https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026

Open Access Earth System Science Data Discussions

RFDTM: A national-scale and wall-to-wall 30 m resolution mangrove sub-canopy topography dataset for New Zealand derived from ICESat-2 ATLAS and multi-band SAR

Yunqiu Wang[1], Jiapeng Huang[1,*], Yue Zhang[1], Yuhao Wang[2], Chunpeng Chen[3], Hongsheng Zhang[4], Bohao He[5], Jihong Chen[6,7], Qingquan Li[8,9], Nan Xu[8,9,*]

[1]School of Geomatics, Liaoning Technical University, Fuxin 123000, China;
[2]School of Earth Sciences and Engineering, Hohai University, Nanjing 211100, China;
[3]State Key Laboratory of Estuarine and Coastal Research, East China Normal University, Shanghai 200241, China;
[4]Department of Geography, The University of Hong Kong, Hong Kong 999077, China;
[5]Department of Electronics, Information and Bioengineering, Polytechnic University of Milan, 20133 Milan, Italy;
[6]College of Management, Shenzhen University, Shenzhen, 518071, China
[7]Shenzhen International Maritime Institute, Shenzhen, 518083, China
[8]MNR Key Laboratory for Geo-Environmental Monitoring of Great Bay Area, Guangdong Key Laboratory of Urban Informatics, Shenzhen Key Laboratory of Spatial Smart Sensing and Services, Shenzhen University, Shenzhen 518060, China;
[9]School of Architecture and Urban Planning, Shenzhen University, Shenzhen 518060, China

Correspondence to: J. Huang (huangjiapeng@lntu.edu.cn); N. Xu (xunan2025@szu.edu.cn)

Abstract. Mangrove sub-canopy topography plays a critical role in coastal hydrological processes, blue carbon storage, ecosystem stability, and inundation vulnerability under sea-level rise. However, existing global Digital Elevation Models (DEMs) often contain large elevation uncertainties and data gaps in mangrove regions because dense canopy cover limits the penetration capability of conventional remote sensing observations, resulting in incomplete and inaccurate representations of sub-canopy terrain. To address this critical data deficiency, we present RFDTM, a large-scale mangrove sub-canopy topography dataset for New Zealand at 30 m spatial resolution generated entirely from publicly available satellite observations. The dataset was developed by integrating ICESat-2 photon-counting LiDAR data with dual-frequency C-band and L-band SAR observations. First, a Hierarchical Multi-Constraint Filtering (HMCF) strategy was employed to extract reliable ground photons and improve the reliability of terrain elevation estimates beneath dense canopies. Subsequently, multi-source terrain and vegetation features were constructed and optimized within a Random Forest regression framework to reconstruct continuous sub-canopy topography and generate the RFDTM product. Validation against airborne LiDAR terrain data across all mangrove regions of New Zealand demonstrates excellent performance, with an R^2 of 0.99, RMSE of 1.01 m, MAE of 0.80 m, and bias of 0.43 m, fully satisfying the accuracy requirements for regional-scale applications. Ablation experiments further confirm the critical contribution of L-band SAR observations, reducing the RMSE from 1.23 m to 1.01 m and substantially enhancing sub-canopy penetration capability. Overall, RFDTM represents the first large-scale mangrove sub-canopy topography product derived solely from open-access satellite data, while the proposed methodology provides a transferable and readily applicable framework for global coastal vulnerability assessment, ecosystem monitoring, and carbon cycle studies.

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026

Open Access Earth System Science Data Discussions

1 Introduction

Mangrove ecosystems are widely distributed in tropical and subtropical coastal zones and serve as critical transitional areas connecting terrestrial and marine environments (Hu et al., 2020). They play essential ecological roles in wave attenuation, carbon sequestration and emission reduction, biodiversity conservation, and coastal stabilization (Prihantono et al., 2022; Hu et al., 2020). Influenced jointly by tidal processes, hydrodynamic conditions, and sedimentary environments, mangrove regions are generally characterized by low and relatively flat terrain but exhibit significant spatial variability in topography (Hu et al., 2020). Such variability exerts a decisive influence on tidal inundation patterns, hydrological connectivity, and the distribution of vegetation types (Xie et al., 2022). Therefore, high-precision mangrove sub-canopy topography data not only provide a fundamental basis for understanding mangrove ecological processes and ecosystem services, but also supply critical topographic information for assessing sea-level rise, storm-surge inundation risks, and environmental changes in coastal zones (Huang et al., 2024b; Xi et al., 2022). However, the canopy structure and complex vegetation height variations of mangroves severely limit the ability of traditional remote sensing techniques to effectively retrieve sub-canopy topography (Li et al., 2021; Li et al., 2023).

Airborne remote sensing data, particularly airborne Light Detection and Ranging (LiDAR), have been widely used for extracting forest and mangrove sub-canopy topography due to their high point density and strong vegetation penetration capability. They are considered an important reference data source for acquiring high-accuracy Digital Terrain Models (DTMs) under complex vegetation conditions (Jiang et al., 2024). However, airborne LiDAR data acquisition is expensive and time-consuming, and is constrained by airspace regulations, meteorological conditions, and operational logistics (Wang et al., 2025b).

Consequently, its spatial coverage is limited, making it difficult to support long-term and continuous monitoring over large-scale mangrove regions (Yin et al., 2024). In contrast, satellite remote sensing provides wide spatial coverage, strong repeat observation, and stable data acquisition, offering a new pathway for large-scale mangrove sub-canopy topography estimation (Pimple et al., 2025). Recent studies show high accuracy of spaceborne LiDAR from ICESat-2 and GEDI in mangroves for ground elevation and canopy height retrieval as key inputs for understory modeling: ICESat-2 validated against airborne LiDAR yields canopy height RMSE = 1.63 m at 10 m segments and 1.37 m at 30 m segments outperforming GEDI's 1.51 m, bathymetry RMSE of 0.25 m, and R^2 = 0.80 for biomass-related heights from nighttime strong-beam data, compared with GEDI's R^2 of 0.76, while GEDI's sampling aids large-scale use (Renshaw et al., 2026; Yu et al., 2026). In particular, recent advances in spaceborne LiDAR have demonstrated significant potential in vegetation structure retrieval and surface elevation estimation (Ma et al., 2019).

Nevertheless, in mangrove environments, several challenges persist for spaceborne LiDAR applications, including sparse and discontinuous sampling due to orbital tracks and limited ground returns beneath canopies, further compounded by strong tidal influences on signal detection. These factors increase the uncertainty of sub-canopy topography retrieval and constrain further improvements in accuracy (Urbazaev et al., 2022; Liu et al., 2021; Choi, 2024; Ma et al., 2019).

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026

In complex forest environments such as mangroves, the limited penetration capability of ICESat-2 and the insufficient accuracy of sub-canopy topography estimation pose significant challenges. Early studies mainly relied on airborne LiDAR data to improve ground elevation estimates; empirical correction approaches were widely applied to mitigate systematic elevation bias in vegetated regions. For example, Zhao et al. (2018) employed linear regression using vegetation height, leaf area index, and slope derived from airborne LiDAR, reducing the root mean square error by approximately 1 m. However, such methods strongly depend on high-cost airborne LiDAR data and do not fully exploit the potential of ICESat-2 photon observations. To address this limitation, Yu et al. (2022) evaluated ICESat-2 ATL08 heights using airborne LiDAR as reference data, demonstrating improved ground elevation estimation with reduced RMSE = 0.96 m. Further, Tian et al. (2021) proposed a multi-constraint filtering strategy that enhances the accuracy of ICESat-2 ground photon extraction by integrating multiple criteria, overcoming the residual errors caused by vegetation and signal noise. Building upon this approach, our previous work applied the Hierarchical Multi-Constraint Filtering (HMCF) scheme to understory elevation point selection, yielding a noticeable improvement in accuracy (Huang et al., 2024a). Collectively, these studies demonstrate the effectiveness of LiDAR-supported correction strategies and advanced filtering methods, while highlighting the need for scalable approaches that can be applied across large mangrove areas.

To reduce reliance on single data sources and improve model generalization, recent studies have increasingly introduced machine learning methods and multi-source remote sensing features. Yu et al. (2024) integrated ICESat-2 canopy samples, Sentinel-2 optical imagery, TanDEM-X elevation data, and climate variables using a Random Forest (RF) model to generate a global mangrove canopy height map at 30 m resolution. Similarly, Zhang et al. (2026) constructed a machine learning framework that incorporated optical variables and canopy height information to correct systematic elevation bias in FABDEM, reducing RMSE by nearly 50%. In addition, Li et al. (2024) and Li et al. (2023) demonstrated that integrating optical, terrain, and canopy features through CNN–LightGBM or BPNN models could significantly improve estimation accuracy in different forest types. Although these studies highlight the benefits of multi-source information, the complementary advantages of different SAR frequency bands in vegetation structural response and ground penetration capability have not yet been fully explored.

Recent studies have also attempted to enhance vegetation penetration capability by incorporating synthetic aperture radar data into estimation frameworks. In particular, C-band synthetic aperture radar data have gradually been introduced to improve the penetration capability of remote sensing observations in vegetated environments. For instance, Huang et al. (2024b) combined ICESat-2 elevation references, SRTM terrain data, and Sentinel-1 C-band backscatter features to retrieve sub-canopy topography using machine learning methods such as RF, achieving RMSE values between 0.847 m and 3.59 m depending on the study area. These results demonstrate the potential of multi-source remote sensing data for improving sub-canopy topography estimation accuracy. Nevertheless, most current studies remain limited to regional case analyses, and large-scale or national-level mangrove sub-canopy topography datasets are still scarce, restricting broader investigations of terrain patterns and related ecological processes.

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026

Despite these advances, several limitations remain in existing studies. Many approaches continue to rely heavily on costly airborne LiDAR, which limits the feasibility of large-scale mangrove sub-canopy topography acquisition. The integration of multi-source remote sensing data is often insufficient, particularly due to the lack of systematic evaluation of SAR data across different frequency bands in terms of their sensitivity to vegetation structure and ground penetration capability. In addition, existing methods for extracting terrain information from spaceborne photon-counting LiDAR data often rely on relatively simple or single-criterion filtering strategies, which are inadequate for handling the complex noise conditions and heterogeneous canopy structures in mangrove environments. This limitation reduces the reliability of terrain control point extraction and constrains overall estimation accuracy. Furthermore, national-scale mangrove sub-canopy topography datasets remain scarce, restricting the analysis of large-scale terrain patterns and associated ecological processes.

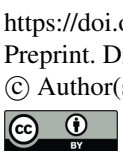
To address the long-standing lack of accurate sub-canopy terrain information in mangrove ecosystems, this study developed the RFDTM dataset, the first national-scale mangrove sub-canopy topography dataset for New Zealand at 30 m spatial resolution generated entirely from publicly available satellite observations. RFDTM integrates ICESat-2 photon-counting LiDAR data with dual-frequency C-band and L-band Synthetic Aperture Radar (SAR) observations to reconstruct continuous terrain surfaces beneath dense mangrove canopies, where conventional global Digital Elevation Models (DEMs) commonly exhibit substantial elevation uncertainty and missing terrain information. A Hierarchical Multi-Constraint Filtering (HMCF) strategy was employed to extract reliable ground photons beneath dense vegetation using multiple physical and geometric constraints, thereby providing robust elevation control points for terrain reconstruction. Furthermore, within a Random Forest regression framework, multi-source remote sensing features related to terrain morphology, vegetation structure, and microwave penetration characteristics were systematically integrated and optimized to generate continuous sub-canopy elevation estimates across all mangrove regions of New Zealand. Here, RFDTM represents the first spatially continuous national-scale mangrove sub-canopy topography dataset for New Zealand derived entirely from openly accessible satellite observations. Validation against airborne LiDAR demonstrates high accuracy and spatial consistency, supporting its applicability from regional to national scales. Moreover, the proposed multi-source fusion framework offers a scalable and transferable approach for sub-canopy topography reconstruction in mangrove ecosystems globally.

2 Study area and data

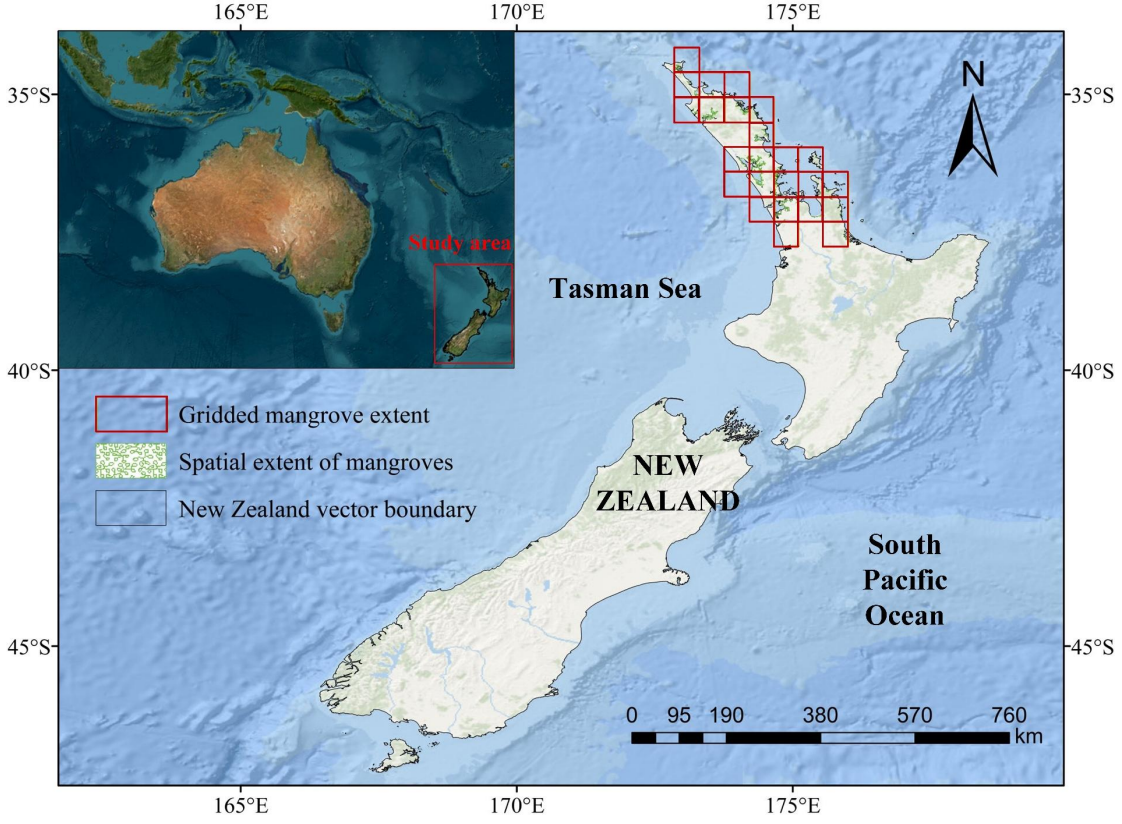
2.1 Study area

The study area is located in New Zealand, spanning approximately 34°–47° S latitude and 166°–178° E longitude, with a total mangrove area of approximately 2–9 km² within the selected regions. The region is low-lying and characterized by mangroves (Avicennia marina subsp. australasica) distributed mainly along estuaries and intertidal zones. The area exhibits a large tidal range, with estuarine and coastal sedimentary environments interspersed, forming typical mangrove wetlands. Mangrove distribution is patchy and spatially heterogeneous, making the region representative for studies of sub-canopy terrain and vegetation structure. The region experiences a temperate oceanic climate, characterized by mean annual temperatures of 14–

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026



16 °C and annual precipitation of 1200–1500 mm, which provide favorable conditions for mangrove development. The study area is shown in Fig. 1.



135 **Fig. 1. Study area. Mangrove extent data were derived from the Global Mangrove Watch (GMW) v3.0 dataset.**

2.2 Data sources

2.2.1 ICESat-2/ATLAS data products

ICESat-2, launched in September 2018 with data available since May 2019 (Popescu et al., 2018), carries the photon-counting Advanced Topographic Laser Altimeter System (ATLAS) for high-precision elevation measurements (Jin et al., 2020).
140 ATLAS operates six laser beams (17 m footprint, 3.3 km cross-track spacing) at a 10 kHz repetition rate from a 496–500 km orbit, enabling continuous global observations of canopy and terrain structure (Molly et al., 2019; Narin et al., 2024; Ranndal et al., 2021; Wang et al., 2024). ICESat-2 provides standard products including ATL03 and ATL08 (Sun et al., 2021; Wu et al., 2021; Li et al., 2022c), and this study utilizes datasets updated in 2020. After quality control (Vernimmen et al., 2020), these data support photon point cloud classification using optimized cloth simulation filtering to distinguish ground, canopy,
145 and noise photons (Christopher et al., 2019; Lin, 2019), thereby enabling sub-canopy topography reconstruction and forest structural analysis (Wang et al., 2023; E et al., 2021; Zhu et al., 2022; Tian et al., 2021; Huang et al., 2022). All data are freely available from the NSIDC website (https://nsidc.org/data/).

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026

2.2.2 ALOS-2 (L-band)

ALOS-2 Advanced Land Observing Satellite-2 launched by the Japan Aerospace Exploration Agency in 2014 carries the Phased Array type L-band Synthetic Aperture Radar-2 PALSAR-2 sensor operating at an L-band wavelength of 23.6 cm. Compared with C- and X-band SAR, L-band signals exhibit stronger penetration capability through vegetation canopies, making ALOS-2 PALSAR-2 particularly suitable for forested and mangrove environments (Arikawa et al., 2014). In this study, L-band SAR data were obtained from the Google Earth Engine platform using the ALOS-2 PALSAR-2 Level-2.1 terrain-corrected backscatter product. The dataset corresponds to observations acquired during the 2020 period January–December 2020 and includes radiometrically calibrated and terrain-corrected sigma nought (σ^0) backscatter coefficients. The data provide dual-polarization measurements HH and HV with a spatial resolution of approximately 25 m depending on the acquisition mode and processing scheme. Both HH and HV polarization channels were utilized to characterize vegetation structure and support terrain-related analysis (Luong et al., 2019). All datasets were accessed and processed via the Google Earth Engine platform (https://code.earthengine.google.com/).

2.2.3 Sentinel data (C-band, Sentinel-2)

The Sentinel satellites, developed by ESA, provide high-resolution Earth observation data. Sentinel-1 operating in Interferometric Wide Swath IW mode provides C-band synthetic aperture radar observations with VV and VH polarizations, enabling all-weather and day-and-night imaging capabilities. Compared with longer-wavelength SAR systems, C-band backscatter is particularly sensitive to vegetation canopy structure and temporal dynamics, making it effective for forest cover monitoring and phenological change detection (Torres et al., 2012). Sentinel-2's Multispectral Instrument (MSI) captures 13 multispectral bands at 10–60 m resolution, with Level-2A products providing 10 m surface reflectance corrected for atmosphere and orthorectified via Sen2Cor (Li et al., 2025), and the datasets updated in 2020 were used in this study. The experimental data were accessed via the GEE platform (https://code.earthengine.google.com/).

2.2.4 Canopy height product

A 30 m global forest canopy height map was produced by integrating Global Ecosystem Dynamics Investigation (GEDI) measurements (April–October 2019) with Landsat ARD time series (Potapov et al., 2021). In this study, the datasets updated in 2020 were used. The accuracy depends on the quality of GEDI observations and the suitability of the optical time series. The data are available via the Google Earth Engine platform (https://code.earthengine.google.com).

2.2.5 GLC_FCS30_2020

GLC_FCS30-2020 is a 30 m global land cover dataset developed by RADI and collaborators, containing 29 classes, with forests further divided into 12 subtypes (Zhang et al., 2021; Zhang et al., 2024), and the dataset updated in 2020 was used in this study. It supports fine-scale land cover analysis and is available at https://doi.org/10.5281/zenodo.3986872.

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026


Open Access Earth System Science Data Discussions

2.2.6 LiDAR DEM data

To validate the derived terrain parameters, high-resolution airborne LiDAR elevation data were obtained from the open data service of Land Information New Zealand. The Northland LiDAR 1 m DEM dataset was used as reference terrain data, and the datasets acquired between December 2018 and February 2020 were selected for analysis. The DEM was derived from airborne laser scanning observations collected using the Leica TerrainMapper-LN system, with an average point density of ~4 points m^{-2} and a reported vertical accuracy of approximately 0.2 m. Elevations referenced to NZVD2016 were converted to EGM2008 to ensure consistency with ICESat-2 observations. The data are available at https://data.linz.govt.nz/data/

2.2.7 Overview of reference DEM data

The reference DEMs used for accuracy assessment of the mangrove sub-canopy topography are summarized in Table 1. These DEMs, differing in data sources, spatial resolution, and processing methods, provide a comprehensive benchmark for evaluating the performance of the RFDTM reconstruction.

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026

190 **Table 1. Overview of the Compared DEM Datasets.**

Comparison data	Data description
FABDEM	FABDEM is a 30 m global bare-Earth elevation dataset generated by removing buildings and forests from Copernicus DEM using a RF approach, and in this study, it was used to derive slope and aspect (Hawker et al., 2022) (https://data.bris.ac.uk/data/dataset/s5hqmjcdj8yo2ibzi9b4ew3sn).
SRTM	The Shuttle Radar Topography Mission (SRTM) provides global elevation data, and the 1 arc-second SRTM dataset is used here as a terrain reference for DTM generation, supporting photon point cloud classification, canopy height modeling, and terrain factor extraction (Li et al., 2022a; Wang et al., 2024) (https://dwtkns.com/srtm30m/).
ASTER GDEM(V3)	The ASTER Global Digital Elevation Model (GDEM) Version 3, derived from ASTER stereo imagery aboard NASA's Terra satellite, provides near-global coverage and improved elevation accuracy, and was used in this study (Chirico et al., 2013) (https://search.earthdata.nasa.gov/search/).
TanDEM-X	TanDEM-X is a bistatic SAR interferometric mission consisting of two nearly identical satellites that generated a high-precision global DEM with an absolute vertical error of about 1 m (Wessel et al., 2021; Huber et al., 2021; Wessel et al., 2018)(https://geoservice.dlr.de/web/dataguide/).
DeltaDTM	DeltaDTM is a global coastal digital terrain model (DTM) optimized for low-lying coastal areas, with a spatial resolution of 10 m. The data were downloaded from the website (https://code.earthengine.google.com).
DiluviumDEM	DiluviumDEM is a globally corrected digital elevation model for coastal terrain, with a spatial resolution of 10 m. It can be accessed from Zenodo (https://zenodo.org/records/8384665).
NASADEM	NASADEM is a reprocessed SRTM dataset that integrates ASTER GDEM, ICESat GLAS, and PRISM data to improve elevation accuracy and reduce data voids (Farr et al., 2007). The dataset is available at https://lpdaac.usgs.gov/products/nasadem_hgtv001/.

Although the multi-source datasets used in this study were acquired at slightly different times between 2018 and 2020, their joint use is considered reasonable. The study area is characterized by relatively stable environmental and vegetation conditions, where terrain structure and canopy distribution exhibit limited short-term variability. All datasets were selected within a consistent temporal window centered around 2020, and the analysis focuses on terrain and structural features that are less

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026

sensitive to temporal differences. In addition, the integration of active and passive remote sensing data provides complementary information, helping to reduce uncertainties. Therefore, the minor temporal inconsistencies are unlikely to introduce significant bias into the results.

3 Methodology

To generate the national-scale mangrove sub-canopy topography dataset RFDTM and effectively address the persistent lack of reliable elevation information beneath dense mangrove canopies, this study implements a specialized data production workflow.In this study, ICESat-2 ATL03 and ATL08 data were preprocessed and co-registered to obtain precise photon point cloud coordinates. The Hierarchical Multi-Constraint Filtering (HMCF) method then removed noise and extracted high-confidence ground photons beneath the canopy using multi-attribute filtering (signal-to-noise ratio, terrain slope, cloud conditions, and photon distribution), providing reliable training samples. A feature system of 57 independent variables was constructed from multi-source data, including FABDEM, ALOS-2 L-band, GEDI (v27), Sentinel-1 C-band, Sentinel-2, and vegetation/water indices (e.g., NDVI, NDWI). All variables were spatially aligned with the photon points via masking and raster row–column alignment. Variables with Variance Inflation Factor (VIF) ≥ 10 were excluded, and importance analysis identified 25 key predictors. These 25 variables were input into a Random Forest (RF) model with ICESat-2 photon elevations as the dependent variable to reconstruct the sub-canopy topography DTM. Finally, the inverted results were mosaicked and clipped to the mangrove extent, and accuracy was evaluated against LiDAR DEM data and existing DTM products. The workflow is shown in Fig. 2.

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026

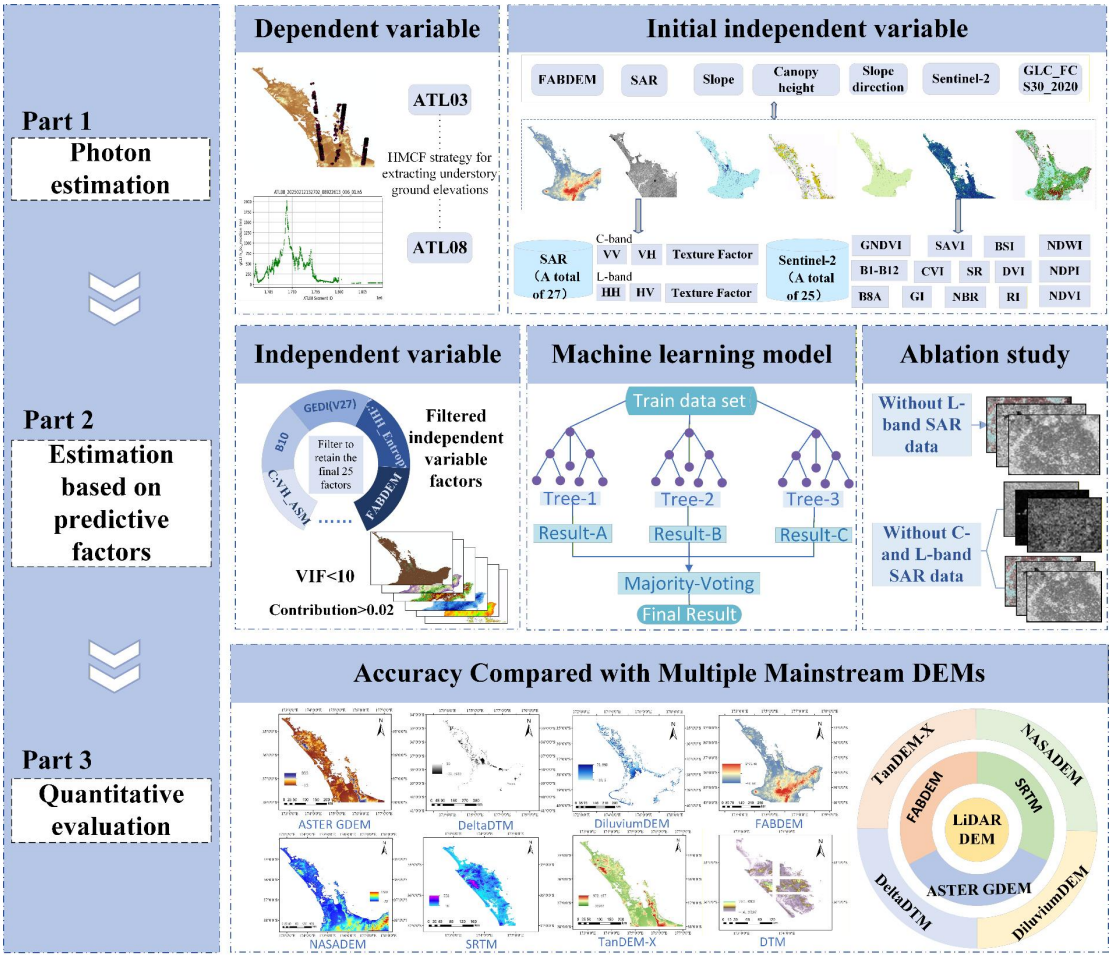


Fig. 2. Workflow of the proposed national-scale mangrove sub-canopy topography estimation framework.

3.1. Data preprocessing

This study applies a HMCF strategy to robustly extract high-quality terrain control points from ICESat-2 ATL03 and ATL08 data under complex mangrove environments. The method integrates ATL03 photon-level observations with ATL08-derived surface classification information through consistency-constrained point-wise matching. Notably, our previous work has already demonstrated the effectiveness of the HMCF scheme for understory elevation point selection, where it achieved a measurable improvement in accuracy (Huang et al., 2024a). Building upon this foundation, HMCF employs a two-stage screening procedure: an initial screening step removes clearly unreliable photons based on signal confidence and classification flags, followed by a fine screening step that progressively eliminates noisy or ambiguous points according to multiple criteria, including photon count per segment, segment size, signal-to-noise ratio, cloud and night flags, and statistical consistency of elevation values. Through this multi-attribute, hierarchical filtering process, high-confidence ground photons beneath the forest canopy are effectively identified and extracted, providing reliable dependent variable samples for subsequent terrain estimation.

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026

For predictor construction, a multi-source remote sensing feature system with 57 variables was established, including terrain factors, optical multispectral characteristics, vegetation indices, and multi-mode SAR features derived from polarization and texture information. The details are shown in Table 2.

Table 2. Overview of the Compared DEM Datasets.

Predictor variables	Detailed information
Terrain Factors	Elevation data were extracted from FABDEM, and terrain derivatives such as slope and aspect were computed to characterize micro-topography variations.
L-band SAR Data	Using ALOS-2 L-band data, HH and HV backscatter coefficients were extracted, and multiple texture features were derived based on the Gray-Level Co-occurrence Matrix (GLCM) to enhance the description of sub-canopy topography structure.
GEDI Forest Structure Parameters	Waveform structure metrics from GEDI (v27) were incorporated to capture the influence of vertical forest structure on terrain estimation.
Land Cover Data	The GLC_FCS30_2020 product was used to obtain land cover types in the study area, providing constraints for mangrove distribution identification and masking.
C-band SAR Data	Sentinel-1 VV and VH backscatter coefficients were used along with computed texture features to improve sensitivity to terrain variations under vegetation cover.
Multispectral Optical Data	Sentinel-2 bands B1–B12 were used to derive multiple spectral indices, including NDVI, NDWI, GNDVI, SAVI, BSI, NDPI, DVI, SR, CVI, RI, NBR, GI, and B8A, representing vegetation growth, water distribution, and bare soil characteristics.

During feature preprocessing, we first reprojected all raster predictors to a unified coordinate reference system and resampled them to a common spatial resolution of 30 m to ensure consistency across multi-source datasets and compatibility with the photon point cloud reference. Bilinear interpolation was applied to continuous variables, while nearest-neighbor resampling was used for categorical data. To ensure spatial consistency, we co-registered all datasets using the DEM as the reference and verified the alignment accuracy based on stable terrain features.

Subsequently, we masked and aligned the 57 predictor variables in raster row–column order to ensure one-to-one spatial correspondence with the photon points. After clipping to the mangrove distribution area, we established a spatial mapping between raster cells and photon positions, achieving precise raster-to-point matching. Finally, we constructed a sample database containing photon elevations as the dependent variable and 57 predictors, providing the foundation for subsequent variable selection, multicollinearity analysis, and RF estimation.

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026

After spatial matching of the 57 predictors, multicollinearity analysis was conducted. The Variance Inflation Factor (VIF) was computed for each variable to assess linear correlations among predictors, as shown in Equation (1):

$$VIF_i = \frac{1}{1-R_i^2} \quad (1)$$

where R_i^2 is the coefficient of determination obtained by performing a multiple regression with the i-th predictor as the dependent variable and the remaining variables as independent variables. When VIF_i exceeds a predetermined threshold (e.g., 5 or 10), the variable is considered to exhibit significant multicollinearity and should be removed. By iteratively eliminating highly collinear variables, feature redundancy is reduced, thereby improving model stability and interpretability.

3.2 Model training

Given the large spatial extent of the study area, a modeling approach with strong generalization capability and scalability is required. RF is therefore considered a suitable choice due to its high adaptability across heterogeneous environments (Zhang et al., 2026). Previous studies have demonstrated that RF–based approaches generally outperform alternative methods in terms of accuracy and robustness, particularly in complex terrain and vegetated environments (Huang et al., 2024d). These characteristics are especially advantageous for sub-canopy topography estimation in mangrove ecosystems, where structural complexity and signal noise are pronounced (Xu et al., 2025). Therefore, RF is adopted in this study as a robust and reliable framework for variable importance evaluation and subsequent modeling. After removing multicollinearity effects, a RF model is constructed to assess variable importance. The RF regression model consists of an ensemble of decision trees, with predictions obtained by averaging the outputs of individual trees, as expressed in Equation (2):

$$\hat{y}(a) = \frac{1}{T}\sum_{t=1}^{T} h_t(a) \quad (2)$$

where T is the number of decision trees, and $h_t(a)$ represents the prediction of the t-th regression tree for sample a. Variable importance is obtained by averaging the contribution of each variable to the reduction in node mean squared error (MSE) across all trees. Combining VIF analysis with variable importance ranking, 25 key factors with significant explanatory power for sub-canopy topography estimation were ultimately selected to form the final input feature set for the RF model.

The predictors included multi-source remote sensing features, covering terrain factors FABDEM elevation, slope, aspect, L-band SAR (ALOS-2) texture metrics (HH_Entropy, HH_Homogeneity, HH_Contrast, HH_Variance; HV_Entropy, HV_Homogeneity, HV_Contrast), C-band SAR (Sentinel-1) backscatter and texture metrics (VV_2020, VV_Dissimilarity, VV_Homogeneity; VH_ASM, VH_Correlation, VH_Variance), optical multispectral bands (Sentinel-2 B1, B5, B8, B9, B10), spectral indices (CVI, GI, SR), and forest height. These predictors jointly capture terrain variation, radar backscatter and texture, vegetation status, and forest vertical structure, enabling multi-scale characterization of spatial heterogeneity in mangrove sub-canopy topography.

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026

In the RF estimation of sub-canopy topography, ground elevations derived from ICESat-2 photon clouds were used as the dependent variable, while the 25 selected predictors served as input features to construct the regression model. All samples were spatially standardized and registered to ensure that each photon location corresponded to all predictor variables, forming a complete training dataset. Bootstrap sampling was applied to generate multiple training subsets with replacement. For a dataset containing N samples, each tree is trained on a bootstrap sample D_t:

$$D_t=\{(x_i,y_i)\}_{i=1}^{N},x_i \sim D \tag{3}$$

where x_i represents the predictor variables and y_i denotes the ground elevation derived from ICESat-2 photons. The RF model predicts ground elevation as the ensemble average of multiple regression trees:

$$\hat{y}(x)=\frac{1}{T}\sum_{t=1}^{T} f_t(x) \tag{4}$$

where $\epsilon_t$represents tree-specific prediction noise, $f_t(x)$ represents the prediction of the t-th regression tree, and T is the total number of trees. At each node, a random subset of m features (m=p/3) is selected to determine the optimal split. The hyperparameters of the RF model were set to commonly used default values without extensive tuning. The optimal split s^*maximizes variance reduction:

$$s^* = arg\ \max_{s}\left[Var(Y_{parent}) - \frac{n_L}{n}Var(Y_L) - \frac{n_R}{n}Var(Y_R)\right] \tag{5}$$

where Y_{parent} is the set of response values in the parent node, Y_L and Y_R are the left and right child nodes after splitting, and n_L,n_R,n are the corresponding sample sizes. RF also allows internal error estimation using OOB samples:

$$OOB - MSE = \frac{1}{N}\sum_{i=1}^{N} (y_i - \hat{y}_{OOB}(x_i))^2, \hat{y}_{OOB}(x_i) = \frac{1}{|\mathcal{T}_i|}\sum_{t\in\mathcal{T}_i} f_t(x_i) \tag{6}$$

where $\mathcal{T}_i$is the set of trees for which $x_i$was not included in the bootstrap sample. The total variance of the ensemble prediction can be expressed as:

$$Var[\hat{y}(x)] = \frac{1}{T^2}\sum_{t=1}^{T} Var[f_t(x)] + \frac{1}{T^2}\sum_{t\neq t'} Cov[f_t(x), f_{t'}(x)] \tag{7}$$

where $f_t(x)$and $f_{t'}(x)$denote predictions of the t-th and t'-th trees, $Var[f_t(x)]$ is the variance of a single tree prediction, and $Cov[f_t(x),f_{t'}(x)]$ represents the covariance between predictions of different trees. This formula quantifies the total variance of the ensemble prediction, including both tree-level variance and correlation effects. This formulation allows quantifying both tree-level variance and correlation effects in ensemble prediction. Finally, the expected prediction error can be decomposed into bias, variance, and irreducible noise:

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026

$$E[(y-\hat{y}(x))^2]=Bias[\hat{y}(x)]^2+Var[\hat{y}(x)]+\sigma_\epsilon^2 \quad (8)$$

where y is the observed value, $\hat{y}(x)$ is the model prediction, $Bias[\hat{y}(x)]^2$ measures the squared systematic error, $Var[\hat{y}(x)]$ denotes the variance of the prediction across different training sets, and σ_ϵ^2 is the irreducible observational noise. This decomposition separates the total expected prediction error into bias, variance, and noise components.

After model training and validation, the optimized RF model was applied to all grid cells across the study area. The point dataset used in this study was divided into a training set and a validation set, with seventy percent of the points allocated to model training and the remaining thirty percent used for validation. Both sets contain substantial number of points that reflect the overall distribution of terrain and vegetation conditions within the study area. The division was performed randomly, ensuring that training and validation points are spatially distributed throughout the region. This design allows the RF model to learn from a wide range of environmental conditions while providing a representative assessment of its predictive performance on independent locations. Predictor values were extracted for each grid cell and input into the model to generate spatially continuous elevation predictions, which were subsequently reconstructed into raster format to produce the final sub-canopy topography DTM.

3.3 Model applications

The goal of this study is to generate a seamless national-scale dataset of mangrove sub-canopy topography by reconstructing the terrain beneath mangrove forests using multi-source satellite data. To achieve this, a targeted modeling strategy was adopted that integrates ICESat-2-derived ground elevation constraints with multi-source remote sensing features to produce the RFDTM dataset.

The specific workflow involved two main stages. First, the preprocessed multi-source features, including terrain variables, Sentinel-2 spectral information, C-band SAR backscatter, and L-band SAR texture features, were input into the trained and optimized Random Forest model. This step utilized high-confidence ground photons extracted by the HMCF strategy as the dependent variable, enabling pixel-level elevation prediction across all mangrove-covered areas and generating spatially continuous estimates of the sub-canopy terrain. Second, these predicted elevations were mosaicked and spatially reconstructed to form a consistent gridded surface representing the mangrove sub-canopy topography at 30 m resolution. By treating ICESat-2 photon-derived elevations as physically constrained “anchor points” and machine learning predictions as spatially continuous “interpolators,” the framework produces a coherent and seamless digital terrain model of the mangrove understory. The final RFDTM product captures terrain variation from tidal creek zones to interior mangrove platforms while effectively reducing vegetation-induced bias and preserving fine-scale microtopographic features. Ultimately, a complete, continuous, and high-accuracy dataset of mangrove sub-canopy topography is generated at the national scale.

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026

Open Access Earth System Science Data Discussions

4 Results

4.1 Qualitative analysis of sub-canopy topography estimation results

Based on the constructed estimation model, a digital terrain model (RFDTM) for the study area was generated, as shown in the results in Fig. 3. The comparison between RFDTM and existing digital elevation model (DEM) products, including FABDEM, SRTM, NASADEM, ASTER GDEM, TanDEM-X, DiluviumDEM, and DeltaDTM, not only highlights the accuracy and applicability of RFDTM but also reveals differences and potential biases among various DEM datasets in terrain representation, providing important insights for subsequent accuracy evaluation and application analysis.

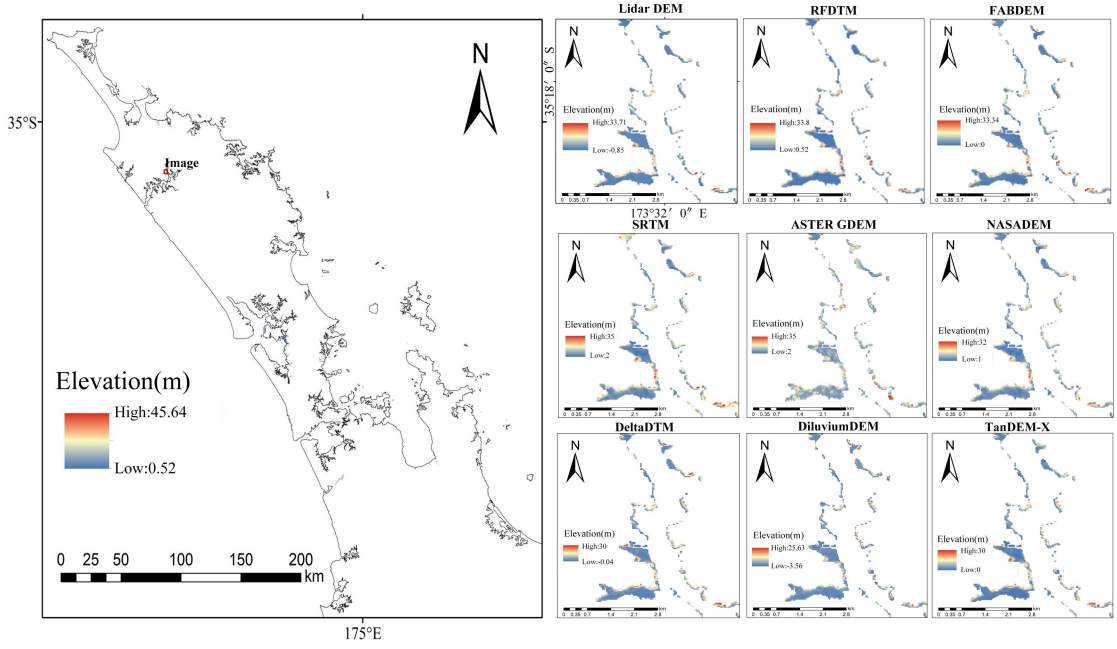


Fig. 3. Comparison of DEM results. Left: overall RFDTM after inversion. Right: comparison of RFDTM with DTMs for two sample areas.

As shown in Fig. 3, the left panel presents the spatial distribution of the inverted mangrove sub-canopy topography at the national scale across New Zealand, revealing a coherent large-scale topographic pattern with clear elevation gradients along the coastal domain. The RFDTM captures the overall geomorphological structure of mangrove systems nationwide, including tidal channels, intertidal flats, and inland transition zones. Low-elevation areas are predominantly distributed along tidal creeks and estuarine margins, while relatively higher elevations appear in interior mangrove platforms, indicating a gradual landward increase in terrain height. The continuity and smoothness of the spatial pattern at the national scale suggest that the estimation framework effectively suppresses noise while preserving meaningful geomorphic features across geographically diverse coastal environments.

The right panel provides a detailed comparison of DEMs within a representative local area embedded in the broader national-scale framework, characterized by complex tidal creek networks. Despite the consistency achieved at the national scale,

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026

substantial differences remain among datasets in their ability to resolve fine-scale terrain structures. Conventional DEM products tend to exhibit oversmoothing or structural distortion, where narrow tidal channels are blurred or partially missing, and branching patterns appear simplified. In contrast, the RFDTM demonstrates a clearer delineation of channel morphology, with sharper boundaries and more continuous network structures. The spatial configuration of dendritic creek systems is more realistically represented, indicating an enhanced capacity to capture micro-topographic variability beneath vegetation cover. This consistency from national-scale patterns to local-scale detail highlights the robustness and scalability of the RFDTM framework for mangrove sub-canopy topography mapping across New Zealand.

4.2 Quantitative assessment of DEMs

Table 3 presents the evaluation results of the proposed RFDTM's accuracy compared to several widely used DEM products, including FABDEM, SRTM, NASADEM, ASTER GDEM, TanDEM-X, DiluviumDEM, and DeltaDTM, based on four statistical indicators: R^2, RMSE, MAE, and bias. These indicators quantify the elevation differences relative to the reference LiDAR terrain data. Additionally, Fig. 4 displays scatter plots that illustrate the relationship between each DEM and the reference elevations, visually highlighting the differences and trends in elevation accuracy among the datasets.

Table 3 Accuracy comparison between the inverted RFDTM and DEMs.

Data	R^2	RMSE(m)	MAE(m)	Bias(m)
RFDTM	**0.99**	**1.01**	**0.80**	**0.43**
FABDEM	0.98	3.87	1.79	1.08
SRTM	0.87	5.28	4.27	4.13
NASADEM	0.89	6.95	4.62	4.28
ASTER GDEM	0.73	9.30	7.87	-7.16
TanDEM-X	0.82	8.64	6.71	-6.28
DiluviumDEM	0.97	3.66	2.25	-0.97
DeltaDTM	0.75	2.73	2.24	1.93

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026

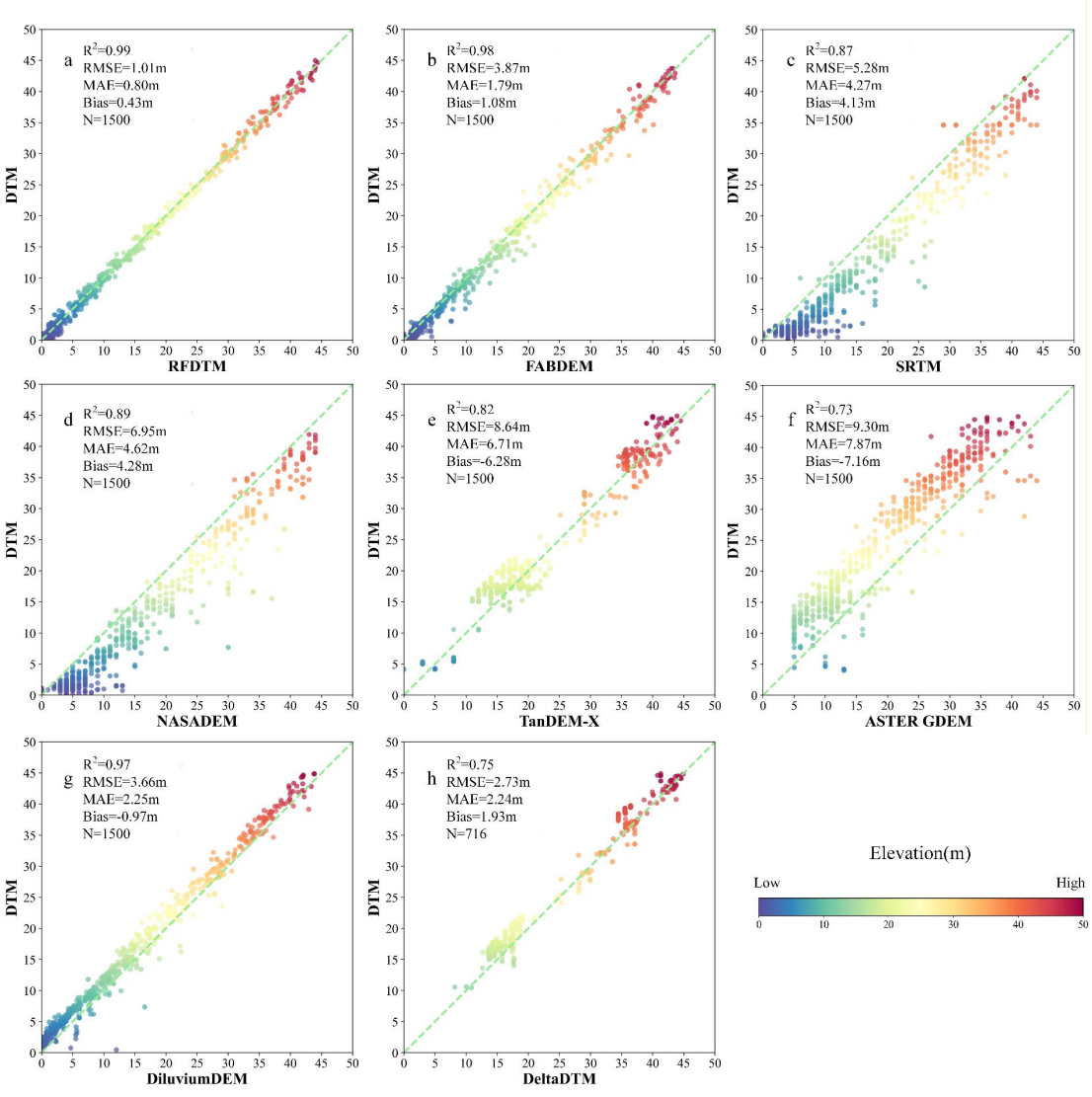


Fig. 4. Accuracy Comparison Scatter Plot. (a) RFDTM scatter plot; (b) FABDEM scatter plot; (c) SRTM scatter plot; (d) NASADEM scatter plot; (e) TanDEM-X scatter plot; (f) ASTER GDEM scatter plot; (g) DiluviumDEM scatter plot; (h) DeltaDTM scatter plot.

The results in Table 3 and Fig. 4, N denotes the number of scatter points. For DeltaDTM, only 716 scatter points due to its random distribution within a small coastal subset of the latitude–longitude range. The RFDTM derived from the estimation model in this study was subjected to a systematic accuracy analysis. RFDTM advantages significant across all indicators: R^2 = 0.99, RMSE = 1.01 m, MAE = 0.80 m, Bias = 0.43 m. High consistency between the inverted DTM and measured terrain, and far superior precision to existing DEM products.

Compared with other DEMs, the superiority of RFDTM is even more pronounced. FABDEM exhibits a relatively high overall correlation R^2 = 0.98 with RMSE and MAE of 3.87 m and 1.79 m, respectively, but it is still clearly inferior to RFDTM. The accuracies of SRTM, NASADEM, ASTER GDEM, and TanDEM-X are lower, with RMSE values ranging from 5.28 to

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026

9.30 m and considerable bias, indicating systematic overestimation or underestimation. These limitations are largely attributable to their data acquisition and processing mechanisms. Radar interferometry and optical stereo matching are strongly affected by vegetation canopy, layover, and shadowing, which reduces vertical precision in forested areas (Huang et al., 2024c; Huang et al., 2025). DiluviumDEM and DeltaDTM show slightly better accuracy RMSE 2.73–3.66 m, likely due to ground filtering or interpolation, yet they still cannot match the fine-scale precision of ICESat-2–derived RFDTM, which directly captures forest sub-canopy topography through photon measurements combined with multi-source predictors (Dusseau et al., 2023).

From the overall comparison in Table 3 and Fig. 4, the RFDTM inverted in this study demonstrates excellent performance in local elevation recovery, overall terrain continuity, and error control. It more accurately reflects the terrain characteristics of the study area, providing a reliable data basis for subsequent terrain analysis and related applications, thereby validating the effectiveness and high-precision advantage of the estimation model.

4.3 Ablation experiment results analysis

Table 4 presents the precision comparison of the ablation experiments on the RFDTM for mangrove sub-canopy topography estimation. The results demonstrate the contribution of multi-frequency SAR data to model accuracy, with the full model incorporating both C- and L-band SAR features showing the best performance in terms of R^2, RMSE, MAE, and Bias. In contrast, the model using only C-band SAR data (without L-band SAR data) exhibited lower accuracy, highlighting the importance of multi-source SAR data fusion. Fig. 5 illustrates the detailed performance of each model, showing the differences in topography estimation under varying SAR data configurations. The scatter plots clearly reveal how the inclusion of L-band SAR data improves the stability and precision of sub-canopy topography estimation compared to the single C-band model.

Table 4. Accuracy comparison of ablation experiments.

Ablation Setting	R^2	RMSE(m)	MAE(m)	Bias(m)
L- and C-band retained	0.99	1.01	0.80	0.43
Without L-band SAR data	0.99	1.23	1.04	0.61
Without C- and L-band SAR data	0.99	1.30	1.05	0.62

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026

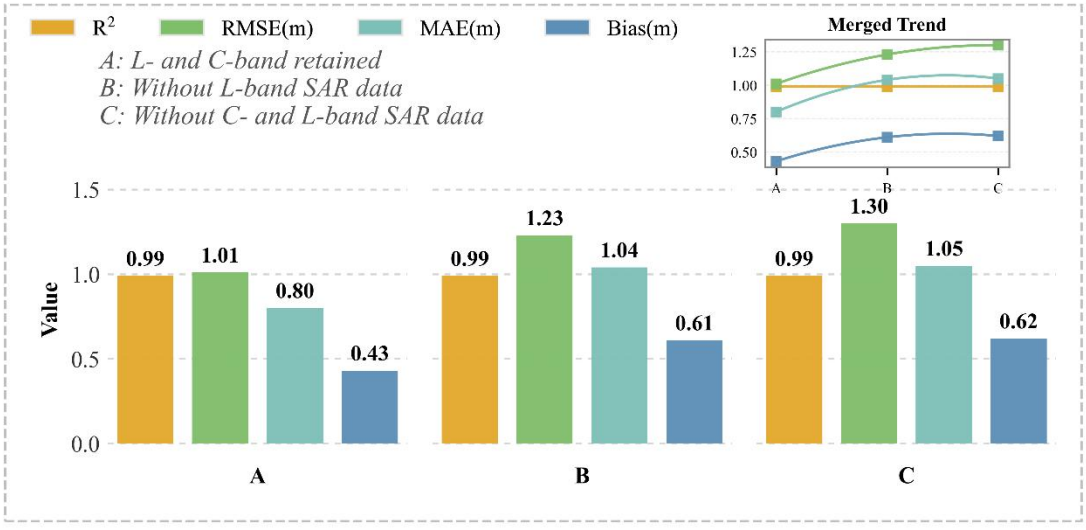


Fig. 5. Comparison of ablation experiment accuracy. Scenario A (L- and C-band), Scenario B (C-band), and Scenario C (no SAR).Four metrics:R², RMSE, MAE, and bias are shown, with the inset merged trend summarizing overall variations.

From the results shown in Table 4 and Fig. 5, it can be seen that all three model schemes achieve a coefficient of determination $R^2 = 0.99$, indicating that the overall explanatory power of the models remains largely consistent across different variable combinations, and the ablation of SAR bands has little effect on overall fit. However, error-based metrics are more sensitive to variable changes and better reflect the actual contribution of different frequency bands to estimation accuracy.

When both L-band and C-band data are retained, the model achieves the lowest RMSE = 1.01 m, MAE = 0.80 m, and Bias = 0.43 m, indicating that multi-frequency SAR data fusion effectively reduces both random errors and systematic bias. After removing the L-band, RMSE increases to 1.23 m, MAE to 1.04 m, and Bias to 0.61 m, showing a significant rise in error, which demonstrates that the L-band plays a key role in enhancing vegetation penetration and improving the extraction accuracy of sub-canopy topography information. Further removal of both L- and C-band data results in RMSE = 1.30 m, MAE = 1.05 m, and Bias = 0.62 m, with the largest model errors observed, indicating that the absence of SAR information weakens the model's ability to capture fine-scale terrain undulations under the canopy (Wang et al., 2025a). This is primarily because radar backscatter is sensitive to both canopy structure and ground surface characteristics, and without SAR data, the model relies solely on optical or DEM-derived predictors, which have limited penetration through vegetation (Luo et al., 2023). Specifically, L-band SAR signals have longer wavelengths, allowing partial penetration of the forest canopy and providing information about sub-canopy terrain morphology, while C-band SAR signals are more sensitive to surface roughness and canopy texture, offering complementary structural cues (Yu et al., 2025).

In summary, although R^2 is insensitive to band ablation, the error metrics clearly demonstrate that multi-frequency SAR data contribute significantly to improving the precision of sub-canopy topography estimation, with the L-band playing a dominant role in capturing sub-canopy elevation variation and the C-band providing synergistic information that enhances model robustness and reduces systematic errors (Gautam et al., 2025; Fu et al., 2025).

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026

5 Discussion

5.1 Advantages of the dataset

The advantages and applicability of RFDTM derived in this study for mangrove areas are reflected in its accuracy relative to multiple DEM products. The accuracy comparison scatter plots are shown in Fig. 6. As shown in Fig. 6, RFDTM significantly outperforms existing DEM products across all evaluation metrics. It achieves an $R^2 = 0.99$, with an RMSE = 1.01 m, an MAE = 0.80 m, and Bias = 0.43 m. In contrast, traditional products such as ASTER GDEM (RMSE = 9.30 m) and SRTM (RMSE = 5.28 m) exhibit pronounced systematic deviations, mainly due to canopy-induced elevation overestimation or underestimation (Han et al., 2021). By directly constraining ICESat-2 photon data and integrating multi-source features, RFDTM effectively mitigates vegetation-induced distortions, enabling more accurate reconstruction of sub-canopy terrain features in complex mangrove ecosystems (Huang et al., 2025). Overall, RFDTM demonstrates both superior accuracy and enhanced stability across varying topographic conditions, making it well suited for ecological monitoring, hydrological modeling, and carbon stock assessment in mangrove-dominated coastal zones, where conventional DEMs often fail to capture subtle yet ecologically significant terrain variations (Ali et al., 2025).

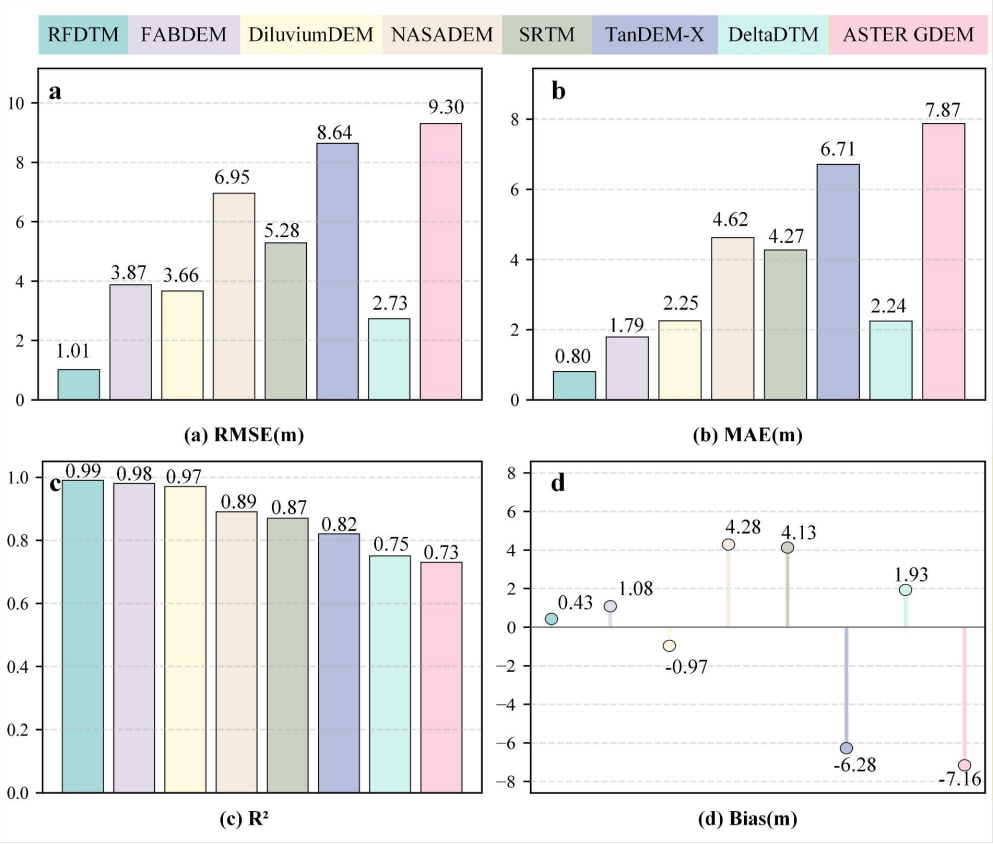
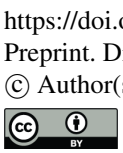


Fig. 6. Accuracy comparison and performance evaluation of DEMs. (a) RMSE bar chart, (b) MAE bar chart, (c) R^2 bar chart, (d) Bias lollipop chart.

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026



Open Access Earth System Science Data Discussions

In representative regions, Fig. 7 shows the elevation maps of RFDTM and LiDAR DEM for areas a and b in the right panels, along with scatter plots comparing the two datasets. The scatter plots demonstrate a high correlation, with $R^2 = 0.99$, and low error levels, with RMSE = 0.94 m, MAE = 0.76 m, and bias = 0.36 m, indicating a high level of agreement between RFDTM and the reference LiDAR DEM. This confirms that the model effectively captures terrain variations beneath the mangrove canopy and preserves fine-scale terrain morphology, rather than merely reproducing general elevation patterns, demonstrating the robustness of the RF-based framework under complex vegetation conditions (Wu et al., 2023; Ming et al., 2025). The improved performance arises from the synergistic integration of multi-source remote sensing data and the nonlinear modeling capability of the random forest. ICESat-2 photon measurements provide reliable ground elevation constraints, while multi-frequency SAR features—particularly the strong penetration of L-band and the structural sensitivity of C-band—contribute complementary information on canopy structure and surface roughness (Nandy et al., 2021). RFDTM also achieves a low overall bias of 0.43 m, reflecting improved stability over conventional DEMs limited by interpolation and indirect terrain representation. By combining direct photon-based constraints with multi-source predictors, RFDTM better resolves sub-canopy terrain variability (Pronk et al., 2024), making it effective for representing subtle topographic variations critical to mangrove ecosystem processes.

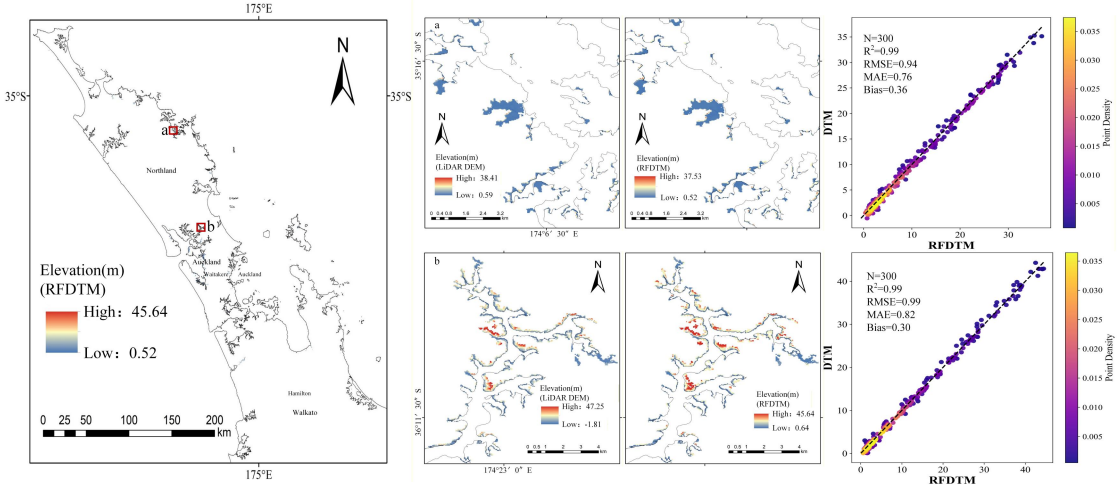


Fig. 7. Mangrove sub-canopy topography profiles in two representative regions. Spatial distribution and validation of mangrove sub-canopy topography derived from the RF model (RFDTM). The left panel shows the study area and the locations of two representative sample regions (a, b). The right panels show the resulting maps of RFDTM and LiDAR DEM for areas a and b, as well as scatter plots that compare them.

RFDTM demonstrates strong potential for regional-to-global mangrove terrain mapping, particularly for sub-canopy topography estimation. Built entirely on spaceborne remote sensing data—including LiDAR, multi-source SAR, and optical imagery—the framework offers broad spatial coverage and long-term observations, enabling high scalability beyond regional applications. Compared to approaches relying on airborne LiDAR or field surveys, RFDTM is more suitable for large-scale mapping, especially in tropical and subtropical mangrove regions where complex canopy structures, tidal dynamics, and muddy surfaces often introduce significant errors in conventional DEMs. By effectively reducing vegetation- and surface-induced biases, the method shows strong applicability across global mangrove environments (Simard et al., 2025). Moreover, its multi-source, machine learning–based design is highly transferable, as it can be readily migrated or retrained using publicly available datasets, providing a practical pathway for large-scale implementation and supporting applications in mangrove ecosystem analysis, coastal geomorphology, and blue carbon assessment.

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026

Open Access Earth System Science Data Discussions

5.2 Uncertainty analysis of our results

Fig. 8 shows the sensitivity of the estimation results to different input variables. The variations in topography estimation highlight how changes in key parameters, such as SAR data sources, model configurations, and processing settings, can impact the accuracy and reliability of the RFDTM beneath the mangrove canopy. Understanding these sensitivities is essential for evaluating the robustness of the derived topography and for improving future model performance.

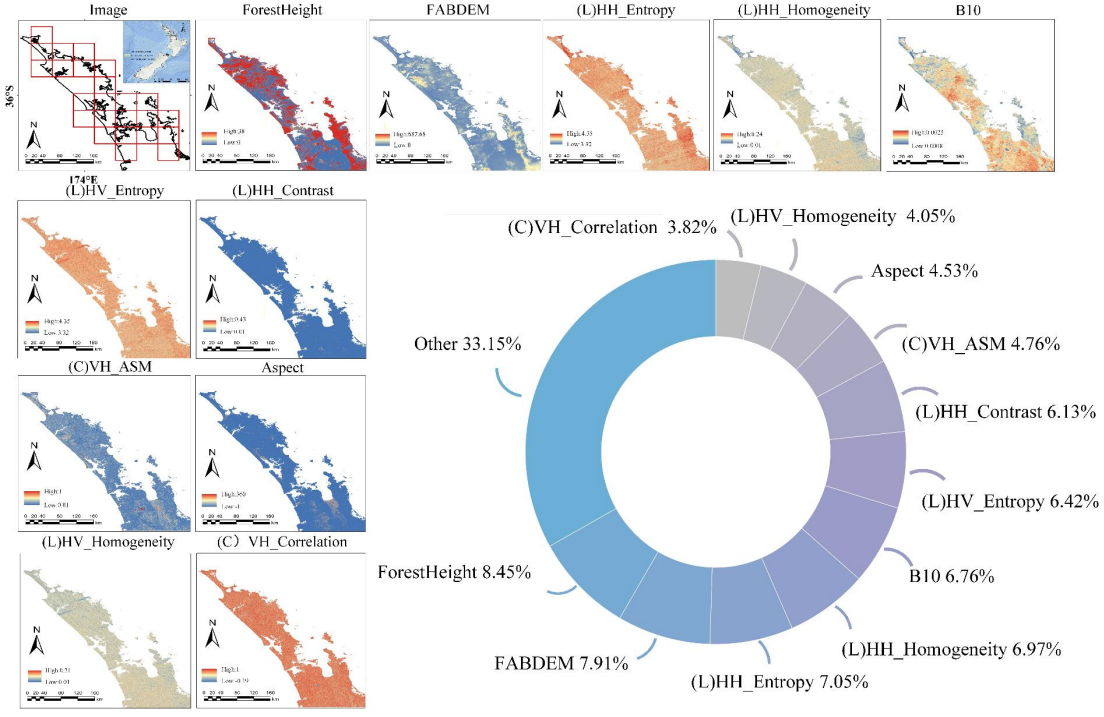
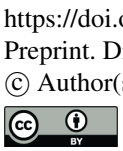


Fig. 8. Variable importance rose chart. The chart shows the relative contribution (in percentage) of each input variable to the RFDTM model (summing to 100%). The top four variables are ForestHeight (8.45%), FABDEM (7.91%), (L) HH_Entropy (7.05%), and (L) HH_Homogeneity (6.97%).

As shown in Fig. 8, variable importance analysis indicates that ForestHeight (8.45%) and FABDEM (7.91%) are the dominant factors, reflecting the key roles of canopy structure and baseline terrain in explaining the complex vegetation–terrain coupling in mangrove areas. Several L-band SAR polarization texture features also contribute substantially (6.13%–7.05%), indicating that their strong penetration capability captures understory structural information and supports sub-canopy microtopography estimation. In contrast, C-band SAR features have lower contributions (3.82%–4.76%) due to their limited penetration and higher sensitivity to canopy structure. Terrain aspect (4.53%) also plays a role, likely by influencing vegetation distribution and local microtopographic conditions. Overall, variable contributions are relatively dispersed, with remaining factors accounting for 33.15%, highlighting that sub-canopy terrain estimation is driven by multi-source data integration with both complementary and redundant information. This underscores that multi-source fusion enhances model stability, while inter-variable correlations may introduce some uncertainty (Fu et al., 2025; Yi et al., 2025).

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026



Open Access Earth System Science Data Discussions

From the perspective of data sources, RFDTM uncertainty is mainly driven by the quality, physical limitations, and inconsistencies of the input and reference datasets. The reference DEMs used for evaluation—including FABDEM, SRTM, NASADEM, TanDEM-X, ASTER GDEM, DiluviumDEM, and DeltaDTM—originate from different sensors and processing workflows, resulting in varying sensitivities to vegetation cover, terrain relief, and signal penetration (Pronk et al., 2024). For instance, C-band products such as SRTM and NASADEM have limited canopy penetration and tend to overestimate elevations in dense mangrove areas, while X-band TanDEM-X and optical-based ASTER GDEM are more affected by surface scattering and noise (Brochado et al., 2024). Although FABDEM and DiluviumDEM reduce vegetation-induced bias through post-processing, residual errors persist due to model assumptions and training data limitations (Vernimmen et al., 2023). These cross-sensor differences lead to heterogeneous error characteristics across DEM products. As shown in Fig. 9, traditional DEMs tend to exhibit broader and skewed error distributions, whereas RFDTM errors are more concentrated and symmetric.

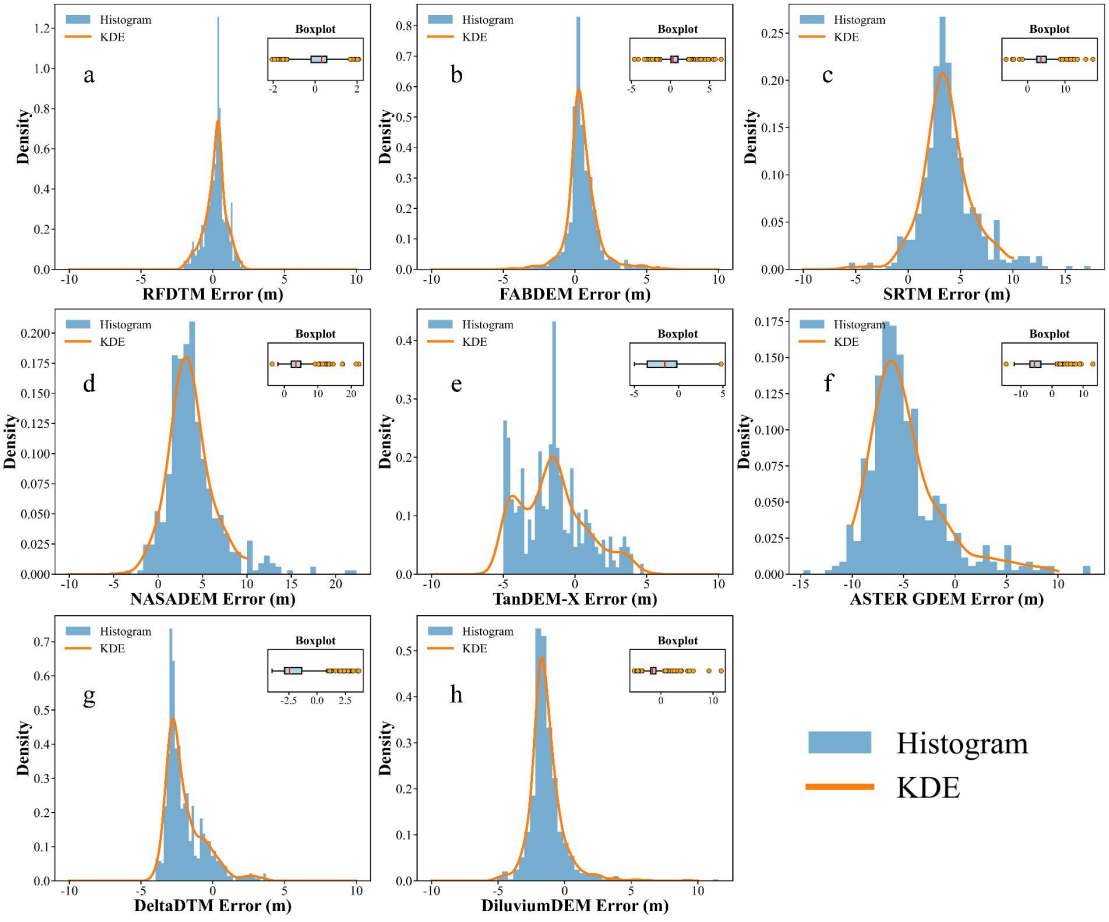


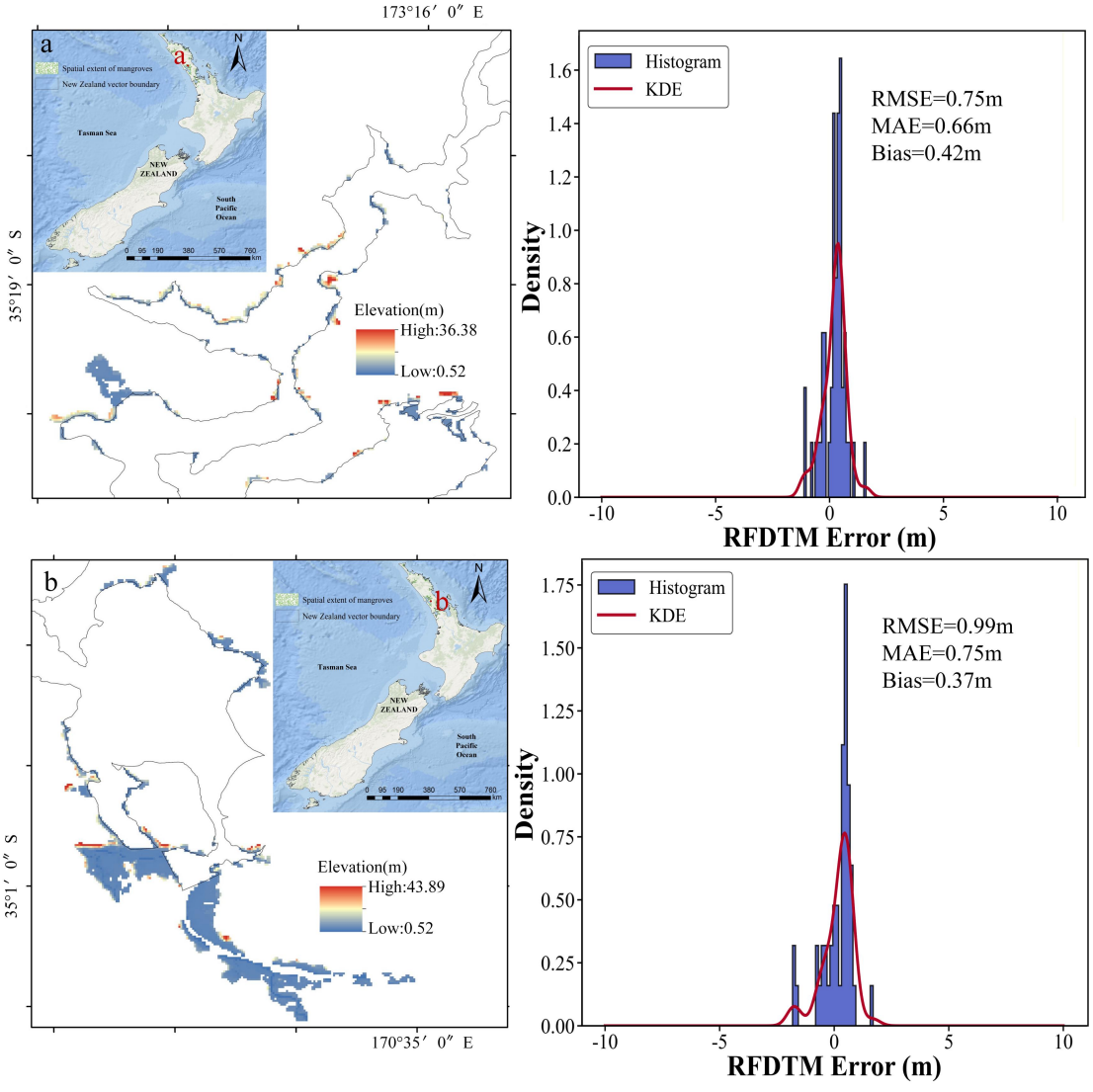
Fig. 9. Error distributions of the evaluated DEMs: (a) RFDTM, (b) FABDEM, (c) SRTM, (d) NASADEM, (e) TanDEM-X, (f) ASTER GDEM, (g) DeltaDEM, and (h) DiluviumDEM. The subplots illustrate differences in error magnitude, spread, and symmetry, with RFDTM showing a more concentrated and symmetric distribution compared to other DEMs. The boxplots illustrate the statistical characteristics of the errors: the box represents the interquartile range (IQR), the horizontal line inside the box indicates the median, the whiskers extend to the most extreme data points within 1.5 times the IQR, and circles denote outliers.

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026

Moreover, differences in data resolution further amplify uncertainty. The evaluated DEMs typically have resolutions ranging from 10 m to 30 m, while RFDTM is generated by fusing multiple predictive variables and validated using ICESat-2 along-track measurements, achieving a unified grid scale. The scale mismatch between point-based reference data and gridded DEMs introduces scale-dependent representation errors, which are particularly pronounced in mangrove environments with fine microtopography and steep terrain variations (Renshaw et al., 2026). Coarser DEMs tend to smooth terrain changes, making it difficult to capture sub-canopy microtopography, whereas higher-resolution DEMs may contain more noise and local artifacts. Consequently, the observed uncertainty reflects not only model performance but also the combined effects of data resolution, spatial sampling, and terrain complexity (Dusseau et al., 2023).

Despite these limitations, the RFDTM error distribution shown in Fig. 9 is more concentrated and symmetric than all reference DEMs, indicating higher stability and lower uncertainty in sub-canopy terrain reconstruction. This advantage can be attributed to the RF model's ability to exploit multi-source complementary information and mitigate bias from individual datasets through ensemble learning.

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026



505

Fig. 10. Comparison of RFDTM error distribution and spatial validation results in two representative regions (a and b). (a) Spatial distribution map (left) and corresponding error statistics distribution (right) for region a; (b) spatial distribution map (left) and corresponding error statistics distribution (right) for region b. The left subplots illustrate the spatial distribution of mangroves and the variation in elevation within the study areas, where the color bar represents the elevation range. The study area boundary and 510 **a geographical location inset are also overlaid to indicate the spatial position of each sample site within the coastal mangrove distribution of New Zealand. The right subplots present the statistical distribution of RFDTM prediction errors, including histograms and kernel density estimation (KDE) curves, which are used to characterize the probability distribution and central tendency of the errors.**

From the perspective of vegetation cover differences shown in Fig. 10, region a shows an RMSE = 0.75 m, MAE = 0.66 m, 515 and bias = 0.42 m. This is likely due to its relatively stable environment, where weaker signal disturbance reduces random errors, although residual inversion-related systematic effects still lead to a slightly higher bias (Colliander et al., 2026). In contrast, region b exhibits an RMSE = 0.99 m, MAE = 0.75 m, and bias = 0.37 m. This may be because stronger multiple

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026

scattering and weaker ground penetration increase data variability and RMSE, while partial error compensation keeps the bias relatively low. Overall, vegetation cover plays a key role in shaping differences in estimation accuracy and error structure between the two regions (Kurum et al., 2021).

From a process-based perspective, tidal dynamics introduce systematic uncertainty in RFDTM estimation in New Zealand mangrove environments by modulating surface scattering, vegetation–water interactions, and elevation reference conditions (Ming et al., 2025). ICESat-2 ATL03/ATL08 observations capture photon returns under varying tidal stages, where inundation shifts the effective reflecting surface from ground to a mixed water–vegetation interface, inducing vertical biases in low-elevation intertidal zones (Renshaw et al., 2026). Hydrodynamic studies further show that tidal levels are moderately correlated with non-tidal residuals, with Kendall's τ ranging from −0.35 to −0.60, tide–surge interactions of 0–10 cm up to approximately 27 cm under extreme conditions, estuarine amplification averaging about 18 cm and reaching up to 32 cm, and dampening effects of 2–6 cm, indicating centimeter- to decimeter-scale variability driven by local morphology (Costa et al., 2023). Consistently, incorporating tidal inundation can constrain elevation uncertainty to the centimeter level, with ICESat-2–Sentinel-2 based estimates achieving RMSE within 7.5 cm against airborne LiDAR, whereas in macro-tidal systems such as the Zambezi Delta, uncorrected tides greater than 2 m may induce decimeter-to-meter scale errors (Lagomasino et al., 2016).

Within this context, multi-source predictors in RFDTM provide a partial compensatory mechanism. L-band SAR retains relatively stable sub-canopy structure due to stronger penetration, while C-band SAR and optical features are more sensitive to surface water dynamics and thus more affected by tidal variability (Hübinger et al., 2026; Battaglia et al., 2025). Through feature fusion, tide-insensitive information can partially offset tide-driven noise; however, because all inputs are acquired under temporally varying tidal conditions, residual tidal effects cannot be fully eliminated.

Although RFDTM achieves high accuracy in this study area, its generalization is constrained by the representativeness of training samples, with reduced performance in regions with differing vegetation or environmental conditions (Huang et al., 2024b; Jin et al., 2018). Variations in remote sensing data quality, tidal regimes, and mangrove structure further affect cross-regional stability (Li et al., 2022b), highlighting the need for expanded multi-region training and transfer learning or region-specific modeling to improve generalization.

5.3 Potential applications

The potential advantages of RFDTM are significant and wide-ranging, making it a valuable tool for both scientific research and practical applications. One of its key strengths is its ability to produce highly accurate terrain models, even in complex environments such as mangrove-dominated coastal zones, where traditional DEMs often struggle (Yao et al., 2024). RFDTM effectively reduces both random and systematic errors, achieving an RMSE of less than 7.5 cm and an R^2 value greater than 0.89 when validated against airborne LiDAR data, thereby improving terrain data precision and enabling more accurate ecological monitoring, hydrological modeling, and carbon stock assessments (Ming et al., 2025). Its high resolution also allows for more detailed habitat mapping, which is critical for understanding the intricate relationships between mangrove ecosystems and surrounding environments.

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026

In addition, RFDTM's integration of multi-source remote sensing data, including ICESat-2 LiDAR, SAR, and Sentinel-2 imagery, enhances its adaptability and robustness. This fusion of data sources allows RFDTM to capture fine-scale terrain variations, even in densely vegetated or flooded areas where other DEMs may fail to provide accurate representations (Yao et al., 2024). The versatility of RFDTM makes it particularly useful for assessing the impacts of climate change and human activities on coastal ecosystems, as well as for supporting disaster management efforts, especially in regions that are difficult to access or monitor through traditional methods.

Furthermore, RFDTM's transferability across different regions and its ability to maintain consistent performance across diverse topographic conditions underscore its potential for global-scale applications. It is not only a powerful tool for mangrove terrain mapping but also an essential resource for ecosystem conservation, land-use planning, and environmental policy-making (Ming et al., 2025). By offering a reliable and comprehensive approach to terrain estimation, RFDTM provides valuable insights into the dynamics of mangrove and coastal ecosystems, supporting sustainable management and protection efforts on both regional and global scales.

6. Data availability

All key independent variables used in this study have been deposited and are openly available from Zenodo at https://zenodo.org/records/20651831.

7. Conclusion

To address the persistent lack of accurate and spatially continuous sub-canopy terrain information in mangrove ecosystems, this study produced RFDTM, the first national-scale 30 m mangrove sub-canopy topography dataset for New Zealand generated entirely from publicly available satellite observations. RFDTM provides a consistent, high-resolution representation of mangrove terrain beneath dense canopies, effectively reducing the elevation uncertainties and terrain omissions commonly found in conventional global Digital Elevation Models (DEMs) over mangrove regions. This is achieved through a large-scale multi-source fusion framework integrating ICESat-2 ATL03/ATL08 photon-counting LiDAR, dual-frequency L-band and C-band SAR observations, Sentinel-2 multispectral imagery, FABDEM, and GEDI v27 products. A Hierarchical Multi-Constraint Filtering (HMCF) strategy was developed to extract reliable ground photons beneath dense vegetation using physical and geometric constraints, including signal-to-noise ratio, terrain slope, cloud conditions, and photon spatial distribution, providing high-confidence elevation control points for terrain reconstruction. Multi-source features capturing terrain morphology, vegetation structure, and microwave penetration characteristics were further constructed and optimized through variance inflation factor (VIF) analysis and Random Forest importance ranking, ensuring physically consistent and stable sub-canopy elevation estimation. The framework was implemented across all mangrove regions of New Zealand to generate a seamless 30 m terrain product with national-scale coverage and strong spatial coherence.

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026

Validation against airborne LiDAR terrain data demonstrates excellent performance, with $R^2 = 0.99$, RMSE = 1.01 m, MAE = 0.80 m, and bias = 0.43 m, confirming the reliability of RFDTM for regional- to national-scale coastal applications. Ablation experiments further indicate that L-band SAR plays a dominant role in sub-canopy terrain retrieval, while C-band SAR provides complementary constraints, reducing RMSE from 1.23 m to 1.01 m and improving reconstruction capability under dense canopy conditions. The framework relies exclusively on openly accessible satellite observations, ensuring strong scalability, reproducibility, and operational feasibility without dependence on airborne data. In conclusion, RFDTM fills a critical coastal topographic data gap by providing a standardized, high-accuracy, and spatially continuous 30 m mangrove sub-canopy terrain product for New Zealand. The proposed methodology offers a transferable framework for large-scale sub-canopy terrain reconstruction, supporting coastal vulnerability assessment, hydrological modeling, ecosystem monitoring, and blue carbon studies across global mangrove systems.

Author contributions

The study was conceptualized by JH and NX. YW developed the methodology, implemented the software, performed the formal analysis, and curated the data. YZ and YHW contributed to methodology development. YW created the figures and tables and wrote the original draft. YZ and HH contributed to visualization. CC, HZ, JC and QL provided resources. JH, NX and QL contributed to project administration, funding acquisition and supervision. All authors contributed to the interpretation of the results, revised the manuscript, and approved the final version.

Competing interests

The authors declare that they have no conflict of interest.

Disclaimer

Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. The authors bear the ultimate responsibility for providing appropriate place names. Views expressed in the text are those of the authors and do not necessarily reflect the views of the publisher.

Acknowledgements

We thank the National Snow and Ice Data Center (NSIDC) for providing ICESat-2 ATL03/ATL08 data (https://nsidc.org/data/). Multi-source remote sensing datasets, including ALOS-2 PALSAR-2, Sentinel-1, Sentinel-2, GEDI canopy height products, and GLC_FCS30-2020 land cover data, were accessed via the Google Earth Engine platform

(https://code.earthengine.google.com/). The GLC_FCS30-2020 dataset is also available at https://doi.org/10.5281/zenodo.3986872. High-resolution airborne LiDAR DEM data used for validation were provided by Land Information New Zealand (LINZ) (https://data.linz.govt.nz/data/). In addition, global DEM products used for comparison include FABDEM (https://data.bris.ac.uk/data/dataset/s5hqmjcdj8yo2ibzi9b4ew3sn), SRTM (https://dwtkns.com/srtm30m/), NASADEM (https://lpdaac.usgs.gov/products/nasadem_hgtv001/), ASTER GDEM (https://search.earthdata.nasa.gov/search/), TanDEM-X (https://geoservice.dlr.de/web/dataguide/), DiluviumDEM (https://zenodo.org/records/8384665), and DeltaDTM (https://code.earthengine.google.com/). Special thanks are extended to all scientists and data managers involved in the acquisition, processing, and dissemination of these datasets, which made this study possible.

Financial support

This work was funded by the National Natural Science Foundation of China [42401488, 42571378, 42571520, 42071351], the National Key Research and Development Program of China [2020YFA0608501, 2017YFB0504204], the Liaoning Province Doctoral Research Initiation Fund Program [2023-BS-202], the Basic Research Projects of Liaoning Department of Education [JYTQN2023202], the Key Laboratory of Land Satellite Remote Sensing Application, Ministry of Natural Resources of the People's Republic of China [KLSMNR-K202309], the Scientific Foundation for Youth Scholars of Shenzhen University [868-000001033431], the Shenzhen Key Laboratory Program [SYSPG20241211173845013], and the Disciplines Breakthrough Project in Aerospace Information and Spatiotemporal Intelligence, MOE, China.

References

Ali, Y. & Rahman, M.M., 2025. Quantifying forest stocking changes in Sundarbans mangrove using remote sensing data. Sci. Remote Sens. (Science of Remote Sensing), 11, 100181. https://doi.org/10.1016/j.srs.2024.100181

Battaglia, M.J. & Bourgeau-Chavez, L.L., 2025. Evaluation of L- and S-Band Polarimetric Data for Monitoring Great Lakes Coastal Wetland Health in Preparation for NISAR. Remote Sens., 17(21), 3506. https://doi.org/10.3390/rs17213506

Brochado, G.T. & Rennó, C.D., 2024. New Method to Correct Vegetation Bias in a Copernicus Digital Elevation Model to Improve Flow Path Delineation. Remote Sens., 16(22), 4332. https://doi.org/10.3390/rs16224332

Chirico, P.G., Malpeli, K.C. & Trimble, S.M., 2013. Accuracy Evaluation of an ASTER-Derived Global Digital Elevation Model (GDEM) Version 1 and Version 2 for Two Sites in Western Africa. GIScience & Remote Sensing, 49(6), 775–801. https://doi.org/10.2747/1548-1603.49.6.775

Choi, C., 2024. Combining TanDEM-X interferometric SAR and GEDI lidar measurements for improving forest height, structure and biomass estimates [Dissertation]. ETH Zurich. https://doi.org/10.57676/v2fg-0z95

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026

Christopher, J., Jeannette, V. & Kelly, M., 2019. Radiometric calibration of a non-imaging airborne spectrometer to measure the Greenland ice sheet surface. Atmos. Meas. Tech., 12(3), 1913–1933. https://doi.org/10.5194/amt-12-1913-2019

Colliander, A., Schwank, M., Zhou, Y., et al., 2026. A review of forward modelling and retrieval approaches for forest soil moisture and vegetation optical depth using L-band radiometry. Remote Sensing of Environment, 334, 115158. https://doi.org/10.1016/j.rse.2026.115158

Dusseau, D., Zobel, Z. & Schwalm, C.R., 2023. DiluviumDEM: Enhanced accuracy in global coastal digital elevation models. Remote Sens. Environ., 298, 113812. https://doi.org/10.1016/j.rse.2023.113812

Farr, T.G. et al., 2007. The shuttle radar topography mission. Rev. Geophys., 45(2), RG2004. https://doi.org/10.1029/2005RG000183

Fu, B. et al., 2025. Estimation of mangrove heights and aboveground biomass using UAV-LiDAR, Sentinel-1 and ZY-3 stereo images. Ecol. Inform., 88, 103160.

Fu, L. et al., 2025. ICESat-2 Performance for Terrain and Canopy Height Retrieval in Complex Mountainous Environments. Remote Sens., 17(11), 1897.

Gautam, A. et al., 2025. Synergistic use of ICESat-2 lidar data and Sentinel-2 imagery for assessing hurricane-driven forest changes. Environ. Monit. Assess., 197(12), 1–25.

Han, H., Zeng, Q. & Jiao, J., 2021. Quality assessment of TanDEM-X DEMs, SRTM and ASTER GDEM on selected Chinese sites. Remote Sens., 13(7), 1304. https://doi.org/10.3390/rs13071304

Hawker, L. et al., 2022. A 30 m global map of elevation with forests and buildings removed. Environ. Res. Lett., 17(2), 024016. https://doi.org/10.1088/1748-9326/ac4d4f

Hu, L. et al., 2020. Advancing the mapping of mangrove forests at national-scale using Sentinel-1 and Sentinel-2 time-series data with Google Earth Engine: A case study in China. Remote Sens., 12(19), 3120. https://doi.org/10.3390/rs12193120

Huang, J., Cao, X., Yu, Y. et al., 2025. Understory Terrain Estimation Based on the Fusion of Multi-Source Remote Sensing Data and Machine Learning Models. IEEE Trans. Geosci. Remote Sens., in press.

Huang, J., Wang, Y. & Yu, Y., 2024a. Multi-Criteria Filtration and Extraction Strategy for Understory Elevation Control Points Using ICESat-2 ATL08 Product. Forests, 15, 2064. https://doi.org/10.3390/f15122064

Huang, J. & Yu, Y., 2024b. Understory Terrain Estimation by Synergizing Ice, Cloud, and Land Elevation Satellite-2 and Multi-Source Remote Sensing Data. Remote Sens., 16, 4770. https://doi.org/10.3390/rs16244770

Huang, J., Yu, Y., 2024c. Vertical accuracy assessment of the ASTER, SRTM, GLO-30, and ATLAS in a forested environment. Forests, 15(3), 426. https://doi.org/10.3390/f15030426

Huang, J., Zhang, Y. & Ding, J., 2024d. Combining LiDAR, SAR, and DEM Data for Estimating Understory Terrain Using Machine Learning-Based Methods. Forests, 15, 1992. https://doi.org/10.3390/f15111992

Huang, J., Xing, Y., Shuai, Y. et al., 2022. A Novel Noise Filtering Evaluation Criterion of ICESat-2 Signal Photon Data in Forest Environments. IEEE Geosci. Remote Sens. Lett., 19, 1–5.

Huber, M. et al., 2021. Shaping the global high-resolution TanDEM-X digital elevation model. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 14, 7198–7212. https://doi.org/10.1109/JSTARS.2021.3107187

Hübinger, C. et al., 2026. Exploring the potential of using L-band InSAR for mapping flooded vegetation in tropical wetlands. Remote Sens. Environ., 332, 115086.

Jiang, Z. et al., 2024. Tree parameter extraction in Fokienia hodginsii plantation based on airborne LiDAR data. Ying Yong Sheng Tai Xue Bao, 35(2), 321–329.

Jin, L. & Huang, H., 2020. Receiver performance and detection statistics of single photon LiDAR. IET Radar Sonar Navig., 14(1), 105–109. https://doi.org/10.1049/iet-rsn.2019.0205

Jin, S. et al., 2018. The Transferability of Random Forest in Canopy Height Estimation from Multi-Source Remote Sensing Data. Remote Sens., 10, 1183. https://doi.org/10.3390/rs10081183

Kurum, M., Kim, S. B., Akbar, R., et al., 2020. Surface soil moisture retrievals under forest canopy for L-band SAR observations across a wide range of incidence angles by inverting a physical scattering model. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14, 1741–1753. https://doi.org/10.1109/JSTARS.2020.2964213

Lagomasino, D. et al., 2016. A comparison of mangrove canopy height using multiple independent measurements from land, air, and space. Remote Sens., 8, 327. https://doi.org/10.3390/rs8040327

Li, B. et al., 2025. Global digital elevation model (GDEM) product generation by correcting ASTER GDEM elevation with ICESat-2 altimeter data. Earth Syst. Sci. Data, 17, 205–220. https://doi.org/10.5194/essd-17-205-2025

Li, L. et al., 2023. Digital elevation model correction method in forest areas considering elevation spatial autocorrelation based on a backpropagation neural network model. J. Geo-Inf. Sci., 25, 935–952.

Li, Q., Wong, F.K.K. & Fung, T., 2021. Mapping multi-layered mangroves from multispectral, hyperspectral, and LiDAR data. Remote Sens. Environ., 258, 112403. https://doi.org/10.1016/j.rse.2021.112403

Li, Q. et al., 2023. Assessment of active LiDAR data and passive optical imagery for double-layered mangrove leaf area index estimation: a case study in Mai Po, Hong Kong. Remote Sens., 15(10), 2551. https://doi.org/10.3390/rs15102551

Li, Q.H. et al., 2024. LightGBM hybrid model based DEM correction for forested areas. PLoS One, 19(10), e0309025. https://doi.org/10.1371/journal.pone.0309025

Li, Y. et al., 2022a. Comparison of Different Transfer Learning Methods for Classification of Mangrove Communities Using MCCUNet and UAV Multispectral Images. Remote Sens., 14, 5533. https://doi.org/10.3390/rs14215533

Li, Y. et al., 2022b. Research Progress and Prospects of Spaceborne Single Photon LiDAR for Shallow Water Bathymetry. Infrared Laser Eng., 1–13.

Li, Z. et al., 2022c. Entering a New Era of InSAR: Advanced Techniques and Emerging Applications. J. Geod. Geoinf. Sci., 5(1), 1–4. https://doi.org/10.11947/j.JGGS.2022.0101

Lin, C., 2019. Improved derivation of forest stand canopy height structure using harmonized metrics of full-waveform data. Remote Sens. Environ., 235, 111436. https://doi.org/10.1016/j.rse.2019.111436

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026

Liu, A., Cheng, X. & Chen, Z., 2021. Performance evaluation of GEDI and ICESat-2 laser altimeter data for terrain and canopy height retrievals. Remote Sens. Environ., 264, 112571. https://doi.org/10.1016/j.rse.2021.112571

Luong, V.N. et al., 2019. Biomass estimation and mapping of Can Gio Mangrove Biosphere Reserve in South of Viet Nam using ALOS-2 PALSAR-2 data. Appl. Ecol. Environ. Res., 17(1).

Luo, Y. et al., 2023. Mapping the forest height by fusion of ICESat-2 and multi-source remote sensing imagery and topographic information: A case study in Jiangxi province, China. Forests, 14(3), 454. https://doi.org/10.3390/f14030454

Ma, Y. et al., 2019. Photon-counting lidar: An adaptive signal detection method for different land cover types in coastal areas. Remote Sens., 11(4), 471. https://doi.org/10.3390/rs11040471

Ming, X. et al., 2025. Coupling ICESat-2 and Sentinel-2 data for inversion of mangrove tidal flat to predict future distribution pattern of mangroves. Int. J. Appl. Earth Obs. Geoinf., 136, 104398.

Molly, E. & Vanessa, M., 2019. NASA's Early Adopter Program Links Satellite Data to Decision Making. Remote Sens., 11(4), 406. https://doi.org/10.3390/rs11040406

Nandy, S., Srinet, R. & Padalia, H., 2021. Mapping forest height and aboveground biomass by integrating ICESat-2, Sentinel-1 and Sentinel-2 data using Random Forest algorithm in northwest Himalayan foothills of India. Geophys. Res. Lett. 48(14), e2021GL093799. https://doi.org/10.1029/2021GL093799

Narin, O.G. et al., 2024. Improving global digital elevation models using space-borne GEDI and ICESat-2 LiDAR altimetry data. Int. J. Digit. Earth 17(1).

Pimple, U. et al., 2025. Tracking mangrove ecosystem dynamics: A remote sensing approach for species classification and conservation assessment. Glob. Ecol. Conserv., e03865.

Popescu, S.C. et al., 2018. Photon counting LiDAR: An adaptive ground and canopy height retrieval algorithm for ICESat-2 data. Remote Sens. Environ. 208, 154–170. https://doi.org/10.1016/j.rse.2018.02.019

Potapov, P. et al., 2021. Mapping global forest canopy height through integration of GEDI and Landsat data. Remote Sens. Environ. 253, 112165. https://doi.org/10.1016/j.rse.2020.112165

Prihantono, J. et al., 2022. Rainfall variability and tidal inundation influences on mangrove greenness in Karimunjawa National Park, Indonesia. Sustainability 14(14), 8948. https://doi.org/10.3390/su14148948

Pronk, M. et al., 2024. DeltaDTM: A global coastal digital terrain model. Sci. Data 11(1), 273. https://doi.org/10.1038/s41597-024-03124-9

Ranndal, H. et al., 2021. Evaluation of a statistical approach for extracting shallow water bathymetry signals from ICESat-2 ATL03 photon data. Remote Sens. 13, 17. https://doi.org/10.3390/rs13010017

Renshaw, M. et al., 2026. Assessing ICESat-2's capability for global mangrove forest canopy measurements. Remote Sens. 18, 117. https://doi.org/10.3390/rs18010117.

Simard, M. et al., 2025. A new global mangrove height map with a 12 meter spatial resolution. Sci. Data 12(1), 15.

Sun, W., Jin, J., Li, G. & Yao, J., 2021. Accuracy evaluation of ICESat-2 laser altimetry satellite for monitoring water levels in Taihu Lake. Sci. Surv. Mapp. 46(11), 6–11.

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026

Tian, X. & Shan, J., 2021. Comprehensive evaluation of the ICESat-2 ATL08 terrain product. IEEE Trans. Geosci. Remote Sens. 59(10), 8195–8209. https://doi.org/10.1109/TGRS.2021.3051086.

Torres, R. et al., 2012. GMES Sentinel-1 mission. Remote Sens. Environ. 120, 9–24. https://doi.org/10.1016/j.rse.2011.05.028.

Urbazaev, M. 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. https://doi.org/10.1016/j.srs.2022.100067

Vernimmen, R., Hooijer, A. & Pronk, M., 2020. New ICESat-2 satellite LiDAR data allow first global lowland DTM suitable for accurate coastal flood risk assessment. Remote Sens. 12, 2827. https://doi.org/10.3390/rs12172827

Wang, C. et al., 2025a. Interpretable multi-sensor fusion of optical and SAR data for GEDI-based canopy height mapping in Southeastern North Carolina. Remote Sens. 17(9), 1536.

Wang, D. et al., 2024. ICESat-2 single photon laser point cloud denoising algorithm based on improved DBSCAN clustering. Earth Planets Space 76, 128. https://doi.org/10.1186/s40623-024-02071-y.

Wang, Y. & Fang, H., 2024. Derivation and evaluation of LAI from the ICESat-2 data over the NEON sites: The impact of segment size and beam type. Remote Sens. 16(16), 3078. https://doi.org/10.3390/rs16163078

Wang, Y. et al., 2025b. Combining airborne LiDAR data and optical imagery for improved national-scale beach topography estimation: A case study in New Zealand. IEEE Trans. Geosci. Remote Sens. 63, 1–23.

Wang, Z. et al., 2023. A methodological framework for specular return removal from photon-counting LiDAR data. Int. J. Appl. Earth Obs. Geoinf. 122, 103387. https://doi.org/10.1016/j.jag.2023.103387

Wessel, B. et al., 2018. Accuracy assessment of the global TanDEM-X digital elevation model with GPS data. ISPRS J. Photogramm. Remote Sens. 139, 171–182. https://doi.org/10.1016/j.isprsjprs.2018.03.004

Wessel, B. et al., 2021. TanDEM-X PolarDEM 90 m of Antarctica: Generation and error characterization. Cryosphere Discuss. 2021, 1–30.

Wu, H., Wang, N. & Guo, Z., 2021. Application of ICESat-2/ATLAS altimetry data in estimating lake water levels of Qinghai Lake. J. Water Resour. Water Eng. 32(5), 11–26.

Wu, Y. et al., 2023. Retrieving sub-canopy terrain from ICESat-2 data based on the RNR-DCM filtering and erroneous ground photons correction approach. Remote Sens. 15, 3904. https://doi.org/10.3390/rs15153904.

X. E, Dai, G. & Wu, S., 2021. Preprocessing and correction methods for ICESat-2 ATL03 data. Infrared Laser Eng. 50(6), 92–102.

Xi, Y. et al., 2022. Quantifying understory vegetation density using multi-temporal Sentinel-2 and GEDI LiDAR data. GIScience Remote Sens. 59(1), 2068–2083. https://doi.org/10.1080/15481603.2022.2140000

Xie, D. et al., 2022. Implications of coastal conditions and sea-level rise on mangrove vulnerability: A bio-morphodynamic modeling study. J. Geophys. Res. Earth Surf. 127(3), e2021JF006301. https://doi.org/10.1029/2021JF006301

Xu, W., Li, J., Peng, D., et al., 2025. Reconstruction of understory terrain based on machine learning combined with GEDI and AW3D30 data. J. Mountain Sci. 22(6), 2159–2176.

https://doi.org/10.5194/essd-2026-356
Preprint. Discussion started: 15 June 2026

Yao, S. et al., 2024. Estimating terrain elevations at 10 m resolution by integrating random forest machine learning model and ICESat-2, Sentinel-1, and Sentinel-2 satellite remotely sensed data. Int. J. Appl. Earth Obs. Geoinf. 132, 104010. https://doi.org/10.1016/j.jag.2024.104010

Yi, Y. et al., 2025. Research on accurate inversion techniques for forest cover using spaceborne LiDAR and multi-spectral data. Forests 16(8), 1215.

Yin, D. et al., 2024. Mangrove tree height growth monitoring from multi-temporal UAV-LiDAR. Remote Sens. Environ. 303, 114002. https://doi.org/10.1016/j.rse.2024.114002

Yu, J. et al., 2022. Accuracy assessment of ICESat-2 ground elevation and canopy height estimates in mangroves. IEEE Geosci. Remote Sens. Lett. 19, 1–5. https://doi.org/10.1109/LGRS.2021.3107440.

Yu, J. et al., 2024. Mapping global mangrove canopy height by integrating Ice, Cloud, and Land Elevation Satellite-2 photon-counting LiDAR data with multi-source images. Sci. Total Environ. 939, 173487. https://doi.org/10.1016/j.scitotenv.2024.173487

Yu, J. et al., 2026. Evaluating and comparing the performance of ICESat-2 and GEDI data for estimating aboveground biomass in mangroves. Int. J. Digit. Earth 19(1), 2627082.

Yu, Y. et al., 2025. Large-scale forest stand height mapping in the northeastern US and China using L-band spaceborne repeat-pass InSAR and GEDI lidar data. Earth Syst. Sci. Data 17(9), 4397–4429. https://doi.org/10.5194/essd-17-4397-2025

Zhang, X. et al., 2021. GLC_FCS30: Global land-cover product with fine classification system at 30 m using time-series Landsat imagery. Earth Syst. Sci. Data 13(6), 2753–2776. https://doi.org/10.5194/essd-13-2753-2021

Zhang, X. et al., 2024. GLC_FCS30D: The first global 30 m land-cover dynamics monitoring product with a fine classification system for the period from 1985 to 2022 generated using dense-time-series Landsat imagery and the continuous change-detection method. Earth Syst. Sci. Data 16(3), 1353–1381. https://doi.org/10.5194/essd-16-1353-2024

Zhang, Y., Huang, J., Liang, L. & Mo, F., 2026. A hierarchical optimization model based on multi-source remote sensing to correct FABDEM. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., in press. https://doi.org/10.1109/JSTARS.2026.3656929.

Zhao, X. et al., 2018. A global corrected SRTM DEM product for vegetated areas. Remote Sens. Lett. 9(4), 393–402. https://doi.org/10.1080/2150704x.2018.1425560.

Zhu, J. et al., 2022. Accuracy assessment of ICESat-2 ATL08 terrain estimates: A case study in Spain. J. Cent. South Univ. 29(1), 226–238. https://doi.org/10.1007/s11771-022-4945-3.