, 2017).

We have shown that in the EAC (using our ref state simulation), ocean surface processes display a wide variety of scales, from weekly changes to seasonal variability in SST (Figure 2A). In the subsurface ocean at 500 m, the most obvious processes are eddy dynamics, which have small scale and fast changes in temperature and EKE (Figures 2C, D). Indeed, inspection of the variability period-wavelength spectrum of EKE at 500 m shows two key features: higher power at the large scales (>100 km and 30–180 days), and a band of increased power stretching from ~350 km/monthly to 30-40 km/daily wavelengths and periods, which likely represents the increased variability associated with the Rossby-mesoscale-submesoscale energy pathway (Figure 3A). Sampling any less frequently than monthly will truncate a large portion of this mesoscale energy pathway. Given that we can quantify the amount of variability with different spatial scales, we can directly describe the portion of the total variability that is captured by repeat sampling at chosen frequencies. In particular, for EKE at 500 m, quarterly sampling captures

~20% of total variability, while increasing to monthly sampling captures ~50%. More rapid sampling rates at fortnightly and weekly frequencies captures ~70% and ~90% of the total EKE variability (Figure 4). Note that since data assimilation systems will not exactly match observations (for e.g. due to observation error), we would only expect these patterns to be approached in the long-term.

While there is no obvious reduction in transect-mean error when sampling is increased from quarterly to fortnightly, reduced RMS is noticeable with weekly sampling (Figures 5A–C). This shows that infrequent observations (i.e. quarterly through to fortnightly subsurface observation repeats) have limited impact on constraining mean subsurface conditions, and highlights the importance of regularly assimilating subsurface observations to maintain an accurate estimate.

This is further confirmed in the vertical transects of temperature RMS, where the highest error region (~250 m) is represented with similar error in quarterly, monthly and fortnightly sampling, but is improved with weekly sampling (cf. Figures 6A–D). Horizontal fields of temperature RMS show that the greatest improvement in representation is indeed in the 250 m depth range, and particularly in the high eddy energy region. This confirms that repeat sampling of at least weekly frequency is required to improve the mean representation of heat in the critical 250 m depth range in the northern upstream region (Figure 7A), separation region (Figure 7E) and southern high eddy region (Figure 7I). Indeed, this confirms the importance of higher repeat frequency observations as suggested by the variability analysis, where sampling at fortnightly–weekly frequencies is required to capture 60–80% of variability in the higher EKE region (Figures 4A, B).

Systemic error is introduced by the data assimilation system and can only be mitigated through improvements to the data assimilation system. This error could present as overly deep eddies (Siripatana et al., 2020), incorrect vertical distribution of temperature and heat content (Gwyther et al., 2022), or baroclinic mode structure (Gwyther et al., 2023). Endemic error results from the resolution of mesoscale processes and is improved by faster or higher resolution sampling. Maps of bias (Figure 8) suggest that the data assimilation process is producing a temperature structure that is too cold between 250–500 m and too warm at 1000 m. The location of the strongest bias suggests that eddies suffer disproportionately from this systemic error. The slight decrease in bias that is observed in the W-12.5 OSSE, particularly for the eddy region, indicates that weekly subsurface temperature observations are playing an important role in repeatedly correcting the vertical structure. The bias-corrected temperature RMS shows that the largest improvement in representation is in the 250 m depth range, which is achieved by taking XBT measurements at faster than monthly frequency. Improvement at 500 m is smaller, though, as most error is concentrated in the 250 m range, this may be acceptable. The improvement in increasing XBT sampling time from quarterly to monthly (see reduction in RMS in the separation zone; Figure 9D) may represent improvements in capturing seasonal-scale processes, such as mean separation latitude (Ypma et al., 2016; Oke et al., 2019), jet core velocity and EKE (Archer et al., 2017). The comparatively larger reduction in RMS observed when moving to fortnightly or weekly sampling likely represents better representation of mesoscale dynamics. This is supported by the frequency analysis (Section 3.1) and by the largest RMS reduction being in the eddy region (Figures 9A, B). In Figure 7, the greatest improvement in RMS error, particularly in the eddy region, is achieved by weekly sampling. In contrast, Figure 9 shows that fortnightly sampling achieves sufficient reduction of bias-corrected RMS error in the eddy region. This indicates that the gain achieved from moving from fortnightly to weekly sampling is through reduction of the bias – likely due to the presence of subsurface observations in each assimilation window.

The large temperature bias (Figure 8) likely results from the data assimilation process itself. This could result from the specification of background error covariance matrix, which controls the influence of observations in the horizontal and vertical (Bannister, 2008) and the weighting of the model forecast (Lee and Huang, 2020). Correctly specifying the background error covariance matrix is a major challenge to accurate data assimilation simulations (see review in Moore et al., 2019). Indeed, several studies have shown that 3-dimensional structure suffers in data assimilation, even with the inclusion of the limited subsurface observations (see for e.g. Zavala-Garay et al., 2012; Pilo et al., 2018; Siripatana et al., 2020; de Paula et al., 2021; Gwyther et al., 2022; Gwyther et al., 2023).

In any case, the result is that subsurface representation needs to be constrained frequently (i.e. weekly), otherwise subsurface structure drifts too far from the truth and any improvement from observations is minimal, though this result may change for different assimilation techniques. This suggests that improvements to assimilation schemes could improve representation of the subsurface structure even in the absence of observations. Furthermore, existing observations (i.e. quarterly subsurface measurements) could have more impact than they currently do, and potentially, sampling on the mesoscale timescale would be sufficient.

5 Conclusion

Operational and research-focussed data assimilation systems benefit greatly from the high spatial resolution, coverage and relatively rapid temporal repeat times of satellite-inferred sea-surface measurements. As a result, these systems can provide accurate and representative estimates of sea surface temperature and height (and hence geostrophic currents). However, subsurface conditions such as temperature and velocities (Gwyther et al., 2022), isothermal slopes and baroclinic mode structure (Gwyther et al.,

2023) are not represented with the same accuracy. These simulation inaccuracies are further compounded when trying to simulate dynamically complex 3-dimensional structures such as eddies (Gwyther et al., 2023). As a result, the assimilation of subsurface observations is critical to improving representation. While some subsurface observations are routinely assimilated into operational ocean models, many of these are sampling sparsely (e.g. Argo floats), or were designed for climatological studies(e.g. the Scripps high resolution XBT program). So, there is some justification for the re-purposing of existing subsurface sampling frameworks, at potentially faster repeat times, so as to be of more use for operational models that estimate and forecastocean conditions.

Our results have showed that frequencies of variability of interest need to be considered when assimilating subsurface observations. We have explored the benefit of assimilating data from an XBT observing network designed to observe climatological-scale processes i.e. heat fluxes through ocean basins over interannual periods. Our results show that assimilating XBT data at fortnightly repeat times would give a better representation of higher frequency processes such as mesoscale eddies. However, we also show that systemic errors in the data assimilation process itself limit the ability to represent accurate vertical structure. Improvements to the assimilation scheme to reduce systemic error and biases would therefore increase the positive benefit obtained from current and future subsurface observation systems.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

DG, MR, CK, SK conceived the study and designed and experiments. DG and CK configured the data assimilating model. DG performed the analyses. DG and MR wrote the first draft of the manuscript. All authors contributed to the manuscript editing and revisions. All authors contributed to the article and approved the submitted version.

Funding

This research and DG were partially supported by Australian Research Council Industry Linkage Grant LP170100498 to MR, SK, and CK.

Acknowledgments

Former model development was supported by Australian Research Council grants DP140102337 and LP160100162 to MR. This research was undertaken with the assistance of resources and services from the National Computational Infrastructure (NCI), under the grant *fu5*, and computations using the computational cluster Katana (https://doi.org/10.26190/669x-a286), supported by Research Technology Services at UNSW Sydney.

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2022.1084784/full#supplementary-material

References

Abernathey, R., and Haller, G. (2018). Transport by lagrangian vortices in the eastern pacific. *J. Phys. Oceanography* 48, 667–685. doi: 10.1175/jpo-d-17-0102.1

Archer, M. R., Roughan, M., Keating, S. R., and Schaeffer, A. (2017). On the variability of the East Australian current: Jet structure, meandering, and influence on shelf circulation. *J. Geophysical Research: Oceans* 122, 8464–8481. doi: 10.1002/2017jc013097

Balmaseda, M. A., Mogensen, K., and Weaver, A. T. (2013). Evaluation of the ECMWF ocean reanalysis system ORAS4. *Q. J. R. Meteorological Soc.* 139, 1132–1161. doi: 10.1002/qj.2063

Bannister, R. N. (2008). A review of forecast error covariance statistics in atmospheric variational data assimilation. I: Characteristics and measurements of forecast error covariances. *Q. J. R. Meteorological Soc.* 134, 1951–1970. doi: 10.1002/qj.339

Brassington, G., Pugh, T., Spillman, C., Schulz, E., Beggs, H., Schiller, A., et al. (2007). BLUElink> development of operational oceanography and servicing in Australia. *J. Res. Pract. Inf. Technol.*. 39(2):151–164.

Brieva, D., Ribbe, J., and Lemckert, C. (2015). Is the East Australian current causing a marine ecological hot-spot and an important fisheries near Fraser island, Australia? *Estuarine Coast. Shelf Sci.* 153, 121–134. doi: 10.1016/j.ecss.2014.12.012

Carton, J. A., Chepurin, G., Cao, X., and Giese, B. (2000). A simple ocean data assimilation analysis of the global upper ocean 1950–95. part I: Methodology. *J. Phys. Oceanography* 30, 294–309. doi: 10.1175/1520-0485(2000)030<0294:asodaa>2.0.co;2

Chamberlain, M. A., Oke, P. R., Fiedler, R. A. S., Beggs, H. M., Brassington, G. B., and Divakaran, P. (2021). Next generation of bluelink ocean reanalysis with multiscale data assimilation: BRAN2020. *Earth System Sci. Data* 13, 5663–5688. doi: 10.5194/essd-13-5663-2021

de Paula, T. P., Lima, J. A. M., Tanajura, C. A. S., Andrioni, M., Martins, R. P., and Arruda, W. Z. (2021). The impact of ocean data assimilation on the simulation of mesoscale eddies at são paulo plateau (Brazil) using the regional ocean modeling system. *Ocean Model.* 167, 101889. doi: 10.1016/j.ocemod.2021.101889

Elsner, J. B., and Tsonis, A. A. (1996). *Singular spectrum analysis: a new tool in time series analysis* (New York: Springer Science & Business Media). doi: 10.1007/978-1-4757-2514-8

Everett, J. D., Baird, M. E., Oke, P. R., and Suthers, I. M. (2012). An avenue of eddies: Quantifying the biophysical properties of mesoscale eddies in the Tasman Sea. *Geophysical Research Letters* 39(16). doi: 10.1029/2012GL053091

Fairall, C. W., Bradley, E. F., Rogers, D. P., Edson, J. B., and Young, G. S. (1996). Bulk parameterization of air-sea fluxes for tropical ocean-global atmosphere coupled-ocean atmosphere response experiment. *J. Geophysical Research: Oceans* 101, 3747–3764. doi: 10.1029/95jc03205

Frenger, I., Gruber, N., Knutti, R., and Münnich, M. (2013). Imprint of southern ocean eddies on winds, clouds and rainfall. *Nat. Geosci.* 6, 608–612. doi: 10.1038/ngeo1863

Gasparin, F., Guinehut, S., Mao, C., Mirouze, I., Rémy, E., King, R. R., et al. (2019). Requirements for an integrated *in situ* Atlantic ocean observing system from coordinated observing system simulation experiments. *Front. Mar. Sci.* 6. doi: 10.3389/fmars.2019.00083

Gwyther, D. E., Keating, S. R., Kerry, C., and Roughan, M. (2023) How does 4DVar data assimilation affect the vertical representation of mesoscale eddies? A case study with observing system simulation experiments (OSSEs) using ROMS v3.9. *Geosci. Model. Dev.* 16:157–178 doi: 10.5194/gmd-16-157-2023

Gwyther, D. E., Kerry, C., Roughan, M., and Keating, S. R. (2022). Observing system simulation experiments reveal that subsurface temperature observations improve estimates of circulation and heat content in a dynamic western boundary current. *Geoscientific Model. Dev.* 15, 6541–6565. doi: 10.5194/gmd-15-6541-2022

Halliwell, G. R., Kourafalou, V., Hénaff, M. L., Shay, L. K., and Atlas, R. (2015). OSSE impact analysis of airborne ocean surveys for improving upper-ocean dynamical and thermodynamical forecasts in the gulf of Mexico. *Prog. Oceanography* 130, 32–46. doi: 10.1016/j.pocean.2014.09.004

Halliwell, G. R., Srinivasan, A., Kourafalou, V., Yang, H., Willey, D., Hénaff, M. L., et al. (2014). Rigorous evaluation of a fraternal twin ocean OSSE system for the open gulf of Mexico. *J. Atmospheric Oceanic Technol.* 31, 105–130. doi: 10.1175/jtech-d-13-00011.1

Kerry, C., Powell, B., Roughan, M., and Oke, P. (2016). Development and evaluation of a high-resolution reanalysis of the East Australian current region using the regional ocean modelling system (ROMS 3.4) and incremental strong-constraint 4-dimensional variational (IS4D-var) data assimilation. *Geoscientific Model. Dev.* 9, 3779–3801. doi: 10.5194/gmd-9-3779-2016

Kerry, C., and Roughan, M. (2020). Downstream evolution of the East Australian current system: Mean flow, seasonal, and intra-annual variability. *J. Geophysical Research: Oceans* 125, e2019JC015227. doi: 10.1029/2019jc015227

Kerry, C., Roughan, M., and Powell, B. (2018). Observation impact in a regional reanalysis of the East Australian current system. *J. Geophysical Research: Oceans* 123, 7511–7528. doi: 10.1029/2017jc013685

Kerry, C., Roughan, M., and Powell, B. (2020). Predicting the submesoscale circulation inshore of the East Australian current. *J. Mar. Syst.* 204, 103286. doi: 10.1016/j.jmarsys.2019.103286

Klocker, A., and Abernathey, R. (2014). Global patterns of mesoscale eddy properties and diffusivities. *J. Phys. Oceanography* 44, 1030–1046. doi: 10.1175/jpo-d-13-0159.1

Lambaerts, J., Lapeyre, G., Plougonven, R., and Klein, P. (2013). Atmospheric response to sea surface temperature mesoscale structures. *J. Geophysical Research: Atmospheres* 118, 9611–9621. doi: 10.1002/jgrd.50769

Lee, J. C. K., and Huang, X.-Y. (2020). Background error statistics in the tropics: Structures and impact in a convective-scale numerical weather prediction system. *Q. J. R. Meteorological Soc.* 146, 2154–2173. doi: 10.1002/qj.3785

Li, J., Roughan, M., and Kerry, C. (2021). Dynamics of interannual eddy kinetic energy modulations in a Western boundary current. *Geophysical Res. Lett.* 48, e2021GL094115. doi: 10.1029/2021gl094115

Li, J., Roughan, M., and Kerry, C. (2022a). Drivers of ocean warming in the western boundary currents of the southern hemisphere. *Nat. Climate Change* 12, 901–909. doi: 10.1038/s41558-022-01473-8

Li, J., Roughan, M., Kerry, C., and Rao, S. (2022b). Impact of mesoscale circulation on the structure of river plumes during large rainfall events inshore of the east australian current. *Front. Mar. Sci.* 9. doi: 10.3389/fmars.2022.815348

Malan, N., Archer, M., Roughan, M., Cetina-Heredia, P., Hemming, M., Rocha, C., et al. (2020). Eddy-driven cross-shelf transport in the East Australian current separation zone. *J. Geophysical Research: Oceans* 125, e2019JC015613. doi: 10.1029/2019jc015613

Moore, A. M., Arango, H. G., Broquet, G., Powell, B. S., Weaver, A. T., and Zavala-Garay, J. (2011). The regional ocean modeling system (ROMS) 4-dimensional variational data assimilation systems: Part I - system overview and formulation. *Prog. Oceanography* 91, 34–49. doi: 10.1016/j.pocean.2011.05.004

Moore, A. M., Martin, M. J., Akella, S., Arango, H. G., Balmaseda, M., Bertino, L., et al. (2019). Synthesis of ocean observations using data assimilation for operational, real-time and reanalysis systems: A more complete picture of the state of the ocean. *Front. Mar. Sci.* 6. doi: 10.3389/fmars.2019.00090

Moore, A. M., Zavala-Garay, J., Arango, H. G., Edwards, C. A., Anderson, J., and Hoar, T. (2020). Regional and basin scale applications of ensemble adjustment kalman filter and 4D-var ocean data assimilation systems. *Prog. Oceanography* 189, 102450. doi: 10.1016/j.pocean.2020.102450

Morris, M., Roemmich, D., and Cornuelle, B. (1996). Observations of variability in the south pacific subtropical gyre. *J. Phys. Oceanography* 26, 2359–2380. doi: 10.1175/1520-0485(1996)026<2359:oovits>2.0.co;2

Oke, P. R., Roughan, M., Cetina-Heredia, P., Pilo, G. S., Ridgway, K. R., Rykova, T., et al. (2019). Revisiting the circulation of the East Australian current: Its path, separation, and eddy field. *Prog. Oceanography* 176, 102139. doi: 10.1016/j.pocean.2019.102139

Oke, P. R., Sakov, P., Cahill, M. L., Dunn, J. R., Fiedler, R., Griffin, D. A., et al. (2013). Towards a dynamically balanced eddy-resolving ocean reanalysis: BRAN3. *Ocean Model.* 67, 52–70. doi: 10.1016/j.ocemod.2013.03.008

Oliver, H., Zhang, W. G., Smith, W. O., Alatalo, P., Chappell, P. D., Hirzel, A. J., et al. (2021). Diatom hotspots driven by Western boundary current instability. *Geophysical Res. Lett.* 48, e2020GL091943. doi: 10.1029/2020gl091943

Pilo, G. S., Oke, P. R., Coleman, R., Rykova, T., and Ridgway, K. (2018). Impact of data assimilation on vertical velocities in an eddy resolving ocean model. *Ocean Model.* 131, 71–85. doi: 10.1016/j.ocemod.2018.09.003

Puri, K., Dietachmayer, G., Steinle, P., Dix, M., Rikus, L., Logan, L., et al. (2013). Implementation of the initial access numerical weather prediction system. *Aust. Meteorological Oceanographic J.* 63, 265–284. puri2013implementation. doi: 10.22499/2.6302.001

Rocha, C., Edwards, C. A., Roughan, M., Cetina-Heredia, P., and Kerry, C. (2019). A high-resolution biogeochemical model (ROMS 3.4+bio_Fennel) of the East Australian current system. *Geoscientific Model. Dev.* 12, 441–456. doi: 10.5194/gmd-12-441-2019

Roemmich, D., and Cornuelle, B. (1990). Observing the fluctuations of gyre-scale ocean circulation: A study of the subtropical south pacific. *J. Phys. Oceanography* 20, 1919–1934. doi: 10.1175/1520-0485(1990)020<1919:otfogs>2.0.co;2

Roemmich, D., Gilson, J., Willis, J., Sutton, P., and Ridgway, K. (2005). Closing the time-varying mass and heat budgets for large ocean areas: The tasman box. *J. Climate* 18, 2330–2343. doi: 10.1175/jcli3409.1

Sandery, P. A., and Sakov, P. (2017). Ocean forecasting of mesoscale features can deteriorate by increasing model resolution towards the submesoscale. *Nat. Commun.* 8:1566. doi: 10.1038/s41467-017-01595-0

Shchepetkin, A. F., and McWilliams, J. C. (2005). The regional oceanic modeling system (ROMS): a split-explicit, free-surface, topography-following-coordinate oceanic model. *Ocean Model.* 9, 347–404. doi: 10.1016/j.ocemod.2004.08.002

Siripatana, A., Kerry, C., Roughan, M., Souza, J. M. A. C., and Keating, S. (2020). Assessing the impact of nontraditional ocean observations for prediction of the East Australian current. *J. Geophysical Research: Oceans* 125, e2020JC016580. doi: 10.1029/2020jc016580

Smith, N. R., Harrison, D., Bailey, R., Alves, O., Delcroix, T., Hanawa, K., et al. (1999). *The role of XBT sampling in the ocean thermal network* (White paper, OceanObs99 Saint Raphaël, France).

Su, C.-H., Eizenberg, N., Steinle, P., Jakob, D., Fox-Hughes, P., White, C. J., et al. (2019). BARRA v1.0: the bureau of meteorology atmospheric high-resolution regional reanalysis for Australia. *Geoscientific Model. Dev.* 12, 2049–2068. doi: 10.5194/gmd-12-2049-2019

Torres, H. S., Klein, P., Menemenlis, D., Qiu, B., Su, Z., Wang, J., et al. (2018). Partitioning ocean motions into balanced motions and internal gravity waves: A modeling study in anticipation of future space missions. *J. Geophysical Research: Oceans* 123, 8084–8105. doi: 10.1029/2018jc014438

Weaver, A., and Courtier, P. (2001). Correlation modelling on the sphere using a generalized diffusion equation. *Q. J. R. Meteorological Soc.* 127, 1815–1846. doi: 10.1002/qj.49712757518

Whiteway, T. (2009) *Australian Bathymetry and topography grid, june 2009*. Available at: http://dx.doi.org/10.4225/25/53D99B6581B9A (Accessed 08-02-2021). Dataset.

Wijffels, S., Roemmich, D., Monselesan, D., Church, J., and Gilson, J. (2016). Ocean temperatures chronicle the ongoing warming of earth. *Nat. Climate Change* 6, 116–118. doi: 10.1038/nclimate2924

Wong, A. P. S., Wijffels, S. E., Riser, S. C., Pouliquen, S., Hosoda, S., Roemmich, D., et al. (2020). Argo data 1999–2019: Two million temperature-salinity profiles and subsurface velocity observations from a global array of profiling floats. *Front. Mar. Sci.* 7. doi: 10.3389/fmars.2020.00700

Young, J., Hobday, A., Campbell, R., Kloser, R., Bonham, P., Clementson, L., et al. (2011). The biological oceanography of the East Australian current and surrounding waters in relation to tuna and billfish catches off eastern Australia. *Deep Sea Res. Part II: Topical Stud. Oceanography* 58, 720–733. doi: 10.1016/j.dsr2.2010.10.005

Ypma, S. L., van Sebille, E., Kiss, A. E., and Spence, P. (2016). The separation of the East Australian current: A Lagrangian approach to potential vorticity and upstream control. *J. Geophysical Research: Oceans* 121, 758–774. doi: 10.1002/2015jc011133

Zavala-Garay, J., Wilkin, J. L., and Arango, H. G. (2012). Predictability of mesoscale variability in the East Australian current given strong-constraint data assimilation. *J. Phys. Oceanography* 42, 1402–1420. doi: 10.1175/jpo-d-11-0168.1

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 23 February 2023
DOI 10.3389/fmars.2023.1085542

Check for updates

OPEN ACCESS

EDITED BY
Shiqiu Peng,
State Key Laboratory of Tropical Oceanography, Chinese Academy of Sciences, China

REVIEWED BY
Toru Miyama,
Japan Agency for Marine-Earth Science and Technology, Japan
Shijian Hu,
Institute of Oceanology (CAS), China

*CORRESPONDENCE
Young Ho Kim
yhokim@pknu.ac.kr

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

RECEIVED 31 October 2022
ACCEPTED 26 January 2023
PUBLISHED 23 February 2023

CITATION
Chang I, Kim YH, Jin H, Park Y-G, Pak G and Chang Y-S (2023) Impact of satellite and regional *in-situ* profile data assimilation on a high-resolution ocean prediction system in the Northwest Pacific. *Front. Mar. Sci.* 10:1085542. doi: 10.3389/fmars.2023.1085542

COPYRIGHT
© 2023 Chang, Kim, Jin, Park, Pak and Chang. 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(s) 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.

Impact of satellite and regional *in-situ* profile data assimilation on a high-resolution ocean prediction system in the Northwest Pacific

Inseong Chang[1], Young Ho Kim[1]*, Hyunkeun Jin[2], Young-Gyu Park[2], Gyundo Pak[2] and You-Soon Chang[3]

[1]Division of Earth Environmental System Science, Pukyong National University, Busan, Republic of Korea, [2]Ocean Circulation Research Center, Korea Institute of Ocean Sciences and Technology, Busan, Republic of Korea, [3]Department of Earth Science Education, Kongju National University, Gongju, Republic of Korea

The impacts of observation data sets on the high-resolution (1/24°) Northwest Pacific prediction system were investigated with the model sensitivity tests. We compared the model experiments assimilating the different combinations of the observation data sets, such as the sea surface height derived from satellite altimetry, sea surface temperature, and *in-situ* profiles, based on the Ensemble Optimal Interpolation. Pseudo-profiles constructed by the method of Cooper and Haines (1996, CH96) were assimilated into the model to assimilate sea surface height data. CH96 applied a conservation principle to derive pseudo-profiles by rearranging preexisting profiles. The comparison of the model experiments suggests that each observation data set enhances the model performance. Especially, the assimilation of the sea surface height reduces the model error by more than 9.81% and 6.44%, respectively, in terms of the root-mean-square error of the ocean temperature and salinity in the subsurface layer. It is interesting that the assimilation of the *in-situ* temperature profiles in the Korean marginal seas contributes to improving the reproducibility of the subsurface temperature and salinity in the East/Japan Sea (EJS) as well as Kuroshio-Kuroshio Extension (K-KE) regions. The improvement in the K-KE region seems to be related to the reproducibility of the Kuroshio axis. As the water mass in the EJS flows into the Pacific Ocean through the Tsugaru Strait, it affects the front of the Sanriku confluence, and it seems to eventually control the Oyashio Current and Kuroshio axis. In conclusion, this study evaluated the contribution of each observation component to ocean analysis in the KOOS-OPEM and confirmed the role of the existing observation networks. This study also suggests that greater attention should be paid to the role of regional ocean observation networks to improve the forecast skill of the ocean prediction system not only in the region but also in the open ocean, such as the Pacific Ocean.

KEYWORDS

satellite altimetry, data assimilation, ensemble optimal interpolation, KOOS-OPEM, regional ocean observing networks

1 Introduction

The Northwest Pacific (NWP) has a complicated ocean circulation system consisting of several major currents including the North Equatorial Countercurrent, Subtropical Countercurrent, Oyashio current (OYC), and the Kuroshio Current (Figure 1). The Kuroshio current is a strong western boundary current, accompanied by energetic variability associated with mesoscale features such as eddies and meander. In addition, four major marginal seas, the South China Sea (SCS), the East China Sea, the Yellow Sea, and the East/Japan Sea (EJS), are connected by four narrow straits: Taiwan, Korea/Tsushima, Tsugaru, and Soya Strait. As ocean currents between the marginal seas distribute heat, salt, and other material through straits, it is important to determine the circulation and variability in the marginal seas and their relation to each other and the NWP (Cho et al., 2009). The Korea Institute of Ocean Science and Technology (KIOST) has developed the Korea Operational Oceanographic System–Ocean Predictability Experiment for Marine environment (KOOS-OPEM, Park et al., 2010; Kim et al., 2021), a high-resolution ocean prediction system for the NWP ocean, to understand the complicated ocean circulation system of the NWP ocean and the relationship between the marginal seas and the NWP and to rapidly respond to extreme marine events and accidents.

Ocean models are imperfect and inevitable sources of uncertainties, such as initial and boundary conditions, model parameterization, and force fields, which may affect their outputs. An efficient strategy to address these uncertainties and improve forecasts is to assimilate available satellite and *in-situ* observation data into ocean models, thereby optimizing the initial conditions of the ocean model. Generally, data assimilation techniques are classified into two types: variational and sequential. One of the most commonly used variational methods is 4D-VAR whereas the Ensemble Kalman Filter (EnKF), introduced by Evensen (1994) and Burgers et al. (1998), is the most commonly used sequential method. However, these methods are limited because they are computationally expensive. Therefore, the ensemble optimal interpolation (EnOI), a simplification of the EnKF method, was proposed by Evensen (2003) and provides a cost-effective alternative to the EnKF. EnOI estimates the background error covariance by using a time-invariant ensemble of model states sampled from long-term model results. EnOI has many advantages, such as inherent multivariate, quasi-dynamically consistent, inhomogeneous, and anisotropic covariance. Accordingly, EnOI has been adopted in many operational ocean forecast systems, such as the data assimilation system of KIOST (DASK, Kim et al., 2015) and Bluelink Ocean Data Assimilation System (BODAS, Oke et al., 2008).

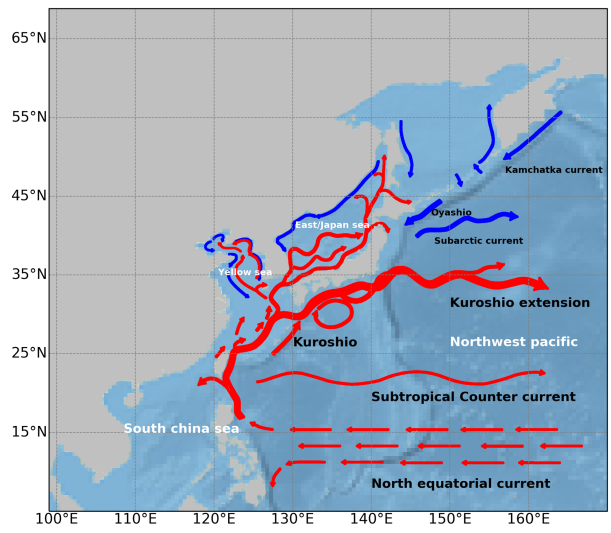


FIGURE 1
Ocean current map in Northwest pacific is from Park et al. (2013).

Traditionally, observational data used in ocean data assimilation include sea surface temperature (SST), temperature/salinity (T/S) profiles, and sea surface height (SSH) derived from satellite altimetry data. SST observations have been widely used in ocean data assimilation because SST is an important ocean variable that connects many processes, such as the air-sea exchange of energy and the formation of water mass in the upper ocean. The assimilation of the T/S profiles improves the representation of seawater density, indicating the mass of water. In particular, the profiles directly affect the ocean heat content (Zhou et al., 2021). The assimilation of SSH improves the ocean surface currents. In addition, these data can reflect the state of the subsurface structure, which provides a physical foundation for improving the temperature and salinity structures of ocean prediction models *via* assimilation (Troccoli and Haines, 1999). However, it is difficult to project satellite altimetry data onto subsurface density structures. In addition, because the assimilation of satellite altimetry data does not impose constraints on the vertical density structure, consistent adjustments of temperature/salinity should be considered to maintain the density structure. However, when the temperature changes due to the assimilation of satellite altimetry data, adjustment of salinity is required to conserve the correct density stratification (Troccoli et al., 2002; Vialard et al., 2003). Therefore, if the data assimilation system is not properly applied, the assimilation of satellite altimetry data can negatively affect the salinity field while improving the temperature (Fu and Zhu, 2011).

Various methods have been proposed to address these problems and effectively assimilate satellite altimetry data. For example, the direct assimilation method for satellite altimetry data using ensemble-based techniques (Parent et al., 2003; Testut et al., 2003; Birol et al., 2005; Oke et al., 2008; Counillon and Bertino, 2009; Zheng and Zhu, 2015) achieved the projection of surface information onto subsurface structures through the inherent multivariate relation derived from the ensembles. Yan et al. (2004) have proposed an assimilation method based on 3D-VAR by developing a statistical relationship between the SSH and the subsurface structure of temperature and salinity. Cooper and Haines (1996) (hereafter CH96) applied the conservation principle to project altimetry information into subsurface structures. CH96 derived pseudo-profiles based on the rearrangement of the preexisting water columns of the model while preserving the water properties. The pseudo-profiles at the specific grid points are constructed by vertically replacing the preexisting profile, which reduces the surface pressure difference between the modeled and observed SSH with conserving the bottom pressure.

In this study, to quantitatively evaluate the contribution of *in-situ* T/S profiles and satellite observation data, such as SST and satellite altimetry, to the ocean prediction system, we conducted a sensitivity

experiment by applying EnOI to KOOS-OPEM. El-Geziry and Bryden (2010) stated that Mediterranean circulation affects not only the Mediterranean basin on a regional scale but also the global circulation due to its effective contribution to the North Atlantic circulation system. Like the Mediterranean Sea, the EJS is a semi-enclosed marginal sea exchanging water mass with the open ocean through narrow straits. This study also aimed to investigate the extent to which assimilation of *in-situ* temperature profile data in Korean marginal seas affects not only the Korean marginal seas including the EJS but also the open ocean such as the NWP. The remainder of this paper is organized as follows: the model configuration, details of the data assimilation system, experimental design, and observation data used in this experiment are described in Section 2 and the results of all experiments are compared with respect to observations in Section 3, and Section 4 discusses the results and summarizes the major conclusions.

2 Materials and methods

2.1 Model and data assimilation system

2.1.1 Model configuration

The KOOS-OPEM is based on the Modular Ocean Model Version 5 (MOM5) developed by the Geophysical Fluid Dynamics Laboratory (GFDL) and includes the Sea Ice Simulator (Winton, 2000). This model solves primitive equations with hydrostatic and Boussinesq approximations using the Arakawa-B grid system (Arakawa and Lamb, 1977). The horizontal domain of this model covers the NWP including the Korean marginal seas and the Yellow and East China Seas, ranging from 5°N to 63°N and 99°E to 170°E, with a horizontal resolution of 1/24° in latitude and longitude. This model has a z-star coordinate system of 51 vertical levels with a finer resolution near the sea surface. K-profile parameterization was employed for the vertical mixing scheme (Troen and Mahrt, 1986; Large et al., 1994). This model uses tidal mixing parameterization as implemented by Simmons et al. (2004), the GM isopycnal mixing scheme for tracer mixing (Gent and Mcwilliams, 1990), and the Smagorinsky biharmonic scheme for horizontal momentum mixing (Griffies and Hallberg, 2000). The model topography was generated by merging the GEBCO 08 data set (http://www.gebco.net) and the Korbathy regional dataset (Seo, 2008).

This model was forced by lateral boundary conditions from the Global Ocean Ensemble Physics Reanalysis, including the daily ocean velocity, temperature, salinity, and SSH provided by the Copernicus Marine Environment Monitoring Service (CMEMS, https://marine.copernicus.eu/). The surface boundary forcings were calculated by applying the bulk formula of Large and Yeager (2004) to the atmospheric variables, including air temperature, specific humidity, surface net solar radiation, surface thermal radiation, snowfall, runoff, mean sea level pressure, total cloud cover, wind velocity, and precipitation from ERA5, provided by the European Center for Medium-Range Weather Forecasts (ECMWF). To reflect runoff and river discharge, climatological data of 40 rivers obtained from RivDIS (Vörösmarty et al., 1998) were inserted into the ocean grid.

2.1.2 EnOI

The basic equation for data assimilation is as follows:

$$\omega^a = \omega^f + \mathrm{K}(d - \mathrm{H}\omega^f) \tag{1}$$

$$\mathrm{K} = \mathrm{P}^f\mathrm{H}^T(\mathrm{H}\mathrm{P}^f\mathrm{H}^T + \mathrm{R})^{-1} \tag{2}$$

where ω^f and ω^a represent the forecast and analytical state vectors, respectively; d is the vector of observations; K is the Kalman gain matrix; H is an operator that transforms from the model grid to the observation grid. Matrices P^f and R represent the background error covariance and observation error covariance, respectively. The superscript T denotes the matrix transpose.

EnOI is based on the work of Evensen (2003) and the analysis approach of Burgers et al. (1998). The practical implementation of the EnOI is similar to the EnKF; however, the EnOI analysis is estimated in a stationary ensemble composed of long-time model integration, and statistical errors do not develop over time. EnOI estimates the background covariance error matrix P^f as follows:

$$\mathrm{P}^f = \alpha\mathrm{A}'\mathrm{A}'^T/(N_e - 1) \tag{3}$$

$$\mathrm{A}' = [\omega_1', \omega_2', \ldots, \omega_{N_e}'] \tag{4}$$

where N_e is the ensemble size, α ($\in$0,1) is introduced as a scaling factor to reduce the variance (Counillon and Bertino, 2009), A' is a stationary and historical ensemble, and ω_i' is the i^{th} model anomaly. In this study, the scaling factor α was considered 0.25 (Oke et al., 2005; Kim et al., 2015). The ensemble members were selected from the monthly mean historical data of the 50-months model integration, and the climatological monthly mean was removed.

An ensemble-based data assimilation system can highly estimate correlations between long-distance points that are likely to be independent of each other when a small number of ensemble members are used (Kim et al., 2008). To reduce this sampling error, we applied localization techniques (Hamill et al., 2001; Houtekamer and Mitchell, 2001; Oke et al., 2007; Kim et al., 2015) of the Schur product (Gaspari and Cohn, 1999). After applying localization, the Kalman gain matrix was expressed as follows:

$$\rho \circ \begin{bmatrix} b_{1,1} & \cdots & b_{i,1} & \cdots & b_{M,1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ b_{1,j} & \cdots & b_{i,j} & \cdots & b_{M,j} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ b_{1,M} & \cdots & b_{i,M} & \cdots & b_{M,M} \end{bmatrix} = \rho_{i,j}[L(x_i, x_j)] \times b_{i,j} \tag{5}$$

$$\mathrm{K} = \rho \circ \mathrm{A}'(\mathrm{HA}')^T(\rho \circ \mathrm{HA}'(\mathrm{HA}')^T + (N_e - 1)R)^{-1} \tag{6}$$

where ρ ° B is calculated by ρ, a function of the distance (L) between x_i and xj. ρ is given as a function of the horizontal and vertical decorrelation distance, which has a value of 0–1. As the distance between two points increases, ρ becomes closer to 0, and when the distance exceeds a certain length, ρ becomes 0. In the ocean prediction model, ρ can be applied to the horizontal and vertical decorrelation distances. In this study, the horizontal and vertical decorrelation distances were 150 km and 100 m, respectively.

2.1.3 Altimetry assimilation method

The method used to assimilate satellite altimetry data was based on the CH96 scheme. To construct the pseudo-profiles, CH96 assumes that the bottom pressure and potential vorticity are preserved. Following CH96, the bottom constraint is as below:

$$g\int_{0}^{-H}\Delta\rho dz = \Delta p_s \tag{7}$$

where $\Delta\rho$, Δp_s, and H are density increment of the water column, change in surface pressure, and depth of the water column, respectively. The change in SSH should be compensated for by the change in the weight of the entire water column. If the model SSH is higher than the observed SSH, the water columns of the model are displaced upward, and some light water masses at the surface are removed and replaced by heavy water masses at the bottom. Similarly, if the model SSH was lower than the observed SSH, the water columns of the model were displaced downward to decrease the density of the water column. The amount of vertical displacement of the water columns was such that the bottom pressure did not change. The displacement Δh is expressed as follows:

$$\Delta h = \frac{\Delta p_s}{g[\rho(0) - \rho(-H)]'} \tag{8}$$

The pseudo-profiles created by CH96 were assimilated into the ocean prediction model instead of SSH observation data through EnOI. The characteristic of CH96 is that convection does not occur because the density structure is preserved, except for the top and bottom, which are removed or added according to the change in surface pressure by conserving the volume of the water mass. In addition, conserving the bottom pressure prevents changes in the bottom torques, thereby decreasing the interactions with steep topography (Alves et al., 2001).

2.2 Experiments and observation

2.2.1 Experimental design

To investigate the impact of the data assimilation of *in-situ* and satellite observations and this indirect assimilation system, we performed six experiments based on observation data from 1993. In all simulations, the model was sequentially integrated forward in time using the initial conditions, which assimilated the SST and T/S profiles from 1990 to 1992. Posterior analysis of the previous cycle was used for the prior initial conditions for each cycle.

For comparison, the first experiment (CTR) was a control experiment without assimilation. The second experiment (EXP01) assimilated the SST. The third and fourth experiments (EXP02 and EXP03) assimilated the SST and T/S profiles. The fifth and sixth experiments (EXP04 and EXP05) assimilated all the variables. To investigate how the assimilation of data obtained in the marginal sea affects the other regions, EXP03 and EXP05 additionally assimilated the Korea Oceanographic Data Center (KODC, https://www.nifs.go.kr/kodc) profile data. Details for the experiments are shown in Table 1.

2.2.2 *In-situ* data

In-situ T/S profile data, which were obtained from various platforms, including conductivity temperature and depth (CTD) sensors, ocean station data (OSD), expendable bathythermographs (XBT), and moored buoys (MRB), were obtained from the World Ocean Database 2018 (WOD). The National Institute of Fisheries Science (NIFS) has produced a KODC database by researching the marine environment, including temperature, as well as other variables such as nutrient concentration and salinity, using CTD at standard observation depths (0, 10, 20, 30, 50, 75, 100, 125, 150, 200, 250, 300, 400, and 500 m), in the seas around the Korean Peninsula six times a year (for even months) from 1961 to the present. The temperature profiles obtained from the KODC were used in this study. KODC salinity data were not used in this study because they contained serious time-dependent bias errors as previously reported by Park (2021). Figure 2A shows the KODC and WOD temperature profile data used when assimilating the profile in February. We used 852 temperature data points collected from 23 stations. These *in-situ* profile data were assimilated into KOOS-OPEM every seven days.

To compare the performance of the vertical profiles of temperature and salinity, the independent T/S profiles (depth above approximately 500 m) of WOD and KODC, not used for assimilation, were used. When selecting the validation data, one-third of the total data was randomly selected. The spatial distribution of the validation profile is shown in Figure 2B.

2.2.3 Satellite data

The National Centers for Environmental Information (NOAA)'s 1/4° daily optimum SST (OISST) generated by interpolating and extrapolating observations from different platforms such as satellites, buoys, ships, and Argo floats were assimilated into KOOS-OPEM every day. The along-track SSH

TABLE 1 Summary of the sensitivity experiment.

	SST	T/S profiles	KODC profiles	SSH (pseudo-profile)
CTR	–	–	–	–
EXP01	o	–	–	–
EXP02	o	o	–	–
EXP03	o	o	o	–
EXP04	o	o	–	o
EXP05	o	o	o	o

SST, sea surface temperature; T/S, temperature/saline; KODC, Korea Oceanographic Data center; SSH, sea surface height; CTR, control; EXP, experiment.

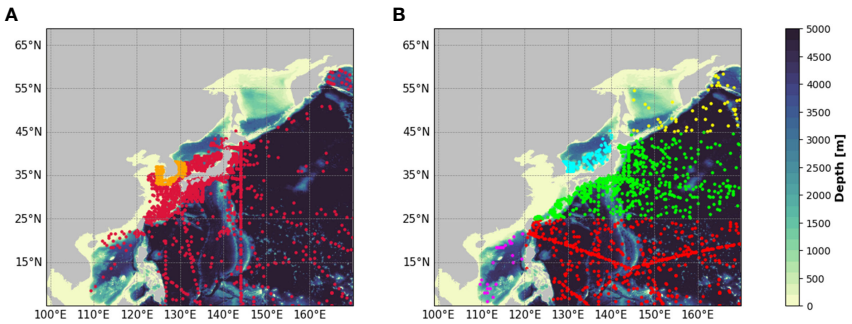


FIGURE 2
Distribution of *in-situ* temperature profiles used in data assimilation and validation. **(A)** temperature profiles used in data assimilation in February 1993. Crimson and orange dots denote profiles taken from World Ocean Database 2018 and Korea Oceanographic Data Center (KODC), respectively. **(B)** temperature profiles used in validation. Red, magenta, yellow, cyan and green dots denote temperature profiles located Northwestern Pacific (NWP), South China Sea (SCS), Oyashio Current (OYC), East/Japan Sea (EJS) and Kuroshio-Kuroshio Extension (K-KE) region, respectively.

data from TOPEX/POSEIDON and ERS-1 were downloaded from CMEMS and used to construct the pseudo-profiles. The original resolution of the along-track SSH data was approximately 7 km; however, we used along-track data subsampled every 50 km to efficiently use computational resources. The pseudo-profiles derived from the along-track data were also assimilated daily.

To compare the performance of SST, the Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA; Good et al., 2020) analysis downloaded from CMEMS was used. These data were generated from satellite and *in-situ* data through high-resolution analysis and intercomparison with spatial and temporal resolutions of 0.054° and daily, respectively. Gridded absolute dynamic topography data were used to compare the performance of the SSH. The temporal and spatial resolution of these data were one day and 1/4°, respectively. Because the gridded SSH data were generated by merging SSH measurements from multiple satellite altimetry, it can be suggested that the generated data are dependent on along-track data. However, as we did not directly assimilate the along-track data but assimilated pseudo-profiles instead, the results are independent of the gridded data.

2.2.4 Statistical metrics

The metrics used to assess assimilation performance were the root-mean-square error (RMSE) and impact of assimilation (IOA), as defined by Chen et al. (2018). These metrics are defined as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(M_i - O_i)^2} \tag{8}$$

$$\mathrm{IOA} = \frac{RMSE_{CTR} - RMSE_{EXP}}{RMSE_{CTR}} \times 100 \tag{9}$$

where M_i is the i^{th} model result of each experiment, O_i is the i^{th} observation, n is the number of values, and the overbar is the average over all periods. $RMSE_{CTR}$ and $RMSE_{EXP}$ denote the RMSEs of the CTR and assimilation experiment with respect to the observation, respectively. IOA is a metric that indicates the improvement of the RMSE compared to CTR. The larger the IOA, the greater the improvement in ocean analysis by the assimilation of each variable.

3 Results

3.1 Sea surface temperature

Figure 3 shows the comparison of the spatial distribution of the RMSE for SST with respect to OSTIA of all experiments. All experiments showed a large RMSE near the Korean and Chinese coastlines. The RMSE of the CTR was large in the middle and high-latitude regions, including the EJS and Kuroshio-Kuroshio Extension (K-KE), and was the largest in K-KE. The RMSEs of all assimilation experiments were greatly reduced in most regions by SST assimilation. In particular, the RMSEs of all experiments were significantly improved in the EJS and K-KE regions with a large RMSE in the CTR. However, the difference in the RMSE of SST is not significant in Figure 3 for the SST is constrained in all experiments by assimilating the OISST. The contributions of *in-situ* T/S profiles and SSH will be shown in next sections.

Table 2 compares the RMSE and IOA for SST by region for all assimilation experiments. The EXP01 of the average IOA was approximately 23.77%, and IOA was the highest in the K-KE region. In the NWP and OYC regions, the RMSE further improved after assimilating the T/S profiles. In the K-KE and EJS regions, the RMSE improved after assimilating the pseudo-profiles; however, in the OYC region, the RMSE increased. In the EJS region, EXP03 and EXP05, in which the profiles of KODC were assimilated, showed a higher IOA than EXP02 and EXP04, and EXP03 which did not assimilate the pseudo-profiles showed a lower RMSE than EXP04. These results indicate that the assimilation of SST satellite data plays a major role in improving the SST field, and that T/S and pseudo-profiles also play a partial role. However, the assimilation of the pseudo-profiles resulted in a slight increase in the RMSE at higher latitudes, such as the OYC region. Moreover, the KODC data reduced the RMSE for SST in the EJS; however, this effect was not observed in other regions.

3.2 Vertical structure

The vertical performance of each experiment was evaluated using WOD18 and KODC data, which were not used for assimilation. To

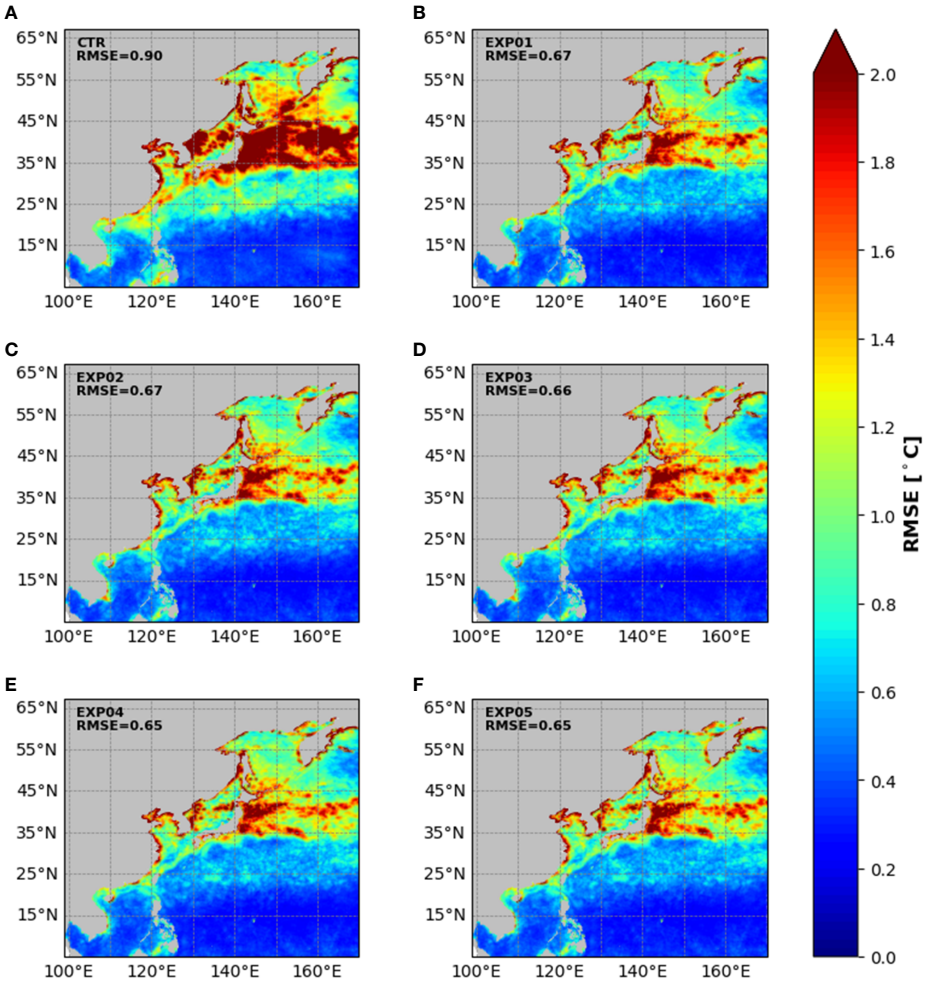


FIGURE 3
Spatial distribution of RMSE for SST with respect to OSTIA SST. Spatial averaged RMSEs of each experiment are shown on each figure. **(A-F)** represent the results of CTR and EXP01-05, respectively. RMSE, root-mean-square error; SST, sea salt temperature; CTR, control; EXP, experiment; OSTIA, Operational Sea Surface Temperature, and Sea Ice Analysis.

investigate regional effects, the domain was separated into five regions: NWP, SCS, OYC, EJS, and K-KE. Each region is denoted by a colored dot in Figure 2B.

Figure 4 shows the vertical profile of the RMSE for both temperature and salinity in each region. In the NWP region, the more observation data are assimilated, the lower the RMSE for the temperature and salinity at the overall depth. However, after assimilating SST, the RMSE for salinity in the subsurface layer increased compared to that of the CTR. Pseudo-profiles were the observational data that contributed the most to improving the RMSE for temperature and salinity in the NWP region. After the assimilation of these data, the RMSE for the temperature and salinity of the subsurface layer decreased substantially. Similar to the NWP region, in the SCS region, the pseudo-profiles contributed the most to reducing the RMSE, and assimilation of KODC data decreased the RMSE for temperature but increased that for salinity. In the OYC region, all experiments showed higher RMSE for temperature than that of the CTR at 50 m. However, in the

TABLE 2 Root-mean-square error (RMSE) averaged in space and time for SST by region from all experiments.

	CTR	EXP01	EXP02	EXP03	EXP04	EXP05
Northwest Pacific	0.45	0.38 (15.56%)	0.37 (17.78%)	0.37 (17.78%)	0.37 (17.78%)	0.37 (17.78%)
South China Sea	0.65	0.53 (18.46%)	0.53 (18.46%)	0.53 (18.46%)	0.51 (21.54%)	0.51 (21.54%)
Oyashio	1.21	0.95 (21.49%)	0.94 (22.31%)	0.94 (22.31%)	0.96 (20.66%)	0.96 (20.66%)
East/Japan sea	1.76	1.22 (30.68%)	1.22 (30.68%)	1.17 (33.52%)	1.18 (32.95%)	1.17 (33.52%)
Kuroshio-Kuroshio Extension	1.47	0.99 (32.65%)	0.99 (32.65%)	0.99 (32.65%)	0.97 (34.01%)	0.97 (34.01%)

The number in the parentheses below RMSE represents the impact of assimilation (IOA).

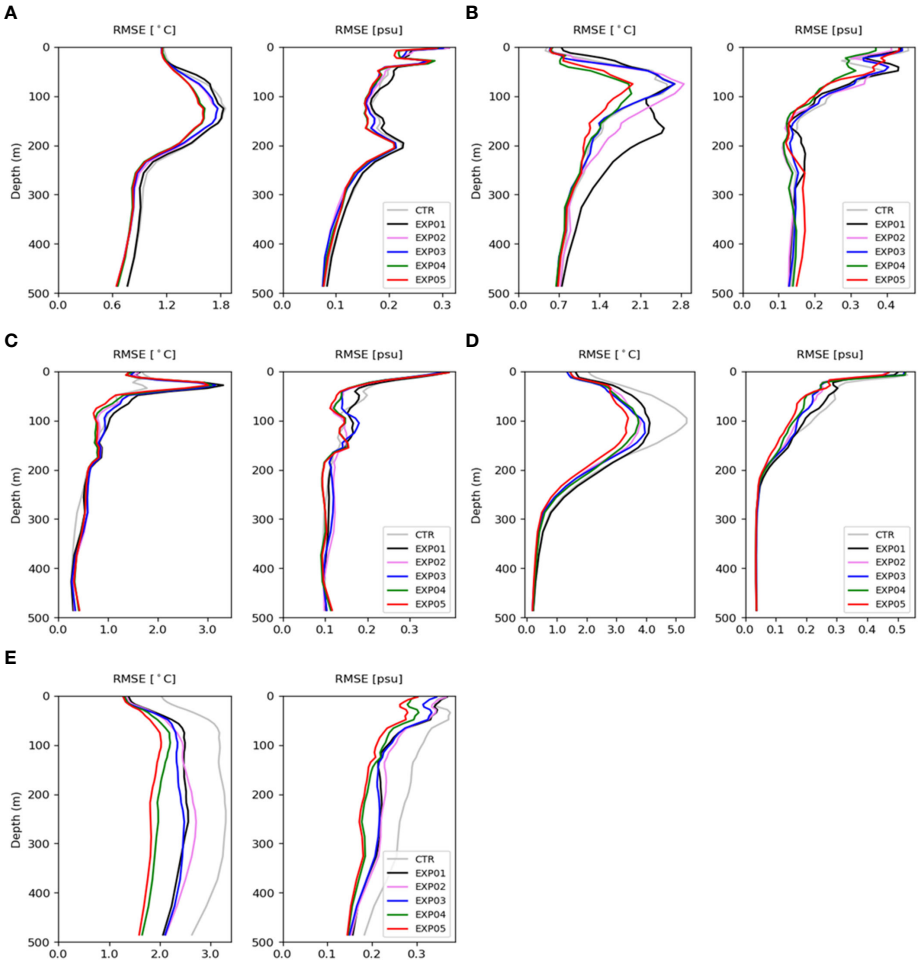


FIGURE 4
RMSE for temperature and salinity profiles at the region of **(A)** Northwestern Pacific (NWP), **(B)** South China Sea (SCS), **(C)** Oyashio Current (OYC), **(D)** East/Japan Sea (EJS), and **(E)** Kuroshio-Kuroshio extension (K-KE). Gray, black, violet, blue, green and red lines denote CTR and EXP01-05, respectively.

subsurface layer, the RMSE decreased after assimilating the pseudo-profiles. The pseudo-profiles also contributed the most to reducing the RMSE in EJS and K-KE regions, and after assimilating KODC data, the RMSE for temperature decreased whereas that for salinity improved in both regions, although the salinity data of KODC were not assimilated. Tables 2 and 3 show the average RMSE of the profiles in Figure 4 by region. As mentioned above, because the RMSE of the other experiments was larger than that of the CTR at 50 m in the OYC, the average RMSE for the entire depth was larger than that of the CTR. These results indicate that pseudo-profiles derived from satellite altimetry data in most areas significantly contribute to improving the vertical structure of temperature and salinity, and the assimilation of KODC data improved the vertical structure of temperature and salinity in the K-KE and EJS regions. (Table 4)

To comprehensively evaluate the contribution of each observation data by region at the subsurface layer, IOA using the RMSE of the experiment according to the depth (100–500 m) with respect to the profile used for each independent validation was calculated and compared by averaging at intervals of 10° for the latitude and longitude.

TABLE 3 RMSE averaged in space and time for temperature by region from all experiments.

	CTR	EXP01	EXP02	EXP03	EXP04	EXP05
Northwest pacific	1.38	1.37 (0.72%)	1.31 (5.07%)	1.31 (5.07%)	1.25 (9.42%)	1.25 (9.42%)
South China Sea	1.46	1.82 (-24.66%)	1.58 (-8.22%)	1.42 (2.74%)	1.19 (18.49%)	1.16 (20.55%)
Oyashio	0.95	1.18 (-24.21%)	1.10 (-15.79%)	1.13 (-18.95%)	1.06 (-11.58%)	1.04 (-9.47%)
East/Japan sea	3.07	2.66 (13.36%)	2.35 (23.45%)	2.34 (23.78%)	2.33 (24.10%)	2.16 (29.64%)
Kuroshio-Kuroshio Extension	2.96	2.25 (23.99%)	2.26 (23.65%)	2.15 (27.36%)	1.90 (35.81%)	1.78 (39.86%)

The number in the parentheses below RMSE represents IOA.

TABLE 4 RMSE averaged in space and time for salinity by region from all experiments.

	CTR	EXP01	EXP02	EXP03	EXP04	EXP05
Northwest Pacific	0.19	0.19 (0.00%)	0.18 (5.26%)	0.18 (5.26%)	0.18 (5.26%)	0.17 (10.53%)
South China Sea	0.24	0.24 (0.00%)	0.23 (4.17%)	0.24 (0.00%)	0.20 (16.67%)	0.23 (4.17%)
Oyashio	0.16	0.17 (-6.25%)	0.16 (0.0)	0.16 (0.0)	0.15 (6.25%)	0.15 (6.25%)
East/Japan sea	0.20	0.18 (10.00%)	0.16 (20.00%)	0.17 (15.00%)	0.16 (20.00%)	0.15 (25.00%)
Kuroshio-Kuroshio extension	0.30	0.26 (13.33%)	0.26 (13.33%)	0.25 (16.67%)	0.23 (23.33%)	0.22 (26.67%)

The number in the parentheses below RMSE represents IOA.

Temperature (Figure 5), when only SST was assimilated (EXP01), had a negative effect on the subsurface layer of most regions, except for some regions in high-latitude regions. EXP02 and EXP03 that assimilated T/S profiles had improved RMSE over EXP01 in the NWP and SCS regions. The assimilation of pseudo-profiles (EXP04 and EXP05) significantly improved the RMSE in most areas, especially in the NWP and regions where the Kuroshio Current passes (120°E–140°E, 25°N–35°N). However, it did not have a significant effect in high-latitude regions (140°E–170°E, 35°N–65°N), except for the EJS. The assimilation of KODC data (EXP03 and EXP05) considerably improved the RMSE not only in the EJS but also in the region where Kuroshio passes and the K-KE region.

In addition, by comparing the IOA for salinity in each experiment (Figure 6), we found that the assimilation of SST did not improve the RMSE for salinity in the subsurface layer in most regions. The assimilation of the T/S profile substantially improved the RMSE in the NWP and SCS regions, and the assimilation of the pseudo-profiles or the temperature substantially improved the RMSE in the regions where the Kuroshio Current passed; however, the effect of the assimilation of the temperature was not substantial in high-latitude

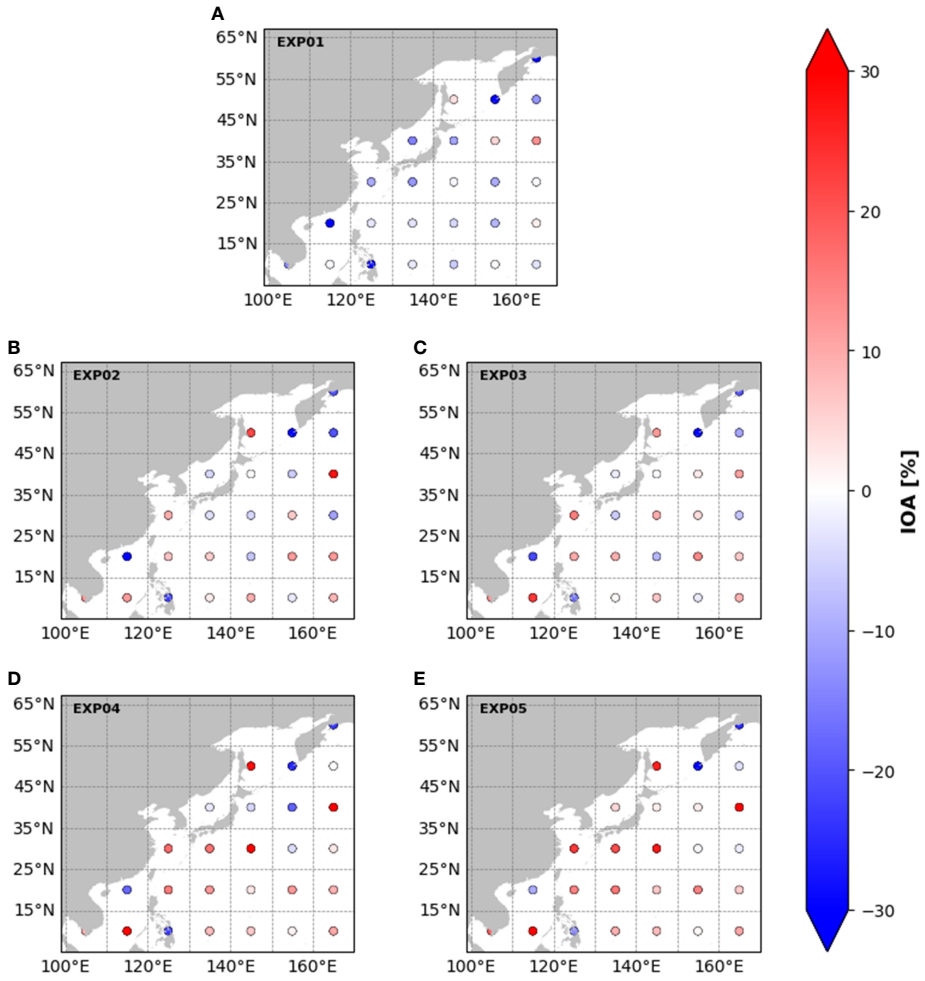


FIGURE 5
Distribution of mean IOA for temperature at subsurface layer. **(A–E)** represent the results of EXP01-05, respectively.

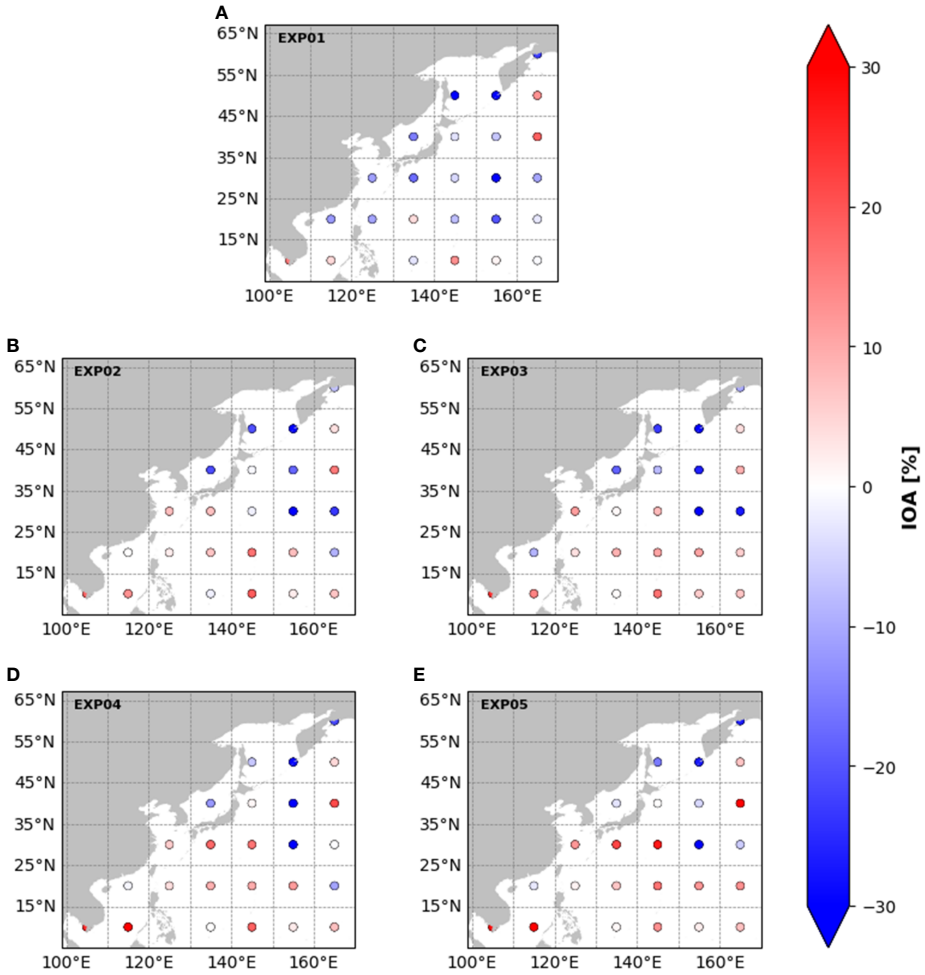


FIGURE 6
Distribution of mean IOA for salinity at subsurface layer. **(A-E)** represent the results of EXP01-05, respectively.

regions. In the assimilation of KODC (EXP03 and EXP05), the RMSE of EXP05, which also assimilated pseudo-profiles, improved in both the EJS and K-KE regions. These results show that the assimilation of SST did not have a significant impact on the subsurface layer whereas that of T/S profiles improved the RMSE in most regions. In addition, the assimilation of pseudo-profiles played a significant role in improving the temperature and salinity at the subsurface layer, similar to the *in-situ* T/S profile data; however, it did not show a significant contribution in high-latitude regions. The assimilation of KODC data, which is Korean marginal sea data, improved the temperature at the subsurface layer in the K-KE and EJS regions, and salinity at the subsurface layer was improved when assimilated with pseudo-profiles.

3.3 Sea surface height

Figure 7 shows the time series of monthly spatial-averaged RMSE for SSH from all experiments according to each region. Pseudo-profiles and SST contributed the most to improving the RMSE for SSH in all regions. SST also improved the RMSE for SSH in all regions. The T/S profiles improved the RMSE for SSH regardless of the season in the NWP region; however, it did not significantly improve the RMSE for SSH in the OYC regions and increased it during the summer season in the K-KE and EJS regions. Moreover, the RMSE of EXP03, which assimilated KODC data, increased from August to November compared to that of EXP02. Notably, the RMSE of EXP05 which assimilated KODC and pseudo-profiles improved from May to September compared to that of EXP04 which did not assimilate KODC data. The data assimilation of KODC data appeared to be more effective in EXP05 with SSH assimilation that in EXP03 without SSH assimilation. The RMSE of EXP05, which assimilated pseudo-profiles, was improved in the NWP during the summer and in the K-KE region from April to June and September to October.

To evaluate the oceanic variability associated with the K-KE region, the Kuroshio axis of the AVISO gridded data and that of each experiment were compared (Figure 8). In January and February, the assimilation effect of each experiment did not appear significant; however, the assimilation effect of each experiment became evident over time, starting in March. The experiments that did not assimilate the pseudo-profiles, including CTR, excessively simulated mesoscale features, such as meandering and eddies. However, experiments that assimilated the pseudo-profiles constrained the features excessively simulated in other experiments. EXP05, which assimilated KODC

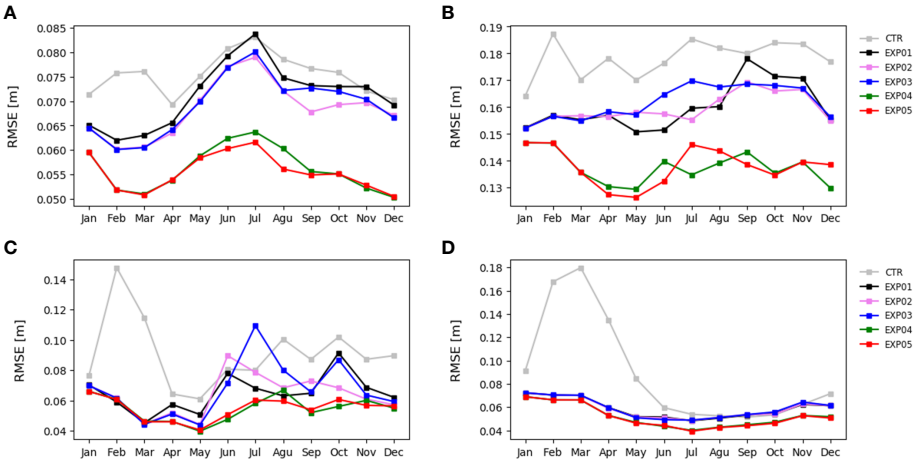


FIGURE 7
Monthly mean RMSE for sea surface height at the region of **(A)** Northwestern Pacific (NWP), **(B)** Kuroshio-Kuroshio extension (K-KE), **(C)** East/Japan Sea (EJS) and **(D)** Oyashio Current (OYC). Gray, black, violet, blue, green and red lines denote CTR and EXP01-05, respectively.

data, better simulated the Kuroshio axis in most months, except for July and August, compared to EXP04.

To evaluate this result more quantitatively, the RMSE for the latitude of the Kuroshio axis with respect to the observational data of each experiment was calculated and compared (Figure 9). The observation data that contributed the most to improving the Kuroshio axis were the pseudo-profiles derived from satellite altimetry data except in January and February. The system seems to become unstable in the beginning as the psudo-profiles from satellite altimetry have been newly assimilated. However, from March, EXP04 and EXP05 show better representation of the Kuroshio axis rather than other experiments. The SST data also contributed to improving the Kuroshio axis to some extent whereas the T/S profiles did not show a significant impact. EXP05, which assimilated all observations presented in this study, including KODC, reduced the RMSE for the latitude of the Kuroshio axis, except for March and the summer

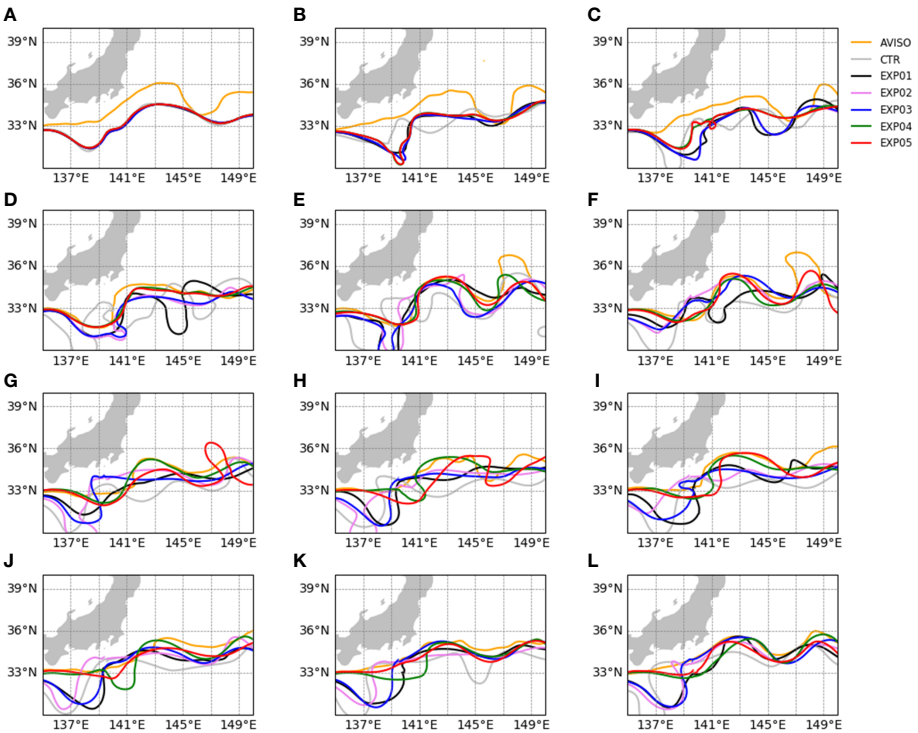


FIGURE 8
Monthly mean Kuroshio axis is denoted by the 0.6m SSH. **A–L** represent results in January-December, respectively. Orange line denotes Kuroshio axis from the AVISO gridded data. Gray, black, violet, blue, green and red lines denote Kuroshio axis of CTR and EXP01-05, respectively.

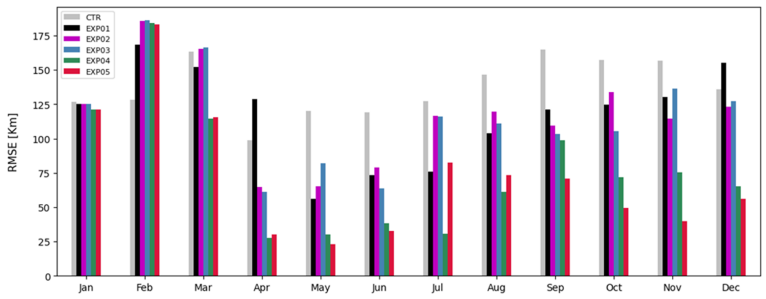


FIGURE 9
Monthly mean RMSEs for the latitude of the Kuroshio axis with respect to the observation data of each experiment. Gray, black, violet, blue, green and red bars denote CTR and EXP01-05, respectively.

season, compared to EXP04, which assimilated all observations except KODC. Table 5 also confirms the contribution of the satellite altimetry and KODC profile data to the significant reduction in RMSE for latitude of the Kuroshio axis. These results suggest that regional ocean observation networks may improve the forecast skill of the ocean prediction system not only in the region but also in the open ocean, such as the Pacific Ocean.

4 Discussion

In this study, sensitivity experiments were conducted to evaluate the impacts of *in-situ* T/S profiles and satellite observation data, including SST and altimetry data, on a high-resolution ocean circulation prediction system, so called the KOOS-OPEM. KOOS-OPEM adopts localized EnOI to assimilate the ocean observation data into the model. The satellite altimetry information was projected into the subsurface layer following CH96 which did not directly assimilate the altimetry but rather pseudo temperature and salinity profiles. The contribution of each observation data was evaluated as follows: IOA of EXP01 for SST data; average of IOA differences of EXP02 and EXP03 with respect to EXP01 for *in-situ* T/S profile data; average of IOA differences between EXP02 and EXP03, and between EXP04 and EXP05 for KODC data; average of IOA differences between EXP02 and EXP04, and between EXP03 and EXP05 for altimetry data.

The comparisons of model experiments suggest that the satellite SST data set has the most contribution (EXP01, 23.77%) to the modeled SST improvement in terms of the RMSE compared to the CTR. Especially, the largest improvement was found in the K-KE region (32.65%). Additionally, assimilating the *in-situ* profiles insignificant impact on the modeled SST performance (EXP02), while the *in-situ* profiles have the greatest influence on the vertical structure of temperature (average 10.26%) and salinity (average 7.50%), especially in the EJS (EXP02 and EXP03). The altimetry assimilation (EXP04 and EXP05) also contributes to improving the subsurface vertical profile structure of ocean temperature (average 12.33%) and salinity (average 10.00%), especially in the K-KE region. It is highlighted that the *in-situ* profiles in the Korean marginal seas provided by KODC have significant impact on the vertical structure of ocean temperature and salinity in not only the EJS but also the K-KE region (EXP03 and EXP05).

The assimilation of SST had a negative effect on both temperature and salinity in the subsurface layer in most areas (Figures 5, 6). In the OYC region, moreover, the SST assimilation seems to increase the RMSE of the temperature around 50 m depth, where the background error variance has a maximum (not shown here). Indeed, it has been observed that the model overestimates the variability of the temperature more than the observation around 50 m depth in the OYC region. The background error covariance calculated from the historical simulation may induce the temperature degradation in this region. In addition, the number of *in-situ* T/S profiles is not sufficient and the contribution of SSH seems to be limited in this region. To resolve the temperature degradation in the OYC region, it is necessary to add a new dataset or improve the model performance, which is left for the next study.

In-situ T/S profile data assimilation was effective in most regions, and pseudo-profile data also substantially contributed to improving both temperature and salinity vertical structures as effectively as the *in-situ* T/S profile data. However, the contribution of pseudo-profiles at high latitudes was lower than at low- and middle-latitudes. In high latitude oceans, where stratification is weak, small changes in surface height are assimilated into large vertical displacements of the water column, which can rather lead to errors (Fox et al., 2000). Therefore, when using the altimetry assimilation method based on CH96, it is necessary to introduce latitude dependency (Vidard et al., 2009) for the next version. As mentioned above, KODC data improved the temperature in the subsurface layer not only in the East Sea but also in the K-KE region; when assimilation was performed using KODC data and pseudo-profiles, the salinity of the subsurface layer improved. Moreover, the assimilation of KODC data affected a large area (from

TABLE 5 RMSE averaged in space and time for the latitude of the Kuroshio axis with respect to AVISO gridded data in each experiment.

	CTR	EXP01	EXP02	EXP03	EXP04	EXP05
RMSE (IOA)	3.40	3.06 (10.00%)	3.30 (3.03%)	3.14 (8.28%)	2.28 (49.12%)	2.18 (55.96%)

The number in the parentheses below RMSE represents IOA.
RMSE, root-mean-square error; IOA, impact of assimilation. CTR, control; EXP, experiment.

120°E to 160°E and 35°N to 45°N). Each observation also contributed to the improvement in the Kuroshio axis. When compared qualitatively, the pseudo-profiles derived from satellite altimetry data constrained mesoscale features, such as meander and eddy that were excessively simulated. When compared quantitatively (Table 5), the pseudo-profiles were the main contributors to the reduction in the RMSE for the latitude of the Kuroshio axis with respect to the AVISO gridded data by an average of 46.89%. SST satellite data also made the second largest contribution, improving the Kuroshio axis by 10.00%. The KODC data also improved the Kuroshio axis by 6.84% when assimilated with the pseudo-profiles.

This study suggests the quantitative impacts of each observation data sets for improvement of the high-resolution ocean prediction system. Notably, pseudo-profiles derived from satellite altimetry data significantly contributed to the ocean analysis field by improving the vertical structure of temperature and salinity. Especially, we clearly showed that the assimilation of the regional ocean observation provided by the Korean regional observation network has non-negligible impacts on the upper layer structure in the open ocean and the representation of the Kuroshio axis. These results highlight the role of regional ocean observation networks in ocean prediction systems to improve the analysis and forecast skills in the open ocean as well as in the region. In particular, it is interesting that data assimilation of the KODC data obtained around Korea peninsula contributes to the improvement of representation of the Kuroshio axis. Figures 10A-D show the upper ocean structures of temperature and current speed between EXP04 and EXP05, respectively, in October. It is noteworthy that the differences in temperature and current between EXP04 and EXP05 was pronounced in the EJS, east of the Tsugaru strait and Kuroshio downstream region rather than in the Kuroshio upstream region. In EXP05, compared to EXP04, the warm water from the Tsugaru Strait extended to the east, and the cold water from the Okhotsk Sea and the Subarctic Pacific was extended to the south. Itoh et al. (2022) analyzed high-resolution observation data and reported that a sharp front often develops in the Sanriku confluence where Tsugaru Warm Current, Oyasio Current, and Kuroshio Current meet. EXP05 may better simulate the representation of the front in the Sanriku confluence by improving the physical properties of the Tsugaru Warm current through the assimilation of KODC data taken around Korean Peninsula. In addition, the better representation of the front may help the Oyashio Current extend to the south where the Kuroshio Current separate from the coast, which affects the fluctuations of the Kuroshio axis. In fact, the IOA from the temperature profiles for EXP04 and EXP05 (Figures 10E, F) shows that the KODC profiles contribute to improving the vertical temperature structure not only around the Korean Peninsula but also in the west of the Tsugaru strait and in the Sanriku confluence. Although the dynamic relationship between the circulation of the Korea marginal seas and Kuroshio was not fully understood in this study, it seems worthwhile for future research. This study also suggests that greater attention should be paid to the role of regional ocean observation networks to improve the forecast skill of the ocean prediction system not only in the region but also in the open ocean, such as the Pacific Ocean.

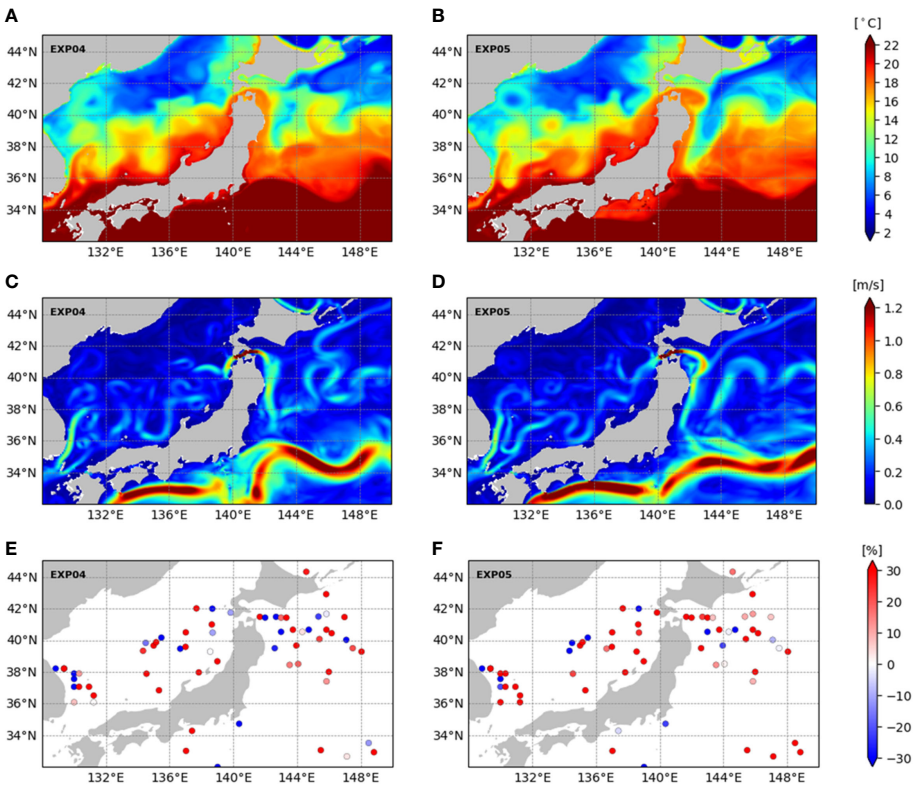


FIGURE 10
Monthly mean temperature (upper panels), current speed (middle panels) averaged over 0 to 100m, and IOA (lower panels) for the temperature profiles over 0 to 100m in EXP04 (left panels) and EXP05 (right panels) in October.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, Further inquiries can be directed to the corresponding author.

Author contributions

IC and YK: conducting the experiment and writing the first manuscript. IC: carrying out experiment and visualization. YK: supervision. YK, HJ and GP: deriving additional analysis ideas. YK, HJ, GP, Y-GP and Y-SC: review and editing. All authors contributed to the article and approved the submitted version.

Funding

This research was part of projects titled 'Investigation and prediction system development of marine heatwave around the Korean Peninsula originated from the subarctic and western Pacific' (20190344) and 'Improvements of ocean prediction accuracy using numerical modeling and artificial intelligence technology' (20180447) funded by the Ministry of Oceans and Fisheries, Korea. This work was supported by Korea Institute of Marine Science and Technology Promotion(KIMST) funded by the Ministry of Oceans and Fisheries, Korea (Development of 3-D Ocean Current Observation Technology for Efficient Response to Maritime Distress, 20210642). YK was supported by by the Pukyong National University Research Fund (CD20191541).

Acknowledgments

We would like to thank Gary Brassington for his valuable discussions and comments.

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Alves, J. O. S., Haines, K., and Anderson, D. L. (2001). Sea Level assimilation experiments in the tropical pacific. *J. Phys. Oceanogr.* 31 (2), 305–323. doi: 10.1175/1520-0485(2001)031<0305

Arakawa, A., and Lamb, V. R. (1977). Computational design of the basic dynamical processes of the UCLA general circulation model. *Gen. Circ. Models atmosphere* 17 (Supplement C), 173–265. doi: 10.1016/B978-0-12-460817-7.50009-4

Birol, F., Brankart, J. M., Lemoine, J. M., Brasseur, P., and Verron, J. (2005). Assimilation of satellite altimetry referenced to the new GRACE geoid estimate. *Geophys. Res. Lett.* 32 (6). doi: 10.1029/2004GL021329

Burgers, G., Van Leeuwen, P. J., and Evensen, G. (1998). Analysis scheme in the ensemble kalman filter. *Monthly weather Rev.* 126 (6), 1719–1724. doi: 10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2

Chen, Y., Yan, C., and Zhu, J. (2018). Assimilation of sea surface temperature in a global hybrid coordinate ocean model. *Adv. Atmospheric Sci.* 35 (10), 1291–1304. doi: 10.1007/s00376-018-7284-6

Cho, Y. K., Seo, G. H., Choi, B. J., Kim, S., Kim, Y. G., Youn, Y. H., et al. (2009). Connectivity among straits of the northwest pacific marginal seas. *J. Geophys. Res.: Oceans* 114 (C6). doi: 10.1029/2008JC005218

Cooper, M., and Haines, K. (1996). Altimetric assimilation with water property conservation. *J. Geophys. Res.: Oceans* 101 (C1), 1059–1077. doi: 10.1029/95JC02902

Counillon, F., and Bertino, L. (2009). Ensemble optimal interpolation: multivariate properties in the gulf of Mexico. *Tellus A: Dynam. Meteorol. Oceanogr.* 61 (2), 296–308. doi: 10.1111/j.1600-0870.2008.00383.x

El-Geziry, T. M., and Bryden, I. G. (2010). The circulation pattern in the Mediterranean Sea: issues for modeller consideration. *J. Operational Oceanogr.* 3 (2), 39–46. doi: 10.1080/1755876X.2010.11020116

Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.: Oceans* 99 (C5), 10143–10162. doi: 10.1029/94JC00572

Evensen, G. (2003). The ensemble kalman filter: Theoretical formulation and practical implementation. *Ocean dynam.* 53 (4), 343–367. doi: 10.1007/s10236-003-0036-9

Fox, A. D., Haines, K., De Cuevas, B., and Webb, D. J. (2000). Altimeter assimilation in the OCCAM global model: Part I: A twin experiment. *J. Mar. Syst.* 26 (3-4), 303–322. doi: 10.1016/S0924-7963(00)00043-9

Fu, W., and Zhu, J. (2011). Effects of sea level data assimilation by ensemble optimal interpolation and 3D variational data assimilation on the simulation of variability in a tropical pacific model. *J. Atmospheric Oceanic Technol.* 28 (12), 1624–1640. doi: 10.1175/JTECH-D-11-00044.1

Gaspari, G., and Cohn, S. E. (1999). Construction of correlation functions in two and three dimensions. *Q. J. R. Meteorol. Soc.* 125 (554), 723–757. doi: 10.1002/qj.49712555417

Gent, P. R., and Mcwilliams, J. C. (1990). Isopycnal mixing in ocean circulation models. *J. Phys. Oceanogr.* 20 (1), 150–155. doi: 10.1175/1520-0485(1990)020<0150

Good, S., Fiedler, E., Mao, C., Martin, M. J., Maycock, A., Reid, R., et al. (2020). The current configuration of the OSTIA system for operational production of foundation sea surface temperature and ice concentration analyses. *Remote Sens.* 12 (4), 720. doi: 10.3390/rs12040720

Griffies, S. M., and Hallberg, R. W. (2000). Biharmonic friction with a smagorinsky-like viscosity for use in large-scale eddy-permitting ocean models. *Monthly Weather Rev.* 128 (8), 2935–2946. doi: 10.1175/1520-0493(2000)128<2935

Hamill, T. M., Whitaker, J. S., and Snyder, C. (2001). Distance-dependent filtering of background error covariance estimates in an ensemble kalman filter. *Monthly Weather Rev.* 129 (11), 2776–2790. doi: 10.1175/1520-0493(2001)129<2776

Houtekamer, P. L., and Mitchell, H. L. (2001). A sequential ensemble kalman filter for atmospheric data assimilation. *Monthly Weather Rev.* 129 (1), 123–137. doi: 10.1175/1520-0493(2001)129<0123

Itoh, S., Tsutsumi, E., Masunaga, E., Sakamoto, T. T., Ishikawa, K., Yanagimoto, D., et al. (2022). Seasonal cycle of the confluence of the tsugaru warm, oyashio, and kuroshio currents east of Japan. *J. Geophys. Res.: Oceans* 127 (8), e2022JC018556. doi: 10.1029/2022JC018556

Kim, Y. H., Hwang, C., and Choi, B. J. (2015). An assessment of ocean climate reanalysis by the data assimilation system of KIOST from 1947 to 2012. *Ocean Model.* 91, 1–22. doi: 10.1016/j.ocemod.2015.02.006

Kim, Y. H., Lyu, S. J., Choi, B. J., Cho, Y. K., and Kim, Y. G. (2008). Implementation of the ensemble kalman filter to a double gyre ocean and sensitivity test using twin experiments. *Ocean Polar Res.* 30 (2), 129–140. doi: 10.4217/OPR.2008.30.2.129

Kim, S. Y., Park, Y. G., Kim, Y. H., Seo, S., Jin, H., Pak, G., et al. (2021). Origin, variability, and pathways of East Sea intermediate water in a high-resolution ocean reanalysis. *J. Geophys. Res.: Oceans* 126 (6), e2020JC017158. doi: 10.1029/2020JC017158

Large, W. G., McWilliams, J. C., and Doney, S. C. (1994). Oceanic vertical mixing: A review and a model with a nonlocal boundary layer parameterization. *Rev. geophysics* 32 (4), 363–403. doi: 10.1029/94RG01872

Large, W. G., and Yeager, S. G. (2004). Diurnal to decadal global forcing for ocean and sea-ice models: The data sets and flux climatologies. *NCAR Tech. Rep. TN-460+STR*, 105–129. doi: 10.5065/D6KK98Q6

Oke, P. R., Brassington, G. B., Griffin, D. A., and Schiller, A. (2008). The bluelink ocean data assimilation system (BODAS). *Ocean Model.* 21 (1-2), 46–70. doi: 10.1016/j.ocemod.2007.11.002

Oke, P. R., Sakov, P., and Corney, S. P. (2007). Impacts of localisation in the EnKF and EnOI: experiments with a small model. *Ocean Dynam.* 57 (1), 32–45. doi: 10.1007/s10236-006-0088-8

Oke, P. R., Schiller, A., Griffin, D. A., and Brassington, G. B. (2005). Ensemble data assimilation for an eddy-resolving ocean model of the Australian region. *Q. J. R. Meteorol. Soc.* 131 (613), 3301–3311. doi: 10.1256/qj.05.95

Parent, L., Testut, C. E., Brankart, J. M., Verron, J., Brasseur, P., and Gourdeau, L. (2003). Comparative assimilation of Topex/Poseidon and ERS altimeter data and of TAO temperature data in the tropical pacific ocean during 1994–1998, and the mean sea-surface height issue. *J. Mar. Syst.* 40, 381–401. doi: 10.1016/S0924-7963(03)00026-5

Park, J. (2021). Quality evaluation of long-term shipboard salinity data obtained by NIFS. *Sea: J. OF THE KOREAN Soc. OF OCEANOGR.* 26 (1), 49–61. doi: 10.7850/jkso.2021.26.1.049

Park, Y. G., Choi, A., Kim, Y. H., Min, H. S., Hwang, J. H., and Choi, S. H. (2010). Direct flows from the ulleung basin into the yamato basin in the East/Japan Sea. *Deep Sea Res. Part I: Oceanographic Res. Papers* 57 (5), 731–738. doi: 10.1016/j.dsr.2010.03.006

Park, K., Park, J. E., Choi, B. J., Byun, D. S., and Lee, E. I. (2013). An oceanic current map of the East Sea for science textbooks based on scientific knowledge acquired from oceanic measurements. *Sea: J. Korean Soc. Oceanogr.* 18 (4), 234–265. doi: 10.7850/jkso.2013.18.4.234

Seo, S. N. (2008). Digital 30sec gridded bathymetric data of Korea marginal seas-KorBathy30s. *J. Korean Soc. Coast. Ocean Engineers* 20 (1), 110–120.

Simmons, H. L., Jayne, S. R., Laurent, L. C. S., and Weaver, A. J. (2004). Tidally driven mixing in a numerical model of the ocean general circulation. *Ocean Model.* 6 (3-4), 245–263. doi: 10.1016/S1463-5003(03)00011-8

Testut, C. E., Brasseur, P., Brankart, J. M., and Verron, J. (2003). Assimilation of sea-surface temperature and altimetric observations during 1992–1993 into an eddy permitting primitive equation model of the north Atlantic ocean. *J. Mar. Syst.* 40, 291–316. doi: 10.1016/S0924-7963(03)00022-8

Troccoli, A., Balmaseda, M. A., Segschneider, J., Vialard, J., Anderson, D. L., Haines, K., et al. (2002). Salinity adjustments in the presence of temperature data assimilation. *Monthly Weather Rev.* 130 (1), 89–102. doi: 10.1175/1520-0493(2002)130<0089

Troccoli, A., and Haines, K. (1999). Use of the temperature–salinity relation in a data assimilation context. *J. Atmospheric Oceanic Technol.* 16 (12), 2011–2025. doi: 10.1175/1520-0426(1999)016<2011

Troen, I. B., and Mahrt, L. (1986). A simple model of the atmospheric boundary layer; sensitivity to surface evaporation. *Boundary-Layer Meteorol.* 37 (1), 129–148. doi: 10.1007/BF00122760

Vialard, J., Weaver, A. T., Anderson, D. L. T., and Delecluse, P. (2003). Three-and four-dimensional variational assimilation with a general circulation model of the tropical pacific ocean. part II: Physical validation. *Monthly Weather Rev.* 131 (7), 1379–1395. doi: 10.1175/1520-0493(2003)131<1379

Vidard, A., Balmaseda, M., and Anderson, D. (2009). Assimilation of altimeter data in the ECMWF ocean analysis system 3. *Monthly Weather Rev.* 137 (4), 1393–1408. doi: 10.1175/2008MWR2668.1

Vörösmarty, C. J., Fekete, B., and Tucker., B. A. (1998). *River discharge database, version 1.1 (RivDIS v1.0 supplement)* [Available online at http://pyramid.sr.unh.edu/csrc/hydro/.]. doi: 10.1594/PANGAEA.859439

Winton, M. (2000). A reformulated three-layer sea ice model. *J. atmospheric oceanic Technol.* 17 (4), 525–531. doi: 10.1175/1520-0426(2000)017<0525

Yan, C., Zhu, J., Li, R., and Zhou, G. (2004). Roles of vertical correlations of background error and T-s relations in estimation of temperature and salinity profiles from sea surface dynamic height. *J. Geophys. Res.: Oceans* 109 (C8). doi: 10.1029/2003JC002224

Zheng, F., and Zhu, J. (2015). Roles of initial ocean surface and subsurface states on successfully predicting 2006–2007 El niño with an intermediate coupled model. *Ocean Sci.* 11 (1), 187–194. doi: 10.5194/os-11-187-2015

Zhou, W., Li, J., Xu, F., Shu, Y., and Feng, Y. (2021). The impact of ocean data assimilation on seasonal predictions based on the national climate center climate system model. *Acta Oceanol. Sin.* 40 (5), 58–70. doi: 10.1007/s13131-021-1732-3

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 05 April 2023
DOI 10.3389/fmars.2023.1125342

Check for updates

OPEN ACCESS

EDITED BY
Jinyu Sheng,
Dalhousie University, Canada

REVIEWED BY
Kyoko Ohashi,
Dalhousie University, Canada
Shangfei Lin,
Hong Kong University of Science and Technology, Hong Kong SAR, China

*CORRESPONDENCE
Wataru Fujimoto
w-fujimoto@classnk.or.jp

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

RECEIVED 16 December 2022
ACCEPTED 13 February 2023
PUBLISHED 05 April 2023

CITATION
Fujimoto W and Ishibashi K (2023) Ensemble-based data assimilation for predictable zones and application for non-linear deep-water waves.
Front. Mar. Sci. 10:1125342.
doi: 10.3389/fmars.2023.1125342

COPYRIGHT
© 2023 Fujimoto and Ishibashi. 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(s) 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.

Ensemble-based data assimilation for predictable zones and application for non-linear deep-water waves

Wataru Fujimoto* and Kinya Ishibashi

Research Institute, Nippon Kaiji Kyokai, Tokyo, Japan

The ensemble-based variational method is easier to implement and parallelize than the adjoint method. For circumstances in which observed data are too limited and sparse for oceanographic data assimilation, the surface wave reconstruction by ensemble adjoint-free data assimilation (SWEAD) method was developed in a previous study. SWEAD generates ensembles of search directions from Fourier modes to numerically differentiate the squared error between observed data and a physical model. However, Fourier modes are global bases and could be redundant for a narrow predictable zone confined by a dispersion relationship. To concentrate ensembles on the predictable zone, we propose using singular value decomposition (SVD) of the approximated Jacobian of the squared error. Here, the Jacobian was first approximated by the linear dispersion relationship and successively updated to consider the non-linearity of the physical system. A new criterion for reusing the ensemble was also devised for this new method, increasing the dimension of search directions. A twin experiment was conducted for non-linear deep-water waves, and the optimization efficiency of the new method—SWEAD using SVD (SWEAD-S)—was significantly greater than that of SWEAD. Expansion of the predictable zone caused by the effect of non-linearity on the wave group velocity is thought to be essential for this improvement.

KEYWORDS

ensemble-based 4DVar, non-linear dispersive wave, singular value decomposition, predictable zone, freak wave, higher order spectral method

1 Introduction

Observed ocean data are often spatially sparse, and the unobserved physical state needs to be estimated from the observed data for a certain period. The four-dimensional variational (4DVar) method estimates the physical state by minimizing the squared error between the value estimated by the physical model and the observed data as a cost function. To solve this least-squares problem, the linear approximation of the cost function must be determined. Methods to do this include the adjoint method, which uses the adjoint

code of the physical model, and the ensemble method, which numerically differentiates the cost function by perturbed ensemble simulation. In recent years, the 4DVar method has been studied using the ensemble method because it does not require adjoint codes, which are expensive to implement, and because parallel computation is straightforward.

The ensemble-based 4DVar (En4DVAR) method utilizes the ensemble members of meteorological forecasts (Liu et al., 2008; Liu et al., 2009). The maximum likelihood ensemble filter (Zupanski, 2005) provides perturbed ensemble members from the square root decomposition of the error covariance matrix. The adjoint-free 4DVar (a4dVar) method alternates perturbation vectors in each iteration of the optimization process (Yaremchuk et al., 2009). The ensemble members are taken from the empirical orthogonal function (EOF) of the model trajectories (Panteleev et al., 2015) or the misfit derived from the model and observed data (Yaremchuk et al., 2016; Yaremchuk et al., 2017). In contrast to meteorological forecasts, a4dVar is generally suitable for oceanographic problems, given that there is no reliable information on their perturbation modes, which have a faster growth rate. a4dVar stacks the perturbed ensemble simulation to construct an approximate of the Hessian matrix of the cost function for efficient optimization. Similar to the Krylov subspace method, a4dVar reinitializes the perturbed ensemble simulation at certain conditions defined for the decay rate of the cost function (Yaremchuk et al., 2016) or the eigenvalues of the Hessian matrix (Yaremchuk et al., 2017).

Fujimoto and Waseda (2020) modified a4dVar and named their modified version SWEAD (surface wave reconstruction by ensemble adjoint-free data assimilation). SWEAD stacks the perturbed ensemble simulation to approximate the Hessian matrix while ensuring conformity to the linear approximation. Eventually, the dimension of the Hessian matrix has no limitations owing to the reinitialization, thereby increasing convergence speed. SWEAD was originally developed to estimate non-linear deep-water waves from observed data and has already been applied to field measurements in the ocean. SWEAD has been used to reconstruct a wave field around an observational tower with stereo camera data (Watanabe et al., 2019), and a non-linear wave group, called the oblique soliton, was captured (Waseda et al., 2021).

In dispersive waves such as deep-water waves, Wu (2004) and Qi et al. (2018) showed that the dispersion relationship confines the predictable zone for a limited amount of observational data. Meanwhile, SWEAD uses Fourier modes to generate perturbed ensembles (see Figure 2 of Fujimoto and Waseda (2020)), which are global bases and would be redundant for the limited predictable zone. Wu (2004) also showed the SVD analysis yields modes that are most sensitive to the cost function. The predictable zone can be evaluated by singular value decomposition (SVD) of the linear dispersion relationship (see Section 2.3.1). To improve optimization efficiency, we proposed using SVD to concentrate ensembles on the predictable zone and accumulate the generated ensembles. Several new techniques for efficient optimization were also devised. We name the method proposed in this study SWEAD-S (SWEAD utilizing SVD).

Section 2 reviews the ensemble methods, including SWEAD, and describes SWEAD-S. A typical example of non-linear dispersive waves is deep-water waves. This study targeted non-linear deep-water waves to demonstrate the performance of SWEAD-S. Section 3 provides a brief background of non-linear deep-water wave studies and the configuration of twin experiments. In Section 4, the performance of SWEAD and SWEAD-S are compared through twin experiments. Finally, Section 5 outlines the findings and directions for future research.

2 Methodology

The 4DVar method minimizes the squared error between a model prediction and observed data as the cost function as:

$$L(\tilde{\mathbf{x}}) = \frac{1}{2}(\tilde{\boldsymbol{A}}(\tilde{\mathbf{x}}) - \tilde{\mathbf{y}})^{*}\mathbf{R}^{-1}(\tilde{\boldsymbol{A}}(\tilde{\mathbf{x}}) - \tilde{\mathbf{y}}) + \frac{1}{2}\tilde{\mathbf{x}}^{*}\mathbf{D}^{-1}\tilde{\mathbf{x}}, \tag{1}$$

where $\tilde{\boldsymbol{A}}(\tilde{\mathbf{x}})$ is the prediction by the physical model, $\tilde{\mathbf{x}}$ is the initial condition, $\tilde{\mathbf{y}}$ is the observed data, and $\mathbf{R}$ and $\mathbf{D}$ are the observational and background error covariance matrices, respectively. To simplify the equation, the variables are scaled as:

$$\mathbf{x} = \mathbf{D}^{-1/2}\tilde{\mathbf{x}}, \mathbf{y} = \mathbf{R}^{-1/2}\tilde{\mathbf{y}}, \boldsymbol{A}(\mathbf{x}) = \mathbf{R}^{-1/2}\tilde{\boldsymbol{A}}(\tilde{\mathbf{x}})\,. \tag{2}$$

The cost function can be reduced to:

$$L(\mathbf{x}) = \frac{1}{2}\| \boldsymbol{A}(\mathbf{x}) - \mathbf{y} \|^2 + \frac{1}{2}\|\mathbf{x}\|^2, \tag{3}$$

where $\|\cdot\|^2$ denotes the L2 norm. We considered only the observational error term (sometimes called misfit) to simplify the discussion:

$$L(\mathbf{x}) = \frac{1}{2}\|\boldsymbol{A}(\mathbf{x}) - \mathbf{y}\|^2 \tag{4}$$

The case where the regularization term is included is explained in Section 2.3.5.

2.1 Fundamentals of ensemble-based 4Dvar

The gradient of the cost function is written as:

$$\nabla L(\mathbf{x}) = \mathbf{A}^{*}(\boldsymbol{A}(\mathbf{x}) - \mathbf{y})\,. \tag{5}$$

where $\mathbf{A}$ denotes the Jacobian matrix of $\boldsymbol{A}(\mathbf{x})$ For a non-linear wave system, its Jacobian $\mathbf{A}$ cannot be expressed analytically and must be obtained numerically. Therefore, the adjoint method requires differentiating all procedures of the physical model and the observational operator, transposing it, and implementing it in a program. On the other hand, the ensemble-based 4DVar method differentiates the cost function numerically by ensemble simulations. Let $\mathbf{V}$ denote a matrix representing the perturbation of the initial values and let $\mathbf{w}_n$ denote weight coefficients for updating the initial values $\mathbf{x}_{n+1}=\mathbf{x}_n+\mathbf{V}\mathbf{w}_n$, where n is an index of iteration. The cost function is rewritten as:

$$L(\mathbf{w}_n) = \frac{1}{2}\|\boldsymbol{A}(\mathbf{x}_n + \mathbf{V}\mathbf{w}_n) - \mathbf{y}\|^2. \quad (6)$$

The perturbations $\boldsymbol{\delta}\mathbf{Y}$ of the physical model $\boldsymbol{A}(\mathbf{x})$ are obtained by comparing an unperturbed simulation and perturbed ensemble simulations:

$$\begin{aligned} &\boldsymbol{\delta}\mathbf{Y}_{(m)} = \frac{\boldsymbol{A}(\mathbf{x}_n + \epsilon\mathbf{v}_{(m)}) - \boldsymbol{A}(\mathbf{x}_n)}{\epsilon} \\ &\boldsymbol{\delta}\mathbf{Y} = (\cdots \boldsymbol{\delta}\mathbf{Y}_{(m)} \cdots), \ \mathbf{V} = (\cdots \mathbf{v}_{(m)} \cdots), \ m = 1 \cdots N_{ens} \end{aligned} \quad (7)$$

ε is a sufficiently small number, such as 0.001. Eq. (6) is summarized as $\boldsymbol{\delta}\mathbf{Y} = \mathbf{A}\mathbf{V}$ in matrix form. Therefore,

$$L(\mathbf{w}_n) \approx \frac{1}{2}\|\boldsymbol{A}(\mathbf{x}_n) + \mathbf{A}\mathbf{V}\mathbf{w}_n - \mathbf{y}\|^2 = \frac{1}{2}\|\boldsymbol{A}(\mathbf{x}_n) + \boldsymbol{\delta}\mathbf{Y}\mathbf{w}_n - \mathbf{y}\|^2 \quad (8)$$

The optimal update $\mathbf{w}_n$ is such that the gradient of its cost function is zero $\nabla L(\mathbf{w}_n) = \mathbf{0}$; therefore, the following equation is solved for w_n:

$$\begin{aligned} &\nabla L(\mathbf{w}_n) = \boldsymbol{\delta}\mathbf{Y}^{\star}(\boldsymbol{A}(\mathbf{x}_n) + \boldsymbol{\delta}\mathbf{Y}\mathbf{w}_n - \mathbf{y}) = \mathbf{0} \\ &\therefore \boldsymbol{\delta}\mathbf{Y}^{\star}\boldsymbol{\delta}\mathbf{Y}\mathbf{w}_n = -\boldsymbol{\delta}\mathbf{Y}^{\star}(\boldsymbol{A}(\mathbf{x}_n) - \mathbf{y}). \end{aligned} \quad (9)$$

The cost function $L(\mathbf{w}_n)$ is optimized in the search subspace of $\mathbf{w}_n$ spanned by $\mathbf{V}$. Hence, $\mathbf{V}$ is crucial for efficient optimization efficiency. As described in the introduction, perturbed ensembles can be generated in several ways. For example, in the a4dVar method, ensembles of search directions are generated based on model trajectories and misfit EOFs. SWEAD uses Fourier modes as perturbations instead of EOFs because it is intended for water waves.

2.2 Summary of a4dVar and SWEAD

Solving Eq. (9) corresponds to the Gauss–Newton method because the Hessian matrix of the cost function is approximated by a product of the Jacobian matrix $\boldsymbol{\delta}\mathbf{Y}^{\star}\boldsymbol{\delta}\mathbf{Y} = \mathbf{V}^{\star}\mathbf{A}^{\star}\mathbf{A}\mathbf{V}$ in the subspace spanned by $\mathbf{V}$. To improve optimization speed, the dimension of subspace $\mathbf{V}$ should be increased. a4dVar and SWEAD stack perturbations to approximate the Hessian matrix and expand the dimension of subspace $\mathbf{V}$. The perturbations generated in previous iterations $\mathbf{V}_{s,n-1}$ and $\boldsymbol{\delta}\mathbf{Y}_{s,n-1}$ are reused, combined with the new perturbations $\mathbf{V}_n$ and $\boldsymbol{\delta}\mathbf{Y}_n$ in the n-th iteration, and stacked $(\mathbf{V}_{s,n-1}|\mathbf{V}_n) \rightarrow \mathbf{V}$ and $(\boldsymbol{\delta}\mathbf{Y}_{s,n-1}|\boldsymbol{\delta}\mathbf{Y}_n) \rightarrow \boldsymbol{\delta}\mathbf{Y}$. Then, Eq. (9) is solved with the stacked $\mathbf{V}$ and $\boldsymbol{\delta}\mathbf{Y}$.

The perturbations for the next iteration $\mathbf{V}_n$ should be orthogonal to the stacked perturbations $\mathbf{V}_{s,n-1}$ to keep $\boldsymbol{\delta}\mathbf{Y}^{\star}\boldsymbol{\delta}\mathbf{Y}$ well conditioned. In other words, $\mathbf{V}_n$ should be drawn from an orthogonal complement $\mathbf{V}_{s,n-1}^{\perp}$ of $\mathbf{V}_{s,n-1}$, which is obtained by the Gram–Schmidt orthogonalization method. Yaremchuk et al. (2009) state that this orthogonalization-optimization process is analogous to the generalized minimal residuals (GMRES) method (Saad and Schultz, 1986), which is a Krylov subspace method. SWEAD uses Fourier modes as the perturbations because they are an orthogonal basis, and Gram–Schmidt orthogonalization is not required. SWEAD uses Fourier modes $\mathbf{V}_n$ different from $\mathbf{V}_{s,n-1}$.

The difference between a4dVar and SWEAD is how $\mathbf{V}_{s,n-1}$ and $\boldsymbol{\delta}\mathbf{Y}_{s,n-1}$ are reused. As shown in Figure 1A, a4dVar reuses all ensembles, i.e., $\mathbf{V} \rightarrow \mathbf{V}_{s,n}$ and $\boldsymbol{\delta}\mathbf{Y} \rightarrow \boldsymbol{\delta}\mathbf{Y}_{s,n}$, but it reinitializes $\mathbf{V}$ and $\boldsymbol{\delta}\mathbf{Y}$ in a certain condition (Yaremchuk et al., 2016; Yaremchuk et al., 2017). This reinitialization corresponds to the restart technique of GMRES to keep $\boldsymbol{\delta}\mathbf{Y}^{\star}\boldsymbol{\delta}\mathbf{Y}$ well conditioned. In contrast, SWEAD reuses some ensembles $\mathbf{V}_{reused}$ and $\boldsymbol{\delta}\mathbf{Y}_{reused}$, conforming to the linear approximation $\boldsymbol{\delta}\mathbf{Y}_{reused} \approx \mathbf{A}\mathbf{V}_{reused}$ from $\mathbf{V}$, then ensembles are stacked $\mathbf{V}_{reused} \rightarrow \mathbf{V}_{s,n}$ and $\boldsymbol{\delta}\mathbf{Y}_{reused} \rightarrow \boldsymbol{\delta}\mathbf{Y}_{s,n}$, as shown in Figure 1B. To check the conformity to the linear approximation, a certain criterion is employed, as described later in Section 2.3.4. The stacking procedure of SWEAD does not limit the dimension of $\mathbf{V}$ and could contribute to faster optimization of the cost function.

2.3 Proposed method: SWEAD-S

The following sections explain what is changed in SWEAD-S from SWEAD, taking deep-water waves as an example.

2.3.1 Predictable zone and singular value decomposition

A wave group conveys wave energy and information, and the wave group velocity determines how far the wave field can be predicted from observed data. Wu (2004) and Qi et al. (2018) analyzed the predictable zone of linear deep-water waves. The predictable zone, which is an area confined by the lowest and fastest wave group velocities, becomes narrower if the measurement period is shorter, or if the directional spread of the wave becomes broader (see Figures 2, 3 of Qi et al. (2018)).

The predictable zone is related to the singular vectors of the Jacobian $\mathbf{A}$, and its SVD is $\mathbf{A}=\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\star}$, where $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices containing the left and right singular vectors. $\boldsymbol{\Sigma}$ is a rectangular diagonal matrix with non-negative real numbers on its diagonal. Let $\mathbf{V}_o$ be the right singular vector corresponding to the kernel space $\mathrm{Ker}(\mathbf{A})$ and $\mathbf{V}_o^{\perp}$ be the right singular vector corresponding to the orthogonal complement space $(\mathrm{Ker}(\mathbf{A}))^{\perp}$. Let $\boldsymbol{\Sigma}_o^{\perp}$ be the diagonal matrix with the singular value corresponding to $\mathbf{V}_o^{\perp}$; $\mathbf{A}=\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\star}$ can be rewritten as $\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}_o^{\perp}\mathbf{V}_o^{\perp}$. The structure of the matrices is illustrated in Figure 2. We assumed that the observed data were sparse, and the physical dimension N_{phys} (column) was larger than the observational dimension N_{obs} (row).

The solution to $\mathbf{A}\mathbf{x} = \mathbf{y}$ is:

$$\mathbf{x} = \mathbf{V}_o^{\perp}\boldsymbol{\Sigma}_o^{\perp\,-1}\mathbf{U}^{\star}\mathbf{y} + \mathbf{V}_o\boldsymbol{\chi}\,. \quad (10)$$

$\boldsymbol{\chi}$ is an arbitrary vector, and $\mathbf{V}_o\boldsymbol{\chi}$ corresponds to an indefinite part of the solution. The term $\mathbf{V}_o^{\perp}\boldsymbol{\Sigma}_o^{\perp\,-1}\mathbf{U}^{\star}\mathbf{y}$ corresponds to a definite part of the solution known as the minimum-norm solution. From the observed data $\mathbf{y}$, only the first term of the above equation can be calculated; the second term is unknown owing to the arbitrary vector $\boldsymbol{\chi}$. Therefore, the subspace spanned by $\mathbf{V}_o^{\perp}$ corresponds to the predictable zone.

If the wave system is linear, then the Jacobian $\mathbf{A}$ is approximated by the linear dispersion relationship as $\mathbf{A}'$. For example, the linear dispersion relationship of deep-water waves is $\omega^2 = gk$, where ω

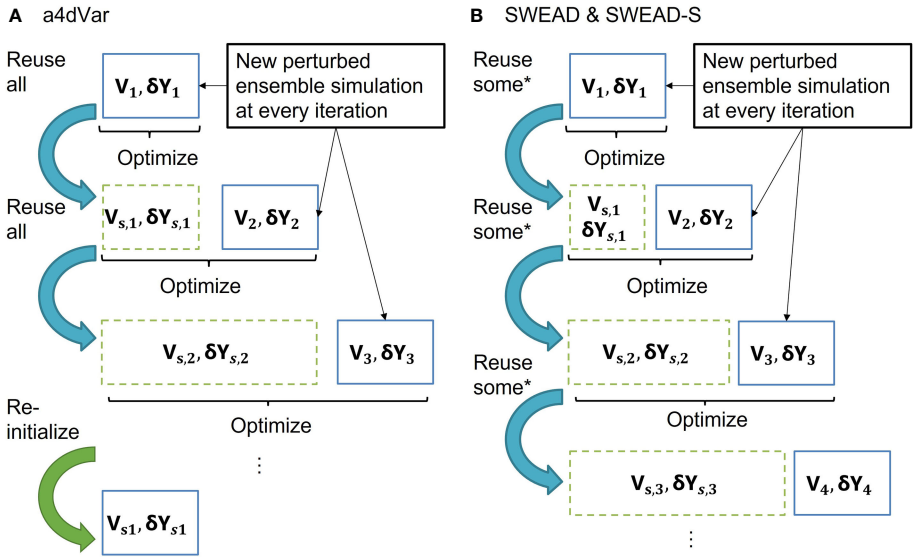


FIGURE 1
Schematic illustration of stacking algorithm for **(A)** a4dVar and **(B)** SWEAD and SWEAD-S. This figure is a modification of Fujimoto and Waseda (2020). © American Meteorological Society. Used with permission.

denotes the angular frequency, g denotes the gravitational acceleration, and k denotes the wavenumber. If the observed data are from a water level gauge and $\tilde{\mathbf{x}}$ is the Fourier coefficient of the initial surface elevation, then $(\mathbf{R}^{(-1/2)}\mathbf{A}'\mathbf{D}^{(1/2)})_{qr} = \exp[i(\omega_r t_q)]$, where q is the index of time and r is the index of the angular frequency and wavenumber. The approximated Jacobian is also decomposed as $\mathbf{A}' = \mathbf{U}\Sigma\mathbf{V}^*$. An example of the right singular vectors $\mathbf{V}_o^{\perp}$ and singular values $\Sigma_o^{\perp}$ are shown in Figure 3. Here, we assumed that the length of the time series of the water level gauge was $25T_p$ and that the spatial domain was $32\lambda_p$, where T_p denoted the peak wave period and λ_p denoted the peak wavelength. This setting is the same as that of the twin experiment of this study, as described later in Section 3.2. The spatial extents of the right singular vectors were limited to $< 25\lambda_p$, as shown in Figure 3D, and depended on their wavenumber components, as shown in Figure 3C.

2.3.2 Generating ensembles in the predictable zone

SWEAD uses Fourier modes, which are global bases and could be redundant for the limited predictable zone. SWEAD-S utilizes SVD to find the most effective perturbations to decrease the cost function.

The gradient is estimated with the approximated Jacobian and its SVD $\mathbf{A}' = \mathbf{U}\Sigma\mathbf{V}^*$ as:

$$\mathbf{A}'^{*}(\boldsymbol{A}(\mathbf{x}_n) - \mathbf{y}) = \mathbf{V}\Sigma^{*}\mathbf{U}^{*}(\boldsymbol{A}(\mathbf{x}_n) - \mathbf{y}). \tag{11}$$

Let a vector $\mathbf{d}$ be defined such that $\mathbf{V}\mathbf{d} = \mathbf{A}'^{*}(\boldsymbol{A}(\mathbf{x}_n) - \boldsymbol{y})$, and then:

$$\mathbf{d} = \Sigma^{*}\mathbf{U}^{*}(\boldsymbol{A}(\mathbf{x}_n) - \mathbf{y}). \tag{12}$$

$\mathbf{d}$ reflects larger singular values of the approximated Jacobian $\mathbf{A}'$ and the misfit and indicates the singular vectors that are most sensitive to the misfit. Therefore, SWEAD-S generates perturbations from the right singular vectors $\mathbf{V}$ corresponding to leading components of $\mathbf{d}$. SWEAD-S automatically neglects $\mathbf{V}_o$, which corresponds to the zero-singular value, and the generated ensembles are limited to the predictable zone.

As mentioned in Section 2.2, the perturbations for the next iteration $\mathbf{V}_n$ should be orthogonal to the stacked perturbations $\mathbf{V}_{s,n-1}$. After projecting the approximated Jacobian $\mathbf{A}'$ onto an orthogonal complement $\mathbf{V}_{s,n-1}^{\perp}$ of the projection operator $\mathbf{V}_{s,n-1}^{\perp}$

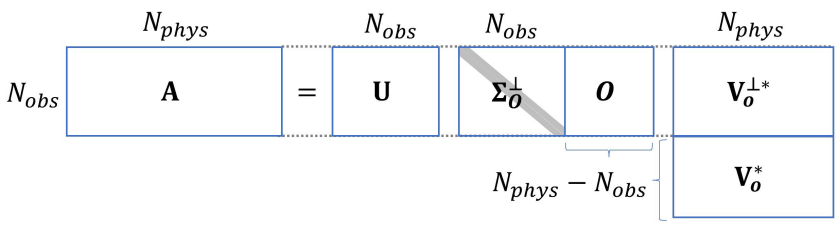


FIGURE 2
SVD of the Jacobian matrix $\mathbf{A} = \mathbf{U}\sum\mathbf{V}^*$.

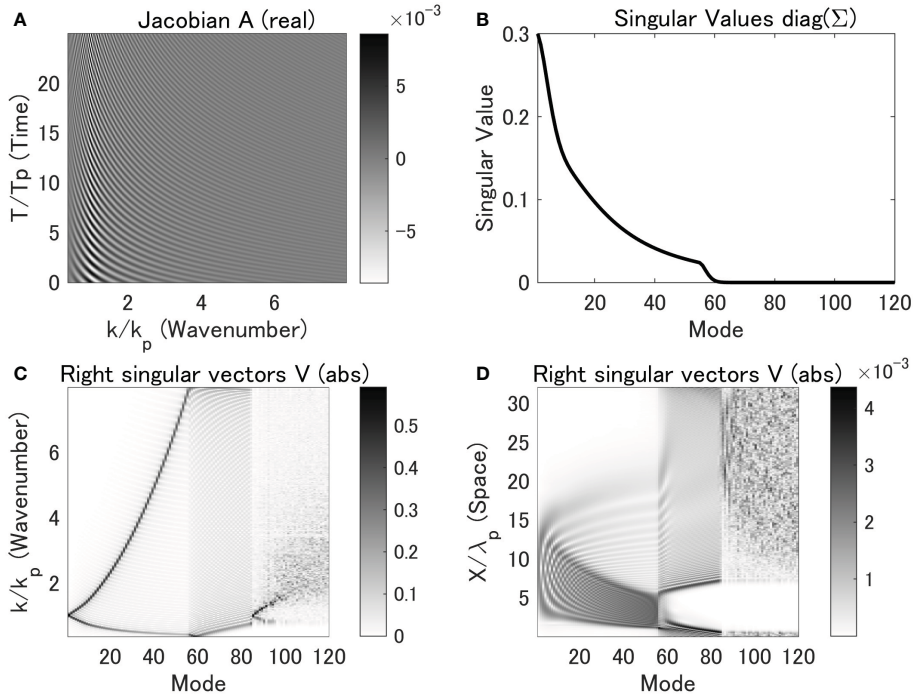


FIGURE 3
Approximated Jacobian obtained from the linear dispersion relation. **(A)** The real part of the approximated Jacobian A′, **(B)** the singular values of the approximated Jacobian, **(C)** the absolute values of the right singular vectors of the approximated Jacobian in the wavenumber domain, and **(D)** the absolute values of the right singular vectors transformed to the spatial domain.

$\mathbf{V}^{\perp\star}_{s,n-1}$, SWEAD-S calculates SVD as described below:

$$\mathbf{A}'\mathbf{V}^{\perp}_{s,n-1}\mathbf{V}^{\perp\star}_{s,n-1} = \mathbf{U}\Sigma\mathbf{V}^{\star},\ \mathbf{d} = \Sigma^{\star}\mathbf{U}^{\star}(\mathbf{A}(\mathbf{x}_n) - \mathbf{y})\,. \tag{13}$$

$\mathbf{V}^{\perp}_{s,n-1}$ is obtained by the Gram–Schmidt orthogonalization method. SWEAD-S selects new perturbations $\mathbf{V}_n$ from $\mathbf{V}$ corresponding to leading components of $\mathbf{d}$.

2.3.3 Updating the approximated Jacobian

If the approximated Jacobian is fixed, the optimization becomes slower because the non-linearity is not reflected in the perturbation generation, as demonstrated in Section 4. SWEAD-S updates the approximated Jacobian sequentially as $\mathbf{A}'\mathbf{V}_n \leftarrow \boldsymbol{\delta}\mathbf{Y}_n$.

From Eq. (2), the amplitudes of the spectral peaks and the rest are all normalized. However, because the non-linear wave interaction is active in a wavenumber range near the spectral peak, the generated perturbations should concentrate on the spectral peak. Therefore, $\mathbf{V}_n$ and the approximated Jacobian $\mathbf{A}'$ are restored to the original scale:

$$\mathbf{V}_{\mathbf{D}n} = \mathbf{D}^{1/2}\mathbf{V}_n, \tilde{\mathbf{A}}' = \mathbf{A}'\mathbf{D}^{-1/2}\,. \tag{14}$$

The approximated Jacobian is projected to the subspace spanned by $\mathbf{V}^{\star}_{\mathbf{D}n}$ with the projection operator $\mathbf{P}_{\mathbf{D}} = \mathbf{V}_{\mathbf{D}n}(\mathbf{V}^{\star}_{\mathbf{D}n}\mathbf{V}_{\mathbf{D}n})^{-1}\mathbf{V}^{\star}_{\mathbf{D}n}$ and updated with $\mathbf{A}'\mathbf{V}_n = \tilde{\mathbf{A}}'\mathbf{V}_{\mathbf{D}n} \rightarrow \boldsymbol{\delta}\mathbf{Y}_n$ as:

$$\begin{aligned}\tilde{\mathbf{A}}' &= \tilde{\mathbf{A}}' + \tilde{\mathbf{A}}'(\mathbf{P}_{\mathbf{D}} - \mathbf{P}_{\mathbf{D}})\\ &= \tilde{\mathbf{A}}' + \tilde{\mathbf{A}}'(\mathbf{V}_{\mathbf{D}n} - \mathbf{V}_{\mathbf{D}n})(\mathbf{V}^{\star}_{\mathbf{D}n}\mathbf{V}_{\mathbf{D}n})^{-1}\mathbf{V}\mathbf{D}^{\star}_n\,.\\ &\leftarrow \tilde{\mathbf{A}}' + (\boldsymbol{\delta}\mathbf{Y}_n - \mathbf{A}'\mathbf{V}_n)(\mathbf{V}^{\star}_{\mathbf{D}n}\mathbf{V}_{\mathbf{D}n})^{-1}\mathbf{V}^{\star}_{\mathbf{D}n}\end{aligned} \tag{15}$$

The second term on the right side of Eq. (11) indicates the updated $\boldsymbol{\delta}\mathbf{Y}_n$–$\mathbf{A}'\mathbf{V}_n$ of the linear approximation $\boldsymbol{\delta}\mathbf{Y} = \mathbf{A}\mathbf{V}$ by changing the control variable $\mathbf{x}_n$ projected to the subspace spanned by $\mathbf{V}_{\mathbf{D}n}$. By multiplying $\mathbf{D}^{1/2}$ by both sides of Eq. (15), we obtain the following equation of updated $\mathbf{A}'$:

$$\mathbf{A} \leftarrow \mathbf{A}' + (\boldsymbol{\delta}\mathbf{Y}_n - \mathbf{A}'\mathbf{V}_n)(\mathbf{V}^{*}_n\mathbf{D}\mathbf{V}_n)^{-1}\mathbf{V}^{*}_n\mathbf{D}\,. \tag{16}$$

2.3.4 Criteria for perturbations to be reused

The Hessian matrix $\mathbf{H}$ of the cost function $L(\mathbf{w}_n)$ is:

$$\mathbf{H} = \boldsymbol{\delta}\mathbf{Y}^{\star}\boldsymbol{\delta}\mathbf{Y} + \sum_{t=1}^{N_T}(A_t(\mathbf{x}_n) - y_t)\mathbf{V}^{\star}\nabla^2\mathbf{A}_t\mathbf{V}, \tag{17}$$

where A_t and y_t denote elements of $\boldsymbol{A}$ and $\mathbf{y}$, respectively, in the time index t. N_T is the total number of time steps. Let $\mathbf{S}$ denote the second term on the right side of the equation; then, $\mathbf{H}=\boldsymbol{\delta}\mathbf{Y}^{\star}\boldsymbol{\delta}\mathbf{Y}+\mathbf{S}$. Ensemble 4Dvar methods such as a4dVar can be regarded as a type of Gauss–Newton method because the second term $\mathbf{S}$ is truncated as $\mathbf{H}=\boldsymbol{\delta}\mathbf{Y}^{\star}\boldsymbol{\delta}\mathbf{Y}$. If $\mathbf{x}_n$ is close to the optimal solution, $A_t(\mathbf{x}_n)$–y_t is small, and this approximation is reasonable. Otherwise, $\mathbf{S}$ should be considered for the optimization.

The following condition, called the secant condition, should hold for the $\mathbf{H}=\boldsymbol{\delta}\mathbf{Y}^{\star}\boldsymbol{\delta}\mathbf{Y}+\mathbf{S}$ and the gradient $\nabla L(\mathbf{x}_n)$:

$$\mathbf{H}\mathbf{w}_n = (\boldsymbol{\delta}\mathbf{Y}^{\star}\boldsymbol{\delta}\mathbf{Y} + \mathbf{S})\mathbf{w}_n = \mathbf{V}^{\star}\nabla L(\mathbf{x}_{n+1}) - \mathbf{V}^{\star}\nabla L(\mathbf{x}_n)\,. \tag{18}$$

The secant condition underlies the derivation of quasi-Newtonian methods such as the Broyden–Fletcher–Goldfarb–Shanno (BFGS) formula (Nocedal and Wright, 2006).

Let us introduce $\mathbf{w}_n^{'}$, which satisfies the following equation:

$$\boldsymbol{\delta Y}^{\star}\boldsymbol{\delta Y}\mathbf{w}_n^{'} \equiv \mathbf{V}^{\star}\nabla L(\mathbf{x}_{n+1}) - \mathbf{V}^{\star}\nabla L(\mathbf{x}_n)\,. \tag{19}$$

SWEAD-S considers $\mathbf{S}$ by comparing $\mathbf{w}_n$ and $\mathbf{w}_n^{'}$.
If $\boldsymbol{\delta Y}^{\star}\boldsymbol{\delta Y}$ and $\mathbf{S}$ are both diagonal matrices, then:

$$\begin{aligned}(\mathbf{w}_n)_i &= \frac{1}{(\boldsymbol{\delta Y}^{\star}\boldsymbol{\delta Y})_{ii}+(\mathbf{S})_{ii}}(\mathbf{V}^{\star}\nabla L(\mathbf{x}_{n+1}) - \mathbf{V}^{\star}\nabla L(\mathbf{x}_n))_i \\ (\mathbf{w}_n^{'})_i &= \frac{1}{(\boldsymbol{\delta Y}^{\star}\boldsymbol{\delta Y})_{ii}}(\mathbf{V}^{\star}\nabla L(\mathbf{x}_{n+1}) - \mathbf{V}^{\star}\nabla L(\mathbf{x}_n))_i\,. \\ \therefore\ \frac{(\mathbf{w}_n)_i}{(\mathbf{w}_n^{'})_i} - 1 &= \frac{(\mathbf{S})_{ii}}{(\boldsymbol{\delta Y}^{\star}\boldsymbol{\delta Y})_{ii}}\,.\end{aligned} \tag{20}$$

$(\mathbf{w}_n)_i/(\mathbf{w}_n^{'})_i - 1$ indicates the ratio of $\mathbf{S}$ compared with $\boldsymbol{\delta Y}^{\star}\boldsymbol{\delta Y}$ in each ensemble dimension. For the assumption $\mathbf{H}=\boldsymbol{\delta Y}^{\star}\boldsymbol{\delta Y}$ to be valid, $\mathbf{S}$ should be suppressed. Therefore, only perturbations satisfying the following criterion are reused in SWEAD-S:

$$\left|\frac{(\mathbf{w}_n)_i}{(\mathbf{w}_n^{'})_i} - 1\right| < e_{TOL}, \tag{21}$$

where e_{TOL} denotes the error tolerance. In SWEAD, the following criterion equation is used:

$$\frac{\left|(\mathbf{w}_n)_i - (\mathbf{w}_n^{'})_i\right|}{\sigma_{wn}} < e_{TOL}\,. \tag{22}$$

σ_{wn} denotes the standard deviation of $\mathbf{w}_n$. Fujimoto and Waseda (2020) did not offer a rationale for the old criterion in Eq. (22), but now the new criterion in Eq. (21) has a rationale.

For $\boldsymbol{\delta Y}^{\star}\boldsymbol{\delta Y}$ and $\mathbf{S}$ to both be diagonal matrices, $\boldsymbol{\delta Y}^{\star}\boldsymbol{\delta Y}$ needs to be an eigenvalue decomposed as $\boldsymbol{\delta Y}^{\star}\boldsymbol{\delta Y} = \mathbf{M}\boldsymbol{\Lambda}\mathbf{M}^{\star}$, where $\boldsymbol{\Lambda}$ denotes the eigenvalue matrix. $\boldsymbol{\delta Y}^{\star}\boldsymbol{\delta Y}$ is a Hermitian matrix and $\mathbf{M}$ is a unitary matrix; then, $\mathbf{M}\mathbf{M}^{\star}=\mathbf{M}^{\star}\mathbf{M}=\mathrm{I}$. Although $\mathbf{S}$ has N_{ens}^2 elements, $\mathbf{w}_n$ and $\mathbf{w}_n^{'}$ are vectors of N_{ens} elements, and only N_{ens} elements of $\mathbf{S}$ can be estimated. We assumed that $\mathbf{M}^{\star}\mathbf{S}\mathbf{M}$ is a diagonal matrix and estimated the N_{ens} diagonal elements using a heuristic approach:

$$\boldsymbol{\delta Y}^{\star}\boldsymbol{\delta Y} + \mathbf{S} = \mathbf{M}\boldsymbol{\Lambda}\mathbf{M}^{\star} + \mathbf{S} = \mathbf{M}(\boldsymbol{\Lambda} + \mathbf{M}^{\star}\mathbf{S}\mathbf{M})\boldsymbol{M}^{\star} \tag{23}$$

After these matrices are replaced as $\boldsymbol{\delta Y}\leftarrow\boldsymbol{\delta Y}\mathbf{M}$, $\mathbf{V}\leftarrow\mathbf{V}\mathbf{M}$, and $\mathbf{S}\leftarrow\mathbf{M}^{\star}\mathbf{S}\mathbf{M}$, Eq. (18) still holds. $\boldsymbol{\delta Y}^{\star}\boldsymbol{\delta Y}$ and $\mathbf{S}$ are already diagonalized, and Eq. (20) is valid.

Note that the difference in gradients was approximated by:

$$\begin{aligned}&\mathbf{V}^{\star}\nabla L(\mathbf{x}_{n+1}) - \mathbf{V}^{\star}\nabla L(\mathbf{x}_n) \\ &= \mathbf{V}^{\star}\mathbf{A}^{\star}(\boldsymbol{A}(\mathbf{x}_{n+1}) - \mathbf{y}) - \mathbf{V}^{\star}\mathbf{A}^{\star}(\boldsymbol{A}(\mathbf{x}_n) - \mathbf{y}) \\ &\approx \boldsymbol{\delta Y}^{\star}(\boldsymbol{A}(\mathbf{x}_{n+1}) - \boldsymbol{A}(\mathbf{x}_n))\,.\end{aligned} \tag{24}$$

2.3.5 Inclusion of the regularization term

Eqs. (9) and (19) are rewritten to include the regularization term:

$$(\boldsymbol{\delta Y}^{\star}\boldsymbol{\delta Y} + \mathbf{V}^{\star}\mathbf{V})\mathbf{w}_n = -\boldsymbol{\delta Y}^{\star}(\boldsymbol{A}(\mathbf{x}_n) - \mathbf{y}) - \mathbf{V}^{\star}\mathbf{x} \tag{25}$$

$$(\boldsymbol{\delta Y}^{\star}\boldsymbol{\delta Y} + \mathbf{V}^{\star}\mathbf{V})\mathbf{w}_{\mathbf{n}}^{'} \equiv \boldsymbol{\delta Y}^{\star}(\boldsymbol{A}(\mathbf{x}_{n+1}) - \boldsymbol{A}(\mathbf{x}_n)) + \mathbf{V}^{\star}\mathbf{V}\mathbf{w}_n\,. \tag{26}$$

Additionally, for the eigenvalue decomposition of $\boldsymbol{\delta Y}^{\star}\boldsymbol{\delta Y}$, the background error covariance matrix is also considered:

$$\boldsymbol{\delta Y}^{\star}\boldsymbol{\delta Y} + \mathbf{V}^{\star}\mathbf{V} = \mathbf{M}\boldsymbol{\Lambda}\mathbf{M}^{\star} \tag{27}$$

$$\boldsymbol{\delta Y} \leftarrow \boldsymbol{\delta Y}\mathbf{M},\ \ \mathbf{V} \leftarrow \mathbf{V}\mathbf{M}\,. \tag{28}$$

The differences between SWEAD and SWEAD-S are summarized in Table 1.

3 Twin experiment for non-linear deep-water waves

As described in Section 2.3.3, SWEAD-S updates the approximation of the Jacobian matrix, starting from the linear dispersion relationship. In this study, we tested SWEAD-S for deep-water waves, which are a typical example of non-linear dispersive waves.

Non-linearity is essential for accurately predicting propagation of deep-water waves (e.g., Mei et al., 2005). The third-order non-linearity increases the propagation speed of Stokes waves. If a is wave amplitude, then the angular frequency of the Stokes wave is $\omega^2=gk(1+1/2\ a^2k^2)$. This equation means that the wave phase velocity ω/k increases with $1/2\ a^2k^2$. The third-order non-linearity could also increase the probability of large waves in irregular wave fields (Janssen, 2003). For irregular waves, H_{m0} denotes the significant wave height, which is a typical wave height corresponding to 4σ, where σ is the standard deviation of the surface elevation. If the crest height of a wave is $> 1.25H_{m0}$, or if the wave height is $> 2H_{m0}$, the wave is commonly called a “freak wave” or “rogue wave” (Haver, 2004). Recent studies on non-linear deep-water waves have been summarized by Waseda (2019).

The higher-order spectral method (HOSM) (Dommermuth and Yue, 1987; West et al., 1987) is a promising method for predicting the propagation of non-linear deep-water waves and has been used in many studies (e.g., Ducrozet et al., 2007; Xiao et al., 2013; Bitner-Gregersen et al., 2020). The advantage of HOSM is that it can be applied to a wave field with a broad spectrum, like real ocean waves.

Some studies (Yoon et al., 2015; Wang and Pan, 2021) have applied the Kalman filter to HOSM. Additionally, other studies (Wu, 2004; Aragh et al., 2008; Blondel-Couprie et al., 2010; Blondel-Couprie et al., 2013; Qi et al., 2016; Köllisch et al., 2018) have applied the variational method to HOSM. The Kalman filter is adequate when the observed data are sufficient, but otherwise it might suffer filter divergence. We adopted the variational method in this study because it is relatively stable, even if the observed data are insufficient.

As described in the next section, HOSM is based on the Taylor expansion of governing equations of water waves and includes many expanded terms, which can make the implementation of the

TABLE 1 Scheme of SWEAD and SWEAD-S.

	Procedure	SWEAD	SWEAD-S
Initialization	Setting the first guess	Set the first guess $\mathbf{x_0}$ assuming that the system is linear. Initialize the misfit $A(\mathbf{x_0})$ - $\mathbf{y}$ by the physical model (HOSM in this study) and estimate the misfit power spectrum	
	Initialization of the approximated Jacobian matrix		Approximated Jacobian $\mathbf{A'}$ is initialized by the linear dispersion relation
Loop until convergence	Ensemble generation	Choose new cosine or sine wave perturbations from wavenumber components around the peak of the misfit power spectrum, except for components already considered in previous perturbations $\mathbf{V}_{s,n-1}$. $\mathbf{V}_n = (\cdots\ v_{n(m)}\ \cdots)$	Conduct SVD of the approximated Jacobian $\mathbf{A'}$. From Eq. (13) (or Eq. (12) in the first iteration), select a few ensembles from leading components of $\mathbf{d}$ and extract them from the corresponding right singular vectors. $\mathbf{V}_n = (\cdots\ v_{n(m)}\ \cdots)$
	Ensemble simulation	Perform the unperturbed and perturbed ensemble simulations and evaluate the respective value. $\mathbf{\delta Y}_{n(m)}=\mathbf{A}(\mathbf{x}_n+\mathbf{v}_{n(m)})-\boldsymbol{A}(\mathbf{x}_n)$ $(m=1...N_{ens})$	
	Updating the approximated Jacobian		Update the approximated Jacobian using Eq. (16)
	Calculate $\mathbf{w}_n'$	If the iteration is after the second, calculate $\mathbf{w}_n'$ using Eq. (26)	
	Choose perturbations to be reused	Choose the perturbations $\mathbf{V}_s$ satisfying the old criterion Eq. (22)	Choose the perturbations $\mathbf{V}_s$ satisfying the new criterion Eq. (21)
	Stacking	Stack the chosen previous perturbations and the newer perturbations $(\mathbf{V}_{s,n-1}\|\mathbf{V}_n)\rightarrow\mathbf{V}$, $(\mathbf{\delta Y}_{s,n-1}\|\mathbf{\delta Y}_n)\rightarrow\mathbf{\delta Y}$	
	Diagonalization		Diagonalize $\mathbf{\delta Y}$ and $\mathbf{V}$ using Eqs. (27) and (28)
	Updating the solution	Solve Eq. (25) and update the initial conditions and perturbations $\mathbf{x}_{n+1}=\mathbf{x}_n+\mathbf{V}\mathbf{w}_n$, $\mathbf{V}\rightarrow\mathbf{V}_{s,n}$, $\mathbf{\delta Y}\rightarrow\mathbf{\delta Y}_{s,n}$	

adjoint method for HOSM difficult. Therefore, SWEAD and SWEAD-S adopt the ensemble-based variational method.

3.1 Wave model: HOSM

Deep-water waves can be regarded as inviscid, irrotational, free-surface flows. The governing equations of deep-water waves are as follows (Zakharov, 1968):

The equation of continuity:

$$\frac{\partial^2 \phi}{\partial z^2} = -\nabla^2 \phi,\ -d \le z \le \eta \tag{29}$$

The bottom boundary condition:

$$\frac{\partial \boldsymbol{\Phi}}{\partial z} = 0\ \ at\ \ z = -\infty \tag{30}$$

Kinematic free surface boundary condition:

$$\frac{\partial \eta}{\partial t} + \nabla \boldsymbol{\Phi} \cdot \nabla \eta = (1 + (\nabla \eta)^2) W\ at\ \ z = \eta \tag{31}$$

Dynamic free surface boundary condition (Bernoulli's law):

$$\frac{\partial \boldsymbol{\Phi}}{\partial t} + \frac{1}{2}(\nabla \boldsymbol{\Phi})^2 + g\eta = \frac{1}{2}(1 + (\nabla \eta)^2) W^2\ at\ \ z = \eta, \tag{32}$$

where

$\eta(x,\ y,\ t)$ is the surface elevation

$\nabla = (\partial_x,\ \partial_y)$: is the nabla for the horizontal axis

$\boldsymbol{\Phi}(x, y, t) = \phi(x, y, z = \eta, t)$ is the velocity potential at the surface

$W = \partial_z \phi(x, y, z = \eta, t)$ is the vertical velocity at the surface

g : is the gravitational acceleration.

In the governing equations, the vertical surface velocity W is unknown. By assuming non-breaking waves, HOSM expands the vertical velocity as:

$$W = \sum_{m=1}^{M} w^{(m)} \tag{33}$$

$$w^{(m)} = \sum_{l=0}^{m-1} \frac{\eta^l}{l!} \frac{\partial^{l+1}}{\partial z^{l+1}}\ \phi^{(m-l)}\ |_{z=0} \tag{34}$$

$$\begin{aligned} &O(1): \phi^{(1)}\ |_{z=0} = \boldsymbol{\Phi} \\ &O(m):\ \phi^{(m)}\ |_{z=0} = -\sum_{l=1}^{m-1} \frac{\eta^l}{l!} \frac{\partial^l}{\partial z^l}\ \phi^{(m-l)}\ |_{z=0}, m \ge 2 \end{aligned} \tag{35}$$

Substituting W with the free-surface boundary conditions in Eqs. (32) and (33) (West et al., 1987):

$$\frac{\partial \eta}{\partial t} = -\nabla \Phi \cdot \nabla \eta + \sum_{m=1}^{M} w^{(m)} + (\nabla \eta)^2 \sum_{m=1}^{M-2} w^{(m)},$$

$$\frac{\partial \Phi}{\partial t} = -\frac{1}{2}(\nabla \Phi)^2 - g\eta + \frac{1}{2}\sum_{m=1}^{M-1} w^{(m)} w^{(M-m)} \quad (36)$$

$$+\frac{1}{2}(\nabla \eta)^2 \sum_{m=1}^{M-3} w^{(m)} w^{(M-2-m)}.$$

HOSM solves these equations under periodic boundary conditions for utilizing the fast Fourier transform (FFT) to evaluate spatial derivatives. Because $w^{(m)}$ consists of M terms of $\phi^{(m)}|_{z=0}$ from Eq. (34), W involves $O(M^2)$ terms. HOSM can represent the free-surface boundary conditions accurately if terms of the fifth and higher orders are included, but this can result in boundary conditions that consist of several tens of terms.

3.2 Configuration of the twin experiments

To compare the performance of SWEAD and SWEAD-S, we conducted a twin experiment similar to that of Fujimoto and Waseda (2020). Time series of surface elevations were extracted from the output of the HOSM simulation initialized by a given spectrum and were contaminated by random noise. Those time series were considered as virtual observed data, which were assimilated into HOSM. Then, the whole wave field was estimated by SWEAD or SWEAD-S, and the true and estimated wave fields were compared.

HOSM generated a freak wave with a crest height of 1.5 H_{m0}. The generated wave field was taken as the truth. The power spectrum is a standard wave spectrum: the JONSWAP spectrum with $\gamma = 3.3$ (Hasselmann et al., 1973). The wave steepness was $H_{m0}k_p/2=0.11$ so that the non-linearity of the wave field was significant; $k_p=2\pi/\lambda_p$ was the peak wavenumber. The computational domain for the initial simulation of the truth was $128\lambda_p$ to suppress the influence of the periodic boundary condition. In the wavenumber region, the computational domain spanned up to $8k_p$ (Tanaka and Yokoyama, 2004). According to Dommermuth (2000), a linear wave field gradually transitions to a non-linearly consistent wave field, including bound waves. The control variable $\mathbf{x}$ was set to the initial value of the water level before the non-linear spin-up ($t=-5T_p$). The time step was set to $\Delta t=T_p/50$, and the fourth-order Runge–Kutta method was used.

The water level time series, including the freak wave, was used as the observed data (Figure 4). The white Gaussian noise was added to the time series, and the standard deviation was 10% of the standard deviation of the original water level time series. To emulate a situation in which the computational domain was redundant when compared with the predictable zone, the computational domain was set to $32\lambda_p$ (Figure 4), which was roughly twice as large as the linearly predictable zone corresponding peak wavenumber (LPZP) of $15\lambda_p$ for the observed time series $25T_p$.

Owing to insufficient observed data, the minimization of the cost function could be unstable. Regularization is a technique to stabilize a solution to an ill-posed problem by constraining the solution with prior information (Tikhonov and Arsenin, 1979). In SWEAD and SWEAD-S, the control variable $\tilde{\mathbf{x}}$ is the Fourier coefficient of the initial wave field in the wavenumber space, and $\mathbf{D}$ is a diagonal matrix with a prior estimation of the power spectrum $S(k)$ in its diagonal components, i.e., $\mathrm{diag}(\mathbf{D})= \alpha S(k)$. α is the regularization parameter. In this twin experiment, $S(k)$ was the JONSWAP spectrum defined above. In reality, $S(k)$ must be obtained by some other means, for example by spectral wave models such as WAVEWATCH III (Tolman, 2016). The regularization parameter α was determined to be $\alpha = 0.001$ by the L-curve method (Hansen, 1992). Data assimilations were performed with 10 realizations of the noise with 10 ensembles $N_{ens}=10$. e_{TOL} was selected as the best value for the old and new criteria: $e_{TOL} = 0.2$ and 0.5, respectively.

4 Results and discussion

The methods SWEAD and SWEAD-S were compared. As shown in Table 2, (a) is the conventional method, SWEAD, (b) is

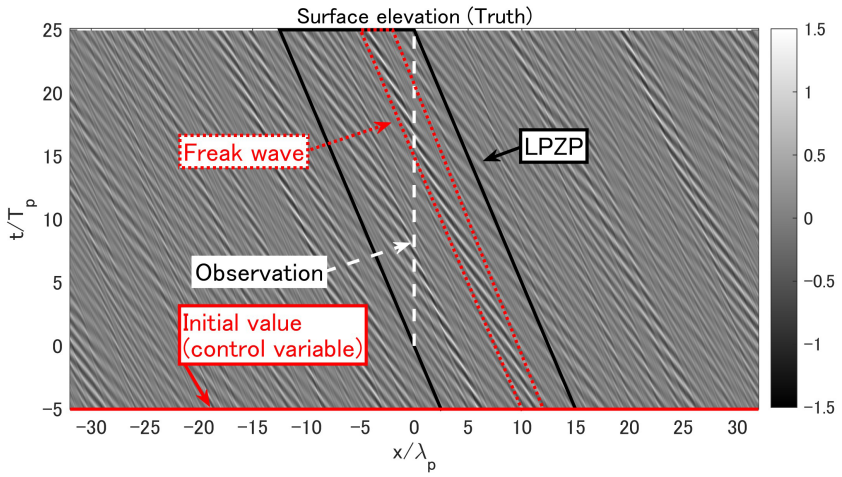


FIGURE 4
Generated truth (contours), measured points in the water level time series (dashed white lines), and linearly predictable zone corresponding peak wavelength (LPZP, solid black line). The wave groups, including the freak wave, are shown as red dotted lines. The initial value was taken as the control variable in the analysis.

the SWEAD-S variant using the new criterion of reusing the perturbations in Eq. (21), and (c–e) are SWEAD-S variants, differing in whether they conduct the Jacobian update in Eqs. (12) and (13) and the diagonalization in Eq. (16).

Figure 5 shows the root mean square error (RMSE) (Figure 5A) and correlation (Figure 5B) within the LPZP for each iteration averaged over the 10 realizations of the white Gaussian noise, and Figure 6 shows the averaged cost function for each iteration.

4.1 Improvements by generating perturbations with SVD

SWEAD (a) and SWEAD-S (c) at the 20th iteration are compared in Figure 7. Because SWEAD (Figure 7A) did not explicitly consider the predictable zone in the ensemble generation, there was innovation, a difference between the analytical value and the linear first guess, beyond the LPZP. In contrast, as shown in Figure 7B, the innovation of SWEAD-S (c) was within the range of the LPZP because the SVD concentrated the search direction within the predictable zone.

The optimization efficiency of SWEAD-S (c) exceeded that of SWEAD (a) by up to approximately 85 iterations (Figure 5). However, SWEAD (a) caught up with the SWEAD-S variant that fixed the approximated Jacobian (c) in approximately 85 iterations (Figure 5). As described next, the method using SVD can be improved by updating the approximated Jacobian.

4.2 Improvements by updating the approximated Jacobian

The SWEAD-S variant with the updated Jacobian (d) optimized more efficiently than that without the updated Jacobian (c) after the 20th iteration (Figure 6). The variant with the updated Jacobian (d) also outperformed SWEAD (a) in terms of efficiency and accuracy (Figure 5). The correlation reached 0.9 in 94 iterations in SWEAD (a), but in 52 iterations in the SWEAD-S variant with the updated Jacobian (d). In other words, the SWEAD-S variant with the updated Jacobian (d) was twice as fast as SWEAD (a).

The reason for this improvement could be that the approximated Jacobian given by the linear dispersion relation (Figure 3) limited the predictable zone of each mode up to $x = 25\lambda_p$, which corresponds to the LPZP. Meanwhile, no such limit appeared in the updated approximated Jacobian (Figure 8). As shown in Figure 8B, the number of non-zero-singular values of the approximated Jacobian obtained after 100 iterations in the SWEAD-S variant with the updated Jacobian (d) was approximately 100. In contrast, that of the approximated Jacobian given by the linear dispersion relation was approximately 50 (Figure 3B). The spread of right singular vectors in the spatial dimension of SWEAD-S (e) increased from the linear dispersion relation, as shown in Figures 3D, 8D. Therefore, updating the approximated Jacobian is essential to expand the predictable zone. The non-linearity increases the wave group velocity, and the predictable zone should expand. It seems that this expansion of the predictable zone was reflected by the updates in the approximated Jacobian.

4.3 Effects of different criteria on reusing perturbations

We confirmed that the old criterion in Eq. (22) resulted in an unstable optimization in SWEAD-S and then devised the new criterion in Eq. (21). Although the old criterion optimized faster than the new criterion in the SWEAD variant with the new criterion (b) (Figure 6), the SWEAD-S variants combining SVD and the new criterion formula (d, e) were the most computationally efficient among all cases.

The SWEAD-S variant with the diagonalization in Eqs. (27) and (28) (e) was optimized faster than the variant without diagonalization (d) at the early stage of the iteration, but variant (e) was caught up later by variant (d). Although the diagonalization was introduced to clarify the rationale of the new criterion equation, it had little advantage for optimization efficiency. As the optimization progresses, term $\mathbf{S}$ decreases, and the criterion equation should become uncritical.

The analytical values obtained after 100 iterations of the SWEAD-S method (e) are shown in Figure 9. As shown in Figure 4, the freak wave was located at approximately $x = 10\lambda_p$ at $t=-5T_p$. In SWEAD-S, the freak wave was reproduced with sufficient accuracy.

5 Conclusion

This study proposes the use of SWEAD-S, which uses the SVD of the approximated Jacobian to generate perturbations only in the predictable region. SWEAD-S updates the approximated Jacobian

TABLE 2 Procedures of SWEAD and some variants of SWEAD-S.

	Ensemble generation	Criterion for ensemble reuse	Jacobian update	Diagonalization
a. SWEAD	Fourier	Old criterion Eq. (22)		
b. SWEAD with the new criterion	Fourier	New criterion Eq. (21)		
c. SWEAD-S without the Jacobian update	SVD: Eqs. (12) and (13)	New criterion Eq. (21)		
d. SWEAD-S without the diagonalization	SVD: Eqs. (12) and (13)	New criterion Eq. (21)	Eq. (16)	
e. SWEAD-S	SVD: Eqs. (12) and (13)	New criterion Eq. (21)	Eq. (16)	Eqs. (27) and (28)

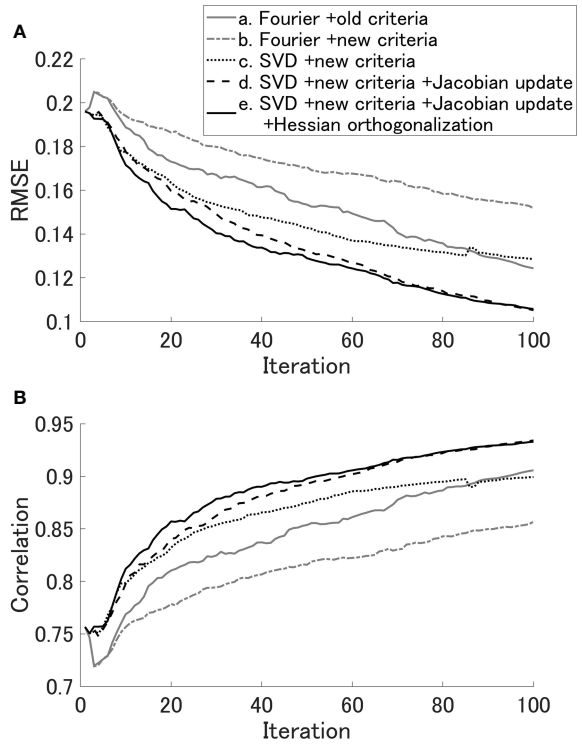


FIGURE 5
Estimation error within the LPZP averaged over the 10 realizations of the white Gaussian noise.

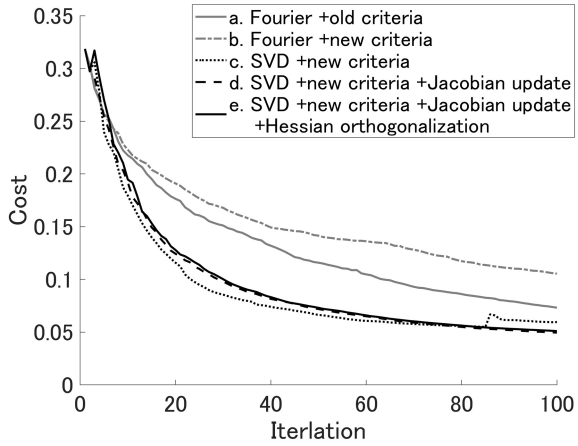


FIGURE 6
Cost function averaged over the 10 realizations of the white Gaussian noise.

for generating ensembles, considering non-linearity. Furthermore, we devised a new criterion equation with a clear rationale for reusing perturbations by referring to the secant condition. This method is relevant where the physical system is weakly non-linear and a linear dispersion relation can roughly approximate the Jacobian. Non-linear deep-water waves are an appropriate example. We tested SWEAD-S using a twin experiment on a large wave called a freak wave, which was generated by HOSM. SWEAD-S reconstructed the freak wave well from only time series data of surface elevation. Furthermore, the optimization speed of SWEAD-S was twice as fast as that of SWEAD. Updating the approximated Jacobian contributes to improving the convergence speed and estimation accuracy by reflecting the expansion of the predictable zone due to non-linearity.

SWEAD-S is not limited to deep-water waves and might apply to other media of weakly non-linear dispersive waves. Nonetheless, strongly non-linear phenomena involving wave breaking are currently unsuitable for SWEAD-S. SWEAD-S is thought to apply to multi-directional waves; Fujimoto and Waseda (2020) demonstrated that SWEAD was applicable to a multi-directional wave field in a $32\lambda_p$ square, where the physical dimension of HOSM was $O(\sim 10^5)$. However, a supercomputer was needed for such a high-dimensional problem, and a further decrease in computational burden is required. Storing the approximated Jacobian requires a large memory, limiting the dimensions of the physical space.

Data assimilation with a phase-resolved non-linear wave model such as HOSM will have many uses in both industry and academia. The wave field around ships or offshore structures could be monitored for marine safety. Additionally, the wave estimation itself could be a tool used for researching wave dynamics both in the ocean and in wave tanks. This study offers a theoretical framework for data

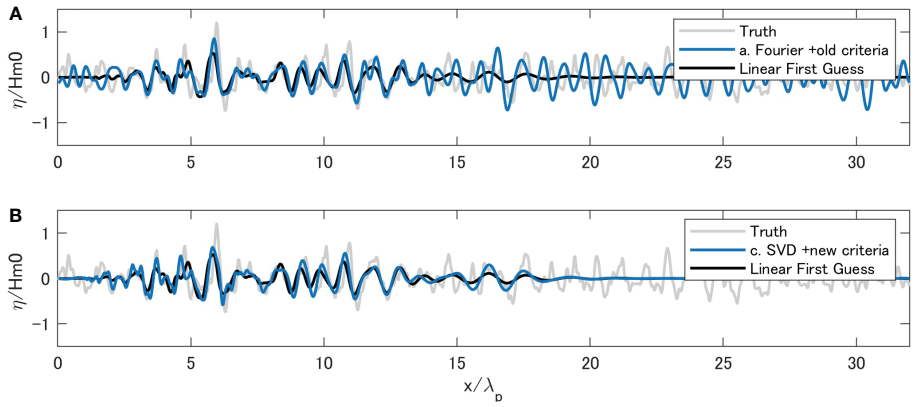


FIGURE 7
The analyses of SWEAD (a) and SWEAD-S variant (c) are compared to the truth and the linear first guess in panel **(A, B)**, respectively.

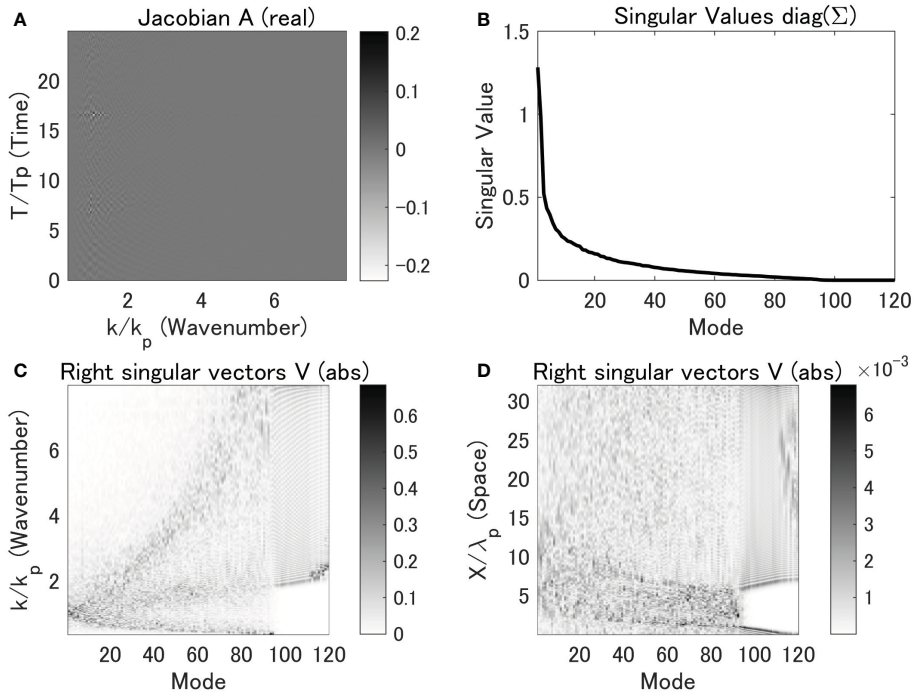


FIGURE 8
Approximated Jacobian obtained from the linear dispersion relation. **(A)** The real part of the approximated Jacobian A′, **(B)** the singular values of the approximated Jacobian, **(C)** the absolute values of the right singular vectors of the approximated Jacobian in the wavenumber domain, and **(D)** the absolute values of the right singular vectors transformed to the spatial domain.

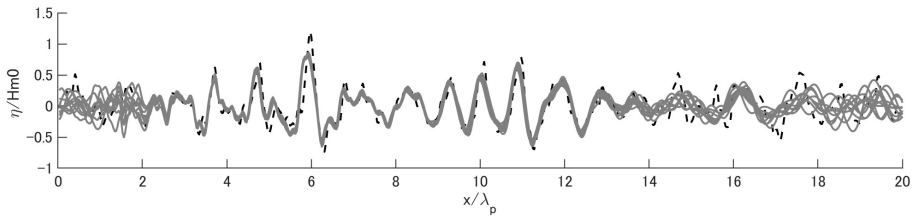


FIGURE 9
Analytical values for the initial value (t = -5T_p) obtained after 100 iterations with SWEAD-S (e). The truth is shown as a black dashed line, and the analytical value is shown as a gray dashed line for each realization. The wave group leading to the freak wave existed around x/λ_p = 11 (Figure 4).

assimilation of deep-water waves. However, SWEAD-S was assessed only *via* simulations; integrating the data assimilation technique and measurements remains challenging. Future studies could develop a method to represent modeling and observational errors in the ocean.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

WF: Investigation, conceptualization, methodology, software, writing-original draft. KI: Review and editing, project administration, resources. All authors contributed to the article and approved the submitted version.

Acknowledgments

The authors acknowledge Takuji Waseda, who provided valuable advice during preparation of the final version of the manuscript. We would like to thank Editage (www.editage.com) for English language editing.

Conflict of interest

WF and KI are employees of Nippon Kaiji Kyokai (ClassNK).

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Aragh, S., Nwogu, O., and Lyzenga, D. (2008). "Improved estimation of ocean wave fields from marine radars using data assimilation techniques," in *The Eighteenth International Offshore and Polar Engineering Conference* (Vancouver, Canada) 8, 565. Available at: http://www.miseagrant.umich.edu/downloads/research/journals/08-330.pdf.

Bitner-Gregersen, E. M., Gramstad, O., Magnusson, A. K., and Sames, P. C. (2020). "Occurrence frequency of a triple rogue wave group in the ocean," in *Proceedings of the ASME 2020 39th International Conference on Ocean, Offshore and Arctic Engineering. Volume 2A: Structures, Safety, and Reliability* (American Society of Mechanical Engineers: Virtual Conference, Online). doi: 10.1115/OMAE2020-19314

Blondel-Couprie, E., Bonnefoy, F., and Ferrant, P. (2010). Deterministic non-linear wave prediction using probe data. *Ocean Eng.* 37, 913–926. doi: 10.1016/j.oceaneng.2010.03.002

Blondel-Couprie, E., Bonnefoy, F., and Ferrant, P. (2013). Experimental validation of non-linear deterministic prediction schemes for long-crested waves. *Ocean Eng.* 58, 284–292. doi: 10.1016/j.oceaneng.2012.10.014

Dommermuth, D. G. (2000). The initialization of nonlinear waves using an adjustment scheme. *Wave Motion* 32, 307–317. doi: 10.1016/S0165-2125(00)00047-0

Dommermuth, D. G., and Yue, D. K. P. (1987). A high-order spectral method for the study of nonlinear gravity waves. *J. Fluid Mech.* 184, 267–288. doi: 10.1017/S002211208700288X

Ducrozet, G., Bonnefoy, F., Le Touzé, D., and Ferrant, P. (2007). 3-d HOS simulations of extreme waves in open seas. *Natural Hazards Earth Syst. Sci.* 7, 109–122. doi: 10.5194/nhess-7-109-2007

Fujimoto, W., and Waseda, T. (2020). Ensemble-based variational method for nonlinear inversion of surface gravity waves. *J. Atmos. Ocean Technol.* 37, 17–31. doi: 10.1175/JTECH-D-19-0072.1

Hansen, P. C. (1992). Analysis of discrete ill-posed problems by means of the l-curve. *SIAM Rev.* 34, 561–580. doi: 10.1137/1034115

Hasselmann, K., Barnett, T. P., Bouws, E., Carlson, H., Cartwright, D. E., Enke, K., et al. (1973). *Measurements of wind-wave growth and swell decay during the joint north Sea wave project (JONSWAP)* (Deutches Hydrographisches Institut). Available at: http://repository.tudelft.nl/view/hydro/uuid%3Af204e188-13b9-49d8-a6dc-4fb7c20562fc/.

Haver, S. (2004). "A possible freak wave event measured at the draupner jacket January 1 1995," in *Rogue waves 2004 : proceedings of a workshop organized by Ifremer and held in Brest, France.* 1–8. Available at: http://www.ifremer.fr/web-com/stw2004/rw/fullpapers/walk_on_haver.pdf.

Janssen, P. A. E. M. (2003). Nonlinear four-wave interactions and freak waves. *J. Phys. Oceanogr.* 33, 863–884. doi: 10.1175/1520-0485(2003)33<863:NFIAFW>2.0.CO;2

Köllisch, N., Behrendt, J., Klein, M., and Hoffmann, N. (2018). Nonlinear real time prediction of ocean surface waves. *Ocean Eng.* 157, 387–400. doi: 10.1016/j.oceaneng.2018.03.048

Liu, C., Xiao, Q., and Wang, B. (2008). An ensemble-based four-dimensional variational data assimilation scheme. part I: Technical formulation and preliminary test. *Mon. Weather Rev.* 136, 3363–3373. doi: 10.1175/2008MWR2312.1

Liu, C., Xiao, Q., and Wang, B. (2009). An ensemble-based four-dimensional variational data assimilation scheme. part II: Observing system simulation experiments with advanced research WRF (ARW). *Mon. Weather Rev.* 137, 1687–1704. doi: 10.1175/2008MWR2699.1

Mei, C. C., Stiassnie, M., and Yue, D. K. P. (2005). *Theory and applications of ocean surface waves* (World Scientific).

Nocedal, J., and Wright, S. J. (2006). "Fundamentals of unconstrained optimization," in *Numerical optimization* (New York, NY: Springer New York), 10–29. doi: 10.1007/978-0-387-40065-5_2

Panteleev, G., Yaremchuk, M., and Rogers, W. E. (2015). Adjoint-free variational data assimilation into a regional wave model. *J. Atmos. Ocean Technol.* 32, 1386–1399. doi: 10.1175/JTECH-D-14-00174.1

Qi, Y., Wu, G., Liu, Y., and Yue, D. K. P. (2018). Predictable zone for phase-resolved reconstruction and forecast of irregular waves. *Wave Motion* 77, 195–213. doi: 10.1016/J.WAVEMOTI.2017.12.001

Qi, Y., Xiao, W., and Yue, D. K. P. P. (2016). Phase-resolved wave field simulation calibration of sea surface reconstruction using noncoherent marine radar. *J. Atmos. Ocean Technol.* 33, 1135–1149. doi: 10.1175/JTECH-D-15-0130.1

Saad, Y., and Schultz, M. H. (1986). GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci. Stat. Comput.* 7, 856–869. doi: 10.1137/0907058

Tanaka, M., and Yokoyama, N. (2004). Effects of discretization of the spectrum in water-wave turbulence. *Fluid Dyn. Res.* 34, 199. doi: 10.1016/J.FLUIDDYN.2003.12.001

Tikhonov, A. N., and Arsenin, V. Y. (1979). Solutions of ill-posed problems. *SIAM Rev.* 21, 266–267. doi: 10.1137/1021044

Tolman, H. L. (2016). *User manual and system documentation of WAVEWATCH III*. (U. S. Department of CommerceNational Oceanic and Atmospheric Administration National Weather ServiceNational Centers for Environmental Prediction University Research Court College Park, MD 20740: Environmental Modeling CenterMarine Modeling and Analysis Branch).

Wang, G., and Pan, Y. (2021). Phase-resolved ocean wave forecast with ensemble-based data assimilation. *J. Fluid Mech.* 918, A19. doi: 10.1017/jfm.2021.340

Waseda, T. (2019). "Nonlinear processes," in *Ocean wave dynamics* (Singapore: World Scientific), 103–161. doi: 10.1142/9789811208676_0004

Waseda, T., Watanabe, S., Fujimoto, W., Nose, T., Kodaira, T., and Chabchoub, A. (2021). Directional coherent wave group from an assimilated non-linear wavefield. *Front. Phys.* 9. doi: 10.3389/fphy.2021.622303

Watanabe, S., Fujimoto, W., Nose, T., Kodaira, T., Davies, G., Lechner, D., et al. (2019). "Data assimilation of the stereo reconstructed wave fields to a nonlinear phase resolved wave model," in *Proceedings of the ASME 2019 38th international conference on ocean, offshore and Arctic engineering* (Glasgow, Scotland, U.K: ASME). doi: 10.1115/OMAE2019-95949

West, B. J., Brueckner, K. A., Janda, R. S., Milder, D. M., and Milton, R. L. (1987). A new numerical method for surface hydrodynamics. *J. Geophys. Res.* 92, 11803–11824. doi: 10.1029/JC092iC11p11803

Wu, G. (2004). *Direct simulation and deterministic prediction of Large-scale nonlinear ocean wave-field* (Boston: Massachusetts Institute of Technology), Vol. 258.

Xiao, W., Liu, Y., Wu, G., and Yue, D. K. P. (2013). Rogue wave occurrence and dynamics by direct simulations of nonlinear wave-field evolution. *J. Fluid Mech.* 720, 357–392. doi: 10.1017/jfm.2013.37

Yaremchuk, M., Martin, P., and Beattie, C. (2017). A hybrid approach to generating search subspaces in dynamically constrained 4-dimensional data assimilation. *Ocean Model. (Oxf)* 117, 41–51. doi: 10.1016/j.ocemod.2017.08.003

Yaremchuk, M., Martin, P., Koch, A., and Beattie, C. (2016). Comparison of the adjoint and adjoint-free 4dVar assimilation of the hydrographic and velocity observations in the Adriatic Sea. *Ocean Model. (Oxf)* 97, 129–140. doi: 10.1016/j.ocemod.2015.10.010

Yaremchuk, M., Nechaev, D., and Panteleev, G. (2009). A method of successive corrections of the control subspace in the reduced-order variational data assimilation*. *Mon. Weather Rev.* 137, 2966–2978. doi: 10.1175/2009MWR2592.1

Yoon, S., Kim, J., and Choi, W. (2015). "An explicit data assimilation scheme for a nonlinear wave prediction model based on a pseudo-spectral method," in *IEEE Journal of Oceanic Engineering.* 1. doi: 10.1109/JOE.2015.2406471

Zakharov, V. E. (1968). Stability of periodic waves of finite amplitude on the surface of a deep fluid. *J. Appl. Mech. Tech. Phys.* 9, 190–194. doi: 10.1007/BF00913182

Zupanski, M. (2005). Maximum likelihood ensemble filter: Theoretical aspects. *Mon. Weather Rev.* 133, 1710–1726. doi: 10.1175/MWR2946.1

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 05 May 2023
DOI 10.3389/fmars.2023.1112025

Check for updates

OPEN ACCESS

EDITED BY
Youmin Tang,
University of Northern British Columbia
Canada, Canada

REVIEWED BY
Zheqi Shen,
Hohai University, China
Li Yineng,
Chinese Academy of Sciences, China
Zhenya Song,
Ministry of Natural Resources, China
Tianyu Zhang,
Guangdong Ocean University, China

*CORRESPONDENCE
Pengfei Lin
linpf@mail.iap.ac.cn
Hailong Liu
lhl@lasg.iap.ac.cn

RECEIVED 30 November 2022
ACCEPTED 06 April 2023
PUBLISHED 05 May 2023

CITATION
Zheng W, Lin P, Liu H, Luan Y, Ma J, Mo H and Liu J (2023) An assessment of the LICOM Forecast System under the IVTT class4 framework.
Front. Mar. Sci. 10:1112025.
doi: 10.3389/fmars.2023.1112025

COPYRIGHT
© 2023 Zheng, Lin, Liu, Luan, Ma, Mo and Liu. 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(s) 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.

An assessment of the LICOM Forecast System under the IVTT class4 framework

Weipeng Zheng[1,2,3], Pengfei Lin[1,2]*, Hailong Liu[1,2,4]*, Yihua Luan[3], Jinfeng Ma[1], Huier Mo[5] and Juan Liu[6]

[1]State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China, [2]College of Earth and Planetary Sciences, University of Chinese Academy of Sciences, Beijing, China, [3]Earth System Numerical Simulation Science Center, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China, [4]Center for Ocean Mega-Science, Chinese Academy of Sciences, Qingdao, China, [5]National Marine Environmental Forecasting Center, Beijing, China, [6]Beijing Institute of Applied Meteorology, Beijing, China

This paper evaluates LFS (LICOM Forecast System) forecasts and compares them with other marine forecast systems under the IVTT (Intercomparison and Validation Task Team) Class 4 framework. LFS delivers real-time daily forecasts driven by the GFS (Global Forecast System) atmospheric analyses and surface forecasts. The nudging method in LFS provides the initial state for forecasting, with only the temperature and salinity restored towards the Mercator PSY4 daily analyses. Assessments show that LFS demonstrates a reasonably good capability in short-term marine environment forecast. For the leading 1-6 days forecasts, the root mean square error (RMSE) ranges between 0.53-0.63°C, 0.57-0.66°C and 0.12-0.13 psu for the sea surface temperature, temperature, and salinity profiles, respectively. The overall performance is comparable to other major marine forecast systems, with a slight advantage in forecasting the temperature and salinity profiles. Different nudging time scales are applied to the upper ocean and deep ocean to preserve the effects of mesoscale processes and correct the large-scale biases in temperature and salinity. However, the absence of other observational constraints, such as the sea level height, significantly affects the regional forecast features. Further analyses are required to improve the performance, and the integration of an assimilation system into LFS is urgently needed.

KEYWORDS

LFS, marine forecast, eddy-resolving, IVTT, ARGO, LICOM

1 Introduction

The safety and efficiency of marine activities, such as marine transportation, oil, and gas industry, military operations, fishery, and marine search and rescue activities, necessitates high-quality ocean reanalysis and short-range predictions of the ocean state at both global and regional scales (Schiller et al., 2008). The first operational global or basin-scale ocean forecast system was developed in the late 1990s when oceanic observations, high-resolution satellite data, high-performance computers, and advanced ocean data assimilation methods were routinely available. Due to the remarkable growth in supercomputing resources and the development of ocean observation systems, the prediction systems of global ocean forecasting were significantly improved from several points of view, including the sensibly increased their resolution, the increased complexity of the models, more processes resolved by the system (Tonani et al., 2015).

Operational marine forecast using global eddy-resolving systems has been conducted since 2010 by three national ocean centers, including the Mercator Océan in France, the Naval Research Laboratory (NRL), and the National Centers for Environmental Prediction (NCEP) in the US (Tonani et al., 2015). In China, the Ocean Forecast System (OFS) has been developed into eddy-resolving in recent years. A global eddy-resolving forecast system based on NEMO (Nucleus for European Modelling of the Ocean, Gurvan et al., 2017) of 1/12° is now operating in the National Marine Environmental Forecasting Center (NMEFC) and a new ocean forecast system based on Mass Conservation Ocean Model (MaCOM) with 10km resolution was published in 2022. The OFS based on the surface wave-tide-circulation coupled ocean model developed by the First Institute of Oceanography (FIO-COM) was published in 2018 (Qiao et al., 2019), which has a resolution of 1/10°. The Institute of Atmospheric Physics of Chinese Academy of Sciences developed a new forecast system named LFS (LICOM Forecast System, Liu et al., 2023) based on the eddy-resolving ocean circulation model - LICOM version 3 (LICOM3, Yu et al., 2018; Lin et al., 2020).

As LFS is a novel ocean forecast system, a comprehensive evaluation is essential. The primary objective of this study is to evaluate the forecast results from LFS against the observations and compare them with other forecast systems based on IVTT Class 4 framework. The IVTT, initiated by GODAE Ocean View, focuses on the inter-comparison between the systems by using a standard set of observations as a proxy for the truth. The IVTT framework includes four classes of comparisons (Class 1-4), with Class 4 contains a set of metrics designed for forecast verification (Ryan et al., 2015). The metrics include bias, the root mean square error (RMSE), the anomaly correlation, and the skill score for global and regional features. Within the framework of IVTT Class4, forecasted physical parameters of SST (sea surface temperature), SSH (sea level height) and the temperature and salinity in the sub-surface can be interpolated and compared to the *in-situ* observations. These are also the method employed in the present study.

This paper is organized as follows. Section 2 details the basic configuration of LFS and methods, while Section 3 presents the evaluation of the results. Concluding remarks and discussions are summarized in section 4.

2 System and methods

2.1 LICOM Forecast System

The LFS, or LICOM Forecast System, is developed based on LICOM3 and incorporates several enhancements to improve its performance. The dynamical framework adopts generalized orthogonal coordinates, and the tripolar grid described by Murray (1996) is applied (Yu et al., 2018). To refine the physical processes, the internal tide parameterization from St. Laurent et al. (2002) and the thickness diffusivities from Ferreira et al. (2005) are introduced (Yu et al., 2017; Li et al., 2020). Additionally, the flux coupler is upgraded to NCAR flux coupler version 7, enhancing the model's flexibility for coupling and facilitating high-resolution simulations (Lin et al., 2016).

The horizontal resolution of the LFS remains consistent with that of LICOM3, featuring 3600×2302 horizontal grids (1/10°) and 55 vertical levels. The average zonal grid distance is approximately 6.9 km globally, about 11 km at the equator, and progressively decreases to 2.7 km at the high-latitude around Antarctic. The eta coordinate is employed vertically; 42 layers configured in the upper 1000 meters. Furthermore, a thermodynamic sea-ice model, CICE4 (Community Ice CodE version 4, Hunke and Lipscomb, 2010), is coupled with the LICOM3 ocean model through the flux coupler. Consequently, the LFS is capable of forecasting sea ice conditions.

In the current version of LFS, the assimilation module is not yet fully developed; therefore, the nudging method is employed to obtain the initial state for forecasting. The simulation, which provides the initial state, is referred to as the analysis experiment (ExpA). Within ExpA, the simulated temperature and salinity in LFS are restored to the temperature and salinity values from Mercator Océan PSY4 analysis (Lellouche et al., 2018), respectively. To preserve the influence of mesoscale eddies, a nudging time scale to 5 days is set for the upper 2000 meters and gradually relaxed to 20 days down to the depth of 5600 meters. Figure 1 displays a comparison of SST and SSH between the LFS and Mercator PSY4 for the period of June 1st to December 31st, 2014. Employing the nudging method, the LFS simulates a spatial distribution of SST that closely resembles that of the Mercator PSY4 analysis (Figures 1A, B). A slightly warmer SST is observed in the Indian Ocean, whereas a cooler SST is evident in the eastern Pacific (Figure 1C). The global standard deviation (STD) is 0.07°C. In ExpA, the simulated SSH also aligns well with the Mercator PSY4 analysis, with a global STD of 0.001 m (Figures 1D–F). The results of ExpA suggest that the nudging method could serve as an effective approach for providing the intimal state necessary for forecasting.

In the forecasting experiment (ExpF), initial values derived from ExpA are utilized, while the atmospheric variables and land

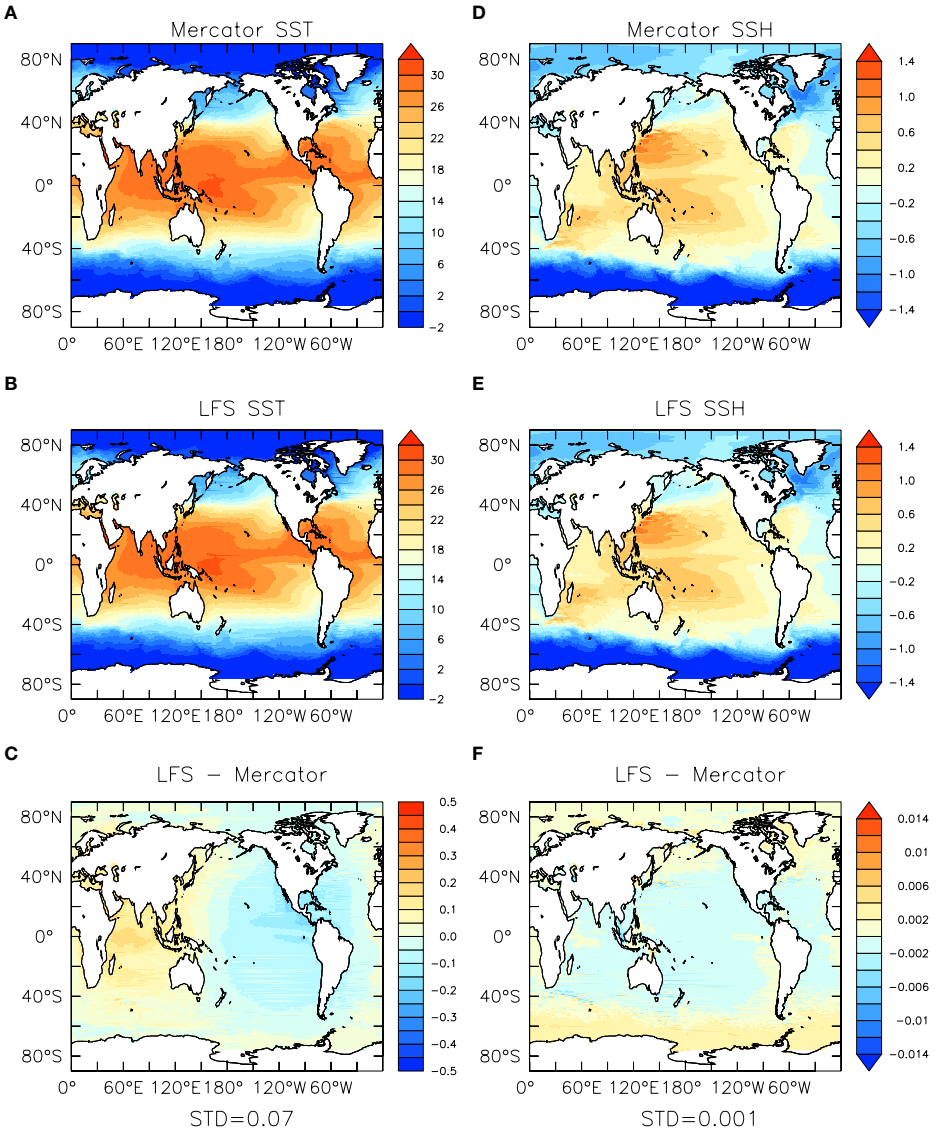


FIGURE 1
(A) The mean sea surface temperature (SST, °C) and **(D)** the mean sea surface height (SSH, m) for the Mercator Océan PSY4 during the 1st Jun – 31st Dec, 2014. **(B, E)** are the same as **(A, D)** but for the analysis experiment ExpA. **(C, F)** show their differences.

surface runoff from the Global Forecast System (GFS) serve as external forcing to drive the LFS. The atmospheric variables include total precipitation, downward and upward shortwave radiation at the surface, downward longwave radiation, sea level pressure, 10-meter zonal/meridional wind component, air temperature and specific humidity at 2 meters. The atmospheric variables and runoff are pre-processed and subsequently read by the atmosphere and land data models in Coupler 7, which then provides the forcing. This forcing is interpolated onto the ocean model grid, and the fluxes are calculated to drive the forecast system. The air-sea fluxes are computed using the Coordinated Ocean Reference Experiments (CORE) bulk formula (Large and Yeager, 2004). In this study, the prediction spans from June 1st to December 31st, 2014, with a forecast time of six days. The forecast variables comprise daily averaged outputs of SST, temperature, salinity, ocean currents, and sea level height.

2.2 Data and methods

In this study, we adopt the metrics defined in Class 4 and validate the LFS by using the observations organized within the IVTT framework. Evaluation metrics include bias, RMSE, anomaly correlation, and forecast skill scores. The IVTT Class 4 reference datasets of 2014 include six forecast systems from five research organizations, which includes the Forecast Ocean Assimilation Model (FOAM) from Met Office, the operational ocean analysis and forecasting systems (PSY3 and PSY4) from Mercator Océan,

the Global Ice Ocean Prediction System (GIOPS) from Environment Canada, the Real-Time Ocean Forecast System (RTOFS) from NCEP/NWS/NOAA, the Ocean Model Analysis and Prediction System (Ocean-MAPS) from Australian Bureau of Meteorology. Details regarding these systems are listed in Table 1. All the systems feature a six-day forecast periods, and the forecasts from June 1st to December 31st are employed for the LFS validation.

The IVTT provides an abundance of *in-situ* drifters and Argo profiles observations. In this study, we evaluate the LFS outputs against IVTT datasets using the following approach. First, the forecast results are extracted and interpolated onto the observation sites from the nearest grid points. Second, metrics for SST statistics are computed directly by comparing the forecasts with drifter measurements. For the temperature and salinity profiles statistics, the following processes is employed: both Argo and forecast data are sampled by depth range according to the prescribed 40 vertical levels. Values within each level are averaged to represent the mean value for that level, and then metrics are calculated by comparing the forecast with the Argo values. This methodology is similar to the approach in Ryan et al. (2015), where 50 levels are used, and the median value servers as the representative. However, this difference does not impact the overall analysis results. The IVTT framework also offers persisted forecasts and the World Ocean Atlas 2001 (WOA2001) as climatological reference, which can be used to assess forecast skills through the Persistence Skill Score (PSS) and Climatological Skill Score (CSS), respectively. The definition of PSS and CSS follows those provided by Ryan et al. (2015).

3 Results

3.1 Sea surface temperature

The sea surface temperatures (SST) from the forecast systems are compared to the *in-situ* drifter observations, as shown in Figure 1. LFS exhibits a warm bias in the global mean SST, which amounts to approximately 0.05°C at the first forecast lead day (day 1). The median warm bias is comparable to those observed in other systems (Figure 2A). This warm bias increases with the forecast time, reaching around 0.1°C on the sixth day (day 6). Such bias may be related to the slightly overestimated incoming solar radiation in the atmospheric forcing, which contributes to the warm bias accumulated during the forecast processes. Despite the warm bias, LFS demonstrates a relatively small SST RMSE within the forecast period among all the forecast systems. The RMSE value is about 0.53°C on day one, increasing to about 0.63°C on day 6 (Figure 2B). The growth rate of RMSE with the forecast time in LFS is similar to those observed in PSY3/4 and GIOPS.

Figure 3 shows the SST bias and RMSE at 1 day lead over the forecast period. All the forecast systems show similar behavior, with relatively small SST biases in summer than increase during late autumn and winter. LFS displays a bias in the middle of the forecast systems in summer, while the bias grows larger in autumn and winter (Figure 3A). It is important to note that since PSY4 analysis data serves as the nudging observations for LFS, the SST RMSE is expected to be similar to PSY4 at day 1. However, the RMSE in LFS is consistently smaller than PSY4 over the forecast period (Figure 3B), partially attributable to the nudging time scale employed in ExpA. Additionally, the differences in external atmospheric forcing and the configurations of the systems may contribute to this feature as well.

The anomaly correlation coefficient (AC) generally represents the predictability of the forecast system, which is commonly used in short-time and seasonal climate predictions. The anomaly correlation coefficient of SST in LFS is about 0.79 on day 1 and 0.72 on day 6, remaining above 0.6 throughout the forecast period as most of the ocean forecast systems (Figure 4A). Over the entire forecast period, the anomaly correlation coefficient of SST in LFS ranges between 0.6 and 0.8. The AC variation in winter (November-December) is more significant than in other seasons (June-October), a feature also observed in other forecast systems (Figure 4B).

TABLE 1 Basic information of the forecast systems analyzed.

	Forecast System	Ocean model	Sea ice model	Horizontal Resolution	Vertical levels	Assimilation	References
1	LFS	LICOM3	CICE4	1/10°	55	Nudging	(Liu et al., 2023)
2	FOAM	NEMO3.2	CICE 4.1	1/4°	75	NEMOVAR based on 3D-VAR	(Blockley et al., 2014)
3	BLK	OFAM (Ocean Forecast Australian Model)	N/A	2°-1/10° (90°E-180°E,16°N-75°S)	51	BODAS (BLUElink Ocean Data Assimilation System)	(Brassington et al., 2012)
4	GIOPS	NEMO3.1	CICE4	1/4°	50	Ocean: SAM2 (Système d'Assimilation Mercator version 2); Sea ice: 3D-Var	(Smith et al., 2016)
5	RTOFS	HYCOM (HYbrid Coordinates Ocean Model)	thermodynamic energy loan ice model	1/12°	32	3D-VAR	(Mehra and Rivin, 2010)
6	PSY3	NEMO3.1	LIM2_EVP	1/4°	50	SAM (SEEK, single evolutive extended Kalman filter); 3D-VAR bias correction	(Lellouche et al., 2013)
7	PSY4			1/12°			

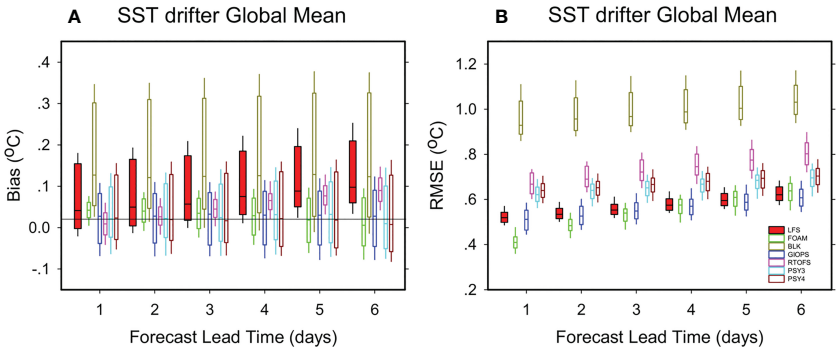


FIGURE 2
Bias and RMSE of the forecast SST against the *in-situ* drifters as a function of lead time. **(A)** Bias; **(B)** RMSE. The boxes show the range of the 95th percentile.

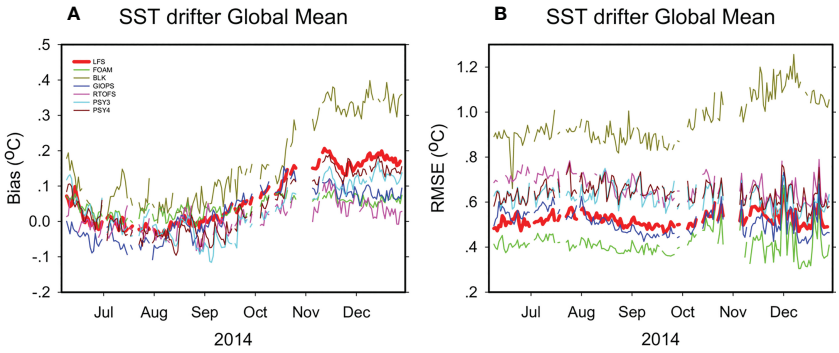


FIGURE 3
The time series of forecast SST at day 1 for **(A)** bias; **(B)** RMSE.

IVTT Class4 provides the two reference datasets, climatological and persisted forecast, from which the CSS and PSS can be obtained by comparing against the WOA2001 climatology and the persisted one-day forecast, respectively. For all forecast systems, the CSS is positive and decreases as the forecast time increases, indicating positive skill against the climatology (Figure 5A). The CSS of SST in LFS is in the middle of all forecast systems, similar to the performance of the anomaly correlation coefficient (Figure 4A), which agrees with the positive relationship between a robust anomaly correlation coefficient and strong positive skill against climatology. For PSS, most forecast systems show an initial weak negative value and a slightly positive trend as the forecast time increases, implying that a 1-day lead forecast is considerable more accurate and skillful than other lead times. The PSS of SST in LFS remains positive over the forecast period (Figure 5B); however, a weak negative trend appears as forecast time increases, potentially

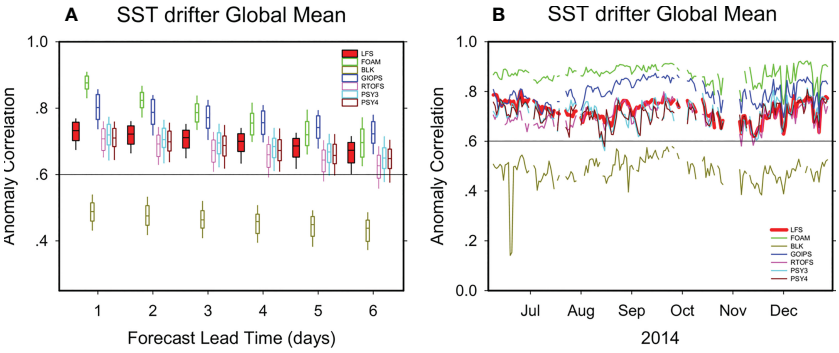


FIGURE 4
(A) The anomaly correlation coefficient and **(B)** the time series of anomaly correlation at day 1 for all the forecast systems.

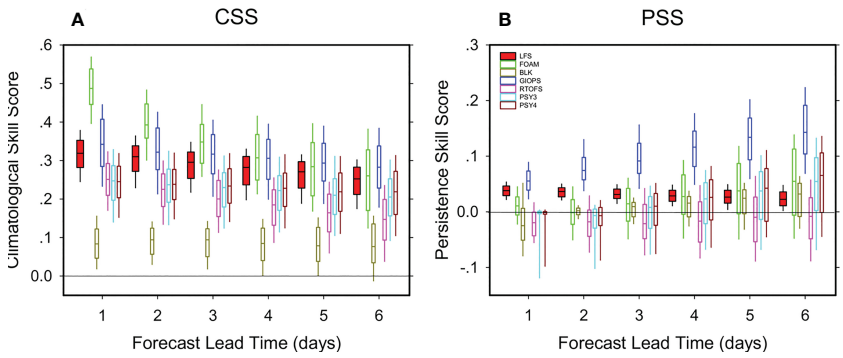


FIGURE 5
(A) The climatological skill score (CSS) and **(B)** persistent skill scores (PSS) of the forecast SST against the SST measurements from the drifters as a function of lead time.

related to the simple nudging method used to generate the initial state for the LFS forecast. The SST information is gradually diminishing as the forecast is integrated over six days.

3.2 Sea level anomaly

The sea level anomaly (SLA) of LFS is computed by subtracting the SSH climatology from ExpA from the forecast data. Figure 6 shows the bias and RMSE of SLA for all the forecast systems. Most forecast systems have a negative SLA bias, with LFS having a median value of approximately −0.06 m. The bias does not exhibit significant changes during the forecast period, consistent with the behavior of other forecast systems (Figure 6A). The SLA RMSE of LFS is around 0.10 m, falling within the range of other forecast systems (0.07–0.15m, Figure 6B). Similar to SST, both SLA RMSE and Bias increase slightly in the winter season (Figure not shown).

There are several possible explanations for the relatively large negative SLA biases in LFS. Firstly, we do not apply nudging method to the SLA, meaning that the constrain of observation is absent in the initial state from ExpA. Secondly, the Exp has only 7 months of simulation, and although the mean SSH pattern closely resemble that of Mercator PSY4, it does not necessarily imply that the climatology is sufficiently accurate or that the SLA is directly comparable to satellite data.

3.3 Temperature and salinity profiles

The global Argo network can provide more than 3000 floats in the world. Due to the operational procedure for Argo floats, the daily number of available profiles in IVTT is approximately 300. These daily profiles are subsequently used to evaluate the forecast results for all the systems. Similar to SST, a warm bias is present in the sub-surface temperature in LFS. Despite the warm bias, the RMSE of the temperature profile in LFS is the smallest among the forecast systems throughout the entire forecast period. The median RMSE value is about 0.57°C on day 1 and 0.66°C on day 6 (Figure 7A). The median RMSE of the salinity profile is between 0.12and 0.13 PSU (Figure 7B). Since the sub-surface ocean evolved slower than the surface, the bias and RMSE show only a slightly positive trend over the forecast period. It is not surprising that the value and trend are similar to the PSY4 since the temperature and salinity from the PSY4 analysis are nudged to the initial state for the LFS forecast.

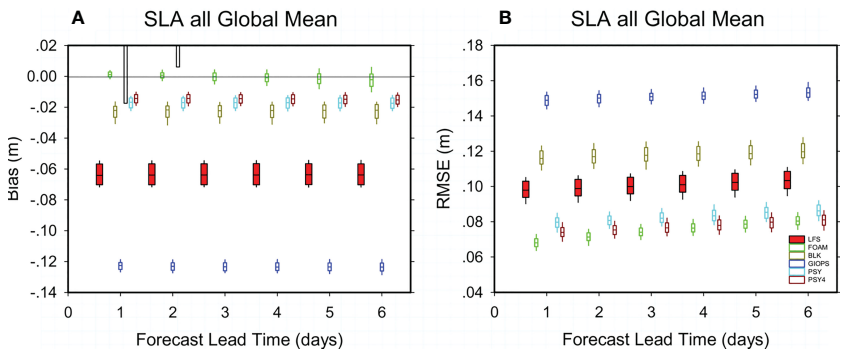


FIGURE 6
Bias and RMSE of the forecast SLA against the satellite altimeters as a function of lead time. **(A)** Bias; **(B)** RMSE. The boxes show the range of the 95th percentile.

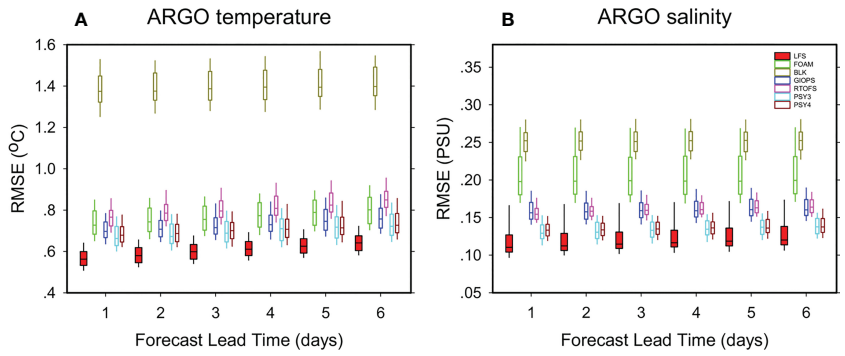


FIGURE 7
The RMSE of the forecast **(A)** temperature and **(B)** salinity against Argo measurements. The boxes show the range of 95th percentile.

Regarding the detailed vertical distribution, all the forecast systems display the most substantial sub-surface temperature errors at the depth of approximately 100 meters (Figure 8A). These errors may be attribute to the strong temperature gradient in the thermocline, which is challenging to simulate in the ocean models. A significant discrepancy exists in the salinity among the forecast systems, particularly in the upper ocean (Figure 8B). LFS has a negative bias of about -0.03 PSU in the upper 50 m. LFS demonstrates a relatively better forecast of the temperature and salinity profiles, as shown in Figure 7, and the RMSE is smaller than most forecast systems.

3.4 Regional features

The regional features are assessed based on the ocean basin division in IVTT Class 4 (Table 2). The global ocean is divided into eight basins with overlapping regions. Then the 1 day lead forecast RMSE of SST, SLA, temperature, and salinity are shown as a function of ocean basins in Figure 9, respectively. Generally, most forecast systems exhibit relatively smaller SST RMSE in the tropical regions of the Atlantic and Pacific but larger values in the North Pacific, North Atlantic, and South Atlantic. The model spread is substantial in the subtropical regions, suggesting model differences in representing the mesoscale processes that influence the forecast results. The performance of LFS is broadly consistent with other systems, showing smaller RMSE in the tropical ocean and larger RMSE in the subtropics for SST (Figure 9A). The situation for SLA mirrors that of SST. The regional accuracy in LFS is also the best in the tropical Atlantic but is poor in the South Atlantic. The North Atlantic and Pacific also have regions with considerable biases (Figure 9B). Although the western boundary currents and mesoscale eddies are reproduced in LFS, the strength and locations are not well matched with the observations, resulting in larger RMSE in subtropical regions.

The RMSE of temperature and salinity, when compared against Argo profiles, shows a different pattern than that of SST. No significant differences can be identified among ocean basins, which may be attributed to the considerable RMSE in the upper ocean surrounding the thermocline in the tropics. LFS is the most accurate system in forecasting the temperature and salinity in nearly all ocean basins (Figures 9C, D), surpassing even the PSY4 system, whose analysis data were used for LFS initial condition by nudging

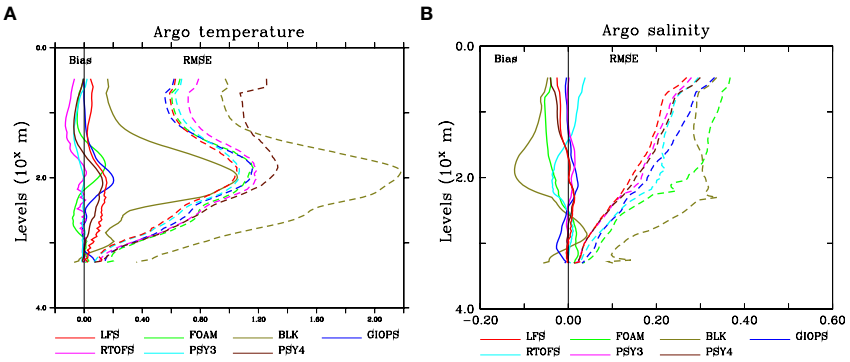


FIGURE 8
The bias and RMSE as a function of depth for the **(A)** temperature and **(B)** salinity on day 1. The vertical axis is logarithmic.

10.3389/fmars.2023.1112025

TABLE 2 IVTT Class 4 ocean basins.

Name	Longitudes	Latitudes
Global Ocean	90°S~90°N	180°W~180°E
North Atlantic	0~70°N	100°W~31°E
Tropical Atlantic	20°S~20°N	70°W~30°E
South Atlantic	60°S~0	70°W~30°E
North Pacific	0~65°N	100°E~77°W
Tropical Pacific	20°S~20°N	90°E~70°W
South Pacific	60°S~0	100°E~70°W
Indian Ocean	40°S~31°N	20°E~120°E
Australasia	70°S~20°N	90°E~180°E

methods. LFS performs better in the open ocean, such as in the south Pacific, than in other basins for both temperature and salinity. The subtropical ocean basins remain a challenge for accurate forecasting.

4 Conclusions and discussions

Within the evaluation framework of IVTT, the LICOM Forecast System (LFS) forecast results are assessed and compared to other marine forecast systems in this study. LFS demonstrates a reasonably good performance in the short-term marine environment forecasting. The global RMSE of SST, when compared against the drifter observations, ranges from 0.53-0.63°C in the forecast period, ranking in the middle of all the forecast systems. Regional decomposition reveals that LFS has better forecast skills in tropical regions, while exhibiting relatively large biases in the subtropics where the eddies are more active. The temperature and salinity profiles in LFS tend to outperform those in other systems. The RMSE of temperature in LFS against the Argo profile in the upper 2000 m ranges from 0.57-0.66°C in the forecast period, whereas the RMSE of salinity in LFS ranges between 0.12 and 0.13 psu. Across all forecast systems, the maximum RMSE of temperature occurs at a depth of around 100 meters, indicating challenges in reproducing the thermocline variations. The largest RMSE of salinity is at the surface, influenced by water mass and heat fluxes exchanges between the atmosphere and the ocean.

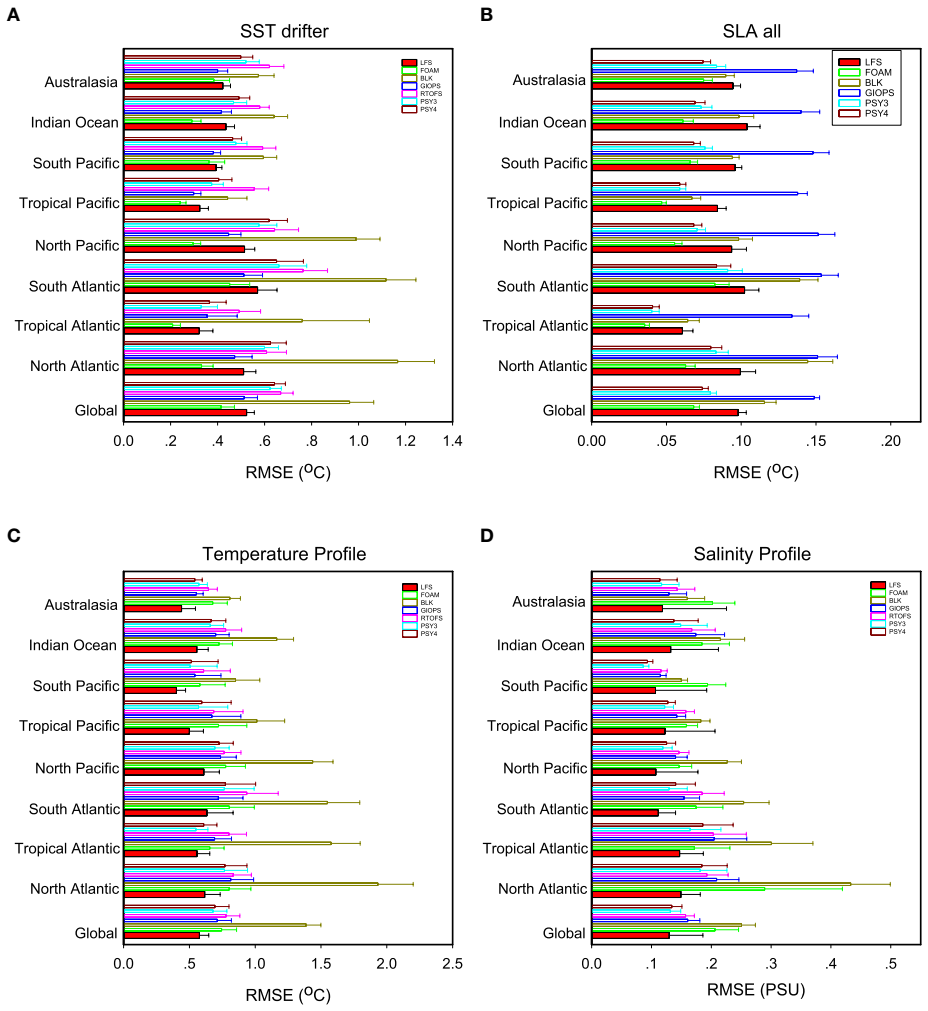


FIGURE 9
Regional RMSE at leading 1-day for **(A)** SST against the in-situ drifters, **(B)** SLA, **(C)** temperature and **(D)** salinity against Argo profiles.

In this iteration of LFS, the initial state was derived by nudging the temperature and salinity of the PSY4, with no additional observational data assimilated into the LFS. The nudging method appears to be a cost-effective approach for obtaining the necessary initial state. In this study, the nudging time scale is set at 5 days for the upper ocean and 20 days for the deep sea. This configuration effectively retains the impacts of mesoscale eddies in the upper ocean while eliminate the weather scale bias, thus favoring the slow growth of biases in the forecasting. Despite the relatively satisfactory performance in capturing the large-scale features, LFS exhibits some defects in predicting the global sea level height anomaly (SLA). The poor performance of SLA may be related to the systematic bias in the climatology, which may be related to unpredictable instabilities. The absence of sea surface height observational data may also have contribution, because nudging the SLA may result in the inconsistency in temperature and salinity that lead to the model's shock, as discussed by Liu et al. (2023). To address these limitations, it is crucial to develop a coordinated assimilation system for LFS, enabling the incorporation of more observational data to provide a more accurate initial condition for the marine forecast.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

WZ, PL, and HL contributed to the development of LFS and the conception of this study. WZ performed the forecast and wrote the first draft of the manuscript. PL and HL wrote sections of the manuscript. YL and JM organized and processed the data. HM and JL performed the statistical analysis. All authors contributed to the article and approved the submitted version.

Funding

This study is supported by the National Key R&D Program for Developing Basic Sciences (2022YFC3104802, 2020YFA0608902), the Special Funds for Creative Research (2022C61540), and the National Natural Science Foundation of China (Grants 41931183, 41976026, 41931182, U2242214).

Acknowledgments

The authors acknowledge the technical support from the National Key Scientific and Technological Infrastructure project "Earth System Science Numerical Simulator Facility" (EarthLab).

Conflict of interest

The authors declare that the research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Blockley, E. W., Martin, M. J., McLaren, A. J., Ryan, A. G., Waters, J., Lea, D. J., et al. (2014). Recent development of the met office operational ocean forecasting system: an overview and assessment of the new global foam forecasts. *Geosci. Model. Dev.* 7 (6), 2613–2638. doi: 10.5194/gmd-7-2613-2014

Brassington, G. B., Freeman, J., Huang, T., Pugh, P. R., Sandery, A., Taylor, A., et al. (2012). "Ocean model analisys and prediction system: version 2," in *CAWCR Technical Report No. 052, The Centre for Weather and Climate Research*.

Ferreira, D., Marshall, J., and Heimbach, P. (2005). Estimating eddy stresses by fitting dynamics to observations using a residual-mean ocean circulation model and its adjoint. *J. Phys. Oceanogr.* 35, 1891–1910. doi: 10.1175/JPO2785.1

Gurvan, M., Bourdallé-Badie, R., Pierre-Antoine, B., Clement, B., Bruciaferri, D., Calvert, D., et al. (2017). Nemo ocean engine: Notes du Pôle de modélisation de l'Institut Pierre-Simon Laplace (IPSL). (27). doi: 10.5281/zenodo.1472492

Hunke, E. C., and Lipscomb, W. H. (2010). "Cice: the Los alamos Sea ice model," in *Documentation and software users manual, version 4.1 (La-Cc-06-012)*, vol. 2010. (Los Alamos: T-3 Fluid Dynamics Group, Los Alamos National Laboratory).

Large, W. G., and Yeager, S. (2004). "Diurnal to decadal global forcing for ocean and Sea-ice models: the data sets and flux climatologies," in *NCAR Technical Note. NCAR/TN-460+STR*. doi: 10.5065/D6KK98Q6

Lellouche, J. M., Greiner, E., Le Galloudec, O., Garric, G., Regnier, C., Drevillon, M., et al. (2018). Recent updates to the Copernicus marine service global ocean monitoring and forecasting real-time 112° high-resolution system. *Ocean Sci.* 14 (5), 1093–1126. doi: 10.5194/os-14-1093-2018

Lellouche, J. M., Le Galloudec, O., Drévillon, M., Régnier, C., Greiner, E., Garric, G., et al. (2013). Evaluation of global monitoring and forecasting systems at Mercator océan. *Ocean Sci.* 9 (1), 57–81. doi: 10.5194/os-9-57-2013

Li, Y., Liu, H., Ding, M., Lin, P., Yu, Z., Yu, Y., et al. (2020). Eddy-resolving simulation of cas-Licom3 for phase 2 of the ocean model intercomparison project. *Adv. Atmospheric Sci.* 37 (10), 1067–1080. doi: 10.1007/s00376-020-0057-z

Lin, P. F., Liu, H. L., Xue, W., Li, H. M., Jiang, J. R., Song, M. R., et al. (2016). A coupled experiment with Licom2 as the ocean component of Cesm1. *J. Meteorol. Res.* 30 (1), 76–92. doi: 10.1007/s13351-015-5045-3

Lin, P. F., Yu, Z. P., Liu, H. L., Yu, Y. Q., Li, Y. W., Jiang, J. R., et al. (2020). Licom model datasets for the Cmip6 ocean model intercomparison project. *Adv. Atmos. Sci.* 37 (3), 239–249. doi: 10.1007/s00376-019-9208-5

Liu, H. L., Lin, P. F., Zheng, W. P., Luan, Y. H., Ma, J. F., Ding, M. R., et al. (2021). A global eddy-resolving ocean forecast system in China – licom forecast system (Lfs). *J. Operational Oceanogr.* 16 (1), 15–27. doi: 10.1080/1755876X.2021.1902680

Mehra, A., and Rivin, I. (2010). A real time ocean forecast system for the north Atlantic ocean. *Terr. Atmos. Ocean. Sci.* 21, 211–228. doi: 10.3319/TAO.2009.04.16.01 (IWNOP)

Murray, R. J. (1996). Explicit generation of orthogonal grids for ocean models. *J. Comput. Phys.* 126 (2), 251–273. doi: 10.1006/jcph.1996.0136

Qiao, F., Wang, G., Khokiattiwong, S., Akhir, M., Zhu, W., and Xiao, B. (2019). China Published ocean forecasting system for the 21st-century maritime silk road on December 10, 2018. *Acta Oceanol. Sin.* 38, 1–3. doi: 10.1007/s13131-019-1365-y

Ryan, A. G., Regnier, C., Divakaran, P., Spindler, T., Mehra, A., Smith, G. C., et al. (2015). Godae oceanview class 4 forecast verification framework: global ocean inter-comparison. *J. Operational Oceanogr.* 8 (sup1), s98–s111. doi: 10.1080/1755876X.2015.1022330

Schiller, A., Mourre, B., Drillet, Y., and Brassington, G. (2008). "An overview of operational oceanography," in *New Frontiers in Operational Oceanography*. Eds. E. Chassignet, A. Pascual, J. Tintoré and J. Verron, 1–26. doi: 10.17125/gov2018.ch01

Smith, G. C., Roy, François, Reszka, M., Colan, D. S., He, Z., Deacu, D., et al. (2016). Sea Ice forecast verification in the Canadian global ice ocean prediction system. *Q. J. R. Meteorol. Soc.* 142 (695), 659–671. doi: 10.1002/qj.2555

St. Laurent, L. C., Simmons, H. L., and Jayne, S. R. (2002). Estimating tidally driven mixing in the deep ocean. https://doi.org/10.1029/2002GL015633. *Geophysical Res. Lett.* 29 (23), 21–1-21-4. doi: 10.1029/2002GL015633

Tonani, M., Balmaseda, M., Bertino, L., Blockley, Ed, Brassington, G., Davidson, F., et al. (2015). Status and future of global and regional ocean prediction systems. *J. Operational Oceanogr.* 8 (sup2), s201–ss20. doi: 10.1080/1755876X.2015.1049892

Yu, Y. Q., Tang, S. L., Liu, H. L., Lin, P. F., and Li, X. L. (2018). Development and evaluation of the dynamic framework of an ocean general circulation model with arbitrary orthogonal curvilinear coordinate, [In Chinese]. *Chin. J. Atmospheric Sci.* 42 (4), 877–889. doi: 10.3878/j.issn.1006-9895.1805.17284

Yu, Z. P., Liu, H. L., and Lin, P. F. (2017). A numerical study of the influence of tidal mixing on Atlantic meridional overturning circulation (Amoc) simulation, [In Chinese]. *Chin. J. Atmospheric Sci.* 41 (5), 1087–1100. doi: 10.3878/j.issn.1006-9895.1805.17284

frontiers | Frontiers in Marine Science

TYPE Original Research
PUBLISHED 09 May 2023
DOI 10.3389/fmars.2023.1084186

Check for updates

OPEN ACCESS

EDITED BY
Zhijin Li,
UCLA, United States

REVIEWED BY
Gilles Reverdin,
Centre National de la Recherche Scientifique (CNRS), France
Simona Masina,
Euro-Mediterranean Center on Climate Change, Italy
Pinqiang Wang,
National University of Defense Technology, China

*CORRESPONDENCE
Xinrong Wu
xrw_nmdis@163.com
Wei Li
liwei1978@tju.edu.cn

RECEIVED 30 October 2022
ACCEPTED 29 March 2023
PUBLISHED 09 May 2023

CITATION
Fu H, Dan B, Gao Z, Wu X, Chao G, Zhang L, Zhang Y, Liu K, Zhang X and Li W (2023) Global ocean reanalysis CORA2 and its inter comparison with a set of other reanalysis products.
Front. Mar. Sci. 10:1084186.
doi: 10.3389/fmars.2023.1084186

COPYRIGHT
© 2023 Fu, Dan, Gao, Wu, Chao, Zhang, Zhang, Liu, Zhang and Li. 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(s) 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.

Global ocean reanalysis CORA2 and its inter comparison with a set of other reanalysis products

Hongli Fu[1], Bo Dan[1], Zhigang Gao[1], Xinrong Wu[1]*, Guofang Chao[1], Lianxin Zhang[1], Yinquan Zhang[1], Kexiu Liu[1], Xiaoshuang Zhang[1] and Wei Li[2]*

[1]Key Laboratory of Marine Environmental Information Technology, National Marine Data and Information Service, Tianjin, China, [2]School of Marine Science and Technology, Tianjin University, Tianjin, China

We present the China Ocean ReAnalysis version 2 (CORA2) in this paper. We compare CORA2 with its predecessor, CORA1, and with other ocean reanalysis products created between 2004 and 2019 [GLORYS12v1 (Global Ocean reanalysis and Simulation), HYCOM (HYbrid Coordinate Ocean Model), GREP (Global ocean Reanalysis Ensemble Product), SODA3 (Simple Ocean Data Assimilation, version 3), and ECCO4 (Estimating the Circulation and Climate of the Ocean, version 4)], to demonstrate its improvements and reliability. In addition to providing tide and sea ice signals, the accuracy and eddy kinetic energy (EKE) of CORA2 are also improved owing to an enhanced resolution of 9 km and updated data assimilation scheme compared with CORA1. Error analysis shows that the root-mean-square error (RMSE) of CORA2 sea-surface temperature (SST) remains around 0.3°C, which is comparable to that of GREP and smaller than those of the other products studied. The subsurface temperature (salinity) RMSE of CORA2, at 0.87°C (0.15 psu), is comparable to that of SODA3, smaller than that of ECCO4, and larger than those of GLORYS12v1, HYCOM, and GREP. CORA2 and GLORYS12v1 can better represent sub-monthly-scale variations in subsurface temperature and salinity than the other products. Although the correlation coefficient of sea-level anomaly (SLA) in CORA2 does not exceed 0.8 in the whole region, as those of GREP and GLORYS12v1 do, it is more effective than ECCO4 and SODA3 in the Indian Ocean and Pacific Ocean. CORA2 can reproduce the variations in steric sea level and ocean heat content (OHC) on the multiple timescales as the other products. The linear trend of the steric sea level of CORA2 is closer to that of GREP than that of the other products, and the long-term warming trends of global OHC in the high-resolution CORA2 and GLORYS12v1 are greater than those in the low-resolution EN4 and GREP. Although CORA2 shows overall poorer performance in the Atlantic Ocean, it still achieves good results from 2009 onward. We plan to further improve CORA2 by assimilating the best available observation data using the incremental analysis update (IAU) procedure and improving the SLA assimilation method.

KEYWORDS

global, ocean, reanalysis, validation, inter-comparison, CORA2

1 Introduction

Ocean reanalysis combines model dynamics with observational information, using data assimilation technology to reconstruct historical and present ocean states. Such a product is important for monitoring the state of the climate and for initializing and validating forecasts. It also has downstream applications, such as driving offline biogeochemical and fishery models, assessing observation networks, and providing lateral boundary conditions for higher-resolution regional ocean general circulation models (Masina and Storto, 2017; Storto et al., 2019b). Therefore, efforts have been made to produce global ocean reanalysis datasets at several institutes, with dozens of products released recently, including GLORYS12v1 (Global Ocean reanalysis and Simulation; Lellouche et al., 2021), ORAS (Ocean ReAnalysis System; Balmaseda et al., 2012; Zuo et al., 2019), C-GLORS (CMCC Global Ocean Reanalysis System; Storto et al., 2014, 2016), GloSea5 (Global Seasonal Forecast System version 5; Blockley et al., 2014), GREP (Global ocean Reanalysis Ensemble Product; Masina et al., 2017), HYCOM (HYbrid Coordinate Ocean Model, Cummings and Smedstad, 2013), ECCO4 (Estimating the Circulation and Climate of the Ocean, version 4; Forget et al., 2015), SODA3 (Simple Ocean Data Assimilation, version 3; Carton et al., 2018), MOVE-G2 (Multivariate Ocean Variational Estimation/Meteorological Research Institute Community Ocean Model - Global version 2; Toyoda et al., 2016), and CORA1 (China Ocean ReAnalysis, version 1; Han et al., 2011, 2013a; 2013b). These global ocean reanalysis products mainly use a global sea ice–ocean coupled model, in which the highest horizontal resolution reaches an eddying-resolving level of 1/12° (such as the models used in HYCOM and GLORYS12v1), and most products assimilate *in-situ* temperature–salinity (T–S) profiles, altimeter sea-level anomaly (SLA), satellite sea-surface temperature (SST), and sea ice concentration (SIC).

Owing to differences in numerical models, assimilation methods, observation data, and atmospheric forcing, there is a diversity in the estimate of three-dimensional ocean state. To identify consistencies and discrepancies among different reanalysis products, it is necessary to carry out an inter-comparison. Similar work has been performed by the CLIVAR/GSOP (Climate and Ocean: Variability, Predictability and Change /Global Synthesis and Observations Panel) and GODAE (Global Ocean Data Assimilation Experiment) communities, and the ORA-IP (Ocean Reanalysis Intercomparison project) Project (Balmaseda et al., 2015; Chevallier et al., 2017; Karspeck et al., 2017; Palmer et al., 2017; Storto et al., 2017; Toyoda et al., 2017; Valdivieso et al., 2017; Uotila et al., 2018; Carton et al., 2019). Furthermore, an eddy-permitting multi-system ensemble reanalysis GREP has been produced in the framework of the Marine Copernicus Service to retain consistent results among different reanalysis products (Masina et al., 2017; Storto et al., 2019a). This dataset offers the possibility of investigating the potential benefits of a multi-system approach and the augmented value of the information on the ensemble spread. The systematic comparison of eddy-permitting global ocean reanalysis products indicates that GREP provides robust conclusions on the recent evolution of oceanic states (Masina et al., 2017).

The China Ocean ReAnalysis (CORA) is supported by the National Marine Data and Information Service (NMDIS). Its first version (CORA1) was released in 2013 (Han et al., 2011, 2013a; 2013b). The NMDIS has now developed a new global reanalysis product, CORA2, by coupling a sea-ice module, adding tidal forcing, enhancing the horizontal (from 25 to 9 km) and vertical resolution (from 35 to 50 layers), and improving atmospheric forcing and data assimilation scheme. Here we introduce the most salient features of CORA2, and present our evaluation of CORA2 through a comparison with CORA1 and several other ocean reanalysis products used by the community. First, we consider GLORYS12v1 and HYCOM, which have an equivalent resolution to CORA2, and we compare these three products to identify their effectiveness in estimating the ocean state under a high-resolution framework. Since the uncertainty of GREP obtained using a low-resolution reanalysis ensemble is consistent with that of high-resolution products (Storto et al., 2019a), we also include GREP in our inter-comparison. ECCO4 uses four-dimensional variational data assimilation (4D-Var) and an expanded set of observational data to modify initial conditions, parameters, and surface forcing fields. These designs are good for conserving ocean momentum, heat, and salt, to provide a dynamically consistent ocean state estimate. Considering its uniqueness, we also included ECCO4 in our inter-comparison. SODA is an ocean reanalysis product with a long history and wide range of applications. Thus, we compare CORA2 and SODA3.

This article is organized as follows. The main characteristics of CORA2 are described in section 2. The data used in assessing CORA2 are introduced in section 3. In sections 4 and 5, we carry out validation, evaluation, and inter-comparison. The summary and conclusions are presented in section 6.

2 Description of reanalysis system CORA2

2.1 Observation data used for assimilation

The assimilated observations include *in-situ* profiles, altimeter SLA, satellite SST, and TPXO8 (TOPEX/POSEIDON global tidal model) surface tidal elevation. The T–S profiles are from the NMDIS archive, World Ocean Database 2018 (WOD 2018; Garcia et al., 2018), Global Temperature and Salinity Profile Project (GTSPP), and the Argo (Array for real-time geostrophic oceanography) Project. The altimeter SLA comes from the gridded AVISO (Archiving, Validation and Interpretation of Satellite Oceanographic data) data, which are part of the Copernicus Marine Environment Monitoring (CMEMS) dataset and merges all the altimetry mission measurements into a daily grid with a spatial resolution of 0.25° (Pujol et al., 2016). The daily NOAA OISSTv2 (Optimum Interpolation Sea Surface Temperature version2) SST data used in CORA1 have been retained in CORA2, with a resolution of 0.25° × 0.25° (Reynolds et al., 2007).

Considering that high-resolution satellite SST data are available, such as ESA CCI (European Space Agency Climate Change Initiative) SST (Merchant et al., 2019) and OSTIA (Operational Sea Surface Temperature and Ice Analysis) SST (Good et al., 2020), the OISSTv2 SST data will be replaced in an updated CORA2 in the future. The TPXO8-atlas with a horizontal resolution of 1/30° is used to generate surface tidal elevation to constrain the MITgcm (MIT General Circulation Model; Egbert and Erofeeva, 2002).

2.2 Ocean and sea-ice models

CORA2 uses version c62h of the MITgcm ocean model, which solves the three-dimensional primitive equations with implicit linear free-surface under the hydrostatic and Boussinesq approximations. The model covers the globe and uses a cube–sphere grid projection, which permits relatively even grid spacing throughout the domain and avoids polar singularities (Adcroft and Campin, 2004; Marshall et al., 1997). Each face of the cube comprises 1,020 × 1,020 grid cells, with a mean horizontal grid spacing of 9 km. The model has 50 vertical levels ranging in thickness from 10 m near the surface to approximately 450 m at a maximum depth of 6,150 m. The topography is from the General Bathymetric Chart of the Ocean (GEBCO08) bathymetry data, with a horizontal resolution of 30 arc-seconds. The time step for model integration is 60 seconds. The model is integrated into a volume-conserving configuration using a finite volume discretization with a C-grid staggering of the prognostic variables. The vertical mixing scheme adopted is the K-profile parameterization (KPP; Large et al., 1994). Horizontal viscosity and diffusivity are parameterized following Griffies and Hallberg (2000). The model employs the quadratic bottom boundary layer drag. The astronomical equilibrium tidal forcing is embodied in the governing equations to simulate the tidal signals (Arbic et al., 2004; Fu et al., 2021). The modeled 3D temperature and salinity are relaxed toward the climatological values from the World Ocean Atlas 2018 (WOA18) with a timescale of approximately 1 year to avoid long-term model drift.

The ocean model is coupled to a dynamic–thermodynamic sea-ice model that computes ice thickness, ice concentration, snow cover, and sea-ice velocity (Zhang et al., 1998). The horizontal grid of the sea-ice model is the same as that of the ocean model. There are momentum, heat, and freshwater flux exchanges between the ocean and sea-ice models. There are seven categories of sea ice in a horizontal grid, which permits an estimate of time-evolving sea-ice thickness distribution. For each category, sea ice is vertically divided into a layer of snow and a layer of ice.

2.3 Data assimilation scheme

The *in-situ* T–S profiles, altimeter SLA, and satellite SST are assimilated using a high-resolution multi-scale data assimilation scheme, which includes four main features. First, the basic data assimilation algorithm is the multi-grid three-dimensional variational (3D-Var) data assimilation scheme used in CORA1 (Li et al., 2008). In the multi-grid 3D-Var, the cost function is first minimized on coarse grids to obtain smooth modes (longwave information), and then the grid resolution increases so that the minimized cost function retrieves oscillatory modes (shortwave information). During the analysis procedure on each grid level, the background error covariance matrix is simplified to the identity matrix. This method can retrieve resolvable information from long and short wavelengths in turn for a given observation network and yield a multi-scale analysis.

Second, it is a high-resolution assimilation. For CORA1, all observations falling in a certain time window are assumed to be located at the analysis time and observation innovations in the cost function is obtained by using three-dimensional spatial interpolation without temporal weight, which might lose some observational signals, especially high-frequency signals (such as diurnal variation). For the high-resolution reanalysis CORA2, this scheme might reduce the quality of small-scale information in the final product. To address this problem, CORA2 uses the First Guess at Appropriate Time (FGAT) approach to enhance the quality of observation innovation in data assimilation to improve the assimilation effect of temporal small-scale signals. The FGAT approach uses the model result with the time nearest to the observation time to compute the observation innovation (Cummings and Smedstad, 2013), which is input into the multi-grid 3D-Var to produce the analysis result. The FGAT approach can help CORA2 reconstruct some deterministic high-frequency variabilities.

Third, the scheme places constraints on the T–S relationship. As in CORA1, CORA2 employs the method proposed by Troccoli et al. (2002), adjusting salinity when temperature measurements are the only available measurements. This constraint ensures that the T–S relationship derived from the model simulation result is essentially conserved during temperature data assimilation. When salinity measurements are available, the model-simulated T–S relationship is adjusted to the observed counterpart by assimilating salinity.

Fourth, the assimilation scheme of gridded daily altimeter SLA data in CORA1 is also retained in CORA2. The assimilated altimeter SLA is given as a daily average, which does not contain tidal information, and is mainly used to optimize meso-scale eddies. The altimeter SLA data are first projected onto the gridded synthetic T–S profiles using the Cooper and Haines (1996) scheme. Then, the synthetic T–S profiles are assimilated to the daily-averaged

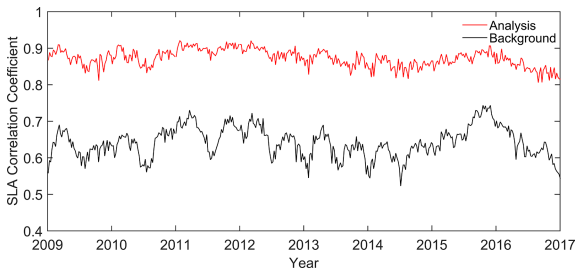


FIGURE 1
Daily spatial correction coefficient of SLA between analysis (red) [background (black)] field and altimeter observation within 50°S–50°N.

background fields using the multi-grid 3D-Var analysis scheme to generate temperature and salinity analysis fields. Figure 1 shows that the assimilation of the altimeter SLA can increase the spatial correlation coefficient of SLA from approximately 0.65 to approximately 0.90. An advantage of this SLA assimilation method is that it relies on the simulated background field and maintains the dynamic consistency of the ocean state. However, a disadvantage is that it cannot explicitly correct for model errors due to model drift.

The high-resolution multi-scale assimilation process is as follows. First, the daily altimeter SLA is converted into synthetic T–S profiles, and the daily-averaged background fields of temperature and salinity are adjusted by using the multi-grid 3D-Var to assimilate those profiles and satellite SST. Second, the multi-grid 3D-Var and FGAT algorithms are used to assimilate *in-situ* temperature profiles to adjust the instantaneous background temperature field, and the T–S relationship constraint is used to complete the adjustment of the instantaneous background salinity field. Finally, the final adjustment of the instantaneous background salinity field is completed by assimilating *in-situ* salinity profiles using the multi-grid 3D-Var and FGAT algorithms. *In-situ* T–S profiles (above 2,000 m) and daily satellite SST are assimilated every day with a 1-day time window. The altimeter-derived T–S profiles above 1,000 m within 50°S–50°N are assimilated every 7 days with a 1-day time window.

Compared with CORA1, an advantage of CORA2 is that it can provide tidal information. To improve tidal accuracy, we employed the nudging method proposed by Fu et al. (2021) to restore the surface tidal elevation of the forecast model toward that of TPXO8 at each integration time. Fu et al. (2021) suggested that this method can not only improve the accuracy of surface tidal elevation but also optimize the subsurface temperature and salinity disruptions caused by tides. Here, we show the amplitude and phase of M_2 and K_1 tidal constituents obtained by using harmonic analysis for the sea-surface height field of CORA2 and the simulation (Figure 2). The comparison between CORA2 and the simulation reveals that tidal assimilation can significantly improve the accuracy of tidal signals. It should be noted that the assimilation of daily altimeter SLA was performed to

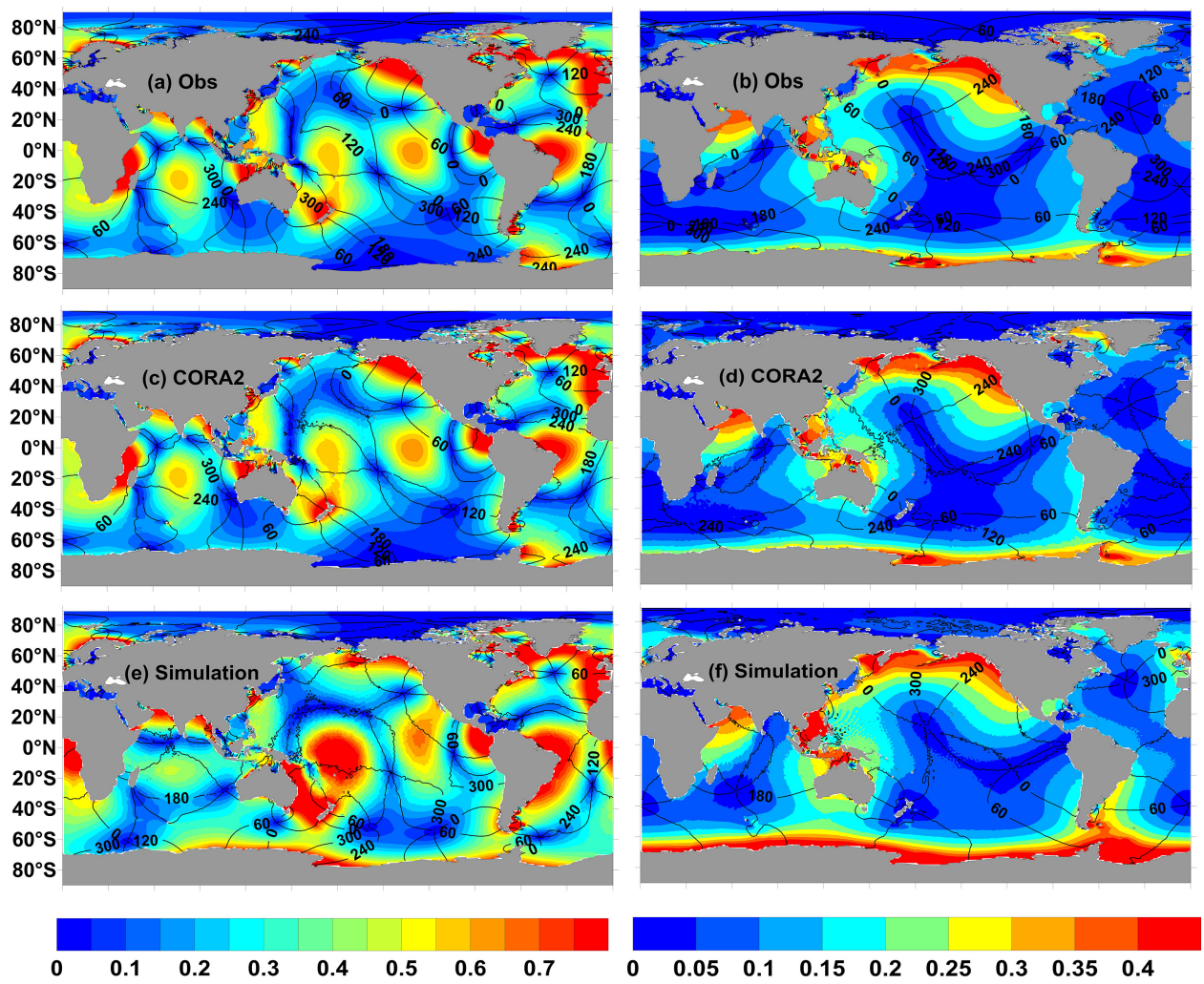


FIGURE 2
M2 **(A, C, E)** and K1 **(B, D, F)** amplitude (shading; units: m) and phase (contour; units: °) of surface tidal elevation in TPXO8 **(A, B)**; a barotropic tide model constrained by observation data), China Ocean ReAnalysis version 2 (CORA2) (**C, D**; with tidal assimilation), and the simulation results (**E, F**; without tidal assimilation).

adjust meso-scale eddies, and the assimilation of surface tidal elevation was performed to improve tidal information accuracy.

2.4 Surface forcing and spin-up

The atmospheric forcing variables include wind at a 10-m height, air temperature and humidity at a 2-m height, total precipitation, and surface downward shortwave and longwave radiative fluxes, which are taken from the Japanese Meteorological Agency reanalyses JRA-25, spanning from 1980 to 2013, and JRA-55, spanning from 2014 to 2019 (Onogi et al., 2007; Kobayashi et al., 2015). The bulk formulae of Large and Pond (1981; 1982) are used to calculate surface fluxes for the open oceans. Surface fluxes over sea ice are calculated based on the method in Parkinson and Washington (1979). Monthly climatology of river runoff is applied along the land mask and treated as freshwater flux (Fekete et al., 2002).

The generation of initial conditions for CORA2 includes the following phases. First, the numerical model is freely integrated for 10 model years starting from the temperature and salinity fields from WOA18, with the climatological atmospheric forcing. This is followed by a 6-year simulation period, driven by the 1980–1985 JRA-25 atmospheric forcing. Then, *in-situ* T–S profiles and satellite SST are assimilated to adjust the model to observations since 1986. Altimeter observation is assimilated from 1997. After that, when the system was integrated until 2009, it is found that there is a slightly large error, with a temperature root-mean-square error (RMSE) of >1.3°C and salinity RMSE of >0.25 psu in the Atlantic Ocean. The error is mainly caused by the following factors: (1) the rationality test of the data range in the *in-situ* profile quality control procedure has a small bug for the Atlantic Ocean; (2) there is an excessive reuse of the high-density *in-situ* profiles; and (3) owing to the overflow of high-salinity water from the Mediterranean Sea in the deep layer, the T–S relationship is relatively complex in the North Atlantic Ocean, and the simple temperature-to-salinity mapping adjustment algorithm proposed by Troccoli et al. (2002) is not applicable. The assimilation scheme is optimized to address the above problems, including through the fine tuning of the *in-situ* profile quality control procedure, the thinning of high-density *in-situ* profiles, and the limiting of the T–S relationship constraint range, for the integration from 2009 onward.

The ocean variables in the CORA2 product, including sea surface height (SSH), 3D temperature, salinity, and current, are saved on a uniform horizontal grid of 0.1° × 0.1° and 50 layers at 3-hour intervals. The tide signals are embodied in the oceanic variables. The derived daily and monthly datasets spanning from 1989 to 2019 are calculated and released on the websites http://mds.nmdis.org.cn and http://www.cmoc-china.cn. In this study, we focus only on the reanalysis products during the Argo-rich period, namely 2004–2019.

3 Other analysis and reanalysis datasets

The reanalysis products GLORYS12v1 (Lellouche et al., 2021), HYCOM (GLBu0.08; Cummings and Smedstad, 2013), GREP (version 2; Storto et al., 2019a), ECCO4 (version 4r4; Forget et al., 2015), and SODA3 (version 3.4.2; Carton et al., 2018) are used here for inter-comparison, and their characteristics are given in Table S1 in the Supplementary Material.

GLORYS12v1 is a global eddy-resolving ocean reanalysis spanning 1993 to 2020 with a horizontal resolution of 0.083°. The surface atmospheric fields are from the European Centre for Medium Range Weather Forecasts (ECMWF) ERA-Interim reanalysis. Reprocessed along-track altimeter SLA from the CMEMS, the Advanced Very High Resolution Radiometer (AVHRR) SST from the NOAA, SIC from the Centre d'Exploitation et de Recherche SATellitaire (CERSAT), and *in-situ* T–S profiles from the CMEMS are jointly assimilated using a Singular Extended Evolutive Kalman (SEEK) filter with a 7-day assimilation cycle. GLORYS12v1 uses the incremental analysis update (IAU) procedure in Bloom et al. (1996) to weaken shocks and spurious waves introduced by the "classical" model correction, where analysis increments would be applied in a one-time step.

The outputs from HYCOM experiment GLBu0.08 versions 19.1 and 19.0, covering the period from October 1992 to December 2012, are used in this study. The horizontal resolution is 0.083°. Surface forcing is from the National Centers for Environmental Prediction (NCEP) Climate Forecast System Reanalysis (CFSR) with a horizontal resolution of 0.3° and temporal resolution of 1 hour. The time window for observation assimilation is 1 day. Along-track satellite SST data are assimilated to maintain a diurnal cycle in the model. The Modular Ocean Data Assimilation System (MODAS) is used to project along-track altimeter SLA to depth in the form of synthetic T–S profiles. The final analysis increments are inserted into the model over a 6-hour time period using the IAU procedure.

Version 2 of GREP is a four-member ensemble reanalysis with a horizontal resolution of 0.25° and 75 standard z-levels spanning 1993 to 2020. All ensemble members use the NEMO (Nucleus for European Modelling of the Ocean) ocean model, which is forced by the ECMWF ERA-Interim atmospheric forcing, albeit with different bulk formulae and employing different observational datasets and data assimilation schemes. A preliminary assessment of GREP indicates that the ensemble mean outperforms all individual members in approaching *in-situ* profiles (Storto et al., 2019a).

Version 4 of the ECCO uses the MITgcm to reconstruct ocean and sea-ice states from 1992 to 2017, with a horizontal resolution of 0.5°. ECCO4 employs a 4D-Var method to modify initial conditions, parameters, and surface forcing fields to minimize analysis-minus-observation misfits in a least-squares sense. Assimilated observation data include SLA from the satellite altimeter, SST from satellite radiometers (AVHRR), sea surface salinity (SSS) from the Aquarius satellite radiometer/scatterometer, ocean bottom pressure (OBP) from the Gravity Recovery and Climate Experiment (GRACE) satellite gravimeter, SIC from satellite radiometers (Special sensor microwave/imager and Special Sensor Microwave Imager Sounder), and *in-situ* T–S profiles from the the World Ocean Circulation Experiment (WOCE), the Global Ocean Ship-based Hydrographic Investigations Program (GO-SHIP), Argo, and so on.

SODA3 (version 3.4.2) reconstructs ocean and sea-ice states from 1980 to 2017. Its horizontal resolution is 0.25°. The surface

forcing is from the ECMWF ERA-Interim reanalysis. SODA3 assimilates WOD T–S profiles, International Comprehensive Ocean-Atmosphere Data Set (ICOADS) *in-situ* observation, and satellite SST (the NOAA Advanced Clear-Sky Processor for Ocean Level 2P SST product) using an optimal interpolation scheme. The IAU procedure is implemented using an update cycle of 10 days.

Objective analysis EN4 ENACT/ENSEMBLES version4 (version 4.2.1) of subsurface temperature and salinity from the Met Office Hadley Centre (Good et al., 2013) is used in the study. It provides gridded data of 1° × 1° × 42 levels and is a monthly complete-spatial-coverage objective analysis. The latest satellite OSTIA SST data with a horizontal resolution of 1/20° (Good et al., 2020) are also used to calculate the RMSE, bias, and correlation coefficient of the SSTs of reanalysis products.

4 Comparison of CORA2 with CORA1

The main differences and improvements of CORA2 with respect to its previous version, CORA1, are described here. We analyze both versions to understand which improvements in the CORA2 system are due to system changes. First, SST RMSEs and biases with respect to the OISST SST in CORA2 are compared with those in CORA1 (Figures 3A, B). The difference in SST biases between CORA2 and CORA1 is not very large, except for a bias excursion in 2013 for CORA1. However, the SST RMSE of CORA2 is significantly smaller than that of CORA1. In CORA1, the satellite OISST SST data were assimilated by using a surface relaxation scheme, and its constraint effect was decided by the relaxation coefficient. In CORA2, the OISST SST is assimilated using the multi-grid 3D-Var method with a 1-day assimilation cycle. The error relative to non-independent observations is mainly decided by the assimilation scheme. Therefore, we suggest that the current satellite SST assimilation scheme in CORA2 is more advantageous for constraining the SST to the observation. Of course, there may also be other factors that cause the reduction of CORA2 SST error, such as the atmospheric forcing change and the resolution improvement.

Figure 3 also shows that the RMSEs and biases in subsurface temperature and salinity with respect to the Argo profiles in CORA2 are smaller than those in CORA1. For the assimilation of *in-situ* T–S profiles, CORA2 and CORA1 use the same basic assimilation method, namely multi-grid 3D-Var; however, compared with CORA1, CORA2 has improved resolution and used the FGAT method to achieve high-resolution assimilation. CORA2 also adds tidal forcing and assimilation to resolve some small-scale internal tidal signals contained in the *in-situ* T–S profiles. These changes may be the main reasons for the improvement in subsurface temperature and salinity accuracy in CORA2.

Compared with CORA1, a significant improvement of CORA2 is the enhancement of the spatial resolution from the eddy-permitting to the eddy-resolving level. To evaluate its ability to reproduce meso-scale eddy signals, the temporal and spatial distributions of eddy kinetic energy (EKE) are calculated based on daily velocity. Figure 4A shows the temporal evolution of the monthly three-dimensional means of EKE over 0–300 m during the period 2010–2017. CORA2 shows higher EKE than CORA1. Figure 4B demonstrates that all the large dynamic systems are well represented by CORA2, including the western boundary currents, Agulhas recirculation, and Antarctic Circumpolar Current (ACC).

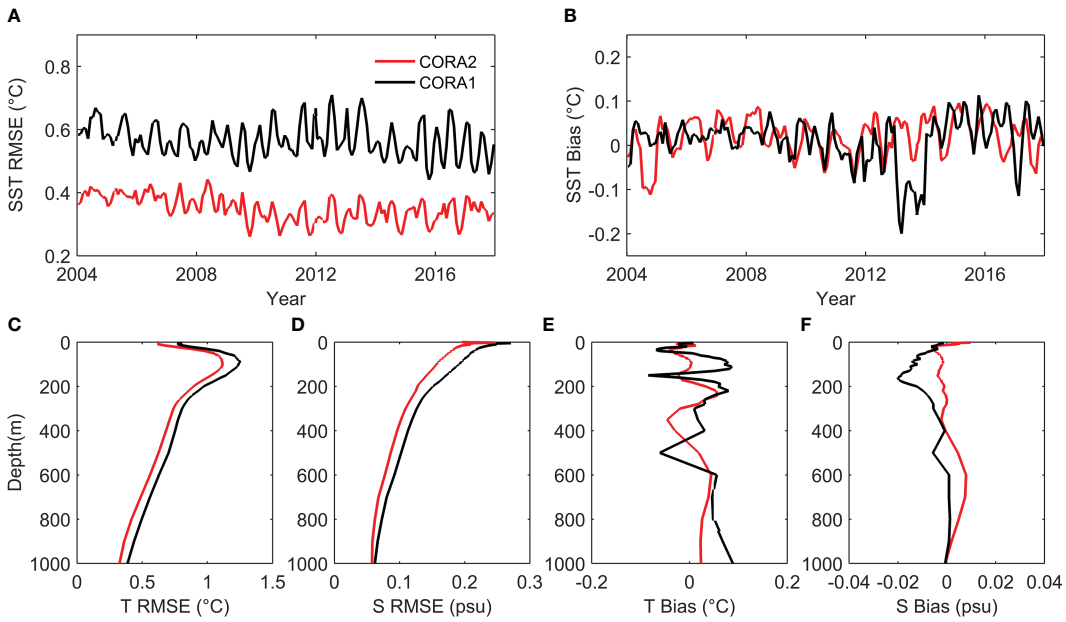


FIGURE 3
Time series of root-mean-square errors (RMSEs) (**A**; units: °C) and biases (**B**; reanalysis minus observation; units: °C) of monthly SST for CORA2 (red) and CORA1 (black) with respect to OISST sea-surface temperature (SST) within 70°S–70°N. Vertical distributions (0–1,000 m) of RMSEs **(C, D)** and biases **(E, F)** of monthly temperature (**C, E**; units: °C) and salinity (**D, F**; units: psu) for China Ocean ReAnalysis version 2 (CORA2) (red) and China Ocean ReAnalysis version 1 (CORA1) (black) with respect to Argo profiles in the global oceans during the period of 2004–2017.

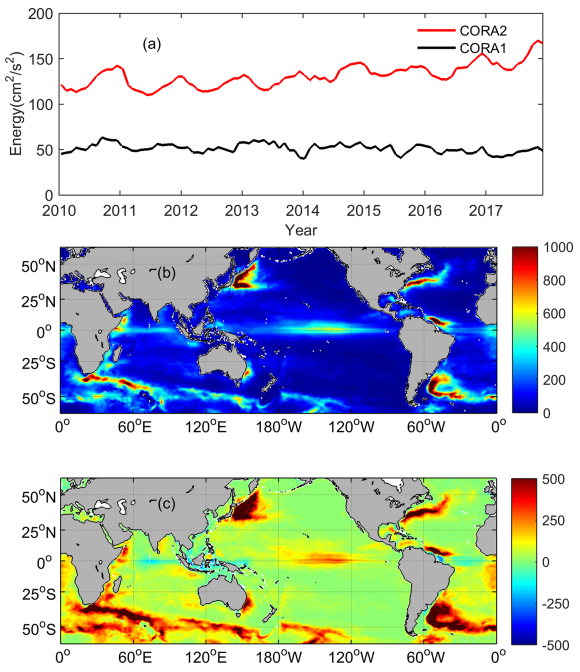


FIGURE 4
(A) Three-dimensional mean of monthly eddy kinetic energy **(EKE)** (cm^2/s^2) at 0–300 m depth for the China Ocean ReAnalysis version 2(CORA2) (red) and China Ocean ReAnalysis version 1 (CORA1) (black). **(B)** Average EKE (cm^2/s^2) at 0-300 m depth during 2010–2017 for CORA2 and **(C)** the differences when compared with CORA1 (CORA2 minus CORA1).

The EKE level in CORA2 is of the same order of magnitude as that of GLORYS12v1, with a similar resolution (Lellouche et al., 2021). A comparison of CORA2 and CORA1 shows an obvious increase in EKE alongside the increased resolution (Figure 4C).

5 Comparison of CORA2 with other reanalyses

We first compare the RMSEs, biases, and correlation coefficients of monthly SSTs for the six reanalysis products (CORA2, GLORYS12v1, HYCOM, GREP, SODA3, and ECCO4). For subsurface temperature and salinity evaluation, we project the monthly reanalyzed 3D temperature and salinity onto *in-situ* profile locations to obtain analysis-minus-observation misfits. We then analyze the time and space errors (RMSE and bias) statistically. The *in-situ* profiles are divided into two groups: the assimilated Argo data and the non-assimilated *in-situ* profiles. The gridded SLA assimilated in CORA2 is also used to evaluate the fidelity of temporal variability in various ocean reanalyses. Considering the different time periods covered by the various products, we evaluate monthly SST, 3D temperature and salinity, and SLA over 2004–2017 for all reanalyses, except for HYCOM (2004–2012). In addition, through comparison with the objective analysis EN4, we also assess variations in monthly steric sea level (including thermosteric and halosteric components) and ocean heat content (OHC) in CORA2, GLORYS12v1, and GREP using the time period of 2004–2019.

5.1 SST

We compare the temperature at the shallowest reanalysis level to the OSTIA SST. Figure 5A shows that the SST RMSE of GREP is the smallest and that of ECCO4 is the largest of the six reanalyses, and that the SST RMSE of CORA2 becomes closer to that of GREP from 2009 onward. SODA3 has a slightly smaller SST RMSE than ECCO4 and a larger SST RMSE than HYCOM and GLORYS12v1. The high accuracy of GREP may be attributed to its ensemble nature (Masina et al., 2017). The daily assimilation of gridded SST in CORA2 greatly constrains the modeled SST to closely match the observations. Although the resolutions of GLORYS12v1 and HYCOM are similar to that of CORA2, the 7-day assimilation cycle in GLORYS12v1 and the along-track SST assimilation in HYCOM provide a relatively weak SST constraint. For SODA3, because the 10-day assimilation cycle is relatively long, the analyzed SST RMSE is relatively large. The largest SST departure of ECCO4 may be related to its assimilation scheme, which tends to maintain the conservation of ocean energy and mass rather than place a mandatory constraint on SST. Compared with the RMSEs, the spread of biases in the six reanalyses is small (Figure 5B). Each reanalysis has a spatial correlation coefficient above 0.99, which means that they can effectively reproduce the spatial structures of SST (Figure 5C).

The spatial patterns of SST RMSE of the six products are similar. Low error mainly occurs in the open seas and high error in coastal waters, western boundary currents, and ACC area (Figure 6). The large errors in the coastal waters, western boundary currents, and ACC area may be associated with the poor representation of strong non-linear dynamic processes and the displacement of SST fronts. Similar to results in the time series of RMSE, the SST RMSEs of CORA2 and GREP are the lowest, being less than 0.3°C in the ocean interior and greater than 0.6°C in the western boundary currents and ACC area. The SST error of ECCO4 is the largest, being less than 0.5°C in the ocean interior and greater than 0.8°C in the western boundary currents and ACC area extending to the coastal waters. The error levels of GLORYS12v1, HYCOM, and SODA3 lie between those of ECCO4 and CORA2/GREP. We also note that the SST RMSE of CORA2 is greater than that of the other products in the Okhotsk Sea, which may be related to the freezing and melting of sea ice. The ratios of the standard deviation of reanalysis SSTs relative to that of the OSTIA SST were also estimated to analyze SST variability (Figure S1 in the Supplementary Material).

5.2 Subsurface temperature and salinity

In this subsection, we use Argo profiles to assess the subsurface performance of CORA2 relative to the other reanalyses. Considering that the use of monthly means does not compromise the skill score statistics during an observation-rich period (Storto et al., 2019a), we focus on the monthly mean fields of temperature and salinity. The vertical distributions of the temperature and salinity RMSEs and biases of the six reanalyses are shown in Figure 7. For all reanalyses, the highest temperature RMSE occurs near the thermocline. The

10.3389/fmars.2023.1084186

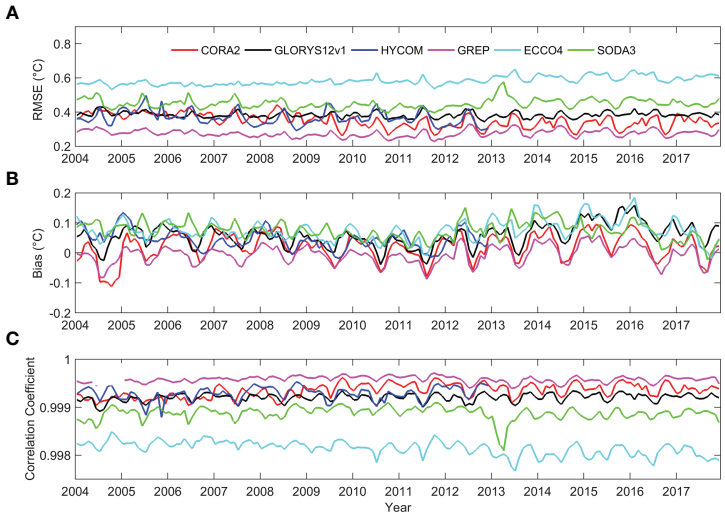


FIGURE 5
Time series of **(A)** root-mean-square errors (RMSEs) (units: °C), **(B)** biases (units: °C; reanalysis minus observation), and **(C)** spatial correlation coefficients of monthly sea-surface temperature (SST) relative to OSTIA SST within 70°S–70°N for CORA2 (China Ocean ReAnalysis version 2) (red), GLORYS12v1 (Global Ocean reanalysis and Simulation) (black), HYCOM (HYbrid Coordinate Ocean Model) (blue), GREP (Global ocean Reanalysis Ensemble Product) (pink), ECCO4 (Estimating the Circulation and Climate of the Ocean, version 4) (cyan), and SODA3 (Simple Ocean Data Assimilation, version 3) (green).

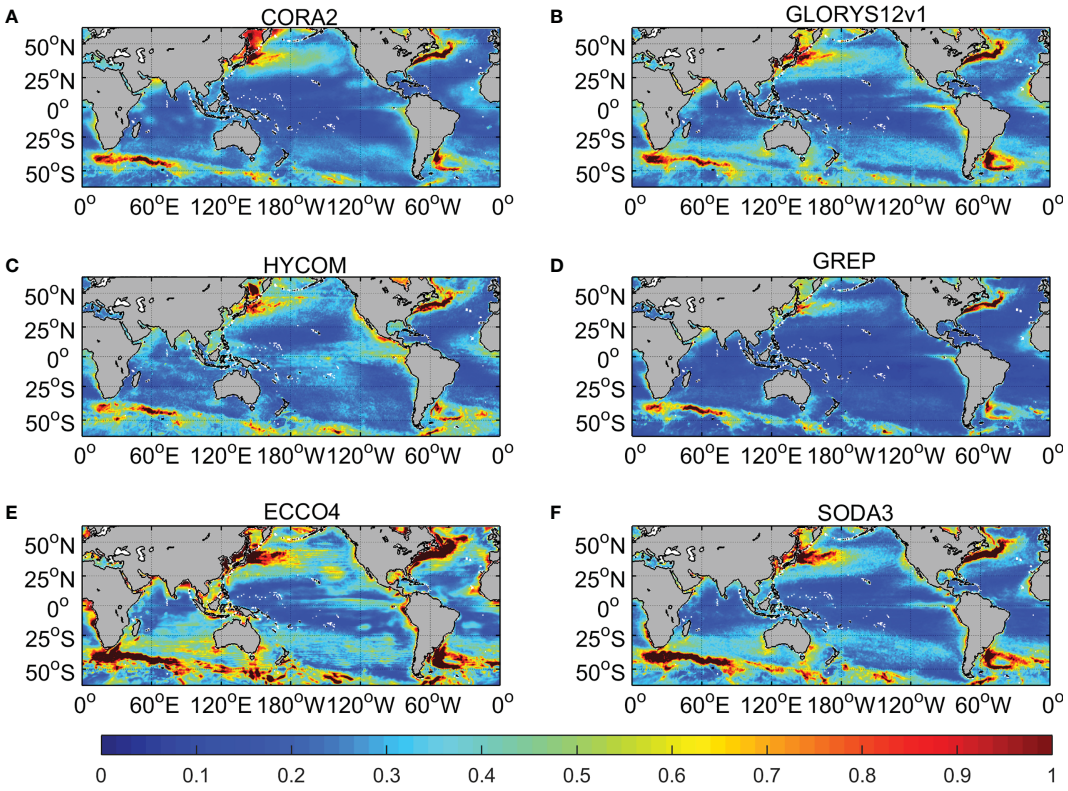


FIGURE 6
Spatial distributions of the root-mean-square error (RMSE) (units: °C) of monthly sea-surface temperature (SST) with respect to OSTIA SST for CORA2 (China Ocean ReAnalysis version 2) **(A)**, GLORYS12v1 (Global Ocean reanalysis and Simulation) **(B)**, HYCOM (HYbrid Coordinate Ocean Model) **(C)**, GREP (Global ocean Reanalysis Ensemble Product) **(D)**, ECCO4 (Estimating the Circulation and Climate of the Ocean, version 4) **(E)**, and SODA3 (Simple Ocean Data Assimilation, version 3) **(F)**. All RMSEs are computed for the period 2004–2017, except for HYCOM (2004–2012).

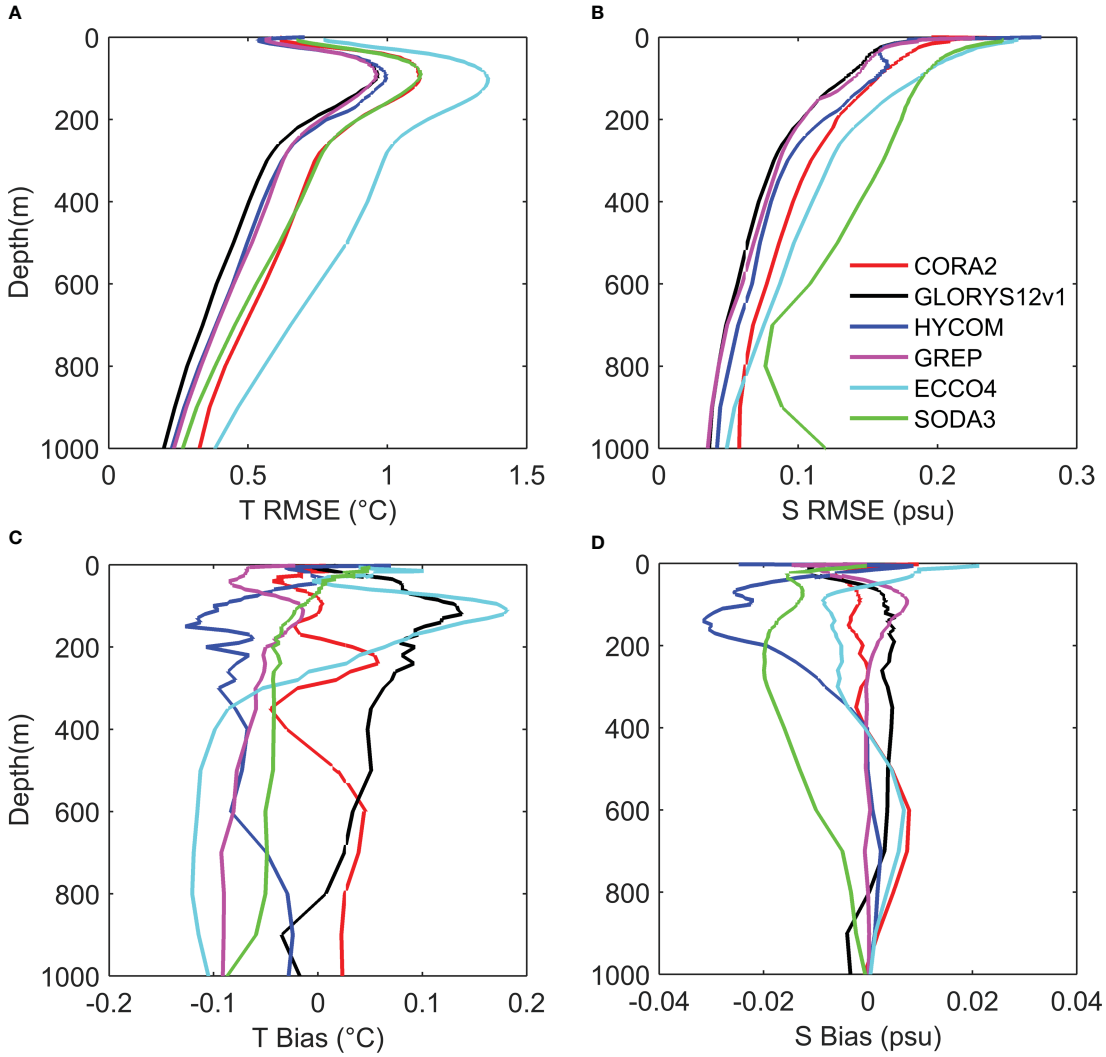


FIGURE 7
Vertical distributions (0–1,000 m) of root-mean-square errors (RMSEs) **(A, B)** and biases (reanalysis minus observation; **C, D**) of monthly temperature (**A, C**; units: °C) and salinity (**B, D**; units: psu) for the six reanalyses [CORA2 (China Ocean ReAnalysis version 2), red; GLORYS12v1 (Global Ocean reanalysis and Simulation), black; HYCOM (HYbrid Coordinate Ocean Model), blue; GREP (Global ocean Reanalysis Ensemble Product), pink; ECCO4 (Estimating the Circulation and Climate of the Ocean, version 4), cyan; SODA3 (Simple Ocean Data Assimilation, version 3), green] with respect to Argo profiles in the global oceans. RMSE and bias are computed between 2004 and 2017 for all reanalyses, except for HYCOM (2004–2012).

temperature RMSE of CORA2 is similar to that of SODA3, lower than that of ECCO4, and larger than those of GLORYS12v1, GREP, and HYCOM. Owing to the uncertainties in the surface freshwater flux and runoff, the highest salinity RMSE occurs near the sea surface. GREP and GLORYS12v1 have the smallest salinity RMSEs and ECCO4 and SODA3 have the largest salinity RMSEs. Compared with the other products, the salinity RMSE of CORA2 is at a medium level. The large salinity RMSE of SODA3 might be related to the large salinity error in the Mediterranean Sea region (Figure S2 in the Supplementary Document).

In the upper ocean, although the temperature bias of CORA2 shows an obvious positive and negative alternation structure, its average is near zero (Figure 7C). The positive and negative alternation may be caused by the misplacement of thermocline depth. Compared with CORA2, the temperatures of GREP, HYCOM, and SODA3 show significant negative biases, while those of GLORYS12v1 and ECCO4 show significant positive biases. For salinity, the biases are relatively small for all the products except SODA3 and HYCOM.

Spatial distributions of RMSEs of the monthly temperature of the six reanalyses averaged over 0–2,000 m are displayed in Figure 8. Temperature RMSEs show similar spatial structures in all six reanalyses. The large errors in strong current areas, for example, the Gulf Stream, Kuroshio, and equatorial currents, may be caused by the misplacement of fronts, eddies, and thermocline. Table 1 shows the RMSEs for the six reanalyses in 12 areas. In the ocean interior, i.e., the North Indian Ocean, South Indian Ocean, Northeast Pacific, South Pacific, South Atlantic, and Southern Ocean, GLORYS12v1, HYCOM, and GREP have errors of between 0.44°C and 0.76°C, while CORA2 and SODA3 have slightly larger errors, of between 0.48°C and 0.82°C. ECCO4 has the largest errors, of between 0.61°C and 1.05°C. In the areas of the Gulf Stream, Kuroshio, and equatorial currents, the RMSEs of

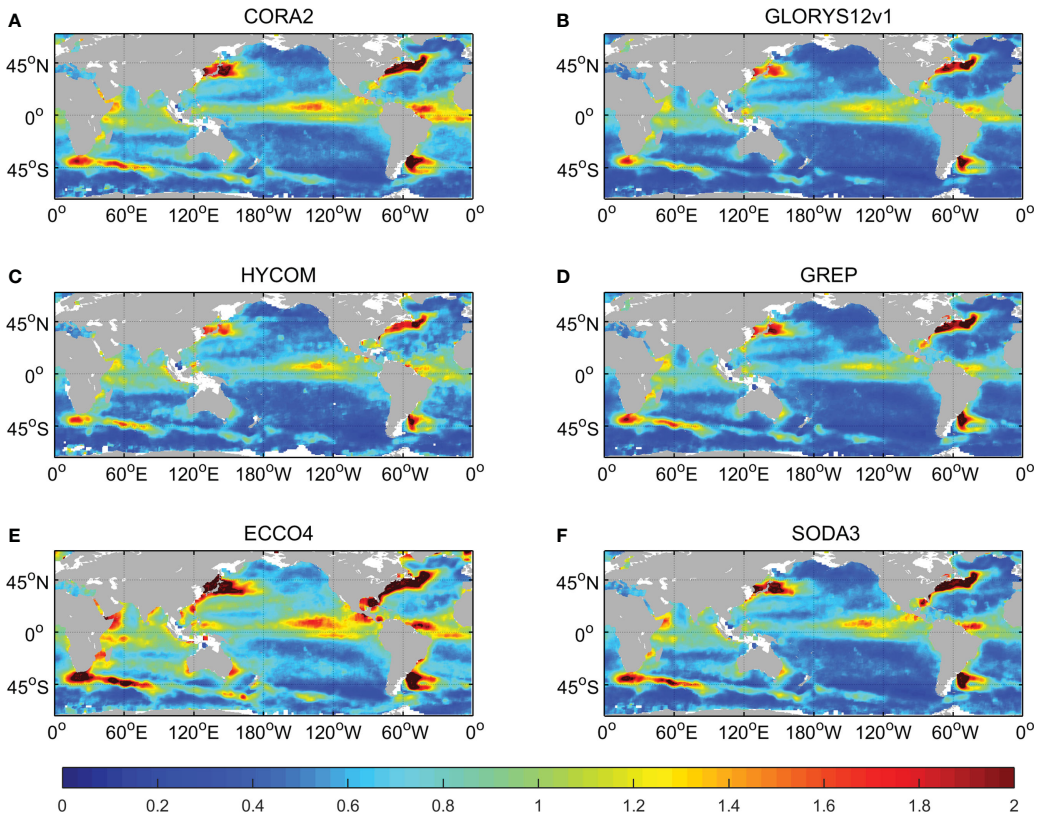


FIGURE 8
Spatial distributions of monthly temperature root-mean-square error (RMSE) (units: °C) of the six reanalyses [CORA2 (China Ocean ReAnalysis version 2), **(A)**; GLORYS12v1 (Global Ocean reanalysis and Simulation), **(B)**; HYCOM (HYbrid Coordinate Ocean Model), **(C)**; GREP (Global ocean Reanalysis Ensemble Product), **(D)**; ECCO4 (Estimating the Circulation and Climate of the Ocean, version 4), **(E)**; SODA3 (Simple Ocean Data Assimilation, version 3), **(F)**] with respect to Argo profiles over 0–2,000 m. RMSE is computed between 2004 and 2017 for all reanalyses, except for HYCOM (2004–2012).

GLORYS12v1, HYCOM, and GREP increase to 0.79–0.98°C, those of CORA2 and SODA3 increase to 0.90–1.27°C, and the RMSE of ECCO4 increases to 1.06–1.52°C. ECCO4 has the largest RMSE, reaching 1.07°C in the global oceans, while CORA2 and SODA3 have medium RMSEs of 0.87°C. HYCOM, GREP, and GLORYS12v1 have the smallest RMSEs, of 0.73–0.75°C, in the global oceans.

Figure 9 is the same as Figure 8, but for salinity. The spatial patterns of salinity RMSE are similar for the six reanalyses, with small errors occurring in the ocean interior and large errors in coastal areas. The large RMSEs are generally associated with the uncertainties of climatological runoff and freshwater flux. Table 1 shows that the salinity RMSE of CORA2 (0.15 psu) in the global oceans is larger than the salinity RMSEs of GLORYS12v1 (0.12 psu), GREP (0.12 psu), and HYCOM (0.13 psu), but smaller than those of ECCO4 (0.17 psu) and SODA3 (0.18 psu). Large RMSEs, of more than 0.20 psu, are found mainly in the North Indian Ocean and Atlantic Ocean for the six reanalyses. In addition, the large salinity RMSE of SODA3 in the Mediterranean Sea is consistent with the result mentioned above (Figure S2 in the Supplementary Document).

Similar to Table 1, Table S2 in the Supplementary Material gives the biases of temperature and salinity for the six reanalyses in the 12 ocean areas. The temperature (salinity) biases in the global oceans are −0.003°C (−0.002 psu), 0.058°C (0.000 psu), −0.052°C (−0.014 psu), −0.060°C (0.001 psu), 0.033°C (−0.001 psu), and −0.024°C (−0.012 psu) for CORA2, GLORYS12v1, HYOM, GREP, ECCO4, and SODA3, respectively. Generally, the absolute value of the bias in the Atlantic Ocean is larger than the absolute bias values in the other regions for all the reanalysis products.

Figure 10 shows the RMSE time series of various reanalyses in the Indian, Pacific, Atlantic, and global oceans. We can see that the temperature RMSEs of all products decrease with time, which may be due to the increasing number of assimilated observations. The accuracy of CORA2 is similar to that of SODA3 in the Pacific, Indian, and Atlantic oceans during 2009–2017. We find that the RMSE in the Atlantic for CORA2 is relatively large before 2009 and sharply declines after we optimized the quality control procedure and assimilation scheme of temperature and salinity for the year 2009 onward. In addition, the salinity error of SODA3 in the global ocean increases rapidly after 2011, which may be caused by the large salinity error in the Mediterranean region.

We chose independent observations from a station at 117.5°E, 19.0°N to further validate the performance of the reanalyses, which comprises the cross-shaped observational array of buoys and moorings in the northern South China Sea deployed by China (Zhang et al., 2016). To better compare the variability at the sub-

TABLE 1 Root-mean-square errors (RMSEs) of monthly reanalyses [CORA2 (China Ocean ReAnalysis version 2), GLORYS12v1 (Global Ocean reanalysis and Simulation), HYCOM (HYbrid Coordinate Ocean Model), GREP (Global ocean Reanalysis Ensemble Product), ECCO4 (Estimating the Circulation and Climate of the Ocean, version 4), SODA3 (Simple Ocean Data Assimilation, version 3)] temperature (T; units: °C) and salinity (S; units: psu) against Argo profiles in the equatorial Indian Ocean (40–100°E, 10°S–10°N), equatorial Pacific (130°E–80°W, 10°S–10°N), equatorial Atlantic (50°W–0°, 10°S–10°N), North Indian Ocean (40–100°E, 10–30°N), Northwest Pacific (120–180°E, 12–50°N), Northeast Pacific (180°E–90°W, 12–50°N), North Atlantic (80°W–0°, 12–50°N), South Indian Ocean (40–120°E, 30–10°S), South Pacific (150°E–80°W, 30–10°S), South Atlantic (50°W–0°, 30–10°S), Southern Ocean (180°E–180°W, 60–30°S), and global oceans.

	CORA2		GLORYS12v1		HYCOM		GREP		ECCO4		SODA3	
	T	S	T	S	T	S	T	S	T	S	T	S
Equatorial Indian Ocean	0.97	0.14	0.89	0.13	0.92	0.14	0.85	0.12	1.07	0.16	0.95	0.14
Equatorial Pacific	0.96	0.14	0.87	0.12	0.88	0.13	0.79	0.11	1.06	0.15	0.90	0.13
Equatorial Atlantic	1.27	0.19	1.00	0.14	1.00	0.14	0.85	0.13	1.21	0.17	0.97	0.14
North Indian Ocean	0.81	0.21	0.71	0.19	0.76	0.23	0.73	0.19	1.05	0.28	0.82	0.23
Northwest Pacific	1.12	0.12	0.92	0.10	0.92	0.10	0.98	0.11	1.52	0.15	1.20	0.12
Northeast Pacific	0.64	0.11	0.55	0.09	0.57	0.10	0.54	0.09	0.81	0.14	0.62	0.10
North Atlantic	1.05	0.24	0.87	0.20	0.90	0.22	0.94	0.21	1.31	0.27	1.12	0.25
South Indian Ocean	0.68	0.13	0.61	0.12	0.64	0.13	0.60	0.12	0.88	0.15	0.67	0.12
South Pacific	0.51	0.10	0.45	0.09	0.48	0.09	0.44	0.08	0.61	0.11	0.48	0.09
South Atlantic	0.69	0.17	0.51	0.11	0.59	0.15	0.47	0.10	0.67	0.15	0.53	0.11
Southern Ocean	0.80	0.13	0.62	0.10	0.66	0.11	0.68	0.10	0.97	0.14	0.81	0.12
Global	0.87	0.15	0.73	0.12	0.75	0.13	0.74	0.12	1.07	0.17	0.87	0.18

Note that statistics are averaged over the period of 2004–2012 for HYCOM and of 2004–2017 for all other reanalyses.

monthly scale, we analyzed daily reanalysis products, i.e., CORA2, GLORYS12v1, GREP, and ECCO4. Figure 11 plots the time series of temperature (Figure 11, left panels) and salinity (Figure 11, right panels) profiles at the station from 1 August 2014 to 28 February 2015 for the observations and for the four reanalyses. The observations exhibit an obvious seasonal variability, with the deepening (August–December 2014) and shoaling (January–February 2015) of the thermocline and halocline, and some sub-monthly-scale disruptions, such as the sinking (day 100) and rising (day 125) of the water column. The four reanalyses can all describe the seasonal variabilities. Owing to their higher resolution, GLORYS12v1 and CORA2 can better depict sub-monthly-scale variability than GREP and ECCO4. GLORYS12v1 has the smallest temperature and salinity RMSEs of 0.78°C and 0.14 psu, respectively. Because of the smoother, warmer, and saltier characteristics of ECCO4, it has the largest RMSEs, exhibiting the worst skill score. Compared with the observations, there are several shocks and spurious waves in the temperature and salinity fields of CORA2; these may be caused by the use of tidal forcing and the assimilation scheme, which simply adds the analysis increment in one step rather than gradually absorbing it.

5.3 Sea level

Monthly sea-level fields in the different reanalysis products were compared with that of the gridded AVISO absolute dynamic topography (ADT) to evaluate their ability to reproduce sea-level variability. CORA2 and HYCOM assimilate the T–S profiles derived by altimeter SLA, while GLORYS12v1 and ECCO4 assimilate along-track altimeter SLA. Thus, the gridded AVISO ADT is not directly assimilated in these reanalyses. Figure 12 shows the spatial distribution of the temporal correlation coefficients of monthly sea level between the reanalyses and satellite observations. GLORYS12v1 and GREP have the highest correlation coefficients, exceeding 0.8 in most regions. HYCOM is comparable to GLORYS12v1 and GREP in the Pacific and Indian oceans, but not in the Atlantic Ocean. CORA2 has higher correlation coefficients in the Pacific Ocean and Indian Ocean and a lower correlation coefficient in the Atlantic Ocean than ECCO4 and SODA3. Considering that the assimilation scheme of SLA used in CORA2 relies on background fields, the large temperature and salinity errors in the Atlantic are probably responsible for the low correlation coefficients. The lowest correlation for SODA3 is mainly due to the absence of sea-level assimilation.

5.4 Steric sea level

Global mean sea level (GMSL) can be decomposed into steric change and mass change, while the steric change can in turn be decomposed into thermosteric and halosteric changes. In this subsection, we compare globally averaged values of steric sea level and its thermosteric and halosteric components in CORA2 with those in the high-resolution reanalysis GLORYS12v1, the ensemble reanalysis GREP, and the objective analysis EN4. The calculation algorithm for steric, thermosteric, and halosteric sea levels is similar to that of Storto et al. (2017), and the global average includes the

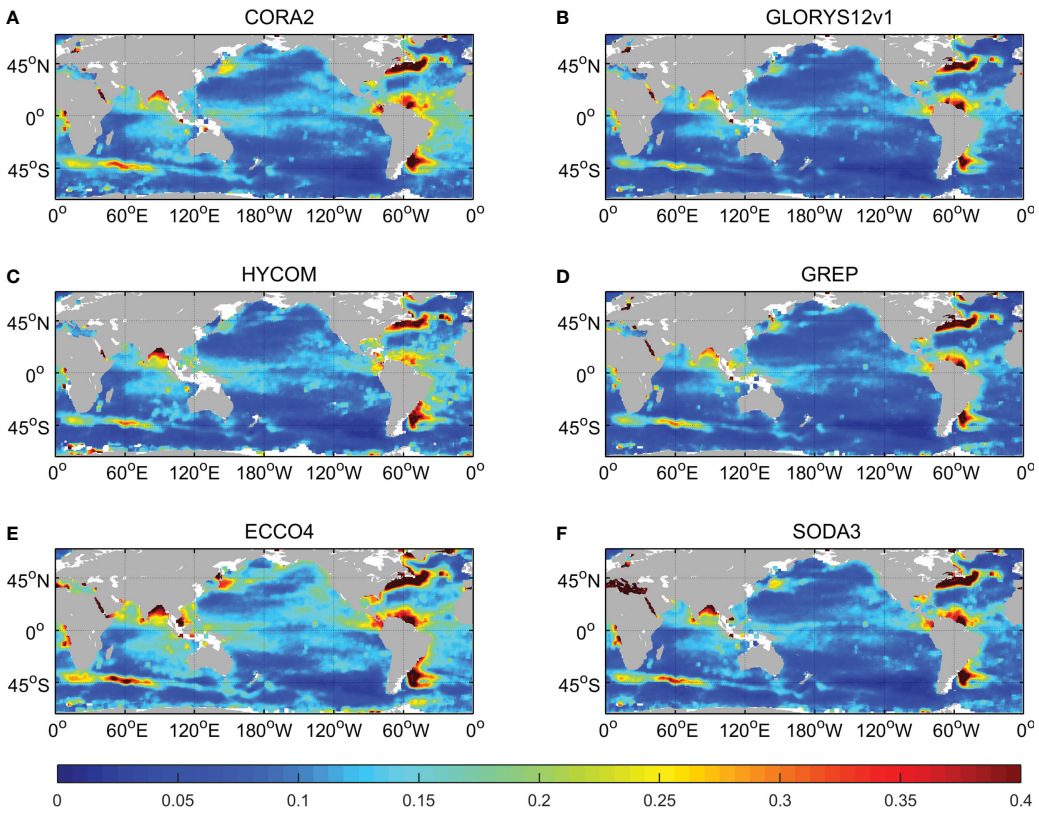


FIGURE 9
Spatial distributions of monthly salinity root-mean-square error (RMSE) (units: psu) of the six reanalyses [CORA2 (China Ocean ReAnalysis version 2), **(A)**; GLORYS12v1 (Global Ocean reanalysis and Simulation), **(B)**; HYCOM (HYbrid Coordinate Ocean Model), **(C)**; GREP (Global ocean Reanalysis Ensemble Product), **(D)**; ECCO4 (Estimating the Circulation and Climate of the Ocean, version 4), **(E)** SODA3 (Simple Ocean Data Assimilation, version 3), **(F)**] with respect to Argo profiles over 0–2,000 m. RMSE is computed between 2004 and 2017 for all reanalyses, except for HYCOM (2004–2012).

upper 1,000-m layer between 60°S and 60°N. Figure 13 shows that the four products satisfactorily capture the global steric sea-level seasonality, while they show large discrepancies in inter-annual variabilities and trends. The time series of CORA2 and GLORYS12v1, which have eddying-resolving resolutions, show more complex structures than those of EN4 and GREP, which have only eddying-permitting resolutions. The steric sea-level trends of EN4, CORA2, GLORYS12v1, and GREP during the period 2009–2019 are 0.80, 1.16, 1.48, and 1.05 mm/year, respectively. Storto et al. (2017) estimated a global steric sea-level trend at full depth during 1993–2010 based on the reanalysis ensemble mean, with the value of 1.02 ± 0.05 mm/year. Compared with Storto et al. (2017), our vertical integration depth is shallower and the time period is more recent, but the total linear trend of the ensemble reanalysis GREP is still very close to their results. The linear trend of CORA2 is closer to that of GREP relative to the other products, meaning a good skill score in terms of reproducing steric sea-level change. The trend of the objective analysis EN4 is smaller than that of ensemble reanalysis GREP; Storto et al. (2017) also suggested that its predecessor EN3 has a smaller linear trend relative to the reanalysis ensemble mean. For the four products, the thermosteric component dominates the change in steric sea-level trend, which is consistent with previous estimates (Storto et al., 2017; Zuo et al., 2017), and their trends range from 1.01 to 1.71 mm/year. The halosteric components exhibit negative trends for the four products, ranging from –0.22 to –0.03 mm/year. The thermosteric and halosteric component trends of CORA2 and GREP are also the closest among the four products.

5.5 Ocean heat content

To assess the capability of the products to describe climate variability, we selected four key areas that are related to the climate indexes for the Indian Ocean Dipole (IOD), El Niño–Southern Oscillation (ENSO), Pacific Decadal Oscillation (PDO), and Atlantic Meridional Overturning Circulation (AMOC): the IOD Western (W) area (western equatorial Indian Ocean), NINO3 (Niño 3 Index) area, Northeast (NE) Pacific, and North (N) Atlantic. Yearly ocean heat content (OHC) anomalies in different regions over 0–300, 0–700, and 0–1,500 m depth ranges were calculated from EN4, CORA2, GLORYS12v1, and GREP (Figure 14).

Figure 14 shows that the four products present similar OHC time series at 0–300 m, including a prominent inter-annual variability in the IOD W and NINO3 areas, an inter-decadal variability in the NE Pacific, and a warming trend in the N Atlantic and global oceans. Similar to previous findings

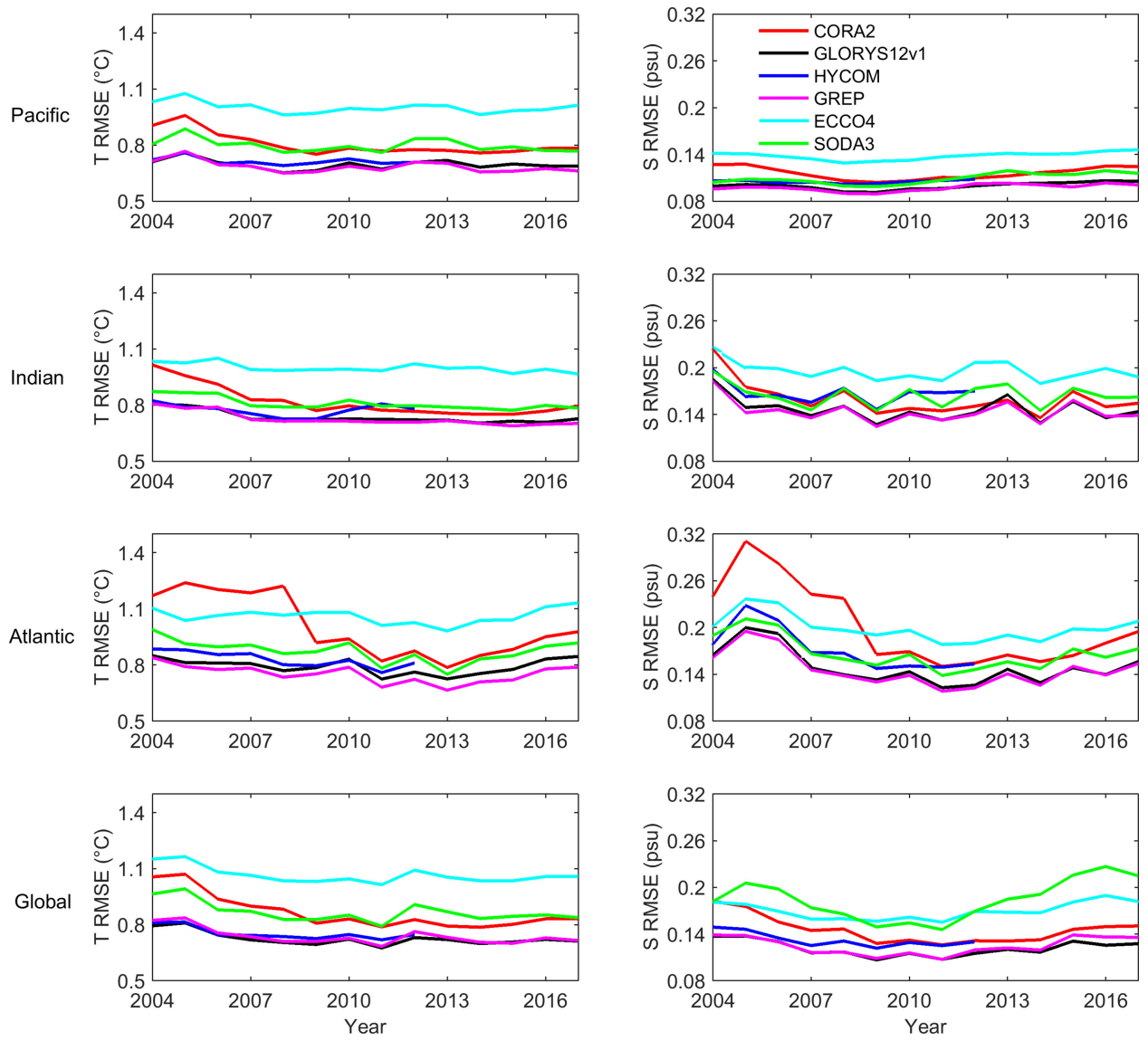


FIGURE 10
Time series of root-mean-square error (RMSE) of the six reanalyses [CORA2 (China Ocean ReAnalysis version 2), red; GLORYS12v1 (Global Ocean reanalysis and Simulation), black; HYCOM (HYbrid Coordinate Ocean Model), blue; GREP (Global ocean Reanalysis Ensemble Product), pink; ECCO4 (Estimating the Circulation and Climate of the Ocean, version 4), cyan; SODA3 (Simple Ocean Data Assimilation, version 3), green] with respect to Argo profiles for temperature (left panels; units: °C) and salinity (right panels; units: psu) over 0–2,000 m in the Pacific, Indian, Atlantic, and global oceans.

(Balmaseda et al., 2015; Palmer et al., 2017; Wang et al., 2018; Storto et al., 2019a), OHC varies on different time scales. The large interannual variabilities in the equatorial Pacific and Indian oceans are associated with the ENSO and IOD, respectively. Three peaks of OHC anomalies in the IOD W area match the IOD warm events in 2006, 2012, and 2015, respectively. Two peaks in the NINO3 area correspond to the strong EI Niño events in 2010 and 2016, respectively. In the NE Pacific, the negative OHC anomalies during 2004–2013 coincide with the cold PDO phase and the positive OHC anomalies during 2014–2019 coincide with the warm PDO phase, which matches previous studies (Palmer et al., 2017). In the N Atlantic, the four products show similar warming trends and local small-scale disturbances. The warming trend reflects a weakening subpolar gyre and a slowdown of the deep western boundary current off Labrador, which is thought to be an indicator of the slowdown of the AMOC (Zhang, 2008; Palmer et al., 2017). The curves of the global mean OHC from the four datasets broadly overlap and represent the global warming trend in the upper ocean. We also found that the warming of the North Atlantic subpolar gyre is enhanced relative to the global oceans, supporting the results of Palmer et al. (2017).

The 0–700 and 0–1,500 m OHC anomalies show similar variabilities to the 0–300 m OHC anomalies. However, the four products start to diverge when vertical integration is carried out at deeper levels. Palmer et al. (2017) suggested that the large spread in the amplitude of OHC anomalies at deeper levels may be caused by

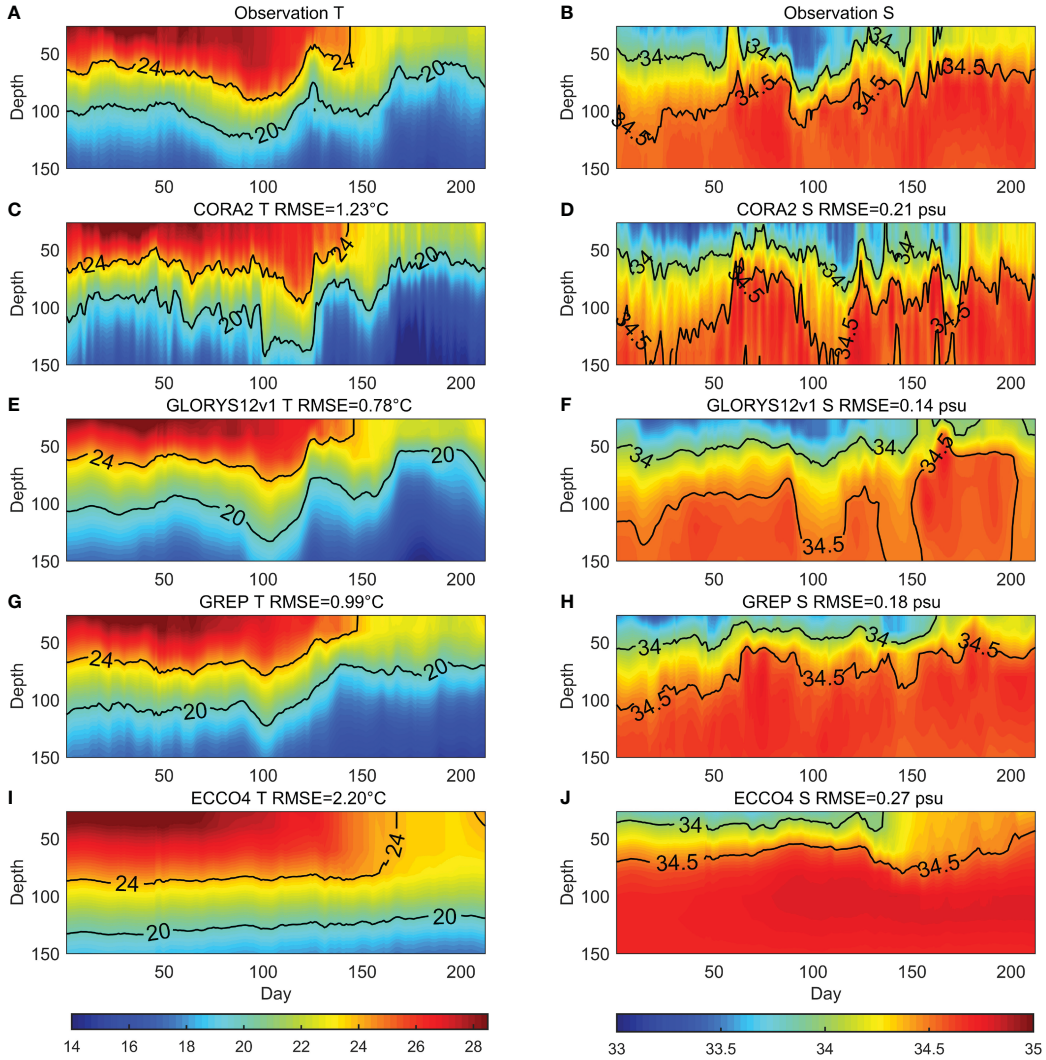


FIGURE 11
Time series of daily temperature (left; units: °C) and salinity (right; units: psu) profiles at a station (117.5°E, 19.0°N) from 1 August 2014 to 28 February 2015 for observations **(A, B)**, CORA2 (China Ocean ReAnalysis version 2) **(C, D)**, GLORYS12v1 (Global Ocean reanalysis and Simulation) **(E, F)**, GREP (Global ocean Reanalysis Ensemble Product) **(G, H)**, and ECCO4 (Estimating the Circulation and Climate of the Ocean, version 4) **(I, J)**. Root-mean-square errors (RMSEs) of temperature and salinity for the four reanalyses are also given.

the lack of observational data. There are more small-scale disturbances in the rising trends of global OHC anomalies for CORA2 and GLORYS12v1 than for EN4 and GREP. Similar results were presented by Storto et al. (2019a), who showed that a regional product with high resolution can exhibit finer-scale structures than a global ensemble mean. The abnormity of the global 0–1,500 m OHC anomalies of CORA2 around 2004 should be treated with caution, and the possible cause requires further analysis. In the N Atlantic, the large deviation of the 0–1,500 m OHC anomalies of CORA2 from the other products before 2009 is consistent with the large RMSEs of temperature and salinity discussed in Section 5.2.

For the global OHC in the upper 1,500 m, the long-term trends over 2004–2019 vary from 0.97 (GREP), 1.03 (EN4), 1.31 (CORA2), to 1.54 × 10^{23} J/decade (GLORYS12v1); most of them are larger than the results of Wang et al. (2018) based on objective analyses during 1998–2012, which vary from 0.81 to 1.0 × 10^{23} J/decade. In addition, the trends of the products with a high resolution (CORA2 and GLORYS12v1) are larger than those with a low resolution (EN4 and GREP).

6 Summary and discussion

We described the China Ocean ReAnalysis version 2 (CORA2), presenting an inter-comparison with its predecessor CORA1 and other popular reanalysis products in terms of observed variables and some climate variabilities. CORA2 is based on the eddy-resolving MITgcm, including interactive sea ice in the high latitudes and tidal forcing. The *in-situ* T–S profiles, daily gridded satellite SLA and SST are assimilated by a high-resolution multi-scale data assimilation method. The surface tidal elevation from TPXO8 is assimilated by the nudging method. The daily satellite SLA can adjust meso-scale eddies, while the TPXO8 data can improve the accuracy of surface and subsurface tidal signals (Fu

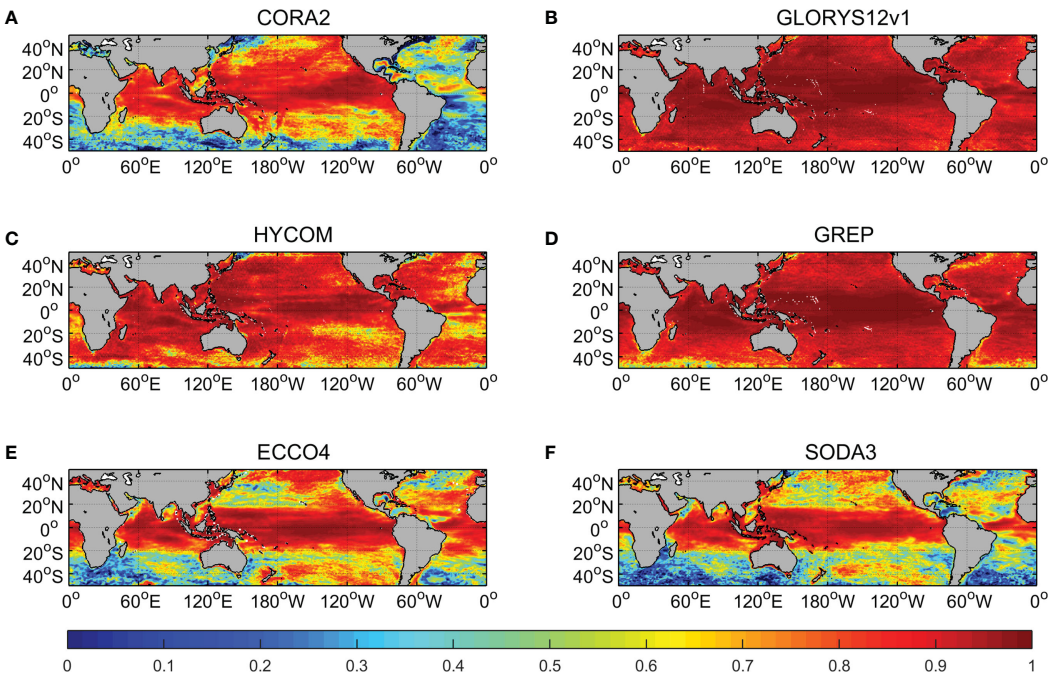


FIGURE 12
Spatial distributions of the temporal correlation coefficient between reanalyses [CORA2 (China Ocean ReAnalysis version 2) **(A)**, GLORYS12v1 (Global Ocean reanalysis and Simulation) **(B)**, HYCOM (HYbrid Coordinate Ocean Model) **(C)**, GREP (Global ocean Reanalysis Ensemble Product) **(D)**, ECCO4 (Estimating the Circulation and Climate of the Ocean, version 4) **(E)**, and SODA3 (Simple Ocean Data Assimilation, version 3) **(F)**] and AVISO altimeter data from Copernicus Marine Environment Monitoring (CMEMS). Statistics are computed using monthly mean sea level during 2004–2017 for all reanalyses, except for HYCOM (2004–2012).

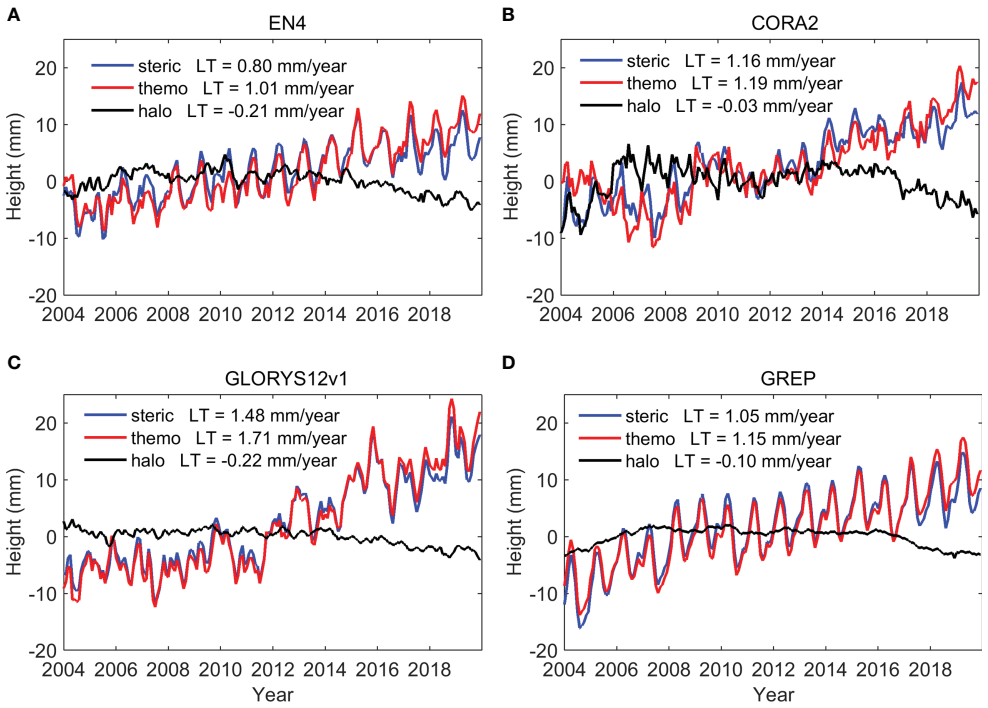


FIGURE 13
Monthly time series of global steric (blue), thermosteric (red), and halosteric (black) sea level (units: mm) for EN4, **(A)** CORA2 (China Ocean ReAnalysis version 2) **(B)**, GLORYS12v1 (Global Ocean reanalysis and Simulation) **(C)**, and GREP (Global ocean Reanalysis Ensemble Product) **(D)** during the period of 2004–2019. LT: linear trend during 2009–2019.

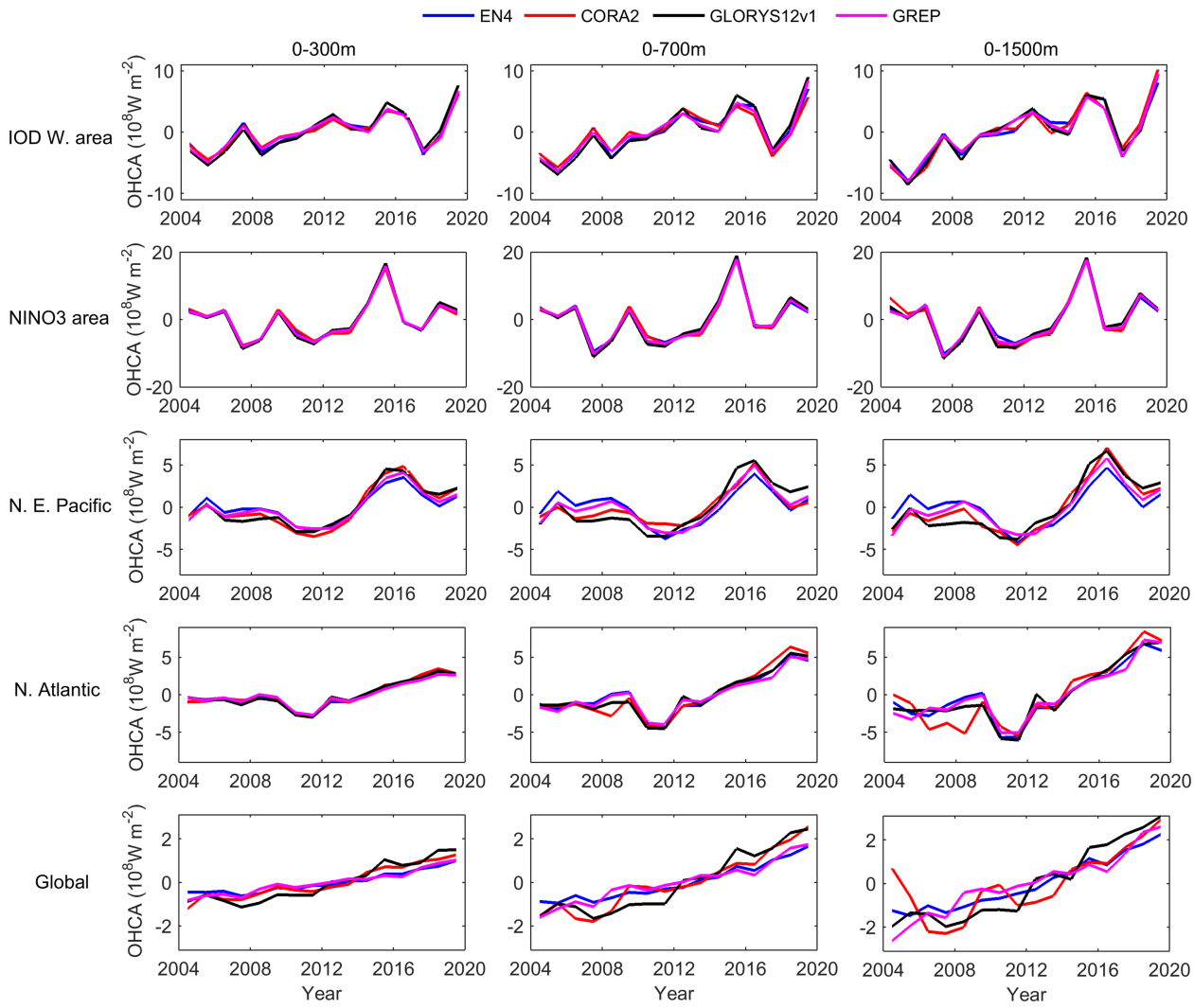


FIGURE 14
Time series of ocean heat content (OHC anomalies (10^8 W m^{-2}) in the Indian Ocean Dipole (IOD) West area (50–70°E, 10°S–10°N), NINO3 area (150–90°W, 5°S–5°N), Northeast Pacific (160–120°W, 20–50°N), North Atlantic (80°W–0°, 20–50°N), and global oceans for EN4 (blue), CORA2 (China Ocean ReAnalysis version 2) (red), GLORYS12v1 (Global Ocean reanalysis and Simulation) (black), and GREP (Global ocean Reanalysis Ensemble Product) (pink) at depths of 0–300 m (left panels), 0–700 m (middle panels), and 0–1,500 m (right panels). OHC anomaly is expressed as the equivalent heating rate in W m^{-2}, relative to the region's ocean surface area.

et al., 2021). The improvement of CORA2 relative to CORA1 and how CORA2 compares with other selected products is presented by analyzing reanalysis misfits to independent and non-independent observations and by comparing the variability of EKE, steric sea level, and OHC. The evaluation results show the advantages and disadvantages of the ocean reanalysis CORA2.

Compared with CORA1, the surface and subsurface T–S errors of CORA2 with respect to non-independent observations are significantly reduced owing to the enhanced resolution, the updated SST assimilation scheme, the use of an FGAT assimilation scheme, and the inclusion of tidal forcing and assimilation. The EKE of CORA2 sharply increases compared with that of CORA1 and is consistent with that of GLORYS12v1, demonstrating that high-resolution reanalyses have a higher EKE than low-resolution ones.

The comparison between the six reanalyses and OSTIA SST reveals that the SST accuracy of CORA2 is more similar to that of GREP and with a smaller error than the other reanalyses since 2009. It is speculated that the high accuracy of CORA2 largely stems from the 1-day assimilation cycle of SST.

Compared with the non-independent Argo profiles, the T–S RMSE of CORA2 is similar to that of SODA3, lower than that of ECCO4, and higher than those of GLORYS12v1 and GREP in most oceans. Although the T–S RMSE of CORA2 is slightly larger in the Atlantic Ocean, it was reduced after we fixed the bug. The CORA2 bias is close to zero, while the other products have some biases, especially for the temperature field. For the variability of subsurface T–S, ECCO4 exhibits a poor performance while GREP shows good seasonal variation; GLORYS12v1 and CORA2 can not only describe the seasonal features but also some sub-monthly-scale fluctuations,

owing to their high resolutions. There are several shocks and spurious waves in the time evolution of CORA2, which may be caused by adding the analysis increment to the background state in a single time step and/or including tidal signals. The large RMSE of ECCO4 may be attributed to the fact that its assimilation method of 4D-Var tends to maintain dynamic consistency.

The SLA of CORA2 is significantly correlated with the altimetry in the Indian and Pacific oceans, but the correlation in the Atlantic is weak. This may be associated with the fact that the SLA assimilation method depends on the background temperature and salinity fields, and their errors in the Atlantic are larger than those in the other two regions. The poor performance of SODA3 is attributed to the absence of altimeter SLA constraints. The seasonal variability of global steric sea level and the proportion of thermosteric and halosteric components can be well described in all the products, while the time series of high-resolution products (CORA2 and GLORYS12v1) show more complex structures than those of low-resolution products (EN4 and GREP). The linear trend of global steric sea level of CORA2 is closer to that of ensemble reanalysis GREP, smaller than that of GLORYS12v1, and larger than that of EN4.

The time series of global OHC anomalies of EN4, CORA2, GLORYS12v1, and GREP show the best agreement, representing the climate variability related to the IOD, ENSO, PDO, and AMOC indices, as well as the global warming trend. The accuracy of the products in representing climate variability gradually decreases with increased depth, owing to the lack of observation constraints. In the Atlantic, the large T–S errors of CORA2 cause the 0–1,500 m OHC to deviate from the average value over 2004–2008, but do not change the overall variation characteristics. The long-term trends of global OHC in the high-resolution reanalyses (CORA2 and GLORYS12v1) are larger than those in the low-resolution reanalyses (EN4 and GREP) and they are also larger than the result of Wang et al. (2018).

It should be noted that a good reanalysis product should use a fixed ocean model and data assimilation scheme with the best available parameterizations, observations, and meteorological forcing, which do not change during the production. However, it is not possible to rerun the entire CORA2 reanalysis according to best practices owing to computational cost, leading to some changes in the CORA2 system during the long integration process: for example, the optimization of the assimilation scheme in 2009 and the change of atmospheric forcing data in 2014. In particular, the optimization of the assimilation scheme has brought obvious improvements to the accuracy of CORA2 in the Atlantic Ocean since 2009.

CORA2 is a complex system, resulting from extensive efforts to combine information and developments from observations, assimilation, and modeling communities. Given the strengths and weaknesses of CORA2 discussed in this paper, key improvements to CORA2 in the future should include the assimilation of the best available observation data (e.g., satellite OSTIA SST, SSS, and sea ice concentration data), the use of the IAU procedure, the improvement of the SLA assimilation method. At the same time, it is necessary to further assess the sea ice, currents, and tides of CORA2 to meet the needs of users in different fields.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

HF integrated the CORA2 reanalysis system, made the products, participated in the validation and inter-comparison, and wrote the original manuscript. BD optimized the reanalysis system and standardized the outputs of the reanalysis products. ZG processed the tide gauge observations and participated in the validation. XW was in charge of the reanalysis system, debugged the assimilation code, performed the assimilation experiments, and modified the manuscript. GC processed the in-situ T–S profiles. LZ participated in the parameterization of the numerical model. YZ participated in the inter-comparison and validation. KL engaged in tide analysis and assimilation. XZ focused on the validation. WL designed the assimilation scheme and coded the original assimilation code. All authors contributed to the article and approved the submitted version.

Acknowledgments

The authors thank the three reviewers GR, SM and PW for their careful reading and for providing very constructive comments that improved the paper. This research was jointly supported by the National Key R&D Program of China (2019YFC1510002, 2021YFC3101603 and 2016YFC1401800) and National Natural Science Foundation of China (41976019).

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmars.2023.1084186/full#supplementary-material

SUPPLEMENTARY TABLE 1
List of five ocean reanalyses used in this study and their main characteristics.

SUPPLEMENTARY TABLE 2
Biases of monthly reanalyses [CORA2 (China Ocean ReAnalysis version 2), GLORYS12v1 (Global Ocean reanalysis and Simulation), HYCOM (HYbrid Coordinate Ocean Model), GREP (Global ocean Reanalysis Ensemble Product), ECCO4 (Estimating the Circulation and Climate of the Ocean, version 4), SODA3 (Simple Ocean Data Assimilation, version 3)] temperature (T; units: °C) and salinity (S; units: psu) against Argo profiles in the equatorial Indian Ocean (40–100°E, 10°S–10°N), equatorial Pacific (130°E–80°W, 10°S–10°N), equatorial Atlantic (50°W–0°, 10°S–10°N), North Indian Ocean (40–100°E, 10–30°N), Northwest Pacific (120–180°E, 12–50°N), Northeast Pacific (180°E–90°W, 12–50°N), North Atlantic (80°W–0°, 12–50°N), South Indian Ocean (40–120°E, 30–10°S), South Pacific (150°E–80°W, 3–10°S), South Atlantic (50°W–0°, 30–10°S), Southern Ocean (180°E–180°W, 60–30°S), and global oceans Note that statistics are averaged over the period 2004–2012 for HYCOM and the period 2004–2017 for other reanalyses.

SUPPLEMENTARY FIGURE 1
Ratios of standard deviation of monthly sea-surface temperature (SST) of CORA2 (China Ocean ReAnalysis version 2) **(A)**, GLORYS12v1 (Global Ocean reanalysis and Simulation) **(B)**, HYCOM (HYbrid Coordinate Ocean Model) **(C)**, GREP (Global ocean Reanalysis Ensemble Product) **(D)**, ECCO4 (Estimating the Circulation and Climate of the Ocean, version 4) **(E)**, and SODA3 (Simple Ocean Data Assimilation, version 3) **(F)** relative to that of OSTIA SST during 2004–2017 (2004–2012 for HYCOM). The SST variances of CORA2 and GREP are the closest to the observation, and the ratios remain around one in most regions. The SST variance of GLORYS12v1 in the Antarctic Circumpolar Current (ACC) region is larger than the observation, and the ratio can reach >1.2. The HYCOM result is the most different from those of the other products, which may be related to its short time period.

SUPPLEMENTARY FIGURE 2
Vertical distribution of root-mean-square errors (RMSEs) **(A)** and biases **(B)** of monthly salinity (units: psu) of SODA3 with respect to Argo profiles in the global oceans, including (red) and excluding (black) the Mediterranean Sea.

References

Adcroft, A., and Campin, J. (2004). Rescaled height coordinates for accurate representation of free-surface flows in ocean circulation models. *Ocean Model.* 7, 269–284. doi: 10.1016/j.ocemod.2003.09.003

Arbic, B., Garner, S., Hallberg, R., and Simmons, H. (2004). The accuracy of surface elevations in forward global barotropic and baroclinic tide models. *Deep-Sea Res. II.* 51, 3069–3101. doi: 10.1016/j.dsr2.2004.09.014

Balmaseda, M., Hernandez, F., Storto, A., Palmer, M., Alves, O., Shi, L., et al. (2015). The ocean reanalyses intercomparison project (ORA-IP). *J. Oper. Oceanogr.* 8 (sup1), s80–s97. doi: 10.1080/1755876X.2015.1022329

Balmaseda, M., Mogensen, K., and Weaver, A. (2012). Evaluation of the ECMWF ocean reanalysis system ORAS4. *Q. J. R. Meteorol. Soc.* 139, 1132–1161. doi: 10.1002/qj.2063

Blockley, E., Martin, M., McLaren, A., Ryan, A., Waters, J., Lea, D., et al. (2014). Recent development of the met office operational ocean forecasting system: an overview and assessment of the new global FOAM forecasts. *Geosci. Model. Dev.* 7 (6), 2613–2638. doi: 10.5194/gmd-7-2613-2014

Bloom, S., Takacs, L., Da Silva, A., and Ledvina, D. (1996). Data assimilation using incremental analysis updates. *Mon. Weather Rev.* 124, 1256–1271. doi: 10.1175/1520-0493(1996)124<1256:DAUIAU>2.0.CO;2

Carton, J., Chepurin, G., and Chen, L. (2018). SODA3: a new ocean climate reanalysis. *J. Climate.* 31 (17), 6967–6983. doi: 10.1175/JCLI-D-18-0149.1

Carton, J., Stephen, G., and Kalnay, E. (2019). Temperature and salinity variability in the SODA3, ECCO4r3, and ORAS5 ocean reanalyses 1993-2015. *J. Climate.* 32, 2277–2293. doi: 10.1175/JCLI-D-18-0605.1

Chevallier, M., Smith, G., Dupont, F., Lemieux, J., Forget, G., Fujii, Y., et al. (2017). Intercomparison of the Arctic sea ice cover in global ocean–sea ice reanalyses from the ORA-IP project. *Clim. Dyn.* 49, 1107–1136. doi: 10.1007/s00382-016-2985-y

Cooper, M., and Haines, K. (1996). Altimetric assimilation with water property conservation. *J. Geophys. Res.* 24, 1059–1077. doi: 10.1029/95JC02902

Cummings, J., and Smedstad, O. (2013). Variational data assimilation for the global ocean. data assimilation for atmospheric. *Oceanic Hydrologic Appl.* 2 (13), 303–343. doi: 10.1007/978-3-642-35088-7_13

Egbert, G. D., and Erofeeva, S. Y. (2002). Efficient inverse modeling of barotropic ocean tides. *J. Atmos. Ocean Tech.* 19, 183–204. doi: 10.1175/1520-0426(2002)019<0183:EIMOBO>2.0.CO;2

Fekete, B., Vörösmarty, C., and Grabs, W. (2002). High-resolution fields of global runoff combining observed river discharge and simulated water balances. *Global Biogeochem. Cy.* 16, 15.1–15.10. doi: 10.1029/1999GB001254

Forget, G., Campin, J., Heimbach, P., Hill, C., Ponte, R., and Wunsch, C. (2015). ECCO version 4: an integrated framework for non-linear inverse modeling and global ocean state estimation. *Geosci. Model. Dev.* 8 (10), 3071–3104. doi: 10.5194/gmd-8-3071-2015

Fu, H., Wu, X., Li, W., Zhang, L., Liu, K., and Dan, B. (2021). Improving the accuracy of barotropic and internal tides embedded in a high-resolution global ocean circulation model of MITgcm. *Ocean Model.* 162, 101809. doi: 10.1016/j.ocemod.2021.101809

Garcia, H., Boyer, T., Locarnini, R., Baranova, O., and Zweng, M. (2018). *World ocean database 2018: user's manual (prerelease)*. Ed. A. Mishonov (Silver Spring, MD: NOAA). Available at: https://www.NCEI.noaa.gov/OC5/WOD/pr_wod.html.

Good, S., Fiedler, E., Mao, C., Martin, M. J., Maycock, A., Reid, R., et al. (2020). The current configuration of the OSTIA system for operational production of foundation Sea surface temperature and ice concentration analyses. *Remote Sens.* 12, 720. doi: 10.3390/rs12040720

Good, S., Martin, M., and Rayner, N. (2013). EN4: quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. *J. Geophys. Res.* 118, 6704–6716. doi: 10.1002/2013JC009067

Griffies, S., and Hallberg, R. (2000). Bibarmonic friction with a smagorinsky-like viscosity for use in Large-scale eddy-permitting ocean models. *Monthly Weather Rev.* 128, 2935–2946. doi: 10.1175/1520-0493(2000)128<2935:BFWASL>2.0.CO;2

Han, G., Fu, H., Zhang, X., Li, W., Wu, X., Wang, X., et al. (2013b). A global ocean reanalysis products in the China ocean reanalysis (CORA) project. *Adv. Atmos. Sci.* 30 (6), 1621–1631. doi: 10.1007/s00376-013-2198-9

Han, G., Li, W., Zhang, X., Li, D., He, Z., Wang, X., et al. (2011). A regional ocean reanalysis system for coastal waters of China and adjacent seas. *Adv. Atmos. Sci.* 28 (3), 682–690. doi: 10.1007/s00376-010-9184-2

Han, G., Li, W., Zhang, X., Wang, X., Wu, X., Fu, H., et al. (2013a). A new version of regional ocean reanalysis for coastal waters of China and adjacent seas. *Adv. Atmos. Sci.* 30 (4), 974–982. doi: 10.1007/s00376-012-2195-4

Karspeck, A., Stammer, D., Köhl, A., Danabasoglu, G., Balmaseda, M., Smith, D., et al. (2017). Comparison of the Atlantic meridional overturning circulation between 1960 and 2007 in six ocean reanalysis products. *Clim. Dyn.* 49, 957–982. doi: 10.1007/s00382-015-2787-7

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., et al. (2015). The JRA-55 reanalysis: general specifications and basic characteristics. *J. Meteor. Soc. Japan.* 93, 5–48. doi: 10.2151/jmsj.2015-001

Large, W., McWilliams, J., and Doney, S. (1994). Oceanic vertical mixing: a review and a model with a nonlocal boundary layer parameterization. *Rev. Geophys.* 32, 363–403. doi: 10.1029/94RG01872

Large, W., and Pond, S. (1981). Open ocean momentum flux measurements in moderate to strong winds. *J. Phys. Oceanogr.* 11 (3), 324–336. doi: 10.1175/1520-0485(1981)011<0324:OOMFMI>2.0.CO;2

Large, W., and Pond, S. (1982). Sensible and latent heat flux measurements over the ocean. *J. Phys. Oceanogr.* 12 (5), 464–482. doi: 10.1175/1520-0485(1982)012<0464:SALHFM>2.0.CO;2

Lellouche, J., Greiner, E., Bourdalle-Badié, R., Garric, G., Melet, A., Drévillon, M., et al. (2021). The Copernicus global 1/12° oceanic and Sea ice GLORYS12 reanalysis. *Front. Earth Sci.* 9. doi: 10.3389/feart.2021.698876

Li, W., Xie, Y., He, Z., Han, G., Liu, K., Ma, J., et al. (2008). Application of the multi-grid data assimilation scheme to the China seas' temper-ature forecast. *J. Atmos. Oceanic Technol.* 25 (11), 2106–2116. doi: 10.1175/2008JTECHO510.1

Marshall, J., Adcroft, A., Hill, C., Perelman, L., and Heisey, C. (1997). A finite volume, incompressible navier–stokes model for studies of the ocean on parallel computers. *J. Geophys. Res.* 102, 5753–5766. doi: 10.1029/96JC02775

Masina, S., and Storto, A. (2017). Reconstructing the recent past ocean variability: status and perspective. *J. Mar. Res.* 75, 727–764. doi: 10.1357/002224017823523973

Masina, S., Storto, A., Ferry, N., Valdivieso, M., Haines, K., Balmaseda, M., et al. (2017). An ensemble of eddy-permitting global ocean reanalyses from the MyOcean project. *Clim. Dyn.* 49, 813–841. doi: 10.1007/s00382-015-2728-5

Merchant, C. J., Embury, O., Bulgin, C. E., Block, T., Corlett, G., Fiedler, E., et al. (2019). Satellite-based time-series of sea-surface temperature since 1981 for climate applications. *Sci. Data.* 6 (223), 2052–4463. doi: 10.1038/s41597-019-0236-x

Onogi, K., Tsutsui, J., Koide, H., Sakamoto, M., Kobayashi, S., Hatsushika, H., et al. (2007). The JRA-25 reanalysis. *J. Met. Soc. Jap.* 85 (3), 369–432. doi: 10.2151/jmsj.85.369

Palmer, M., Roberts, C., Balmaseda, M., Chang, Y., Chepurin, G., Ferry, N., et al. (2017). Ocean heat content variability and change in an ensemble of ocean reanalyses. *Clim. Dyn.* 49, 909–930. doi: 10.1007/s00382-015-2801-0

Parkinson, C., and Washington, W. (1979). A large-scale numerical model of sea ice. *J. Geophys. Res.* 84, 311–337. doi: 10.1029/JC084iC01p00311

Pujol, M., Faugère, Y., Taburet, G., Dupuy, S., Pelloquin, C., Ablain, M., et al. (2016). DUACS DT2014: the new multi-mission altimeter data set reprocessed over 20 years. *Ocean Sci.* 12, 1067–1090. doi: 10.5194/os-12-1067-2016

Reynolds, R., Smith, T., Liu, C., Chelton, D., Casey, K., and Schlax, M. (2007). Daily high-resolution-blended analyses for sea surface temperature. *J. Cli.* 20 (22), 5473–5496. doi: 10.1175/2007JCLI1824.1

Storto, A., Alvera-Azcárate, A., Balmaseda, M., Barth, A., Chevallier, M., Counillon, F., et al. (2019b). Ocean reanalyses: recent advances and unsolved challenges. *Front. Mar. Sci.* 6. doi: 10.3389/fmars.2019.00418

Storto, A., Masina, S., Balmaseda, M., Guinehut, S., Xue, Y., Szekely, T., et al. (2017). Steric sea level variability, (1993–2010) in an ensemble of ocean reanalyses and objective analyses. *Clim. Dyn.* 49, 709–729. doi: 10.1007/s00382-015-2554-9

Storto, A., Masina, S., and Dobricic, S. (2014). Estimation and impact of nonuniform horizontal correlation length scales for global ocean physical analyses. *J. Atmos. Ocean Technol.* 31, 2330–2349. doi: 10.1175/JTECH-D-14-00042.1

Storto, A., Masina, S., and Navarra, A. (2016). Evaluation of the CMCC eddy permitting global ocean physical reanalysis system (C-GLORS 1982–2012) and its assimilation components. *Q. J. R. Meteorol. Soc.* 142, 738–758. doi: 10.1002/qj.2673

Storto, A., Masina, S., Simoncelli, S., Iovino, D., Cipollone, A., Drevillon, M., et al. (2019a). The added value of the multi-system spread information for ocean heat content and steric sea level investigations in the CMEMS GREP ensemble reanalysis product. *Clim. Dyn.* 53, 287–312. doi: 10.1007/s00382-018-4585-5

Toyoda, T., Fujii, Y., Kuragano, T., Kamachi, M., Ishikawa, Y., Masuda, S., et al. (2017). Intercomparison and validation of the mixed layer depth fields of global ocean syntheses. *Clim. Dyn.* 49, 753–773. doi: 10.1007/s00382-015-2637-7

Toyoda, T., Fujii, Y., Yasuda, T., Usui, N., Ogawa, K., Kuragano, T., et al. (2016). Data assimilation of sea ice concentration into a global ocean-sea ice model with corrections for atmospheric forcing and ocean temperature fields. *J. Oceanogr.* 72 (2), 235–262. doi: 10.1007/s10872-015-0326-0

Troccoli, A., Balmaseda, M., Segschneider, J., Vialard, J., Anderson, D., Haines, K., et al. (2002). Salinity adjustments in the presence of temperature data assimilation. *Mon. Wea Rev.* 130, 89–102. doi: 10.1175/1520-0493(2002)130<0089:SAITPO>2.0.CO;2

Uotila, P., Goosse, H., Haines, K., Matthieu, C., Antoine, B., Clément, B., et al. (2018). An assessment of ten ocean reanalyses in the polar regions. *Clim. Dyn.* 52, 1613–1650. doi: 10.1007/s00382-018-4242-z

Valdivieso, M., Haines, K., Balmaseda, M., Chang, Y., Drevillon, M., Ferry, N., et al. (2017). An assessment of air–sea heat fluxes from ocean and coupled reanalyses. *Clim. Dyn.* 49, 983–1008. doi: 10.1007/s00382-015-2843-3

Wang, G., Cheng, L., Abraham, J., and Li, C. (2018). Consensuses and discrepancies of basin-scale ocean heat content changes in different ocean analyses. *Clim. Dyn.* 50, 2471–2487. doi: 10.1007/s00382-017-3751-5

Zhang, R. (2008). Coherent surface-subsurface fingerprint of the Atlantic meridional overturning circulation. *Geophys. Res. Lett.* 35, L20705. doi: 10.1029/2008GL035463

Zhang, H., Chen, D., Zhou, L., Liu, X., Ding, T., and Zhou, B. (2016). Upper ocean response to typhoon kalmaegi, (2014). *J. Geophys. Res. Oceans.* 121, 6520–6535. doi: 10.1002/2016JC012064

Zhang, J., Hibler, W., Steele, M., and Rothrock, D. (1998). Arctic Ice-ocean modeling with and without climate restoring. *J. Phys. Oceanogr.* 28, 191–217. doi: 10.1175/1520-0485(1998)028<0191:AIOMWA>2.0.CO;2

Zuo, H., Balmaseda, M., and Mogensen, K. (2017). The new eddy-permitting ORAP5 ocean reanalysis: description, evaluation and uncertainties in climate signals. *Clim. Dyn.* 49, 791–811. doi: 10.1007/s00382-015-2675-1

Zuo, H., Balmaseda, M., Tietsche, S., Mogensen, K., and Mayer, M. (2019). The ECMWF operational ensemble reanalysis-analysis system for ocean and sea ice: a description of the system and assessment. *Ocean Sci.* 15 (3), 779–808. doi: 10.5194/os-15-779-2019

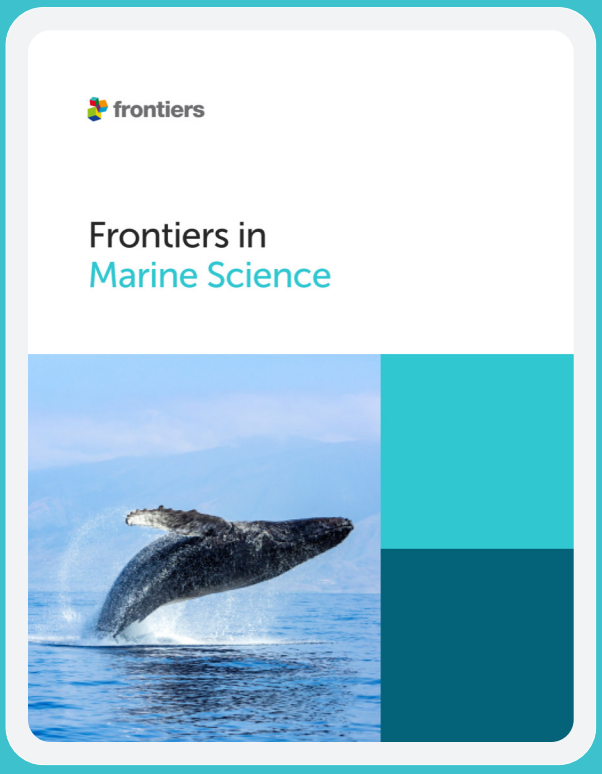
Frontiers in Marine Science

Explores ocean-based solutions for emerging global challenges

The third most-cited marine and freshwater biology journal, advancing our understanding of marine systems and addressing global challenges including overfishing, pollution, and climate change.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact