The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume L-4/W1-2026
Free and Open Source Software for Geospatial (FOSS4G) 2026 – Academic Track "Bridging Geospatial Technology and Humanity",
30 August–5 September 2026, Hiroshima, Japan

Regional 10-m Mapping of Forest Foliage Height Diversity in Hokkaido, Japan, Using GEDI and Foundation-Model Satellite Embeddings

Tetsu Sasaki[1], Narumasa Tsutsumida[1,2]

[1] Graduate School of Science and Engineering, Saitama University, Japan - sasaki.t.131@ms.saitama-u.ac.jp
[2] Space Data Frontiers Research Center, Fujitsu Research, Fujitsu Limited, Japan - rsnaru.jp@gmail.com

Keywords: GEDI, FHD, Satellite Embedding, Open Geospatial Data, Forest Structure Mapping, LightGBM.

Abstract

Forest vertical structural diversity reflects ecosystem complexity and biodiversity potential. Foliage height diversity (FHD), an indicator measured by the Global Ecosystem Dynamics Investigation (GEDI), describes how foliage is distributed across canopy height layers. However, GEDI LiDAR samples forests within discontinuous footprints rather than continuous coverage. To produce a wall-to-wall, 10 m-resolution FHD map of Hokkaido, Japan, we propose a LightGBM regression framework trained on approximately 150,000 GEDI footprints collected from May to October 2023. Predictors include conventional remote sensing and environmental layers (Sentinel-1 C-band SAR, Sentinel-2 optical imagery, AW3D30 elevation, and CGLS forest type) alongside a foundation-model-derived feature set: Satellite Embedding V1, a 64-dimensional representation from AlphaEarth Foundations. We compared three models: (1) conventional predictors only, (2) embeddings only, and (3) combined predictors and embeddings. The embedding-only approach clearly outperformed the conventional baseline, while the combined model achieved the best overall performance (RMSE = 0.303; $R^2 = 0.507$). These results suggest that the embeddings capture information relevant to forest vertical structure, providing complementary value beyond standard predictors. The final 10 m FHD map revealed spatially coherent patterns across Hokkaido, suggesting its potential utility for regional assessments of ecosystem complexity and biodiversity. Our findings indicate that foundation-model-derived satellite embeddings can substantially improve GEDI-based forest structure estimation when integrated with open geospatial datasets in a reproducible mapping workflow.

1. Introduction

Forest vertical structural diversity is a key indicator of ecosystem complexity, habitat heterogeneity, biodiversity potential, and carbon dynamics (Ehbrecht et al., 2021; Schneider et al., 2020). The vertical distribution of foliage and canopy elements within forests influences light availability, microclimatic conditions, and habitat space, thereby shaping ecological niches for a wide range of organisms. Previous studies have suggested that vertical forest structure is one of the major determinants of species diversity and can strongly explain biodiversity patterns, especially for taxa such as primates and amphibians (Oliveira and Scheffers, 2019). Foliage height diversity (FHD) derived from vertical vegetation profiles is widely used to quantify this structural heterogeneity as an indicator of vertical canopy complexity. Higher FHD values generally indicate that foliage is distributed across a wider range of height layers, whereas lower values suggest a simpler vertical structure. FHD is commonly derived from LiDAR-based vertical vegetation profiles, which provide detailed information on canopy structure across height layers.

Airborne laser scanning (ALS) provides accurate three-dimensional information on forest structure, but its limited spatial coverage and high observation costs hinder large-scale monitoring. The Global Ecosystem Dynamics Investigation (GEDI) mission has enabled global sampling of forest vertical structure using spaceborne LiDAR (Dubayah et al., 2020), while its footprint-based sampling produces spatially discontinuous observations. GEDI records the return waveforms of laser pulses reflected from the canopy surface to the ground, enabling vertical structure metrics to be calculated. For example, the GEDI Level 2A product provides relative height metrics, while the Level 2B product provides derived canopy structure metrics such as FHD (Dubayah et al., 2020). GEDI observations are footprint-based and acquired along orbital tracks, resulting in spatially discontinuous data. Therefore, although GEDI-derived FHD is useful for quantifying forest vertical structural diversity, additional geospatial predictors are necessary to infer continuous FHD patterns spatially at a regional scale.

Machine learning techniques with satellite remote sensing data can complement GEDI's spatial discontinuity by providing repeated wall-to-wall observations over broad regions. Synthetic aperture radar (SAR) and multispectral optical imagery are commonly used for this purpose. In particular, Sentinel-1 SAR and Sentinel-2 multispectral imagery are open datasets with 10-m spatial resolution and global coverage. Previous studies have demonstrated that GEDI-derived canopy height can be mapped using machine learning with satellite data. For example, Lang et al. (2023) estimated GEDI-derived canopy height at 10 m resolution using Sentinel-2 spectral reflectance and vegetation indices with a Random Forest model. Schwartz et al. (2024) also used Sentinel-1 and Sentinel-2 data with a U-Net model to estimate canopy height at 10 m resolution in the Landes region of France. These studies indicate that machine-learning models can learn relationships between sparse GEDI observations and satellite-derived predictors.

Although satellite-based canopy height estimation has been studied extensively, regional-scale estimation of FHD remains limited. Compared with canopy height, FHD represents the internal vertical distribution of foliage and is therefore more difficult to infer from surface observations alone. Burns et al. (2024) predicted FHD using Sentinel-1 and Sentinel-2 data, but the spatial resolution was 1 km, which may be too coarse to represent fine-scale forest structural heterogeneity. The relationship between remote sensing indicators and vegetation or species diversity depends strongly on image spatial resolution (Sang et al., 2024). From this perspective, estimating forest structural diversity at a fine spatial resolution is important for evaluating forest structure and ecosystem functions at management-relevant scales. Tsutsumida et al. (2023) demonstrated a 10-m mapping approach for forest vertical structure attributes around Mt. Washibetsu, Hokkaido, by integrating GEDI, Sentinel-1, Sentinel-2, and a Random Forest model. Their study mapped FHD, showing the feasibility of estimating forest vertical structure from fused satellite datasets. However, the study was conducted in a limited area, and they noted that their model accuracy remained insufficient. These studies suggest that Sentinel-1 and Sentinel-2 are useful predictors, but also that such conventional satellite features alone may be insufficient for accurate and scalable FHD estimation over heterogeneous forest landscapes.

In recent years, geospatial foundation models have provided a new opportunity for regional-scale environmental mapping. Satellite Embedding V1, generated by AlphaEarth Foundations (AEF) (Brown et al., 2025), is a geospatial foundation model developed by Google DeepMind to produce compact representations of Earth's surface from multiple Earth observation and geospatial data sources, including optical imagery, radar observations, Landsat-derived and Sentinel-derived information, and other geospatial datasets. AEF provides a 64-dimensional learned representation of Earth's surface at 10 m resolution. Unlike manually designed spectral indices or radar metrics, such embeddings can encode complex spatial, spectral, temporal, and contextual information from multiple Earth observation sources. While the use of embeddings may be useful for estimating forest vertical structural diversity, such an approach has not yet been explored for FHD mapping.

This study proposes a framework for estimating regional-scale forest vertical structural diversity by extending the GEDI footprint-based FHD across space. Specifically, we integrate GEDI-derived FHD, conventional geospatial variables, and Satellite Embedding V1 within a LightGBM regression framework for 10-m FHD mapping across Hokkaido, northern Japan. Hokkaido is suitable for evaluating the proposed approach because it contains extensive forested areas with diverse topographic, climatic, and management conditions, including mountainous forests, plantations, natural broad-leaved forests, and mixed forests (Kira, 1991). Such environmental heterogeneity may provide an appropriate testbed for assessing whether learned satellite embeddings can improve FHD estimation across different forest conditions.

2. Materials and Methods

The overall workflow (Figure 1) of this study consists of five main steps. First, GEDI Level 2B footprints acquired from May to October 2023 were collected and filtered to obtain reliable FHD samples over forested areas. Second, multi-source wall-to-wall geospatial predictor layers were prepared using open satellite datasets. These predictors included conventional geospatial variables derived from Sentinel-1, Sentinel-2, AW3D30, and forest type, as well as the 64-dimensional Satellite Embedding V1 generated by AEF. Third, predictor values were sampled at the locations of the filtered GEDI footprints to construct point-based modeling datasets. Each sample consisted of a GEDI-derived FHD value as the response variable and the corresponding conventional and/or embedding-derived features as explanatory variables. Fourth, three LightGBM regression models were trained and compared to evaluate the contribution of Satellite Embedding V1: Model 1 using conventional geospatial features only, Model 2 using Satellite Embedding V1 only, and Model 3 using both feature groups. Finally, the best-performing model was applied to the Hokkaido-wide predictor layers to generate a continuous wall-to-wall FHD map.

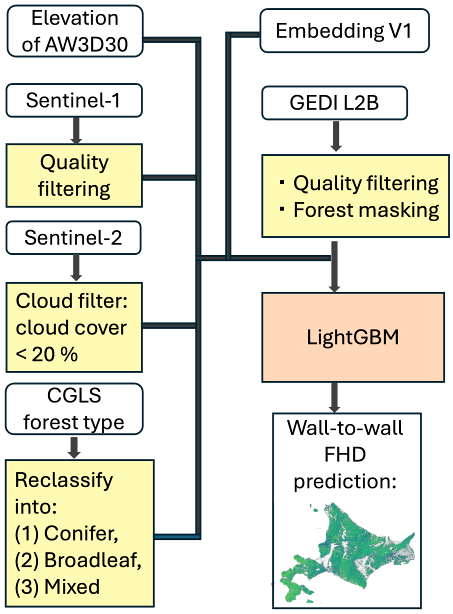

Figure 1. The workflow of this study

2.1 Study area

The study area is Hokkaido, the northernmost main island of Japan. Figure 2 shows the location of the study area together with the GEDI footprint samples described in the following sections. The region includes diverse topographic, climatic, and management conditions (Kira, 1991). These heterogeneous environmental conditions provide an appropriate testbed for assessing whether open satellite and geospatial datasets, including Sentinel-1, Sentinel-2, elevation, forest type information, and foundation-model-derived satellite embeddings, can explain spatial variation in GEDI-derived forest vertical structure.

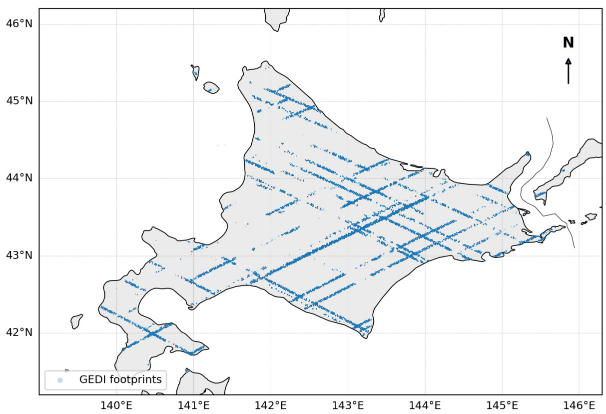

Figure 2. Study area and GEDI footprints in Hokkaido.

2.2 Data

We used GEDI-derived forest structural information as the target variable and combined it with multiple open satellite datasets. We developed an overall data design to construct a reproducible, multi-source geospatial dataset that captures forest canopy condition, terrain, and land-cover characteristics at GEDI footprint locations.

The target variable in this study is FHD, derived from the GEDI Level 2B product (Dubayah et al., 2020). GEDI Level 2B provides footprint-level canopy structure and biophysical metrics derived from full-waveform LiDAR observations, including FHD. In this study, GEDI footprints were used as point-based samples, and the corresponding FHD value was assigned as the response variable for LightGBM.

We integrated multiple open geospatial datasets for FHD mapping. Sentinel-1 synthetic aperture radar data were used to incorporate microwave backscatter information at a 10-m spatial resolution. Sentinel-1 is an active C-band SAR sensor, and its observations are less affected by cloud cover and solar illumination conditions than optical imagery. C-band backscatter is influenced by surface roughness, canopy moisture, and scattering from vegetation components such as leaves, branches, and stems. We used backscatter intensity of VV and VH polarization. Sentinel-1 provides both ascending and descending orbit observations; however, we found that the ascending observations contained a large number of missing values over the study area. Therefore, only descending orbit data were used.
Sentinel-2 optical imagery (Drusch et al., 2012) was used to capture spectral information related to vegetation conditions. Sentinel-2 provides optical bands at 10, 20, and 60 m spatial resolutions. We used the following 12 bands: B1 (coastal aerosol), B2 (Blue), B3 (Green), B4 (Red), B5 (Red Edge 1), B6 (Red Edge 2), B7 (Red Edge 3), B8 (NIR1), B8A (Red Edge 4), B9 (water vapor), B11 (SWIR1), B12 (SWIR2). All reflectance were resampled at 10-m resolution. Normalized Difference Vegetation Index (NDVI) (Equation 1) and Enhanced Vegetation Index (EVI) (Equation 2) were also prepared as predictor variables:

$$NDVI = \frac{NIR - Red}{NIR + Red}, \tag{1}$$

$$EVI = \frac{2.5(NIR - Red)}{NIR + 6Red - 7.5Blue + 1}, \tag{2}$$

Where *NIR* represents the intensity of near-infrared wavelengths (B8), *Red* represents the intensity of red wavelengths (B4) and *Blue* represents the intensity of blue wavelengths (B2). These Sentinel-2 features provide information on canopy greenness, vegetation vigor, and moisture-related surface conditions (Drusch et al., 2012), which are expected to be associated with forest structure.

Topographic information was derived from JAXA's AW3D30 digital elevation model (Tadono et al., 2014), which provides global elevation information at 30 m spatial resolution. Elevation is an important environmental variable because forest structure is often influenced by altitude and associated environmental gradients. (Tadono et al., 2014)
Forest type information was also included. The Copernicus Global Land Service (CGLS) forest type product (Buchhorn et al., 2020), provided at 100 m spatial resolution, was used to distinguish broad forest categories and reclassified into three categories (coniferous forest, broad-leaved forest, and mixed forest).

In addition to the above conventional geospatial features, this study used Satellite Embedding V1 generated by AEF (Brown et al., 2025). Each 10-m pixel is represented by a 64-dimensional embedding vector that summarizes annual surface conditions and temporal variations around that pixel. In this study, the embedding vectors were used as predictors for FHD estimation. For each GEDI footprint, the corresponding 64 embedding dimensions were sampled. These embedding dimensions were prepared as inputs of LightGBM.

2.3 Geospatial Preprocessing

GEDI footprint samples were collected using Google Earth Engine (GEE) (Gorelick et al., 2017). We collected GEDI shots from May to October 2023 during the leaf-on season for deciduous trees in Hokkaido. The spatial distribution of the GEDI footprint samples used in this study is shown in Figure 2. Initial filtering was applied during data collection to retain reliable GEDI observations. Shots were retained only when they were acquired by strong beams, had no degradation flag, had sensitivity values of 0.9 or higher, and satisfied the quality flag condition. The remaining GEDI footprints were then filtered using the ESA WorldCover forest class (Zanaga et al., 2022) so that the modeling dataset focused on forested areas. Additional quality control was applied after the initial extraction to reduce elevation-related errors in the GEDI footprints. First, shots with invalid elevation values or invalid relative height metrics were removed. Outliers in the elevation of the lowest detected waveform mode (elev_lowestmode) found in GEDI L2B data were then identified using a rolling-window filter with a window size of seven shots and a threshold of two standard deviations. After this filtering, a geoid correction was applied to derive GEDI-based ground elevation (gedi_dem). As an independent elevation consistency check, mean AW3D30 elevation values were extracted within a 12.5 m buffer around each footprint. Finally, only shots with an absolute difference of 20 m or less between the GEDI-based ground elevation and the AW3D30-derived reference elevation were retained. After these filtering steps, approximately 150,000 GEDI footprints were used for the subsequent analysis.

The predictor datasets were prepared mainly through GEE. In this study, the geospatial predictors were Sentinel-1, Sentinel-2, elevation derived from AW3D30, forest type information, and Satellite Embedding V1. For each dataset, raster layers covering Hokkaido were prepared according to their temporal characteristics, spatial resolution, and data format. These layers were created as wall-to-wall predictors so that they could be used consistently for both point-based model training and subsequent regional-scale wall-to-wall prediction.

For Sentinel-1, C-band synthetic aperture radar data were used to derive microwave backscatter predictors. Sentinel-1 Ground Range Detected images were obtained from May to October 2023 and filtered to the Interferometric Wide (IW) swath mode. Images were further separated by orbit direction, using ascending and descending passes, and only images containing VV or VH transmitter–receiver polarizations were retained. For each orbit direction and polarization, seasonal composite images were generated over the target period. VV- and VH-based

backscatter layers and their ratio were then used as Sentinel-1 predictors. Although both ascending and descending orbit observations were initially considered, ascending observations contained many missing values over the study area. Therefore, only descending orbit predictors were retained for model training.
For Sentinel-2, surface reflectance images from May to October 2023 were obtained from the Sentinel-2 SR Harmonized collection. Images were filtered by cloud cover, and only scenes with less than 20 % cloudy pixels were used. NDVI and EVI were calculated for each image from the red, blue, and near-infrared bands, and a median composite was generated for the season. These optical predictors were used to represent vegetation greenness and canopy condition during the same seasonal period as the GEDI observations.

Elevation was derived from JAXA's AW3D30 digital elevation model (Tadono et al., 2014). The AW3D30 image collection was accessed through GEE, and the DSM band was selected as the elevation layer.
Forest type information was derived from the CGLS forest type product (Buchhorn et al., 2020). The 2019 annual product was used because it was the most recent available CGLS forest type dataset. The forest type class band was reclassified into broad forest categories for this study. Classes 1 (evergreen needleleaf) and 3 (deciduous needleleaf) were grouped as coniferous forest, classes 2 (evergreen broadleaf) and 4 (deciduous broadleaf) were grouped as broad-leaved forest, and class 5 (mixed forest types) was assigned to mixed forest. This reclassified forest type layer was used both as a categorical predictor and for the forest-type-specific evaluation of model performance.

Satellite Embedding V1 was processed in the same general framework as the other raster predictors. The annual embedding image for 2023 was obtained through GEE. Since Satellite Embedding V1 provides a 64-dimensional vector for each pixel, each embedding dimension was treated as an independent continuous predictor variable.

After the individual datasets were processed, Sentinel-1, Sentinel-2, AW3D30-derived elevation, CGLS forest type, and Satellite Embedding V1 were prepared as Hokkaido-wide wall-to-wall predictor layers. These layers were created so that they could be used consistently for both point-based model training and subsequent regional-scale prediction. Predictor values were then extracted at the locations of the filtered GEDI footprints. For each footprint, the predictor values were combined with the corresponding GEDI-derived FHD value, producing a footprint-based modeling dataset for training and validation. In parallel, the same Hokkaido-wide predictor layers were retained for later application of the trained model to generate a continuous wall-to-wall FHD prediction map.

For model evaluation, we used five-fold spatial block cross-validation. The study area was divided into 40-km grid blocks, and all samples within the same block were assigned to the same fold. This block-based split was used to reduce spatial leakage between nearby GEDI footprints and to evaluate model performance under more spatially independent conditions. Final evaluation metrics were calculated from pooled out-of-fold predictions across the five validation folds.

2.4 Methodology

Based on the multi-source dataset described in the section 2.2 and section 2.3, FHD estimation was formulated as a supervised regression task. Each GEDI footprint was treated as an observation sample. The GEDI-derived FHD value was used as the response variable, modeled by two distinct types of predictors as explanatory variables: conventional geospatial variables and Satellite Embedding V1 extracted at the same location. To evaluate the contribution of Satellite Embedding V1, we compared three model configurations, labeled as Models 1, 2, and 3 throughout this paper.

Model 1: Geospatial variables only (19 features)
This model uses only the conventional geospatial variables described in Section 2.2, which include 3 Sentinel-1 bands, 12 Sentinel-2 bands, 2 vegetation metrics derived from Sentinel-2, elevation derived from AW3D30, and forest type. This model represents the baseline approach.

Model 2: Satellite Embedding V1 only (64 embeddings)
This model uses only the 64-dimensional Satellite Embedding V1 features to evaluate whether the learned embedding representation alone can serve as an effective predictor of forest vertical structural diversity.

Model 3: Combined features (19 features and 64 embeddings)
This configuration combines the conventional geospatial variables with the 64-dimensional Satellite Embedding V1 features to evaluate whether the embedding provides complementary information beyond conventional predictors.

All three feature configurations were trained using the same regression framework by LightGBM because it is well-suited for datasets with heterogeneous feature types (e.g., continuous and categorical) (Ke et al., 2017). LightGBM can model nonlinear relationships and interactions among variables without requiring strong assumptions about the distribution of the predictors. This property is useful for multi-source remote sensing data, where the relationships among satellite signals, environmental conditions, learned embeddings, and forest structure are often nonlinear (Zhang and Ni, 2023).

Using the same model architecture for Model 1, Model 2, and Model 3 ensures that differences in predictive performance can be attributed mainly to the feature configuration rather than to the difference in the learning algorithm. Hyperparameters for training were set to the learning rate (0.05), number of leaves (128), feature fraction (0.8), bagging fraction (0.8), bagging frequency (1), and minimum number of samples per leaf (500).

2.5 Evaluation metrics

Model performance was assessed using several regression metrics. The root mean squared error (RMSE) was used as the primary accuracy metric because it represents prediction error in the same unit as the target variable and penalizes large errors. The mean absolute error (MAE) was also calculated to provide a more robust measure of average prediction error. In addition, the coefficient of determination (R^2) was used to assess the proportion of variance in GEDI-derived FHD explained by the model. Mean bias error (MBE) was calculated to evaluate systematic overestimation or underestimation. In this study, MBE was defined as the mean difference between predicted and observed FHD values. A positive MBE indicates that the model tends to overestimate FHD, whereas a negative MBE indicates that the model tends to underestimate FHD. These metrics were calculated from the pooled out-of-fold predictions across all five folds.

In addition to the overall evaluation, model performance was assessed separately by forest type. The forest type classes were derived from the CGLS forest type (Buchhorn et al., 2020) product and grouped into coniferous forest, broad-leaved forest, and mixed forest. For each forest type, RMSE, NRMSE, R^2, MAE, and MBE were calculated using the validation samples belonging to that class. This stratified evaluation was conducted to examine whether prediction accuracy differed among forest types.

2.6 Wall-to-Wall prediction

The trained regression model was applied to the Hokkaido-wide raster predictor layers to generate a continuous FHD prediction map. This step was conducted not only to produce a spatially continuous representation of forest vertical structural diversity but also to qualitatively examine whether the trained model produced reasonable regional patterns beyond the GEDI footprint locations.
In the wall-to-wall prediction step, predictor values were prepared for each raster pixel, and the trained model was used to estimate FHD values across the study area. The resulting map was used to visually assess whether predicted FHD patterns were spatially coherent and broadly consistent with expected environmental gradients, such as mountainous forest regions and human-influenced landscapes.

3. Results

3.1 Error Assessment

Table 1 summarizes the prediction performance of the three feature configurations of Model 1, Model 2, and Model 3. The results indicate that using Satellite Embedding V1 substantially improved prediction accuracy compared to conventional geospatial features. Model 2 achieved the second-best performance and reduced RMSE by 0.062 and improved R^2 by 0.226 compared with Model 1. This result indicates that the Satellite Embedding V1 contains substantial information related to forest vertical structural diversity, even without explicitly using conventional geospatial variables. Model 3 achieved the best overall performance, reducing RMSE by 0.066 and improving R^2 by 0.237, compared with Model 1. This demonstrates that the characteristics of Satellite Embedding V1 are not consistent with those of conventional geospatial variables, providing further informative aspects of vertical structures.

All three models showed nearly zero negative MBE values, indicating that the models did not exhibit strong systematic over- or underestimation of FHD at the overall dataset level. The reduction in both RMSE and MAE from Model 1 to Model 3 suggests that integrating embedding features improved the stability of FHD prediction while maintaining low bias.

	RMSE	R^2	MAE	MBE
Model 1	0.369	0.270	0.226	-0.005
Model 2	0.307	0.496	0.214	-0.004
Model 3	0.303	0.507	0.212	-0.007

Table 1. Error performance of LightGBM using conventional 19 geospatial variables (Model 1), Satellite Embedding V1 (Model 2), and their combination (Model 3).

Figure 3 shows the relationship between GEDI-observed FHD and model-predicted FHD for the three models. Model 2 showed a stronger alignment with the 1:1 line than Model 1. The higher FHD samples were distributed more closely along the diagonal, especially for samples with observed FHD values between approximately 2.5 and 3.3. The scatter pattern was also more compact than in Model 1, consistent with the improvements in RMSE, MAE, and R^2 shown in Table 1. This result visually confirms that Satellite Embedding V1 captured FHD-related spatial information more effectively than the conventional feature set alone. The visual difference between Model 2 and Model 3 was smaller than that between Model 1 and Model 2, suggesting that most of the performance gain was achieved by introducing Satellite Embedding V1, while the additional contribution of conventional geospatial variables was modest.

Across all three models in Figure 3, prediction errors were larger with low observed FHD. In particular, observed FHD values below approximately 2.0 were often overestimated. This pattern suggests that low-complexity forest cases may be relatively difficult to distinguish using the current predictor variables. In contrast, the models performed more consistently in the high-density region of the dataset, where observed FHD values were concentrated around approximately 2.7 to 3.2.

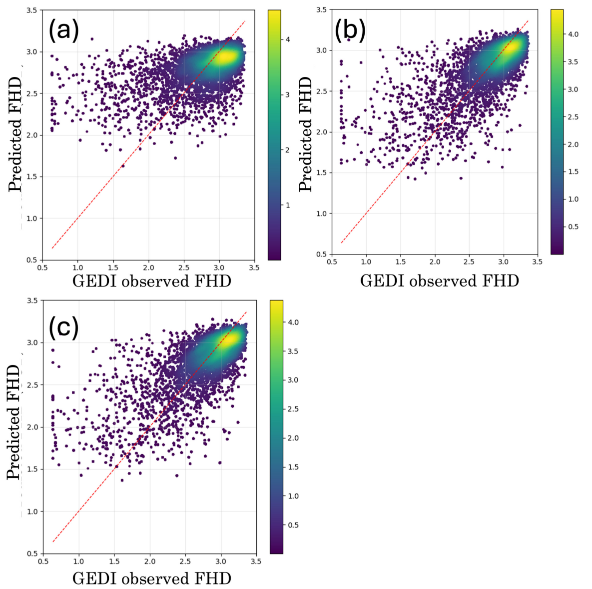

Figure 3. Scatter plots of GEDI-observed FHD and predicted FHD for the three feature configurations by LightGBM using: (a) conventional 19 geospatial variables (Model 1), (b) Satellite Embedding V1 (Model 2), and (c) their combination (Model 3). The red dashed line represents the 1:1 line, and the color scale indicates point density.

To further examine the behavior of the best-performing model (Model 3), the prediction error was evaluated across forest types (Table 2). Coniferous forests had the lowest RMSE and MAE, suggesting a relatively simple vertical structure. For broad-leaved forests, while RMSE and MAE were higher than those for coniferous forests, R^2 was higher, suggesting that the model captured a greater proportion of the variation in FHD. For mixed forests, the RMSE and MAE were the largest; however, mixed forests showed the highest R^2 value. One possible interpretation is that mixed forests have a wider range of vertical structural conditions, making absolute prediction more difficult while still allowing the model to capture relative spatial differences. All forest types showed negative MBE values close to

https://doi.org/10.5194/isprs-archives-L-4-W1-2026-267-2026 | © Author(s) 2026. CC BY 4.0 License.

zero, indicating that the models did not exhibit strong systematic over- or underestimation across any forest type.

	RMSE	R^2	MAE	MBE
Conifer	0.260	0.340	0.177	-0.015
Broadleaf	0.315	0.490	0.225	-0.003
Mixed	0.327	0.571	0.246	-0.044

Table 2. Error measures of FHD estimation of LightGBM (Model 3) by forest types.

3.2 Wall-to-Wall FHD Map

Figure 4 (upper full-Hokkaido map) shows that relatively high FHD values were widely observed in forested mountain regions, whereas lower values appeared in flatter or more human-influenced landscapes. Figure 4 also includes enlarged maps for two representative regions. The right zoom panels show the Hidaka Mountains, a steep mountainous region with extensive forested slopes and high-elevation ridgelines. In this region, lower FHD values were observed along ridgelines and high-elevation zones, whereas relatively higher values appeared on surrounding forested slopes. These patterns are consistent with known elevation-related ecological gradients in mountain forests (Ameztegui et al., 2021). The left zoom panels show the area around New Chitose Airport, representing a human-influenced lowland landscape where airport facilities, urban areas, agricultural land, and fragmented forests occur in close proximity. Lower FHD values appeared close to non-forest and human-influenced land-cover types, while relatively higher values were observed in surrounding forested areas. These spatial patterns suggest that the model captured meaningful variation in forest vertical structure associated with both elevation-driven ecological gradients and human-influenced landscapes.

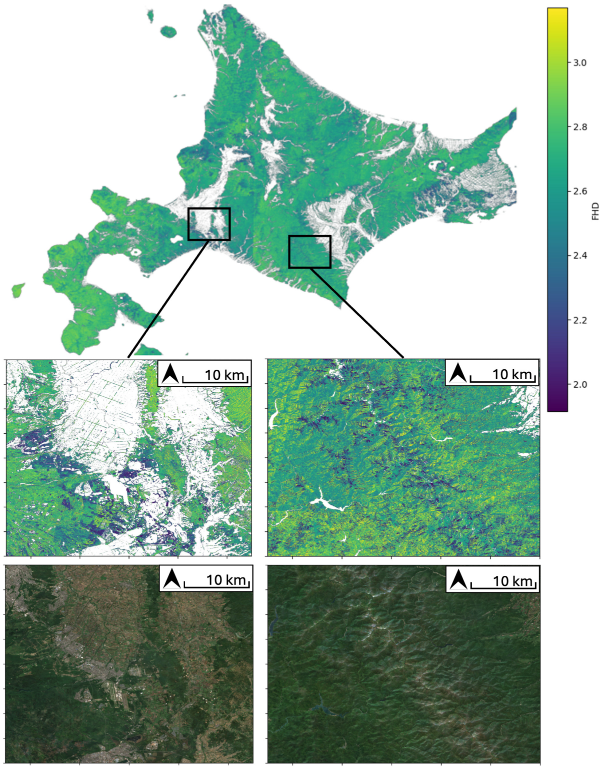

Figure 4. (upper) Wall-to-wall predicted FHD map across Hokkaido by LightGBM using geospatial variables and Satellite Embedding V1. (middle) Magnified views of two representative regions: the right map shows the Hidaka Mountains, and the left one shows the area around New Chitose Airport. (bottom) The corresponding Sentinel-2 true-color images.

4. Discussion

Conventional geospatial variables, including Sentinel-1, Sentinel-2, terrain, and forest types, are physically interpretable and have been widely used in remote sensing studies (Lang et al., 2023; Schwartz et al., 2024; Burns et al., 2024). However, such variables may not fully capture the complex spatial, spectral, temporal, and contextual patterns underlying forest structure. The improvement from Model 1 to Model 2 suggests that Satellite Embedding V1 provides FHD-relevant information beyond conventional predictors. Because FHD reflects the vertical distribution of foliage within the canopy rather than a simple surface property, the embedding may encode indirect structural signals related to canopy condition, forest type, phenology, disturbance history, and environmental gradients. Model 3 achieved the best overall performance, although the improvement from Model 2 to Model 3 was smaller than that from Model 1 to Model 2. This indicates that Satellite Embedding V1 already captures a large portion of the useful information contained in conventional predictors, while conventional variables still add modest complementary ecological context. Therefore, Satellite Embedding V1 should not be regarded simply as a replacement for conventional features, but as a strong general-purpose representation that can be combined with interpretable geospatial variables.

The use of precomputed embeddings also has practical significance. In this study, the original foundation model was not fine-tuned on GEDI data. Instead, the embedding layers were sampled at GEDI footprint locations and used as predictors in a LightGBM regression model. This demonstrates that foundation-model-derived geospatial representations can be incorporated into conventional machine-learning workflows without requiring users to train large deep learning models. Such a workflow is particularly useful for regional-scale mapping because it connects spatially discrete GEDI LiDAR observations with wall-to-wall satellite-derived predictors.

The wall-to-wall prediction map (Figure 4) showed spatially coherent FHD patterns across Hokkaido, including the two enlarged examples in the Hidaka Mountains and around New Chitose Airport. In the Hidaka Mountains, lower FHD values along ridgelines and high-elevation zones were consistent with elevation-related ecological gradients in mountain forests. Around New Chitose Airport, lower FHD values appeared near urban, agricultural, and fragmented landscapes. These examples suggest that the proposed method can transform spatially discrete GEDI observations into continuous information that reflects both elevation-driven ecological gradients and human-influenced land-cover conditions. Such information may be useful for regional forest structure assessment, particularly in heterogeneous landscapes where field-based observation alone is spatially limited.

Despite the improved performance, prediction errors remained. The scatter plots (Figure 3) showed that lower observed FHD values were often overestimated, whereas samples in the intermediate-to-high FHD range were predicted more consistently. This tendency may reflect the distribution of the training samples and the indirect nature of the predictor variables. As shown by the point-density patterns in Figure 3, most GEDI footprints were concentrated in common intermediate-to-high FHD conditions. The model may have learned these dominant forest structural patterns more strongly and tended to regress rare low-FHD cases toward the mean. In addition, FHD represents the vertical distribution of foliage within the canopy, whereas the predictor variables mainly describe surface reflectance, radar backscatter, terrain, forest type, and learned spatial context. Therefore, forests with relatively simple vertical structures may still be predicted as having moderate FHD when their surface or contextual characteristics resemble more common forest conditions.

Error characteristics across forest categories (Table 2) showed a distinct difference. The relatively low R^2 for coniferous forests, despite their low RMSE and MAE, may be related to the narrower range of observed FHD values within this forest type. Coniferous forests may include relatively homogeneous stands, such as plantations or uniformly managed forests, which can reduce within-class variability in vertical structure. When the observed variance is limited, the coefficient of determination becomes more sensitive to small prediction errors, even if the absolute error remains low. In contrast, mixed forests showed the highest R^2 but also the largest RMSE. This may indicate that mixed forests contain a wider range of vertical structural conditions.

Future work should address these remaining errors from several perspectives. First, the prediction of low-FHD stands and structurally complex mixed forests should be improved. Second, the interpretability of embedding features should be examined more carefully. Although Satellite Embedding V1 improved accuracy, it remains unclear which aspects of the embedding representation contributed most strongly to FHD estimation. Feature importance analysis, SHAP-based interpretation, or comparison with physically interpretable predictors may help clarify this point. Third, the generalizability of the proposed workflow should be tested in other ecological regions because Hokkaido represents only one case study with cool-temperate and boreal-like forest conditions. Finally, comparison with other geospatial foundation models or task-specific deep learning approaches would help clarify the broader potential and limitations of learned satellite representations for forest vertical structure mapping.

Overall, the results suggest that Satellite Embedding V1 is a useful off-the-shelf representation for regional-scale forest structure mapping. By combining GEDI-derived FHD, open geospatial datasets, and a standard machine-learning model (LightGBM), the proposed workflow extends spatially sparse LiDAR observations into continuous FHD estimations while remaining compatible with reproducible geospatial processing.

5. Conclusions

This study developed a regional-scale framework for estimating forest vertical structural diversity across Hokkaido, Japan, by integrating GEDI-derived FHD with multi-source open geospatial datasets. GEDI provides valuable vertical structure measurements, but its footprint-based sampling is spatially discontinuous. To address this limitation, this study used GEDI-derived FHD as training data and predicted continuous FHD patterns using wall-to-wall geospatial predictors.

The results demonstrated the effectiveness of Satellite Embedding V1 for FHD estimation. The embedding-only model outperformed the conventional-feature baseline, and the combined model achieved the best performance. These findings suggest that the learned geospatial representation generated by AEF contains information relevant to forest vertical structure and provides complementary value to conventional remote sensing variables. Because the embedding was used as an off-the-shelf raster feature without fine-tuning the original foundation model, the proposed workflow can reduce the burden of manual feature engineering while remaining practical for large-area geospatial analysis.

Overall, this study demonstrates that combining GEDI LiDAR observations, open geospatial datasets, machine-learning models, and foundation-model-derived satellite embeddings provides a promising approach for reproducible regional-scale mapping of forest vertical structural diversity. By extending spatially discrete LiDAR observations into continuous FHD maps, the proposed workflow may support regional assessments of canopy structural complexity, which is relevant to ecosystem complexity, habitat heterogeneity, biodiversity potential, and carbon dynamics.

References

Ameztegui, A., Rodrigues, M., Gelabert, P. J., Lavaquiol, B., Coll, L., 2021. Maximum height of mountain forests abruptly decreases above an elevation breakpoint. *GIScience & Remote Sensing*, 58(3), 442–454.

Brown, C. F., Kazmierski, M. R., Pasquarella, V. J., Rucklidge, W. J., Samsikova, M., Zhang, C., Shelhamer, E., Lahera, E., Wiles, O., Ilyushchenko, S. et al., 2025. Alphaearth foundations: An embedding field model for accurate and efficient global mapping from sparse label data. *arXiv preprint arXiv:2507.22291*.

Buchhorn, M., Smets, B., Bertels, L., De Roo, B., Lesiv, M., Tsendbazar, N., Herold, M., Fritz, S., 2020. Copernicus global land service: Land cover 100m: collection 3: epoch 2019: Globe.

Burns, P., Hakkenberg, C. R., Goetz, S. J., 2024. Multi-resolution gridded maps of vegetation structure from GEDI. *Scientific Data*, 11(1), 881.

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P. et al., 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote sensing of Environment*, 120, 25–36.

Dubayah, R., Blair, J. B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S. et al., 2020. The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. *Science of remote sensing*, 1, 100002.

Ehbrecht, M., Seidel, D., Annighöfer, P., Kreft, H., Köhler, M., Zemp, D. C., Puettmann, K., Nilus, R., Babweteera, F., Willim, K. et al., 2021. Global patterns and climatic controls of forest structural complexity. *Nature communications*, 12(1), 519.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote sensing of Environment*, 202, 18–27.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.-Y., 2017. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 30.

Kira, T., 1991. Forest ecosystems of east and southeast Asia in a global perspective. *Ecological research*, 6(2), 185–200.

Lang, N., Jetz, W., Schindler, K., Wegner, J. D., 2023. A high-resolution canopy height model of the Earth. *Nature Ecology & Evolution*, 7(11), 1778–1789.

Oliveira, B. F., Scheffers, B. R., 2019. Vertical stratification influences global patterns of biodiversity. *Ecography*, 42(2), 249–249.

Sang, Y., Gu, H., Meng, Q., Men, X., Sheng, J., Li, N., Wang, Z., 2024. An Evaluation of the Performance of Remote Sensing Indices as an Indication of Spatial Variability and Vegetation Diversity in Alpine Grassland. *Remote Sensing*, 16(24), 4726.

Schneider, F. D., Ferraz, A., Hancock, S., Duncanson, L. I., Dubayah, R. O., Pavlick, R. P., Schimel, D. S., 2020. Towards mapping the diversity of canopy structure from space with GEDI. *Environmental Research Letters*, 15(11), 115006.

Schwartz, M., Ciais, P., Ottlé, C., de Truchis, A., Vega, C., Fayad, I., Brandt, M., Fensholt, R., Baghdadi, N., Morneau, F. et al., 2024. High-resolution canopy height map in the Landes forest (France) based on GEDI, Sentinel-1, and Sentinel-2 data with a deep learning approach. *International Journal of Applied Earth Observation and Geoinformation*, 128, 103711.

Tadono, T., Ishida, H., Oda, F., Naito, S., Minakawa, K., Iwamoto, H., 2014. Precise global DEM generation by ALOS PRISM. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2, 71–76.

Tsutsumida, N., Kato, A., Osawa, T., Doi, H., 2023. Mapping forest vertical structure attributes with gedi, sentinel-1, and sentinel-2. *IGARSS 2023-2023 IEEE International Geoscience and Remote Sensing Symposium*, IEEE, 538–541.

Zanaga, D., Van De Kerchove, R., Daems, D., De Keersmaecker, W., Brockmann, C., Kirches, G., Wevers, J., Cartus, O., Santoro, M., Fritz, S. et al., 2022. ESA WorldCover 10 m 2021 v200.

Zhang, D., Ni, H., 2023. Inversion of forest biomass based on multi-source remote sensing images. *Sensors*, 23(23), 9313.