



## OPEN ACCESS

## EDITED BY

Parvez Rana,  
Natural Resources Institute Finland (Luke),  
Finland

## REVIEWED BY

Mcebisi Qabaqaba,  
University of Pretoria, South Africa  
Poonam Seth Tiwari,  
Indian Space Research Organisation, India

## \*CORRESPONDENCE

Abid Nazir,  
✉ gee2018abid@gmail.com

RECEIVED 15 October 2025

REVISED 03 January 2026

ACCEPTED 07 January 2026

PUBLISHED 28 January 2026

## CITATION

Nazir A, Hanan NP, Gilani H and Shrestha HL  
(2026) Species level mapping of forest canopy  
height in Nepal using GEDI with Sentinel-1  
and Sentinel-2.  
*Front. Remote Sens.* 7:1725509.  
doi: 10.3389/frsen.2026.1725509

## COPYRIGHT

© 2026 Nazir, Hanan, Gilani and Shrestha. 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.

# Species level mapping of forest canopy height in Nepal using GEDI with Sentinel-1 and Sentinel-2

Abid Nazir<sup>1\*</sup>, Niall P. Hanan<sup>1</sup>, Hammad Gilani<sup>2</sup> and  
Him Lal Shrestha<sup>3</sup>

<sup>1</sup>New Mexico State University, Plant and Environmental Sciences, Las Cruces, NM, United States,

<sup>2</sup>National University of Sciences and Technology, School of Interdisciplinary Engineering and Sciences,  
Islamabad, Pakistan, <sup>3</sup>Department of Environmental Science and Engineering, Kathmandu University,  
Kathmandu, Nepal

Forest canopy height mapping is critical for mapping and modeling biogeophysical and ecological factors, including forest aboveground biomass, carbon reserves, forest carbon emissions, habitat diversity, forest degradation, and restoration success. The Global Ecosystem Dynamics Investigation (GEDI) is a spaceborne Light Detection and Ranging (LiDAR) sensor designed specifically to collect data on forest ecosystems worldwide. However, the information obtained by GEDI is not wall-to-wall, requiring data fusion approaches to map spatially continuous canopy heights. This study, for the first time, presents canopy height models for the entire country of Nepal based on interpolated GEDI tree heights fusing Sentinel-2 multispectral imagery with Sentinel-1 synthetic aperture radar (SAR), creating species-specific continuous canopy height models for Nepal at 10 m resolution. Forest plot field data, collected from a nationwide campaign, provided data on species identity, which was used for species mapping and accuracy evaluation. Differences in canopy-architecture and leaf-level traits mean that species-specific models are needed to interpolate GEDI tree heights using the Sentinel optical and SAR data. The national forest height map was compared with an independent set of GEDI data (RMSE = 2.4 m,  $R^2 = 0.92$ , intercept (c) = 0.53 m and slope (m) = 0.98) and fully independent field data (RMSE = 3.7 m,  $R^2 = 0.74$ , c = 4.1 m, and m = 0.89). The developed forest type map and canopy height models have the potential to aid in both operational monitoring and hindcasting of historical forest height and its dynamics. Local and national forest management initiatives and international climate and sustainable development projects require this kind of capacity.

## KEYWORDS

ecosystem, forest, lidar, optical remote sensing, SAR, tree species, vertical structure

## 1 Introduction

Forests are critical for maintaining the Earth's biosphere, regulating temperature, and providing various ecosystem services (Bonan, 2008; Brocknerhoff et al., 2017). Growing forests and plantations are generally major sinks of carbon (Dixon et al., 1994; Pan et al., 2011; Pugh et al., 2019), while anthropogenic management and disturbance can convert forest sinks into major sources, primarily due to deforestation for agriculture and wood extraction (Houghton, 1999). Measurement of forest vertical structure is necessary for

assessing various forest processes, such as monitoring carbon emissions, sequestration, degradation, and restoration success (Goetz et al., 2015). Additionally, accurate metrics of forest canopy height and spatiotemporal variations facilitate effective forest management and policymaking (Xi et al., 2022).

Traditional forest inventories provide precise data on forest density, canopy cover, and canopy height within individual forest plots. Still, such field-based measurements are time-consuming and labor-intensive, which limits them to local-scale assessments. Remote sensing technologies, however, are being effectively utilized to monitor forests and measure carbon storage and timber volume across landscape, regional and global scales (Goetz et al., 2015), and to model other important ecosystem variables such as biodiversity and primary production (Goetz and Dubayah, 2011; Asner et al., 2012). Passive remote sensing generally lacks vertical structure information compared to active Light Detection and Ranging (LiDAR) technologies which have revolutionized our ability to measure vegetation structure (Pang et al., 2016; McManamon, 2019; Coops et al., 2021; Burns et al., 2024).

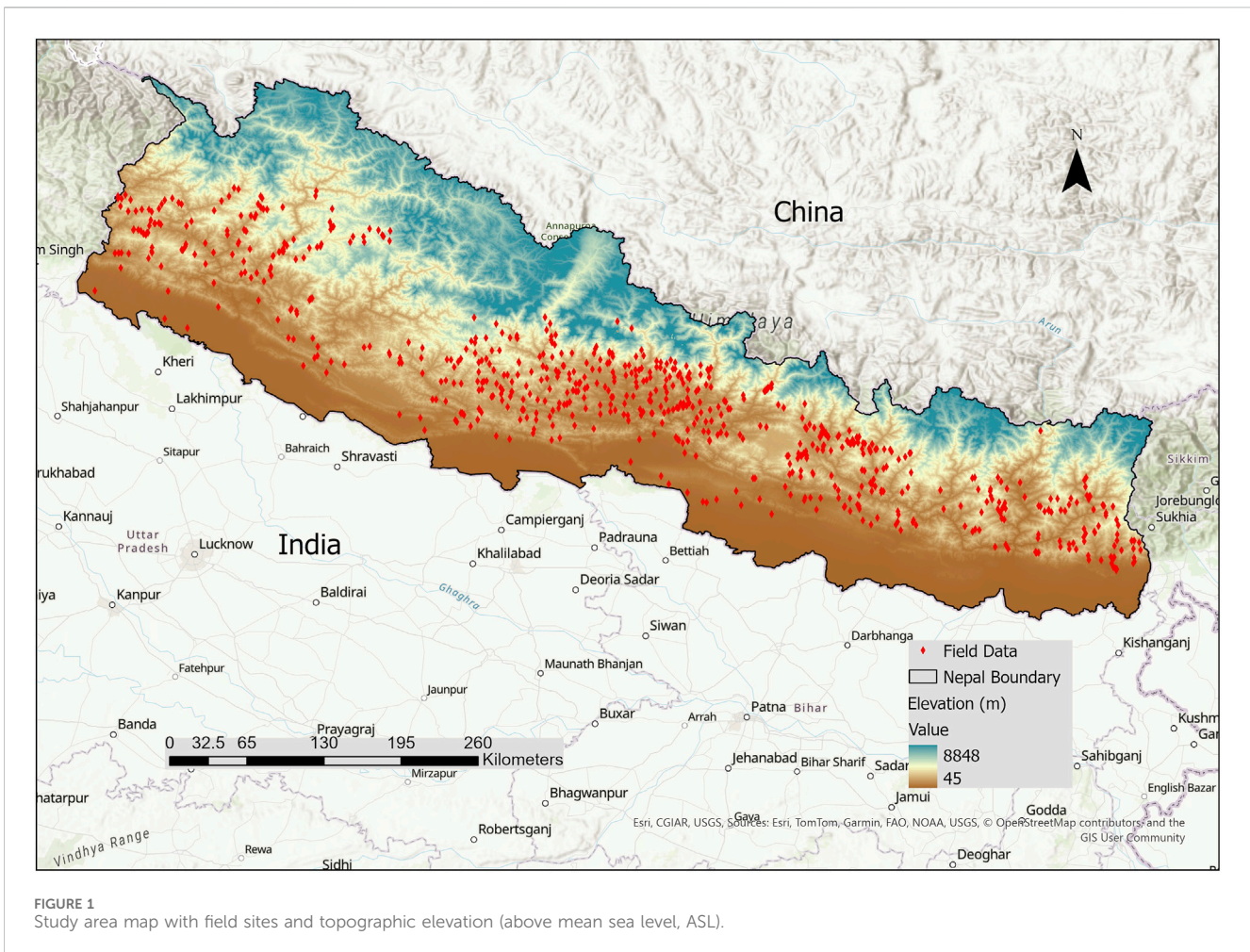
For example, Airborne Laser Scanning (ALS) and terrestrial LiDAR technologies have been effectively employed for mapping vegetation structure (Pascual et al., 2023). However, both ALS and terrestrial-based approaches remain costly and logistically impractical for large-scale applications. The NASA Geoscience Laser Altimeter System (GLAS) aboard ICESat was the first satellite mission capable of measuring global vertical structure (Lefsky, 2010), followed by the Advanced Topographic Laser Altimeter System (ATLAS) onboard ICESat-2, which further enhanced canopy height retrievals from space (Lefsky, 2010; Simard et al., 2011; Popescu et al., 2018). Despite their capabilities, these instruments were initially designed for cryosphere applications, primarily to monitor ice sheets and glaciers rather than vegetation (Montesano et al., 2015; Hansen et al., 2016). The Global Ecosystem Dynamics Investigation (GEDI) mission was launched in December 2018, becoming the first spaceborne LiDAR specifically designed for forest vegetation monitoring using full-waveform technology (Patterson et al., 2019; Dubayah et al., 2020). Mounted on the International Space Station (ISS), GEDI collects footprint-based LiDAR observations between 51.6°N and 51.6°S latitudes, constrained by the ISS's orbital path (Rishmawi et al., 2022). The mission aims to improve our understanding of forest structure, support above-ground biomass (AGB) estimation in temperate and tropical forests, inform land surface carbon budgeting, and analyze the role of vegetation structure in biodiversity (Dubayah et al., 2020; Pascual et al., 2023). GEDI generates data products derived from LiDAR waveforms, including canopy height, cover, and ultimately, estimates of above-ground biomass density (AGBD) (Shendryk, 2022).

Several limitations constrain the practical utility of standard GEDI data products. GEDI's footprint-level observations represent only a sparse sampling (approximately 4%, Dubayah et al., 2020) of the land surface, leaving large areas unobserved (Zhao et al., 2024). Consequently, GEDI sampling may not be ideal for monitoring localized or infrequent forest dynamics, such as selective logging, or for assessment of forest structure at finer spatial scales, such as in forest management units or logging concessions (Potapov et al., 2021).

However, integrating GEDI data with wall-to-wall spatiotemporal optical observations from platforms like Landsat, Sentinel-2 (S2), and Sentinel-1 (S1) facilitates the extrapolation of GEDI samples for broader and more continuous monitoring of vegetation structure and can enhance the precision of ecosystem mapping, especially in remote regions (Potapov et al., 2021; Xi et al., 2022). For instance, growing season S2 images have been used to classify tree species in several studies (Nelson, 2017; Bolyn et al., 2018; Wessel et al., 2018), and fusion with S1 SAR data can enhance accuracy (Xi et al., 2022). Wall-to-wall optical remote sensing datasets have been widely used to extend LiDAR-based sampling for forest structure and biomass estimation across diverse forest types and geographical regions (Saarela et al., 2018; Healey et al., 2020; Potapov et al., 2021). Footprint-level biomass estimates from GEDI and canopy structure measurements from GLAS have been extrapolated using MODIS data at the national (Chi et al., 2015), continental (Baccini et al., 2008), and global scales (Simard et al., 2011; Potapov et al., 2021). At local (Duncanson et al., 2010) and biome-level scales (Tyukavina et al., 2015) Landsat imagery has supported the extrapolation of GLAS-based canopy height and biomass data (Hudak et al., 2002; Hansen et al., 2016). Additionally, long-term studies of forest extent and structural change at the regional (Potapov et al., 2019) and national levels (Matasci et al., 2018) have been facilitated through the integration of Landsat and Airborne Laser Scanning (ALS) data. Forest structural and physiological parameters can also be derived from Landsat-based surface phenology observations (Tyukavina et al., 2015; Potapov et al., 2019).

Despite the availability of wall-to-wall interpolated GEDI-based global biomass datasets, their accuracy remains limited in regions lacking local calibration and validation (Nazir et al., 2025). Prior studies using optical datasets for tree height extrapolation (e.g., Baccini et al., 2008; Simard et al., 2011; Chi et al., 2015; Potapov et al., 2021) tend to overlook species-specific structural, phenological, and physiological traits, which limits aggregated forest structure estimates. Global forest canopy height maps with a spatial resolution of 10 and 30 m are produced using GEDI canopy height data in combination with optical imagery (Potapov et al., 2021; Lang et al., 2023). However, SAR data was not employed in any of these investigations. The link between canopy height, tree species type, and SAR backscatter is mediated by forest structure (Sothe et al., 2022). Additionally, spatial heterogeneity in species composition, elevation-driven ecological gradients, and limited road networks contribute to inconsistent sampling and data gaps. These factors collectively hinder the development of accurate, temporally and spatially continuous forest inventory datasets, particularly in mountainous landscapes (Xu et al., 2025). Consequently, many countries are reluctant to adopt these products for policy formulation, carbon stock estimation, and forest management interventions due to concerns over reliability and contextual relevance (Mitchell et al., 2017).

In this study, we develop new nation-scale high-resolution (10m) tree species maps for Nepal, leveraging extensive field measurements and Sentinel-2 (optical) phenological information. Mapping forest composition would assist land managers in identifying ecologically and economically valuable species, guide reforestation and conservation efforts, and contribute to monitoring of forest health and tree species distribution under changing climatic conditions (Fremout et al., 2020). We then used GEDI mean tree



height data with Sentinel-1 (SAR) and Sentinel-2 data to build independent species-specific models for tree heights at a similar scale for the entire country. Integrating both species and structural information into national forest inventories will enhance the precision of forest resource assessments, supporting evidence-based policy (such as Reducing Emissions from Deforestation and Forest Degradation (REDD+) and Measurement, Reporting, and Verification (MRV)) and land-use planning. As a mountainous country with diverse forest types that range from tropical lowlands to alpine zones, Nepal's forest ecosystems support rich biodiversity and provide critical ecosystem services, including carbon sequestration, water regulation, and livelihoods for local communities (Bhusal et al., 2016). To our knowledge, this study is the first to provide geospatially consistent high resolution species distributions and canopy heights for the country of Nepal.

## 2 Materials and methods

The workflows shown in [Supplementary Figure S1](#) integrate data from 980 field sites, 99,000 GEDI canopy height estimates, and 34 tiles of Sentinel-1 and Sentinel-2 images from 2017 to 2020 that we use to map and validate species composition, and map and validate mean canopy heights.

### 2.1 Study area

This study focuses on the country of Nepal ([Figure 1](#)), located between longitudes  $80^{\circ}04'E$  and  $88^{\circ}12'E$ , and latitudes  $26^{\circ}22'N$  and  $30^{\circ}27'N$ , covering a total area of 147,516 km<sup>2</sup>. Positioned in the central southern region of the Hindu Kush Himalayan range, Nepal borders India to the south, east, and west, with China to the north ([Uddin et al., 2018](#)). Elevations range from 45 m in the southern lowlands to 8,848 m at Mount Everest, the world's highest mountain peak ([Uddin et al., 2015](#)). Nepal's diverse landscapes encompass a variety of climates associated primarily with elevation, ranging from subtropical to warm temperate, and cold alpine and high elevation tundra ([Dahal and Hasegawa, 2008](#)). Notably, about 80% of annual precipitation occurs in the warmer months from June to September across most of Nepal ([Dahal and Hasegawa, 2008](#)).

According to national forestry statistics, Nepal's forested land area comprises 59,600 km<sup>2</sup>, accounting for 40.36% of the land area ([Nepal Forest Resource Assessment, 2015](#)). An additional 6,500 km<sup>2</sup>, or 4.38%, is classified as other wooded land (OWL). Within the forested area, 17.32% (10,300 km<sup>2</sup>) is designated as protected zones, while 82.68% (49,300 km<sup>2</sup>) lies outside protected areas under private or local government management ([Nepal Forest Resource Assessment, 2015](#)). The physiographic region of the Middle Mountains constitutes 37.80% of the total forest area, followed by

the High Mountains and High Himalayan regions (32.25%), the Churia region (20.04%), and the Terai region (6.90%) (Nepal Forest Resource Assessment, 2015). Approximately 443 tree species exist in this region, belonging to 239 genera and 99 families. Estimates indicate that Nepal's forests store ~1,055 million tons of carbon, equivalent to an average of ~177 tons per hectare, but with considerable spatial variability and differences among forest types and management histories (Nepal Forest Resource Assessment, 2015).

## 2.2 Forest inventory data

The forest inventory dataset used in this study comprises 1,010 sample plots collected between 2014 and 2016. These data serve as the foundation for training and validating machine learning models for tree species classification and canopy height prediction. Each circular field plot is (490–706 m<sup>2</sup> in area; 12.5–15 m radius) was georeferenced and richly annotated with attributes including tree species, diameter at breast height (DBH), tree height, stand density, topographic variables (elevation, slope, and aspect), and forest management type. This comprehensive ground-truth dataset provided critical reference information to support model calibration for species composition and height estimation across the diverse forest types in Nepal, with data subsets withheld from the calibration procedures to enable *post hoc* and independent validation of estimates (see below).

### 2.2.1 Pre-processing forest inventory data

We first filtered the field-based forest inventory plots to remove plots that were harvested or disturbed during the 5-year interval between the forest inventory survey (2014–2020) and the satellite imagery acquisition (2017–2020). We removed 30 field plots by virtually revisiting all plots using the high-resolution images available in Google Earth to identify those plots impacted by stand replacement disturbances in the years following field measurements.

We utilized field data to achieve two goals (identification of dominant tree species and mean height). To achieve the height mapping goal, we utilized the remaining 980 samples, providing independent validation of GEDI height estimates. However, to achieve the species mapping goal, we selected a subset of sample plots that represented the six dominant tree species (defined when a single species represents >70% of stems within the field plot) and a mixed forest class when no one species is above the 70% threshold (Table 1). Uneven distribution of forest types in the field sample resulted in a skewed probability distribution that favors the six dominant forest species over the mixed forest class (Maxwell et al., 2018). To address this imbalance, we down sampled the original 980 field plots to 533 to ensure balanced representation across classes. The mono-dominant forest classes included two evergreen needle-leaf classes (*Pinus roxburghii* and *Pinus wallichiana*), two evergreen broad-leaf classes (*Rhododendron arboreum* and *Schima wallichii*), and two deciduous broad-leaf classes (*Alnus nepalensis* and *Shorea robusta*). For tree species mapping, we randomly selected 75% (i.e., 399) of the forest inventory sample plots to train the species classification model, reserving 25% (134) for validation (Supplementary Figure S1).

TABLE 1 Field based forest plots (plot areas 500 m<sup>2</sup> depending on tree stem density) in Nepal (N = 533), classified by dominant species for plots where the species represented >70% occurrence of the plot and a mixed species class with no single species considered dominant. These datasets are used for model training for forest type mapping and height estimation. Measurements of all individuals within each field plot include tree species, stem density, stem diameters, height, and canopy area.

Species name	PFT type	Leaf type	Total samples
<i>Pinus roxburghii</i>	Needle leaf	Evergreen	115
<i>Pinus wallichiana</i>	Needle leaf	Evergreen	44
<i>Rhododendron arboreum</i>	Broad leaf	Evergreen	24
<i>Schima wallichii</i>	Broad leaf	Evergreen	29
<i>Shorea robusta</i>	Broad leaf	Deciduous	121
<i>Alnus nepalensis</i>	Broad leaf	Deciduous	55
Mixed species	Mixed leaf	Mixed	145

## 2.3 Satellite data

### 2.3.1 Optical data – Sentinel-2

We utilized S2 MSI surface reflectance (SR) images with <10% cloud cover mosaicked across 34 tiles covering Nepal from April 2017 to April 2020 available on the Google Earth Engine platform. The S2 Level-1C SR is an orthoimage that provides bottom-of-atmosphere reflectance in 10 spectral bands (Liu et al., 2024). This imagery includes supplementary quality indicators for cloud and snow probability, as well as aerosol optical thickness data, at a 60 m resolution (Liu et al., 2024). For each S2 scene, the QA60 bitmask band was applied to mask both cirrus and opaque clouds (Liu et al., 2024). We computed spectral signatures for each tree species using 10 spectral bands (B2–B8A, B11, and B12) and phenology using two vegetation indices: the Normalized Difference Vegetation Index (NDVI) (Tucker, 1979) and the Enhanced Vegetation Index (EVI) (Supplementary Figures S2, S3) (Huete et al., 1997). The NDVI measures chlorophyll reflectance and enhanced absorption to detect photosynthesizing plants (Tucker, 1979). In contrast, the EVI is more responsive to variations in canopy structures across different species, providing greater sensitivity in areas with dense vegetation (Huete et al., 1997).

### 2.3.2 SAR data – Sentinel-1

We collected the S1 time series data for April 2017–April 2020 from Google Earth Engine, covering ascending and descending flight paths across the 34 tiles. S1 provides high spatial and temporal precision in SAR data, at 10 m spatial resolution (Mullissa et al., 2021). The Google Earth Engine offers post-processed Analysis Ready Data (ARD) (Mullissa et al., 2021), including noise removal, radiometric terrain normalization, topographic correction, and speckle filtering. Since SAR signals are unaffected by clouds (Rüetschi et al., 2017), we utilized all seasons data from S1, mirroring the S2 time series approach. Additionally, we computed the cross-ratio (CR) from cross and co-polarized backscatter, a measure that is sensitive to forest phenology, biomass, and vegetation water content (Toan et al., 1992; Frison et al., 2018; Vreugdenhil et al., 2018).

The utilization of temporal S1 and S2 for mapping forest characteristics at the local scale builds on similar earlier work

(Lechner et al., 2022; Liu et al., 2024). We established three time-scales (annual, seasonal, and monthly) to enhance classification accuracy using aggregated temporal images. The seasonal dataset was divided into four seasons, reflecting the vegetation phenology in the study area: winter (December to February), spring (March to May), summer (June to August), and autumn (September to November) (Praticò et al., 2021; Liu et al., 2024). Using the temporal aggregation algorithm in Google Earth Engine, the median value for each pixel across these time scales was calculated independently.

### 2.3.3 LiDAR data – GEDI

The GEDI L2A product delivers footprint-level relative height (RH) estimates within a footprint diameter of 25 m. GEDI utilizes full-waveform LiDAR technology to calculate normalized cumulative waveform energy and fit Gaussian distributions to waveform peaks, thereby capturing vertical structural details (Dubayah et al., 2020). The GEDI RH metrics, derived from percentiles of return waveform energy, provide a thorough overview of forest vertical structure (Marselis et al., 2020). Canopy height percentiles (RH 1–100) are presented at 1% intervals (Beck et al., 2021). For example, relative height at the 98th percentile (RH98) is often used for measuring tree canopy height (Schneider et al., 2020). This approach enables the determination of relative heights for both the ground and canopy tops for each footprint (Marselis et al., 2020).

For tree canopy height mapping in Nepal's forests, we obtained the latest version of the GEDI L2A product ([https://lpdaac.usgs.gov/products/gedi02\\_av002/](https://lpdaac.usgs.gov/products/gedi02_av002/); Version 2; accessed in November 2024). We utilized tree canopy top height via the RH98 parameter to minimize noise that can be present in the RH100 parameter. We implemented filtering processes to ensure that the canopy height data at the footprint level is reliable (as described in Li et al., 2017; Beck et al., 2021; Liu et al., 2021). First, we removed erroneous waveform data using the quality flag filters (selecting only those footprints where “quality\_flag” = 1, “delta\_time” < 2024 and “beam\_flag” < (Full power = 0101, 0110, 1000, 1011)). Finally, we filtered the footprints based on the criteria: “sensitivity” > 0.98 and “degrade\_flag” = 0 (Liu et al., 2021; Urbazaev et al., 2022). Moreover, we filtered GEDI data to exclude noise-related heights above 43 m (removing 197 GEDI footprints across the full Nepal study region), aligning our results with the maximum tree height observed in the field. Following GEDI data retrieval and quality control we had a total of 99,000 GEDI footprints covering the entire geographic domain of Nepal.

### 2.3.4 DEM, climate, and forest cover datasets

Topography and climate provide additional input data for species and canopy height mapping because of their significant correlation with vegetation distribution (Pasquarella et al., 2018). The Shuttle Radar Topography Mission (SRTM) generates a ~30 m digital elevation model (DEM) (Farr et al., 2007) that we used to extract topographic variables (elevation, slope, and aspect) over Nepal. Climate data were sourced from Terra Climate, developed by the University of California, Merced (Abatzoglou et al., 2018). This platform provides spatial climate layers at a resolution of 4600 m, including monthly averages for minimum and maximum temperatures, as well as accumulated precipitation.

We computed the average minimum and maximum temperatures, and precipitation amounts for each pixel from 2010 to 2020 and resampled the layers to a 10 m resolution.

We elected not to use pre-existing forest cover classifications (e.g., ESA Forest Cover map, Zanaga et al., 2022) because of misclassification of forest extent that is common across much of Nepal (Nepal Forest Resource Assessment, 2015). In this study, we used independent forest/non-forest assessments in 4030 locations, together with the optical and SAR remote sensing datasets, to develop a new forest cover map (with 95% accuracy) using the RF model with two classes (forest and non-forest) (see details in the supplementary section: “forest cover” and [Supplementary Table S3](#)).

## 2.4 Tree species classification

Two machine-learning algorithms, Random Forest (RF) (Breiman, 2001; Belgiu and Drăgu, 2016) and Classification and Regression Trees (CART) (Loh, 2011), were implemented within the Google Earth Engine platform to perform national-scale tree-species classification across Nepal. Both models were trained on 75% of the field-collected observations and evaluated using the remaining 25% for independent validation ([Supplementary Figure S1](#)). We elected to test both RF and CART models because prior studies (Praticò et al., 2021; Liu et al., 2024; Lechner et al., 2022) suggest that both are effective in different situations and this allowed us to test and select the best-performing models. After applying each algorithm to the multi-source remote-sensing predictor dataset, we compared their performance using cross-validated overall accuracy (kappa) and class-specific producer and user accuracies. Although both approaches produced reliable classification outputs, the model demonstrating superior predictive accuracy was selected to generate the final species distribution map. This comparative framework not only leverages the strengths of each algorithm but also ensures that the results are underpinned by the most robust and accurate modeling approach.

CART is based on a single-tree decision, like the RF classifier (Loh, 2011). CART creates a structure called a decision tree by iteratively splitting the data. The attribute that specifies the split partitions of data into subsets at each tree node according to the normalized information gains. The characteristic with the most significant normalized information gain is used to make final judgments (Friedl and Brodley, 1997; Loh, 2011). The Google Earth Engine only permits the adjustment of two parameters for CART: the maximum number of nodes and the minimum leaf number (Praticò et al., 2021). RF is a non-parametric ensemble learning system that uses classification trees. It is suitable for high-dimensional data with correlated characteristics, but sensitive to overfitting (Breiman, 2001; Belgiu and Drăgu, 2016). Each tree was developed using a distinct subsample of training data (Breiman, 2001; Lyons et al., 2018; Meyer and Edzer, 2021). We employed all 770 predictor features extracted from S1 and S2, topography and climate data, and used Google Earth Engine's hyperparameter optimizer to identify the best parameter settings for each algorithm ([Supplementary Table S1](#)). For the RF model, we grew 100 trees with a minimum node size of five and considered ≈25 randomly selected predictors at each split. We tuned the maximum number of nodes and the minimum leaf population

for the CART model through iterative trial-and-error, evaluating model accuracy under each parameter combination. Model performance was then validated against the independent field dataset, allowing us to compare classification accuracy using independent data not used in the model calibration step. Finally, we generated variable-importance and variable-contribution graphs for species and canopy-height mapping (Supplementary Figures S4, S5).

## 2.5 Tree height prediction

We employed the same parameter optimization for RF and CART regression models for tree height mapping at the national scale in Nepal. A standard definition of a “tree” in carbon accounting is a woody plant with a minimum height of 2 m (Melo et al., 2023). In this context, woody vegetation that stands at least 2.3 m tall is classified as a tree. Following this definition, and consistent with findings that GEDI is unable to differentiate tree heights less than ~2.3 m due to the 15.6 ns pulse width (Burns et al., 2024; Cho et al., 2025; Li et al., 2024; Yu et al., 2024; Zhu et al., 2022). We designated forest height as zero for GEDI samples with RH98 < 2.3 m. The model calibration incorporated SRTM elevation and multitemporal S1 and S2 data as independent variables. GEDI RH98 served as the dependent variable in the models. For calibration, we extracted random sample sets of training data from each species mask within the GEDI data range. This process resulted in seven distinct training sets for the seven forest classes, averaging 14,142 GEDI samples per class (Supplementary Table S4).

## 2.6 Accuracy assessments

The accuracy of the tree species classification was evaluated using 134 independent samples, which comprised 25% of the total samples in the study area. We employed a confusion matrix constructed from estimated class area proportions to evaluate Kappa statistics, overall accuracy, user, and producer accuracies for tree species mapping.

The accuracy of the tree height model was evaluated using three types of datasets, 1) k-fold cross-validation, 2) an independent 30% GEDI dataset (Stojanova et al., 2010; Gupta and Sharma, 2022), and 3) field data. In each evaluation case, we determined the optimal model by computing the performance measures, coefficient of determination ( $R^2$ ), and root mean square error (RMSE). Furthermore, we validated the accuracy of tree height using an ordinary least squares (OLS) regression approach with 980 independent validation field plots.

# 3 Results

## 3.1 Tree species mapping

We evaluated the classification accuracies of both models and found that RF (Kappa: 0.63, overall accuracy: 0.7) outperformed CART (Kappa: 0.54, overall accuracy: 0.65) (Table 2; Supplementary Figure S2). Needle-leaf and broad-leaf species were effectively

separated because of their spectral and phenological variations, which were captured using multitemporal data (Supplementary Figures S2, S3; Tables 2, 3). By contrast, increased classification errors occurred between phenologically and physiologically related species within the broadleaf and needleleaf plant functional types (PFT), due to spectral-temporal similarity (e.g., *P. roxburghii* and *P. wallichiana*; Supplementary Figure S2). Error rates in mixed-species forests likely relate to the relative abundance of the constituent species increasing, for example, the probability of confusion with the *Pinus* family and *S. robusta* that are common in mixed forest stands (Table 2). However, there are shifts between mixed and pure stands in terms of producer and user accuracy, particularly for minor tree species (Table 2).

The forest types and distributions across Nepal are shown in Figure 2. Our estimated total forest area (56,029 km<sup>2</sup>) aligns with the NFI data (59,600 km<sup>2</sup>) (Nepal Forest Resource Assessment, 2015). Furthermore, the mapped tree species distributions (Figure 2) effectively represent expected species composition in the study area, with *S. robusta* common in the southern lowlands (16,036 km<sup>2</sup>), while *P. roxburghii* (13,790 km<sup>2</sup>) and *P. wallichiana* (4,028 km<sup>2</sup>) are common at higher altitudes (Figure 5; Table 3; Supplementary Table S4). However, large areas of forest in Nepal are not dominated by a particular species. They are classified as mixed forest (15,106 km<sup>2</sup>), particularly in the middle mountain regions of the country (Figures 2, 5). Tree species (e.g., *Rhododendron species* and *S. wallichii*) commonly found in Nepal are often classified in the mixed class because they tend to be less dominant (i.e., <70% stem density).

## 3.2 Tree height mapping

We evaluated the significance of S1 and S2 for height estimation using held-out GEDI samples (Supplementary Figure S5). S1 and S2 performed better for height estimation by fusing GEDI height data. We compared the model test efficiency under various feature combinations as an alternative experimental setup for both RF and CART models. The best height estimations were obtained when surface reflectance bands and indices were combined; all other settings produced either a significant underprediction of tall heights or an overprediction of low heights. We compared the predictions of RF and CART models with the RH98 values from the GEDI validation dataset. CART estimates short and tall heights more effectively compared to the RF. We selected the CART model due to better height estimation (RMSE of 2.4 m and  $R^2$  of 0.92 with a slope of 0.97 and intercept of 0.53 m; Figure 3) compared to RF estimations (RMSE of 2.7 m and  $R^2$  of 0.89 with a slope of 0.95 and intercept of 0.6 m; Supplementary Figure S6). In the CART model predictions, there is no bias for height values between 15 and 30 m, which comprise 80% of the GEDI footprints. Predictions for short stature canopies begin at approximately 2.3 m (the minimum GEDI prediction), while the CART model shows a slight tendency to underestimate tall trees (those between 35 and 43 m).

We validated predicted heights with an independent subset of the field data, which demonstrates a strong correlation with our predicted height using CART ( $R^2 = 0.74$ , RMSE = 3.7 m,  $c = 0.89$ , and  $m = 4.1$  m) (Figure 4). The maximum and average heights of trees in our calibration set varies by tree species, with the mean

TABLE 2 Random Forest (RF) model tree species classification error matrix using independent validation samples.

Species name	<i>Pinus roxburghii</i>	<i>Pinus wallichiana</i>	<i>Rhododendron arboreum</i>	<i>Schima wallichii</i>	<i>Shorea robusta</i>	<i>Alnus nepalensis</i>	Mixed species	Total	User accuracy
<i>Pinus roxburghii</i>	16	1			2	1	5	25	0.64
<i>Pinus wallichiana</i>		7			2		2	11	0.63
<i>Rhododendron arboreum</i>			4			1	1	6	0.66
<i>Schima wallichii</i>				4			2	6	0.66
<i>Shorea robusta</i>		1			19		4	24	0.79
<i>Alnus nepalensis</i>	1					5	1	7	0.71
Mixed species	4	2	2	2	4	1	40	55	0.72
Total	21	11	6	6	27	8	55	134	
Producer accuracy	0.76	0.63	0.66	0.66	0.70	0.62	0.72		0.70
Kappa value									0.63

TABLE 3 Tree species minimum, mean, and maximum height in field-based sample plots and in the mapped dataset.

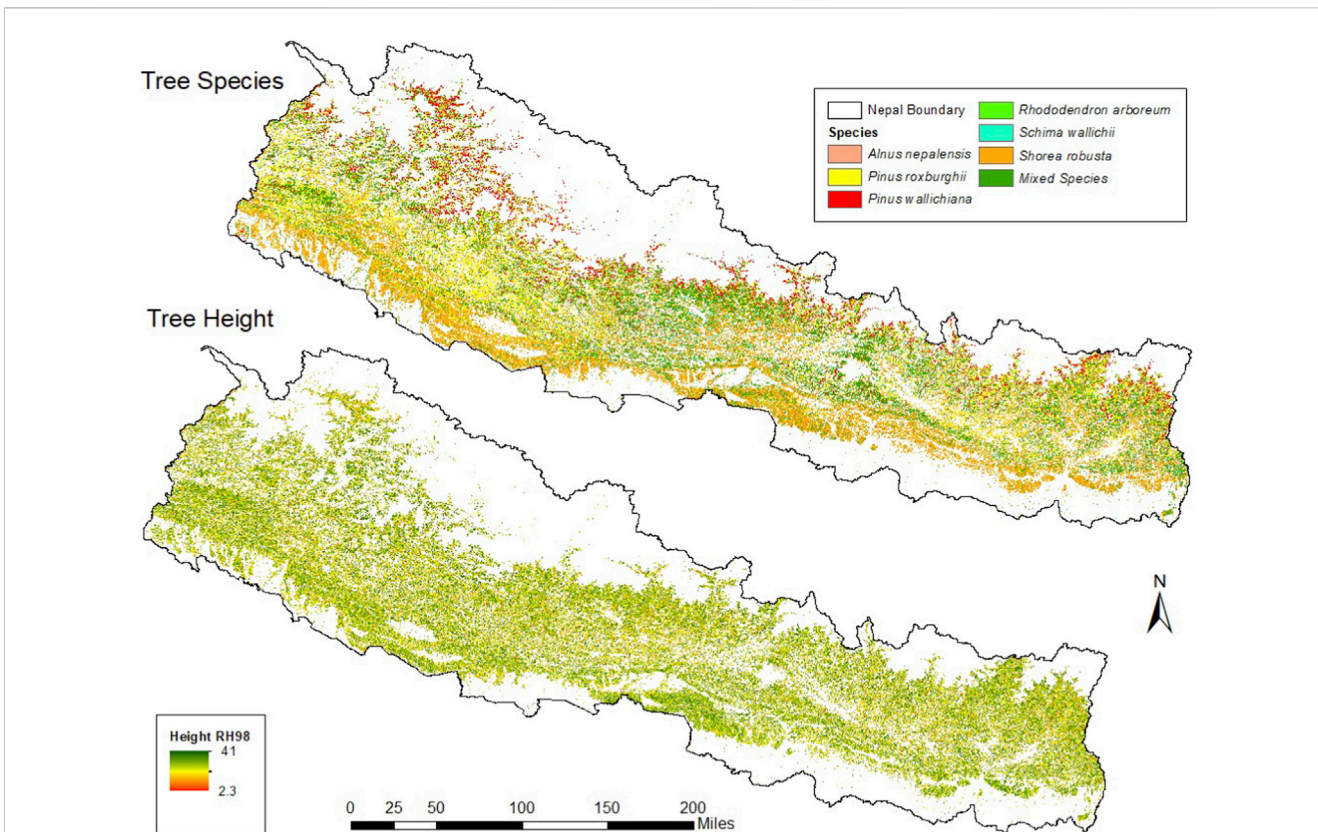
Species name	Field data statistics				Predicted estimates			
	Samples max height (m)	Samples mean height (m)	Samples min height (m)	Samples mean elevation (m)	Mapped max height (m)	Mapped mean height (m)	Mapped min height (m)	Species mapped area (km <sup>2</sup> )
<i>Pinus roxburghii</i>	42.2	17.1	2.9	1670	41	19.8	3.5	13790
<i>Pinus wallichiana</i>	41.9	15.2	3.1	2542	41	18.5	3.7	4028
<i>Rhododendron arboreum</i>	33.7	14.6	3.7	2060	39.7	18.6	3.6	1002
<i>Schima wallichii</i>	35.7	11.7	2.2	1142	38.8	13.8	2.3	995
<i>Shorea robusta</i>	41.3	10.9	1.5	820	40.1	14.5	5.8	16036
<i>Alnus nepalensis</i>	33.8	15.5	4.7	1695	35.4	19.2	3.9	5072
Mixed species	42.2	14.1	2.4	NA	40.9	17.3	3.5	15106

height ranging from 10.9 to 17.1 m and the maximum height spanning from 33.7 to 42.2 m (Table 3). We observed a larger prediction bias for short stature forests, with GEDI tending to overestimate slightly in heights between 2.3 and 10 m (Figures 2, 4; Supplementary Figure S6).

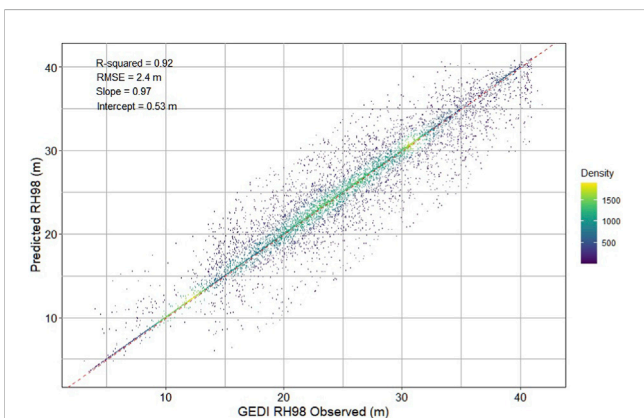
Before applying the height model for all of Nepal, it is critical to confirm that the GEDI L2A product accurately represents the height of the forest canopy. We assessed GEDI metrics overall and within each species mask. Across all of Nepal, we observed a mean of 19.9, a minimum of 2.3, and a maximum of 43 m. At species-level (comparing field site observations to GEDI estimates) we

observed that the GEDI mean height was overestimated by 2–5 m compared to the NFI data (Table 3). For instance, for *S. robusta* and *A. nepalensis* the mean height is ~4 m overestimated (Table 3). The overestimation may be attributed to data noise, poor training samples, faulty model assumptions, and sensor resolution limitations.

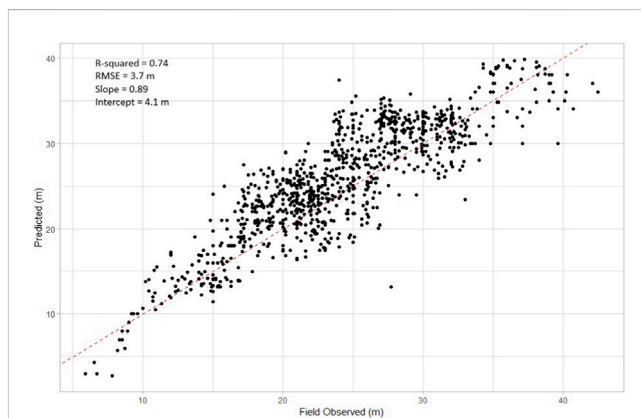
Lastly, we produced height projections for the nation’s forest areas (Figure 5). We found that ~60% of tree heights fall within the range of 15–30 m (Figures 4, 5). The tree species *P. roxburghii* occupies the largest area (2,091 km<sup>2</sup>) above 35 m, followed by mixed tree species (1,141 km<sup>2</sup>) (Figure 5). We discovered that the height



**FIGURE 2** Tree species and mean canopy height predictions for Nepal. Tree species was predicted using Sentinel-1 SAR and Sentinel-2 visible-near infrared imagery based on 533 training sites for six dominant species and a mixed class (Table 1–3). Tree heights for each forest type were then estimated from GEDI footprints based on a CART model. White areas within Nepal represent high mountain tundra and non-vegetation areas in the north and low elevation farmlands in the south, where forest vegetation is absent.



**FIGURE 3** CART model scatter plot using independent GEDI height validation data (utilizing OLS regression). This data was not used in the model calibration. Color represents density of points, yellow represents higher density and blue represents low density of points. The red dotted line represents the 1:1 line.



**FIGURE 4** Field data height validation (independent field data: N = 980) using an OLS regression model. The red dotted line represents the 1:1 line.

distribution of forest stands varied by species type. Most short stature stands are of *S. robusta* (152 km<sup>2</sup>) and mixed tree species (102 km<sup>2</sup>) (Figure 5). This signifies the establishment of new plantations, with

many younger stands found at heights below 10 m. We noted that post-disturbance and new plantation areas (height <5 m) are around 700 km<sup>2</sup>, compared to the 430 km<sup>2</sup> reported by the NFI datasets (Nepal Forest Resource Assessment, 2015).

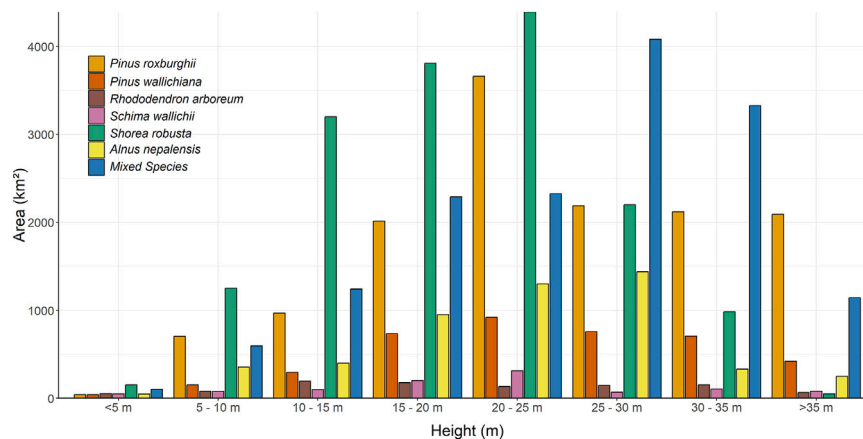


FIGURE 5  
Mapped area for each forest type in different height classes (5 m intervals).

## 4 Discussion

This work makes available two datasets characterizing the forested ecosystems of Nepal: a new forest type map, mapping forests dominated (>70% occurrence) by six species and a “mixed forest” class, and a forest height map differentiating the height distributions of different forest types across the sub-tropical lowlands to the high mountains of Nepal. For all forest types, our findings showed the potential of using field-based forest plot measurements, with GEDI tree heights product, optical and SAR data (from the Sentinel satellites) to map forest types and forest canopy height on a large scale. Mapping forest canopy heights by dominant species is novel, providing significant new insight into ecosystem distributions and dynamics across Nepal, and species-specific information on forest resources and other forest-related ecosystem functions.

We employed a novel technique to map tree height by integrating optical, SAR, and LiDAR data. Using field-based forest plot measurements (NFI data), we developed a tree species map to train species-specific independent height models, recognizing that different the relationship between forest structure and optical and SAR imagery will likely be different for each forest type based on plant functional type (e.g., needleleaf versus broad leaf) and phenological traits (evergreen versus deciduous). In this study, the different forest types have varying average and maximum heights, which cannot be adequately represented using a single model. For example, *S. robusta* in southern low-elevation areas has a lower average height than *Pinus* genera in high-elevation areas. Training different models for each species can address this issue.

The S2 10 m bands are crucial for obtaining spectral and phenological information; however, utilizing 20 m bands can enhance the precision of classification results. For instance, the NIR, SWIR, and red-edge bands are significant for the spectral signatures of conifer and broadleaf tree species (Persson et al., 2018; Wessel et al., 2018). Seasonal variations may influence the spectral response (Persson et al., 2018); however, adding phenology as a separate band can account for these variations. Our investigation reveals that the prediction improves with the number of S1 and

S2 layers provided as inputs. The S1 bands proved to be the most promising predictors in our case for reducing errors in mountainous locations. We discovered that using two S1 band combinations produced better results than using only a single S1 band. Thus, the model’s prediction ability is enhanced by adding more input bands. A model trained only on S1 is sufficient to estimate forest height in locations where, for example, S2 data may not be available due to high levels of cloud cover. Our results were consistent with studies that demonstrated the key roles of VV and VH in estimating forest canopy height (Sothe et al., 2022). A model trained just with optical data can produce accurate predictions, but it saturates in regions with mountain shadows and can be impacted by persistent cloud in many regions, impacting the reliability of earlier estimates for Nepal (Potapov et al., 2021; Lang et al., 2023).

CART height predictions reveal higher errors for the hilly areas and short-height forests (Figure 3). These errors are linked to GEDI overestimation in the hilly areas and the footprint geolocation error (~10 m) (Chen et al., 2024). For example, GEDI RH98 values in short canopies overestimate actual heights, while underestimating in taller canopies, consistent with earlier study results (Morin et al., 2019; Potapov et al., 2021; Lang et al., 2023). In this study, the height overestimation in short canopies and underestimation in tall canopies might be in part related to an uneven distribution of reference height data, with only ~1% of the total footprints in the test dataset <5 m and ~2% > 35 m. Future versions of the CART model can examine other methods to lessen this impact, such as using a weighted cost function or an alternative sampling procedure among the height data.

The GEDI LiDAR mission provides a relatively dense distribution of height samples in forests between 51.6 °S and 51.6 °N. However, depending on the desired spatial resolution, interpolation strategies, typically using complementary satellite imagery, are needed to provide wall-to-wall canopy height models (Potapov et al., 2021; Lang et al., 2023). The GEDI data is a valuable resource for large-scale CHM mapping since it offers several benefits over alternative techniques (Potapov et al., 2021). We compared the CART model canopy height predictions (using S2 and S1 datasets trained with GEDI measurements) to two distinct global canopy height products: P21 (Potapov et al., 2021) and L23

(Lang et al., 2023) for three sample locations in Nepal. Our locally calibrated CART model shows improved error metrics (RMSE = 3.27 m) compared to the two-height models (P21 = 9.9 m and L23 = 8.2 m) validated using field measurements. Furthermore, both P21 and L23 models exhibit a saturation effect around 15–20 m (Supplementary Figure S7). A visual comparison demonstrates that the P21 and L23 models predict higher spatial variability within forest stands, often associated with topography (slope and aspect; Supplementary Figure S7). These findings highlight the importance of the scale at which the analysis was performed. P21 and L23 are the results of global models that were trained using GEDI measurements as reference height and optical data (Landsat-8 for P21, S2 for L23), thus providing optimal global estimates but not necessarily working well at local and national scales. Tree species based canopy height modeling is an optimal option to map tree height variance at local to national scale with higher accuracy and lower bias.

The GEDI mission offers a high-precision LiDAR assessment of forest attributes. Image-based interpolation, however, needs to consider local and regional characteristics that may require local calibration of predictive models. The primary challenge in classifying tree species is the similar spectral and temporal characteristics shared among them. Tree species often exhibit comparable phenology due to similarities in timing and duration of leaf initiation, production and abscission. For instance, deciduous broadleaf species exhibit strong seasonality compared to other broadleaf species (Supplementary Figures S2, S3). It is challenging to distinguish between broadleaf species when they are found in mixed-species stands. This difficulty is most noticeable in areas where complex mixed broadleaf stands are prevalent. High mountains also have fewer independent ground verification data due to challenging terrain and inaccessibility.

## 5 Conclusion

Forest canopy height is an important measure for measuring forest biomass, carbon storage, and both structural diversity and bio-diversity. GEDI has collected a significant amount of global LiDAR sample point data. However, GEDI data is not wall-to-wall, so strategies for interpolation of sample data are needed. Optical and SAR data provide correlative surface information and have several benefits for interpolation, particularly wall-to-wall coverage, high spatial resolution, numerous spectral bands, and frequent measurements allowing phenological representation. Here, we developed a forest type map and canopy height model for the entire country of Nepal, combining a large and comprehensive field measurement campaign with the different surface feature information provided by GEDI, optical imagery, and SAR data. To accomplish large-scale forest canopy height estimates at regional scales, future research can use GEDI and ICESat-2 data to extract ground sample points, merge them with optical remote sensing and SAR imagery, and build machine-learning models. Improving the accuracy of forest canopy height estimates enables an improved assessment of biomass and carbon storage. Through these developments, we can better protect and manage valuable forest resources, promote the growth of a green economy, significantly aid in tackling climate change problems, and accomplish sustainable development objectives.

## Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: <https://zenodo.org/records/18308466>.

## Author contributions

AN: Conceptualization, Formal Analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review and editing. NH: Conceptualization, Data curation, Funding acquisition, Project administration, Resources, Supervision, Validation, Writing – review and editing. HG: Conceptualization, Data curation, Resources, Validation, Writing – review and editing. HS: Data curation, Resources, Validation, Writing – review and editing.

## Funding

The author(s) declared that financial support was received for this work and/or its publication. This work was supported by the NASA GEDI Science Team Grant (Grant # 80NSSC21K0201).

## Acknowledgements

The authors acknowledge the Savanna Lab Team, the Nepal forest department for supplying NFI data, and the LPDAAC Earth Resources Observation and Science Center for supplying GEDI L2A, and Sentinel 1 and 2.

## Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author NH declared that they were an editorial board member of Frontiers at the time of submission. This had no impact on the peer review process and the final decision.

## Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

## 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

- Abatzoglou, J. T., Dobrowski, S. Z., Parks, S. A., and Hegewisch, K. C. (2018). TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. *Sci. Data* 5 (1), 1–12. doi:10.1038/sdata.2017.191
- Asner, G. P., Mascaró, J., Müller-Landau, H. C., Vieilledent, G., Vaudry, R., Rasamoelina, M., et al. (2012). A universal airborne LiDAR approach for tropical forest carbon mapping. *Oecologia* 168 (4), 1147–1160. doi:10.1007/S00442-011-2165-Z/FIGURES/8
- Baccini, A., Laporte, N., Goetz, S. J., Sun, M., and Dong, H. (2008). A first map of tropical Africa's above-ground biomass derived from satellite imagery. *Environ. Res. Lett.* 3 (4), 45011. doi:10.1088/1748-9326/3/4/045011
- Beck, J., Wirt, B., Armston, J., Hofton, M., Scott, L., and Tang, H. (2021). "Global Ecosystem Dynamics Investigation (GEDI) level 02 user guide".
- Belgiu, M., and Drăgu, L. (2016). Random Forest in remote sensing: a review of applications and future directions. *ISPRS J. Photogrammetry Remote Sens.* 114, 24–31. doi:10.1016/j.isprsjprs.2016.01.011
- Bhusal, J., Bhusal, J. K., Chapagain, P. S., Regmi, S., Gurung, P., Zulkafli, Z., et al. (2016). Mountains under pressure: evaluating ecosystem services and livelihoods in the upper Himalayan Region of Nepal. *Int. J. Ecol. Environ. Sci.* 42 (3), 217–226.
- Bolyn, C., Michez, A., Gaucher, P., and Lejeune, P. (2018). Gestion des ressources forestières et des milieux naturels Michez, Peter; Université de Liège-ULiège >. Ingénierie des biosystèmes (Biose) >. Gestion des ressources forestières et des milieux naturels Gaucher, Philippe; Université de Liège-ULiège >. Ingénierie des biosystèmes (Biose) >. Gestion des ressources forestières et des milieux naturels Lejeune, and Stéphanie; Université de Liège-ULiège >. Echanges Eau-Sol-Plantes Bonnet. 2018. "Forest Mapping and Species Composition Using Supervised per Pixel Classification of Sentinel-2 Imagery. *Biotechnol. Agronomie, Société Environnement* 22 (3), 172–187. doi:10.25518/1780-4507.16524
- Bonan, G. B. (2008). Forests and climate change: forcings, feedbacks, and the climate benefits of forests. *Science* 320 (5882), 1444–1449. doi:10.1126/SCIENCE.1155121/SUPPL\_FILE/BONAN\_SOM.PDF
- Breiman, L. (2001). Random forests. *Mach. Learn.* 45 (1), 5–32. doi:10.1023/A:1010933404324/METRICS
- Brockerhoff, E. G., Barbaro, L., Castagneryol, B., Forrester, D. I., Gardiner, B., González-Olabarria, J. R., et al. (2017). Forest biodiversity, ecosystem functioning and the provision of ecosystem services. *Biodivers. Conserv.* 26 (13), 3005–3035. doi:10.1007/S10531-017-1453-2
- Burns, P., Hakkenberg, C. R., and Goetz, S. J. (2024). Multi-Resolution gridded maps of vegetation structure from GEDI. *Sci. Data* 11 (1), 1–14. doi:10.1038/s41597-024-03668-4
- Chen, R., Wang, X., Liu, X., and Wang, S. (2024). Optimizing GEDI canopy Height estimation and analyzing error impact factors under highly complex terrain and high-density vegetation conditions. *Forests* 15 (11), 2024. doi:10.3390/f15112024
- Chi, H., Sun, G., Huang, J., Guo, Z., Ni, W., and Fu, A. (2015). National Forest aboveground biomass mapping from ICESat/GLAS data and MODIS imagery in China. *Remote Sens.* 7 (5), 5534–5564. doi:10.3390/RS7055534
- Cho, M.-S., Roy, D. P., Kashongwe, H. B., Yan, L., and Shen, M. (2025). Assessment and improvement of GEDI canopy Height estimation in tropical and temperate forests. *Sci. Remote Sens.* 11, 100221. doi:10.1016/j.srs.2025.100221
- Coops, N. C., Tompalski, P., Goodbody, T. R. H., Queinnec, M., Luther, J. E., Bolton, D. K., et al. (2021). Modelling lidar-derived estimates of Forest attributes over space and time: a review of approaches and future trends. *Remote Sens. Environ.* 260, 112477. doi:10.1016/j.rse.2021.112477
- Dahal, R. K., and Hasegawa, S. (2008). Representative rainfall thresholds for landslides in the Nepal Himalaya. *Geomorphology* 100 (3–4), 429–443. doi:10.1016/j.geomorph.2008.01.014
- Dixon, R. K., Brown, S., Houghton, R. A., Solomon, A. M., Trexler, M. C., and Wisniewski, J. (1994). Carbon pools and flux of global Forest ecosystems. *Science* 263 (5144), 185–190. doi:10.1126/SCIENCE.263.5144.185
- Dubayah, R., Hofton, M., Blair, J., Armston, J., Tang, H., and Luthcke, S. (2020). GEDI L2A elevation and height metrics data global footprint level V001. *NASA EOSDIS Land Processes Distributed Active Archive Center (DAAC) data set (2020): GEDI02\_A-001*.
- Duncanson, L. I., Niemann, K. O., and Wulder, M. A. (2010). Estimating Forest canopy height and terrain relief from GLAS waveform metrics. *Remote Sens. Environ.* 114, 138–154. doi:10.1016/j.rse.2009.08.018
- Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Riley, D., Scott, H., et al. (2007). The shuttle radar topography Mission. *Rev. Geophys.* 45 (2). doi:10.1029/2005RG000183
- Fremout, T., Thomas, E., Gaisberger, H., Van Meerbeek, K., Muenchow, J., Briers, S., et al. (2020). Mapping tree species vulnerability to multiple threats as a guide to restoration and conservation of tropical dry forests. *Glob. Change Biol.* 26 (6), 3552–3568. doi:10.1111/GCB.15028;JOURNAL:JOURNAL:13652486;WGROU:STRING:PUBLICATION
- Friedl, M. A., and Brodley, C. E. (1997). Decision tree classification of land cover from remotely sensed data. *Remote Sens. Environ.* 61 (3), 399–409. doi:10.1016/s0034-4257(97)00049-7
- Frison, P. L., Fruneau, B., Kmiha, S., Soudani, K., Dufrère, E., Le Toan, T., et al. (2018). Potential of Sentinel-1 data for monitoring temperate mixed Forest phenology. *Remote Sens.* 10 (12), 10. doi:10.3390/RS10122049
- Goetz, S., and Dubayah, R. (2011). Advances in remote sensing technology and implications for measuring and monitoring Forest carbon stocks and change. *Carbon Manag.* 2 (3), 231–244. doi:10.4155/CMT.11.18
- Goetz, S. J., Hansen, M., Houghton, R. A., Walker, W., Laporte, N., and Busch, J. (2015). Measurement and monitoring needs, capabilities and potential for addressing reduced emissions from deforestation and Forest degradation under REDD+. *Environ. Res. Lett.* 10 (12), 123001. doi:10.1088/1748-9326/10/12/123001
- Gupta, R., and Sharma, L. K. (2022). Mixed tropical forests canopy height mapping from spaceborne LiDAR GEDI and multisensor imagery using machine learning models. *Remote Sens. Appl. Soc. Environ.* 27, 100817. doi:10.1016/j.rsa.2022.100817
- Hansen, M. C., Potapov, P. V., Goetz, S. J., Turubanova, S., Tyukavina, A., Krylov, A., et al. (2016). Mapping Tree Height Distributions in Sub-Saharan Africa using Landsat 7 and 8 data. *Remote Sens. Environ.* 185, 221–232. doi:10.1016/j.rse.2016.02.023
- Healey, S. P., Yang, Z., Gorelick, N., and Simon, I. (2020). Highly local model calibration with a new GEDI LiDAR asset on Google Earth engine reduces Landsat Forest height signal saturation. *Remote Sens.* 12 (17), 2840. doi:10.3390/RS12172840
- Houghton, R. A. (1999). The annual net flux of carbon to the atmosphere from changes in land use 1850–1990. *Tellus B* 51 (2), 298–313. doi:10.1034/j.1600-0889.1999.00013.x
- Hudak, A. T., Lefsky, M. A., Cohen, W. B., and Berterretche, M. (2002). Integration of lidar and landsat ETM+ data for estimating and mapping Forest canopy height. *Remote Sens. Environ.* 82 (2–3), 397–416. doi:10.1016/S0034-4257(02)00056-1
- Huete, A. R., Liu, H. Q., Batchily, K., and Van Leeuwen, W. (1997). A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote Sens. Environ.* 59 (3), 440–451. doi:10.1016/S0034-4257(96)00112-5
- Lang, N., Jetz, W., Schindler, K., and Jan, D. W. (2023). A high-resolution canopy height model of the Earth. *Nat. Ecol. and Evol.* 7 (11), 1778–1789. doi:10.1038/s41559-023-02206-6
- Lechner, M., Dostálová, A., Hollaus, M., Atzberger, C., and Immitzer, M. (2022). Combination of Sentinel-1 and Sentinel-2 data for tree species classification in a central European Biosphere Reserve. *Remote Sens.* 14 (11), 2687. doi:10.3390/RS14112687
- Lefsky, M. A. (2010). A global Forest canopy height map from the moderate resolution imaging spectroradiometer and the geoscience laser altimeter System. *Geophys. Res. Lett.* 37 (15). doi:10.1029/2010GL043622
- Li, A., Dhakal, S., Glenn, N. F., Spaete, L. P., Shinneman, D. J., Pilliod, D. S., et al. (2017). Lidar aboveground vegetation biomass estimates in shrublands: prediction, uncertainties and application to coarser scales. *Remote Sens.* 9 (9), 903. doi:10.3390/RS9090903
- Li, Yi, Gao, S., Fu, H., Zhu, J., Hu, Q., Dong, Z., et al. (2024). Error analysis and accuracy improvement in Forest canopy height estimation based on GEDI L2A product: a case Study in the United States. *Forests* 15 (9), 1536. doi:10.3390/f15091536
- Liu, A., Cheng, X., and Chen, Z. (2021). Performance evaluation of GEDI and ICESat-2 laser altimeter data for terrain and canopy height retrievals. *Remote Sens. Environ.* 264, 112571. doi:10.1016/j.rse.2021.112571
- Liu, P., Ren, C., Wang, Z., Jia, M., Yu, W., Ren, H., et al. (2024). Evaluating the potential of Sentinel-2 time series imagery and machine learning for tree species classification in a mountainous Forest. *Remote Sens.* 16 (2), 16. doi:10.3390/RS16020293
- Loh, W. Y. (2011). Classification and regression trees. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* 1 (1), 14–23. doi:10.1002/WIDM.8

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/frsen.2026.1725509/full#supplementary-material>

- Lyons, M. B., Keith, D. A., Phinn, S. R., Mason, T. J., and Elith, J. (2018). A comparison of resampling methods for remote sensing classification and accuracy assessment. *Remote Sens. Environ.* 208, 145–153. doi:10.1016/j.rse.2018.02.026
- Marselis, S. M., Abernethy, K., Alonso, A., Armston, J., Baker, T. R., Bastin, J. F., et al. (2020). Evaluating the potential of full-waveform lidar for mapping pan-tropical tree species richness. *Glob. Ecol. Biogeogr.* 29 (10), 1799–1816. doi:10.1111/GEB.13158
- Matasci, G., Hermosilla, T., Wulder, M. A., White, J. C., Coops, N. C., Hobart, G. W., et al. (2018). Large-Area mapping of Canadian boreal Forest cover, height, biomass and other structural attributes using landsat composites and lidar plots. *Remote Sens. Environ.* 209, 90–106. doi:10.1016/j.rse.2017.12.020
- Maxwell, A. E., Warner, T. A., and Fang, F. (2018). Implementation of machine-learning classification in remote sensing: an applied review. *Int. J. Remote Sens.* 39 (9), 2784–2817. doi:10.1080/01431161.2018.1433343
- McManamon, P. F. (2019). *LiDAR technologies and systems*. SPIE.
- Melo, J., Baker, T., Nemitz, D., Quegan, S., and Ziv, G. (2023). Satellite-Based global maps are rarely used in Forest reference levels submitted to the UNFCCC. *Environ. Res. Lett.* 18 (3), 034021. doi:10.1088/1748-9326/ACBA31
- Meyer, H., and Edzer, P. (2021). Predicting into unknown space? Estimating the area of applicability of spatial prediction models. *Methods Ecol. Evol.* 12 (9), 1620–1633. doi:10.1111/2041-210X.13650
- Mitchell, A. L., Rosenqvist, A., and Mora, B. (2017). Current remote sensing approaches to monitoring Forest degradation in support of countries measurement, reporting and verification (MRV) systems for REDD+. *Carbon Balance Manag.* 12 (1), 1–22. doi:10.1186/S13021-017-0078-9/TABLES/2
- Montesano, P. M., Rosette, J., Sun, G., North, P., Nelson, R. F., Dubayah, R. O., et al. (2015). The uncertainty of biomass estimates from modeled ICESat-2 returns across a boreal Forest gradient. *Remote Sens. Environ.* 158, 95–109. doi:10.1016/j.rse.2014.10.029
- Morin, D., Planells, M., Guyon, D., Villard, L., Mermoz, S., Bouvet, A., et al. (2019). Estimation and mapping of Forest structure parameters from open access satellite images: development of a generic method with a Study case on coniferous plantation. *Remote Sens.* 11 (11), 1275. doi:10.3390/RS11111275
- Mullissa, A., Vollrath, A., Odongo-Braun, C., Slatger, B., Balling, J., Gou, Y., et al. (2021). Sentinel-1 SAR backscatter analysis ready data preparation in Google Earth engine. *Remote Sens.* 13 (10), 1954. doi:10.3390/RS13101954
- Nazir, A., Hanan, N. P., Yu, Q., and Gilani, H. (2025). Enhancing GEDI above ground biomass density estimates in contrasting forests of Pakistan. *For. Ecol. Manag.* 587, 122747. doi:10.1016/j.foreco.2025.122747
- Nelson, M. (2017). Evaluating multitemporal Sentinel-2 data for Forest mapping using random Forest. Available online at: <https://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-146657>.
- Nepal Forest Resource Assessment (2015). *State of Nepal's forests*. Forest Resource Assessment (FRA).
- Pan, Y., Birdsey, R. A., Fang, J., Houghton, R., Kauppi, P. E., Kurz, W. A., et al. (2011). A large and persistent carbon sink in the world's forests. *Science* 333 (6045), 988–993. doi:10.1126/SCIENCE.1201609/SUPPL\_FILE/PAPV2.PDF
- Pang, Y., Li, Z., Ju, H., Lu, H., Jia, W., Lin, Si, et al. (2016). LiChy: the ca's LiDAR, CCD and hyperspectral integrated airborne observation System. *Remote Sens.* 8, 398. doi:10.3390/rs8050398
- Pascual, A., Guerra-Hernández, J., Armston, J., Minor, D. M., Duncanson, L. I., May, P. B., et al. (2023). Assessing the performance of NASA's GEDI L4A footprint aboveground biomass density models using National Forest inventory and airborne laser scanning data in Mediterranean Forest ecosystems. *For. Ecol. Manag.* 538, 120975. doi:10.1016/j.foreco.2023.120975
- Pasquarella, V. J., Holden, C. E., and Woodcock, C. E. (2018). Improved mapping of Forest type using spectral-temporal landsat features. *Remote Sens. Environ.* 210, 193–207. doi:10.1016/j.rse.2018.02.064
- Patterson, P. L., Healey, S. P., Ståhl, G., Saarela, S., Holm, S., Andersen, H. E., et al. (2019). Statistical properties of hybrid estimators proposed for GEDI—Nasa's global ecosystem dynamics investigation. *Environ. Res. Lett.* 14 (6), 065007. doi:10.1088/1748-9326/AB18DF
- Persson, M., Lindberg, E., and Reese, H. (2018). Tree species classification with multitemporal Sentinel-2 data. *Remote Sens.* 10 (11), 1794. doi:10.3390/RS10111794
- Popescu, S. C., Zhou, T., Nelson, R., Neuenschwander, A., Sheridan, R., Narine, L., et al. (2018). Photon counting LiDAR: an adaptive ground and canopy Height Retrieval Algorithm for ICESat-2 data. *Remote Sens. Environ.* 208, 154–170. doi:10.1016/j.rse.2018.02.019
- Potapov, P., Tyukavina, A., Turubanova, S., Talero, Y., Hernandez-Serna, A., Hansen, M. C., et al. (2019). Annual continuous fields of woody vegetation structure in the lower Mekong Region from 2000–2017 landsat time-series. *Remote Sens. Environ.* 232, 111278. doi:10.1016/j.rse.2019.111278
- Potapov, P., Li, X., Hernandez-Serna, A., Turubanova, S., Alexandra, T., Hansen, M. C., et al. (2021). Tropical Forest canopy structure and change assessment using landsat, Gedi, and airborne lidar data. *Int. Geoscience Remote Sens. Symposium (IGARSS)*, 666–669. doi:10.1109/IGARSS47720.2021.9554814
- Praticò, S., Solano, F., Fazio, S. Di, and Modica, G. (2021). Machine learning classification of Mediterranean Forest habitats in Google Earth engine based on seasonal Sentinel-2 time-series and input image composition optimisation. *Remote Sens.* 13 (4), 586. doi:10.3390/RS13040586
- Pugh, T. A. M., Lindeskog, M., Smith, B., Poulter, B., Arneth, A., Haverd, V., et al. (2019). Role of Forest regrowth in global carbon sink dynamics. *Proc. Natl. Acad. Sci. U. S. A.* 116 (10), 4382–4387. doi:10.1073/PNAS.1810512116/SUPPL\_FILE/PNAS.1810512116.SD02.TXT
- Rishmawi, K., Huang, C., Schleweweis, K., and Zhan, X. (2022). Integration of VIIRS Observations with GEDI-Lidar measurements to Monitor Forest Structure Dynamics from 2013 to 2020 across the Conterminous United States. *Remote Sens.* 14 (10), 2320. doi:10.3390/rs14102320
- Rüetschi, M., Schaepman, M. E., and Small, D. (2017). Using multitemporal Sentinel-1 C-Band backscatter to monitor phenology and classify deciduous and coniferous forests in Northern Switzerland. *Remote Sens.* 10 (1), 55. doi:10.3390/RS10010055
- Saarela, S., Holm, S., Healey, S. P., Andersen, H. E., Petersson, H., Prentius, W., et al. (2018). Generalized hierarchical model-based estimation for aboveground biomass assessment using GEDI and landsat data. *Remote Sens.* 10 (11), 1832. doi:10.3390/RS10111832
- Schneider, F. D., António, F., Hancock, S., Duncanson, L. I., Dubayah, R. O., Pavlick, R. P., et al. (2020). Towards mapping the diversity of canopy structure from space with GEDI. *Environ. Res. Lett.* 15 (11), 115006. doi:10.1088/1748-9326/AB9E99
- Shendryk, Y. (2022). Fusing GEDI with Earth observation data for large area aboveground biomass mapping. *Int. J. Appl. Earth Observation Geoinformation* 115, 103108. doi:10.1016/J.JAG.2022.103108
- Simard, M., Pinto, N., Fisher, J. B., and Baccini, A. (2011). Mapping Forest canopy height globally with spaceborne lidar. *J. Geophys. Res. Biogeosciences* 116 (G4), 4021. doi:10.1029/2011JG001708
- Sothe, C., Gonsamo, A., Lourenço, R. B., Kurz, W. A., and Snider, J. (2022). Spatially continuous mapping of Forest canopy height in Canada by combining GEDI and ICESat-2 with PALSAR and sentinel. *Remote Sens.* 14 (20), 5158. doi:10.3390/RS14205158
- Stojanova, D., Panov, P., Gjorgjioski, V., Kobler, A., and Džeroski, S. (2010). Estimating vegetation height and canopy cover from remotely sensed data with machine learning. *Ecol. Inf.* 5 (4), 256–266. doi:10.1016/J.ECOINF.2010.03.004
- Toan, T. Le, Beaudoin, A., Riom, J., and Guyon, D. (1992). Relating Forest biomass to SAR data. *IEEE Trans. Geoscience Remote Sens.* 30 (2), 403–411. doi:10.1109/36.134089
- Tucker, C. J. (1979). Red and photographic infrared Linear combinations for monitoring vegetation. *Remote Sens. Environ.* 8 (2), 127–150. doi:10.1016/0034-4257(79)90013-0
- Tyukavina, A., Baccini, A., Hansen, M. C., V Potapov, P., V Stehman, S., Houghton, R. A., et al. (2015). Aboveground Carbon Loss in Natural and Managed Tropical forests from 2000 to 2012. *Environ. Res. Lett.* 10 (7), 74002. doi:10.1088/1748-9326/10/7/074002
- Uddin, K., Shrestha, H. L., Murthy, M. S. R., Bajracharya, B., Shrestha, B., Gilani, H., et al. (2015). Development of 2010 national land cover database for the Nepal. *J. Environ. Manag.* 148, 82–90. doi:10.1016/J.JENVMAN.2014.07.047
- Uddin, K., Matin, M. A., and Maharjan, S. (2018). Assessment of land cover change and its impact on changes in soil erosion risk in Nepal. *Sustainability* 10 (12), 4715. doi:10.3390/SU10124715
- Urbazaev, M., Hess, L. L., Hancock, S., Sato, L. Y., Jean, P. O., Thiel, C., et al. (2022). Assessment of terrain elevation estimates from ICESat-2 and GEDI spaceborne LiDAR missions across different land cover and Forest types. *Sci. Remote Sens.* 6, 100067. doi:10.1016/J.SRS.2022.100067
- Vreugdenhil, M., Wagner, W., Bauer-Marschallinger, B., Pfeil, I., Teubner, I., Rüdiger, C., et al. (2018). Sensitivity of Sentinel-1 backscatter to vegetation dynamics: an Austrian case Study. *Remote Sens.* 10 (9), 1396. doi:10.3390/RS10091396
- Wessel, M., Brandmeier, M., and Tiede, D. (2018). Evaluation of different machine learning algorithms for scalable classification of tree types and tree species based on Sentinel-2 data. *Remote Sens.* 10 (9), 1419. doi:10.3390/RS10091419
- Xi, Z., Xu, H., Xing, Y., Gong, W., Chen, G., and Yang, S. (2022). Forest Canopy height mapping by synergizing ICESat-2, Sentinel-1, Sentinel-2 and topographic information based on machine learning methods. *Remote Sens.* 14 (2), 364. doi:10.3390/RS14020364
- Xu, W., Cheng, Y., Luo, M., Mai, X., Wang, W., Zhang, W., et al. (2025). Progress and limitations in Forest carbon stock estimation using remote sensing technologies: a comprehensive review. *Forests* 16 (3), 449. doi:10.3390/F16030449
- Yu, Q., Ryan, M. G., Ji, W., Prihodko, L., Anchang, J. Y., Kahiu, N., et al. (2024). Assessing Canopy height measurements from ICESat-2 and GEDI orbiting LiDAR across six different biomes with G-LiHT LiDAR. *Environ. Res. Ecol.* 3 (2), 025001. doi:10.1088/2752-664X/AD39F2
- Zanaga, D., Van De Kerchove, R., Daems, D., De Keersmaecker, W., Brockmann, C., Kirches, G., et al. (2022). ESA WorldCover 10 m 2021 v200. doi:10.5281/zenodo.7254221
- Zhao, Y., Du, S., Li, K., Jiang, J., Guo, Q., and Xiao, W. (2024). Estimation of canopy height based on multi-source remote sensing data using Forest structure aided sample selection. *Int. J. Remote Sens.* 45 (7), 2235–2268. doi:10.1080/01431161.2024.2326537
- Zhu, X., Nie, S., Wang, C., Xi, X., Lao, J., and Dong, Li (2022). Consistency analysis of Forest height retrievals between GEDI and ICESat-2. *Remote Sens. Environ.* 281, 113244. doi:10.1016/j.rse.2022.113244