Remote Sensing of Environment 345 (2026) 115591

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Beyond RMSE: The GEDI imputed waveform product minimizes artificial homogeneity evident in existing global forest height maps

Eugene Seo[a,*], Sean P. Healey[b], Zhiqiang Yang[b], John D. Armston[c], Ralph O. Dubayah[c], Simon Ilyushchenko[d], Tiago De Conto[c], Matthew G. Betts[a]

[a] Department of Forest Ecosystems and Society, Oregon State University, Corvallis, OR 97331, USA
[b] USDA Forest Service, Riverdale, UT 84405, USA
[c] Department of Geographical Sciences, University of Maryland, College Park, MD 20737, USA
[d] Google Inc., Mountain View, CA 94043, USA

ARTICLE INFO

Keywords:
GEDI
GEDI L4D imputed waveform
Canopy height maps
Forest vertical structure
Forest height variability
RMSE
Population-level accuracy

ABSTRACT

The models that produce most existing forest height maps minimize mean individual prediction errors quantified by metrics such as Root Mean Square Error (RMSE). However, when predictor data do not explain all height variability in the training sample, this objective function leads to prediction toward the mean, distorting population-level prediction of height range and variability. Several existing forest height maps have used training data from NASA's Global Ecosystem Dynamics Investigation (GEDI) lidar mission, and while the mission's retrievals are subject to measurement error, we used GEDI to: 1) evaluate population-level errors of three prominent global height maps; and 2) produce new maps designed to better represent the full height distribution. Specifically, we introduce the GEDI L4D Imputed Waveform product, which was created using a nearest neighbor algorithm. This model assigned a high-quality waveform, represented by 11 Relative Height metrics and other measurements, to each 30 m pixel in the tropical and temperate domain. Covariates included synthetic Landsat bands derived from time series fit to all clear imagery. While previous maps systematically homogenized tree heights over large regions, L4D did not, and tests verified that L4D preserved appropriate covariance among GEDI-derived forest vertical structure metrics. Validation shows that L4D top height RMSE values were up to approximately 1.1 m greater than other maps. Applications such as predicting timber yield, modeling species distributions, and simulating fire spread may benefit from more demographically accurate tree height maps. Moving forward, attention to population-level accuracy metrics may increase the practical value of remotely sensed maps.

1. Introduction

Forest canopy height is a key covariate for numerous ecosystem services, including timber volume, biomass storage, and habitat suitability; maps of forest height are therefore a crucial component of global forest monitoring. In the absence of globally comprehensive ground-based forest inventory data, high-quality measurements from space-borne lidar instruments have provided calibration data for wall-to-wall maps of height and other vertical structure metrics (Simard et al., 2011). NASA's Global Ecosystem Dynamics Investigation (GEDI; Dubayah et al., 2020) lidar mission has contributed to the calibration of a number of recent global height maps, including maps evaluated here from Potapov et al. (2021), Meta (Tolan et al., 2024), and Lang et al. (2023). Throughout this paper, quality-filtered GEDI waveforms are treated as reference data against which the accuracy of derived wall-to-wall maps may be evaluated. Beyond pixel-level accuracy, the population sampled by GEDI is taken (with specific minimum sample number thresholds) as the best available globally consistent approximation of “true” forest height distributions. We proceed with this approximation while recognizing that remotely sensed measurements have inherent uncertainty, including saturation effects, which may compromise population-level accuracy metrics, and that the accuracy of GEDI height measurements is lower than airborne laser scanning (Roy et al., 2021; Dhargay et al., 2022; Mandl et al., 2023).

* Corresponding author.
E-mail addresses: seoe@oregonstate.edu (E. Seo), sean.healey@usda.gov (S.P. Healey), zhiqiang.yang@usda.gov (Z. Yang), armston@umd.edu (J.D. Armston), dubayah@umd.edu (R.O. Dubayah), simonf@google.com (S. Ilyushchenko), tiagodc@umd.edu (T. De Conto), matt.betts@oregonstate.edu (M.G. Betts).

https://doi.org/10.1016/j.rse.2026.115591
Received 3 April 2026; Received in revised form 2 July 2026; Accepted 23 July 2026
Available online 31 July 2026

While it is conventional to report map accuracy in terms of the magnitude of the average residual error via metrics such as Root Mean Square Error (RMSE), the sheer density of 25 m footprint measurements available from GEDI (this study uses 8.5 billion high-quality shots) now also enables evaluation of the parameters of the predicted population against a densely sampled presumed reference population. While low mean prediction error is clearly desirable, predicting a population with the correct mix of high and low values can also be important. For example, GEDI's spatially incomplete sample of waveforms has been used with the height-structured Ecosystem Demography model (Ma et al., 2023) to constrain represented ecosystem processes. While wall-to-wall height maps would expand calibration options for this type of model, biases realized in the predicted population – for example, an inaccurate range of heights – may result in misleading ecological conclusions. Forest structure metrics involving canopy height and cover also support fuel characterization used in fire science (Picotte et al., 2019; Heisig et al., 2025). Models which exaggerate uniformity of tree heights may affect mapped fuel classifications and any subsequently modeled fire behavior.

There is reason to suspect that many available global forest height maps, including the three mentioned above, exhibit an artificial concentration of predictions around the mean of measured remote sensing data. First, empirical case studies demonstrate that underprediction of high height values and overprediction of low values is common in mapped populations (Moudrý et al., 2024; Chen et al., 2025). Second, models which generate predictions without attention to prediction distribution are mathematically likely to predict toward the mean if minimization of unexplained error is an objective. To the degree that predictor data do not differentiate tree heights, models ranging in form from least squares regression to complex neural networks solving for parameters that minimize prediction error will tend to predict toward a mean value. Third, models developed using limited and imbalanced sample data may fail to capture the full range of the forest height distribution. The artificial homogeneity introduced in the resulting map may degrade intended applications.

The newly released GEDI L4D Imputed Waveform product (Seo et al., 2025) was specifically formulated to avoid distortions of the predicted population with respect to the local sample of high-quality GEDI waveforms. The product provides 11 imputed GEDI L2A Relative Height (RH) metrics, corresponding L2B canopy cover, and L4A aboveground biomass density (AGBD) at 30 m resolution. Hereafter, we refer to these 11 imputed RH metrics as the "imputed waveform," as they represent the vertical structure captured by GEDI waveforms. A k-nearest neighbor with $k = 1$, referred to here as the Most Similar Neighbor imputation approach (MSN; Valbuena et al., 2018), was used to assign a nearby high-quality GEDI waveform to every terrestrial 30 m pixel according to a distance function calculated in a multivariate space defined by corresponding Landsat time series data (as described in Section 2.1). If GEDI provides an approximately representative sample in a given modeling area, the density of GEDI shots and Landsat data should covary such that each "neighbor" (reference observation) has approximately the same number of "nearest" pixels, resulting in a distribution of predictions that should reflect the distribution, across all of GEDI's measured variables, implied by an independent hold-out set of waveforms. Additionally, since all waveform properties are imputed in aggregate, the covariance measured by GEDI among fields such as canopy top height and canopy cover should be preserved in the imputed L4D maps (Moeur and Stage, 1995).

In this paper, we evaluate both GEDI L4D and the three existing global height maps mentioned above globally in terms of RMSE and four metrics of population-level fidelity to GEDI's original height sample:

1. Population dispersion: Difference between the predicted and reference canopy height Inter-Decile Range (IDR). IDR is a measure of local height variation defined by the difference in meters between the local 10th and 90th percentile heights for a given area;
2. Population distribution: Kolmogorov-Smirnov (KS) test for the difference between the predicted canopy height distribution of the wall-to-wall map and the distribution of local canopy heights recorded by an independent sample of GEDI data;
3. Population-level covariance: Preservation of covariance among relative height (RH) metrics derived from waveform lidar. Since only GEDI L4D (among the maps considered here) predicted multiple RH values, this metric was evaluated only for that product;
4. Population-level correlation: Preservation of the GEDI-sampled correlation between canopy height and aboveground biomass density (AGBD) in the mapped population. As above, this could only be applied to L4D.

Good performance related to IDR requires accurate prediction of extreme values (measured at the 10th and 90th percentiles) while the KS test requires an accurate mapped representation of the continuous distribution of predictions in GEDI's sample. The covariance and correlation dimensions extend the evaluation to properties among multiple variables. This last metric is relevant when multiple maps are combined for a given application. For instance, the habitat-related application described by Healey et al. (2006) focused on forests that met mapped criteria for both percent cover and basal area. In this study, we evaluated practical tradeoffs among these emergent elements of accuracy and RMSE. As space lidar missions provide increasingly comprehensive samples of forest structure, evaluation of the population-level accuracies of alternative products may stimulate more deliberate consideration of which elements of map accuracy may be most relevant for specific applications.

2. Methods

The GEDI L4D Imputed Waveform product provides wall-to-wall 30-m GEDI waveform predictions for the year 2023 and is generated using the MSN algorithm operating upon multi-spectral and topographic inputs. Details of the imputation procedure are provided in Section 2.1, and the population-level evaluation methods for assessing the globally imputed products are presented in Section 2.2.

2.1. Most similar neighbor (MSN) imputation for GEDI L4D

MSN is equivalent to the 1-nearest neighbor ($1-$NN) algorithm. The k-NN framework is a non-parametric machine learning method used for both classification and regression (Bishop, 2006). Predictions are made by identifying the most similar observed samples on the basis of ancillary data and typically averaging their response values without assuming a parametric model. The choice of k, the number of neighbors, is the primary hyperparameter in k-NN and influences predictive performance. In general, aggregating information from multiple neighbors ($k > 1$) reduces mean prediction error by capturing overall trends and reducing noise. However, as k increases, predictions are increasingly pulled toward the mean of the response variable, causing the distribution of predictions to diverge from that of the observations and compressing variation across the full range of values (Hudak et al., 2008). In contrast, MSN ($k = 1$) better preserves the full range of the data because each prediction inherits the value of the single most similar observed sample; no averaging occurs. As described above, MSN also tends to yield a prediction distribution closer to that of the training data, provided the training data adequately represent the landscape in the covariate space assembled for imputation. However, this advantage comes with the risk of overfitting to outliers in the samples, which we discuss later as part of the strengths and limitations of the MSN approach. To ensure high-quality data for the development of GEDI L4D, extensive quality filtering was applied to ensure inclusion only of high-quality GEDI shots (Seo et al., 2025); we recognize that our results may be sensitive to the stringency of the filtering criteria. The quality-filtering criteria applied are as follows:

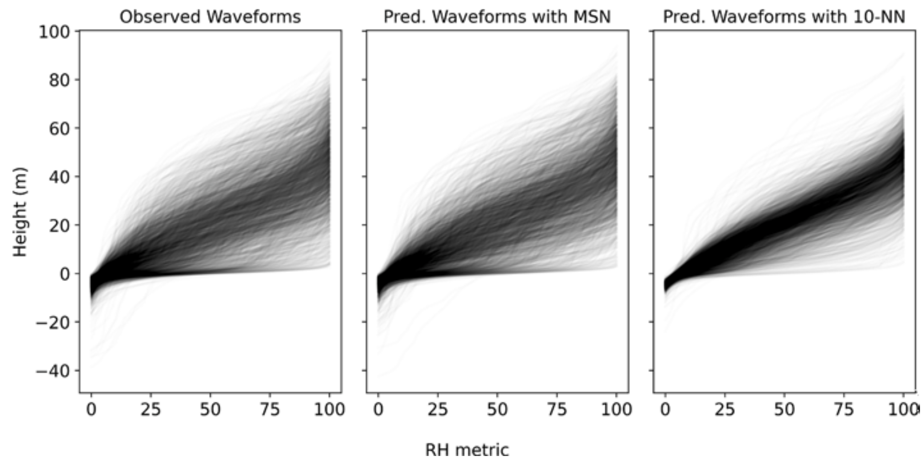

Fig. 1. The leftmost figure displays the waveforms collected by GEDI within a 10 × 10-km tile in Olympic National Park, Washington State, USA, while the remaining figures show imputed waveforms generated using different values of *k* (e.g., the number of nearest neighbors). This area was chosen for illustration purposes because of its large range of canopy heights.

Table 1
L4D (MSN), a *k*-NN model variant (10-NN), and other global canopy-height maps evaluated in this study.

	L4D (MSN)	10-NN	Potapov et al., 2021	Lang et al., 2023	Tolan et al., 2024
Approach	*k*-nearest neighbor ($k = 1$)	*k*-nearest neighbor ($k = 10$)	Ensembles of bagged regression trees	Ensembles of convolutional neural networks	Self-supervised learning and convolutional neural networks
Input data	CCDC	CCDC	Landsat analysis-ready	Sentinel-2 imagery	Maxar imagery
Referenced GEDI data	All RH metrics	All RH metrics	RH95	RH98	RH95
Output resolution	30 m	30 m	30 m	10 m	1 m
Target year	2023	No map	2019	2020	2017–2020
Target evaluation metrics	RMSE, IDR, KS test, Covariance/correlation	Covariance/ correlation	RMSE, IDR, KS test	RMSE, IDR, KS test	RMSE, IDR, KS test

1. Quality filtering based on V002 GEDI Level-2A flags: We downloaded GEDI shots for which the GEDI L2A flag 'quality_flag' was 1 (valid surface returns with sensitivity >0.9), and the 'degrade_flag' was one of 0, 10, 20, 30, 3, 13, 23, 33 (not acquired under significantly degraded geolocation conditions), and the 'leaf-off-flag' was either 0 (leaf-on) or 255 (unknown).
2. Quality filtering based on a list of sub-orbit granules excluded from the GEDI L4B product: The GEDI Science Team provides a JSON file for filtering out sub-orbit granules affected by anomalies (Dubayah et al., 2023). We excluded GEDI shots identified in this JSON file.
3. Quality filtering based on grid tiles: We filtered out GEDI shots that fall in grid tiles entirely covered by water.
4. Quality filtering based on covariate availability: We filtered out GEDI shots that had no valid covariates. The covariates used in this study are described in Section 2.1.2.
5. Quality filtering using a local outlier detection algorithm: We developed a local outlier detection algorithm to identify GEDI shots with abrupt spike values that were inconsistent with the surrounding spatial context. For each tile, we fitted a local spatial smoothing model using geocoordinates and selected environmental covariates, and used it to filter out shot-level outliers with large discrepancy from expected spatial patterns.

It should be noted that MSN offers a computational advantage over more complex approaches. Prediction is a straightforward function of multivariate distance, and while large-area prediction does require substantial parallel processing, the absence of iterative tuning and/or ensemble decision-making makes global processing less intensive than some other machine learning tools.

2.1.1. *Response variables of GEDI L4D*

MSN is particularly advantageous when the response is multivariate and the objective is to preserve not only each variable's marginal distribution but also the covariance and correlation structure among variables, given a large and representative reference dataset (Eskelson et al., 2009). In the GEDI L4D product, the set of outputs taken directly from the nearest neighbor includes GEDI waveforms (RH metrics) along with associated estimates of canopy cover and aboveground biomass density. Similar to using larger *k* in *k*-NN, fitting generalized predictive models to multivariate responses can disrupt the observed covariance structure, especially when atypical but ecologically important values are smoothed or poorly represented. MSN avoids this by transferring a complete set of jointly observed attributes from a single nearest-neighbor waveform at the expense of potentially noisier predictions.

2.1.2. *Environmental covariates*

Under the *k*-NN assumption in which observations with similar covariates are expected to have similar response values, it is essential to define the input feature space to be informative for the response variable. For the GEDI L4D product, which provides continuous 30-m predictions of GEDI L2A waveform relative height (RH) metrics, L2B canopy cover, and L4A AGBD, we used Landsat time-series data derived from Continuous Change Detection and Classification (CCDC; Zhu and Woodcock, 2014; Zhu et al., 2015; Gorelick et al., 2023) as environmental covariates. CCDC fits harmonic functions to multi-year Landsat observations, enabling the generation of synthetic reflectance values for any timestamp. To characterize land surface properties and compute similarity between locations for predicting GEDI data, we used synthetic Landsat reflectance bands (blue, green, red, near-infrared, and two shortwave infrared channels; Zhu et al., 2015) from two phenologically important dates: (1) "greenup," marking the start of the growing season,

Table 2
Validation design by each evaluation metric.

	MSN	10-NN	External maps	Notes
RMSE	20% reference waveforms (04/2022–03/2023)	NA	20% reference waveforms Potapov et al. & Tolan et al. (04/2019–03/2020) Lang et al. (01/2020–12/2020)	Evaluated for 10-km illustration tile and globally for 10-km tiles with at least 50% forest coverage (ESA map)
IDR	(10-km illustration tile) IDR over training waveforms vs IDR over 20% reference waveforms (global) Mapped IDR over forest areas vs distribution of training forest waveforms	NA	(10-km illustration tile) IDR over training waveforms vs IDR over 20% reference waveforms (global) Mapped IDR over forest areas vs distribution of training forest waveforms	Evaluated for 10-km illustration tile and globally for 10-km tiles with at least 50% forest coverage (ESA map)
KS	Distribution of reference waveform heights vs distribution of predictions at the same locations (04/2022–03/2023)	NA	Distribution of reference waveform heights vs distribution of predictions at the same locations Potapov et al. & Tolan et al. (04/2019–03/2020) Lang et al. (01/2020–12/2020)	Evaluated for 10-km illustration tile and globally for 10-km tiles with at least 50% forest coverage (ESA map)
Covariance	Covariance among RH metrics in the reference dataset vs. covariance across the waveforms imputed for the locations of the reference dataset	Covariance among RH metrics in the reference dataset vs. covariance across the waveforms imputed for the locations of the reference dataset	NA (single variable)	Evaluated for the entire Amazon ecosystem
Correlation	Correlation between RH98 and biomass in the reference dataset vs. correlation across the waveforms imputed for the locations of the reference dataset	Correlation between RH98 and biomass in the reference dataset vs. correlation across the waveforms imputed for the locations of the reference dataset	NA (single variable)	Evaluated globally for 10-km tiles with at least 50% forest coverage (ESA map)

and (2) "peak," representing maximum greenness. Both dates were identified on a 10 km tile basis using Moderate Resolution Imaging Spectroradiometer (MODIS) phenology (Zhang, 2024). In addition to the 12 synthetic Landsat bands, we included the SRTM-derived Continuous Heat-Insolation Load Index (CHILI; Theobald et al., 2015) to capture physiographic variation. In total, 13 environmental covariates were extracted for each high-quality GEDI waveform at its location and year, and for all prediction locations globally in the year 2023 to generate GEDI L4D outputs. The covariates were standardized as a preprocessing step using standard scaling.

2.1.3. *Distance metric*

Among several distance metrics available for calculating similarity in the feature space, we selected the Bray–Curtis metric, implemented in Python's SciPy library, based on preliminary evaluations of mean residual error across representative regions with different vegetation types. The Bray–Curtis metric achieved lower mean residual error than the other distance metrics across the tested regions, while showing KS test performance similar to that of the other metrics. Unlike the Bray–Curtis formulation often used for species-composition counts (Bray and Curtis, 1957), the SciPy implementation is applicable to any real-valued data. It expresses the L1 (Manhattan) distance, which is the absolute difference across all dimensions, as a proportion of the total magnitude by dividing the sum of absolute differences by the sum of absolute values, yielding a relative measure of dissimilarity:

$$d_{bc}(X,Y)=\frac{\sum_{i=1}^{m}\left|x_i-y_i\right|}{\sum_{i=1}^{m}|x_i|+\left|y_i\right|} \tag{1}$$

where X and Y are the feature vectors of two locations, and m is the number of feature dimensions. The Bray–Curtis distance is bounded between 0 (identical) and 1 (completely dissimilar).

2.1.4. *Spatial domain of the MSN model*

An MSN model was applied to each 10 km × 10 km tile globally, defined using the EASE-Grid 2.0 projection, to generate the global GEDI L4D product. The choice of tile size was based on two considerations: (1) prior study by Healey et al. (2020), which suggested that highly local models produce less bias at the low and high ends of forest structures, and 2) the need to ensure sufficient high-quality GEDI data for imputation. To impute 30 m GEDI data within a focal tile, the MSN model used observed GEDI shots not only from the focal tile but also from its eight adjacent tiles, thereby alleviating visible discontinuities along tile boundaries. The average number of high-quality shots across all 10-km modeling tiles was 7787. By default, all eight adjacent tiles were included; however, when the number of observed data points was insufficient, the reference domain was incrementally expanded until at least 2500 GEDI shots were available. Some isolated islands for which sufficient high-quality shots could not be obtained even from surrounding areas were excluded from GEDI L4D data production.

2.2. *Evaluation methods*

A central objective here is to assess, for GEDI L4D and other products, how closely the distribution of mapped values reflects the distribution of the data observed by GEDI. Specifically, RMSE was augmented by four population-level error metrics, which are described in the following subsections.

2.2.1. *Root mean square error (RMSE)*

RMSE is a commonly used metric for evaluating prediction quality (Bishop, 2006). RMSE is computed from residuals—the difference between predicted and observed values for each sample—and measures the average magnitude of prediction error. For the GEDI L4D application, RMSE can be applied to any predicted variable, including RH metrics, canopy cover, and AGBD. For example, RMSE for a given RH metric within a focal 10-km tile is defined as:

$$\mathrm{RMSE_{RHx}}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\mathrm{RHx_i}^{\mathrm{pred}}-\mathrm{RHx_i}^{\mathrm{obs}}\right)^2} \tag{2}$$

where RHx_i^{pred} is the predicted RHx metric for the i-th sample, RHx_i^{obs} is the corresponding observed value, and n is the total number of validation samples within the focal tile.

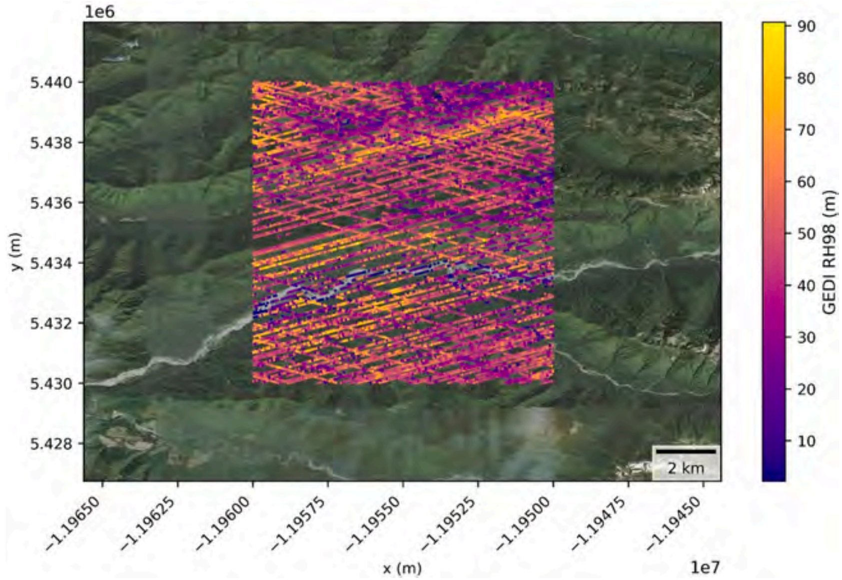

Fig. 2. A 10 km × 10 km tile in Olympic National Park, Washington State, USA, showing all quality-filtered GEDI shots. Colors represent the RH98 values (m) of the shots.

2.2.2. Inter-decile range (IDR)

To quantify variability from low to high height values, IDR was used to measure the spread (m) between the 10th and 90th percentile heights within an area of interest:

$$IDR = P_{90} - P_{10} \tag{3}$$

where P_{90} is the value at the 90th percentile of the data group and $P_{10}$is the value at the 10th percentile. Lower values indicate clustering of heights near the mean. The difference between the predicted and observed (from independent nearby GEDI waveforms) IDR values was used to assess similarity in distributional characteristics:

$$IDR_{diff} = IDR^{pred} - IDR^{obs} \tag{4}$$

A negative IDR_{diff} indicates that the canopy height predictions are more tightly clustered than GEDI's top-of-canopy height measurements. In contrast, a positive IDR suggests that predictions are more widely dispersed than the observations. An IDR_{diff} value close to zero does not guarantee that the two distributions are identical, as two distributions may share the same IDR yet differ if one is shifted relative to the other. To account for this, we also examine the difference in median values (i.e., 50th percentile, P_{50}) as a complementary measure:

$$P_{50\,diff} = P_{50}{}^{pred} - P_{50}{}^{obs} \tag{5}$$

2.2.3. Kolmogorov–Smirnov test (KS test)

To further evaluate the similarity between predicted and observed distributions, the classical Kolmogorov–Smirnov (KS) test was applied. The KS test is based on empirical cumulative distribution functions (CDFs) and provides two key values: the KS statistic (D) and the *p*-value (p). The statistic measures the maximum vertical distance between the two CDFs. For example, the KS statistic for predicted and observed *RHx* distributions within a 10-km tile is defined as:

$$D = \max_{x=RHx} |F^{pred}(x) - F^{obs}(x)| \tag{6}$$

where $F^{pred}(x)$ and $F^{obs}(x)$ are the empirical CDFs of the predicted and observed *RHx* values, respectively. A smaller *D* indicates closer agreement between the distributions, whereas a larger *D* suggests greater divergence.

The associated p-value is then used to test the null hypothesis that the two samples are drawn from the same underlying distribution. The p-value reflects how likely it is to observe a difference as large as the one measured if the predicted and observed distributions were actually identical. A smaller p-value indicates that such a difference is unlikely to occur by chance, suggesting a significant difference between the distributions. In this study, a significance level (α) of 0.05 was used, meaning that cases with $p < 0.05$ are considered statistically different, whereas cases with $p \geq 0.05$ indicate no significant difference. The KS statistics and corresponding *p*-values were obtained using the two-sample KS test implemented in the SciPy (Python) statistical library.

2.2.4. Covariance and correlation

A distinguishing feature of the GEDI L4D product is its ability to represent forest vertical structure through coherent series of RH (relative height) metrics drawn from a nearby high-quality GEDI shot. To evaluate the assumption that MSN conserves important covariances among multiple imputed variables, we tracked predicted and observed (via independent GEDI data not included as “neighbors”) covariances among all RH variables across the entire Amazonian forests (Wagner et al., 2025). This comparison was limited to a single ecosystem, albeit a large one, so that covariance patterns would not be disrupted by sharp ecological boundaries. Fig. 1 illustrates how GEDI waveforms predicted by the two *k*-NN algorithms, MSN and 10-NN, differed from the observed waveforms. The figure shows that the shape of the predicted waveforms, and consequently their covariance, can differ substantially from those derived from the direct GEDI observations.

The covariance between two RH metrics, RH_i and RH_j, is computed as:

$$\mathrm{COV}(\mathrm{RH_i}, \mathrm{RH_j}) = \frac{1}{\mathrm{N}-1}\sum_{\mathrm{k}=1}^{\mathrm{N}}(\mathrm{RH_{k,i}} - \overline{\mathrm{RH_i}})(\mathrm{RH_{k,j}} - \overline{\mathrm{RH_j}}) \tag{7}$$

where *N* is the total number of samples, $RH_{k,i}$ and $RH_{k,j}$ are the *RH* values of the *k*-th sample for metrics *i* and *j*, and $\overline{RH_i}$ and $\overline{RH_j}$ are their respective means. We then compare the covariance matrices derived from observed and predicted RH metrics to evaluate whether the covariance structure is maintained.

We also evaluated whether the correlation between AGBD and RH98 represented in the GEDI product was preserved across MSN and 10-NN. We selected RH98 because it is a primary determinant of AGBD (Kellner et al., 2023). To reduce the influence of noise and outliers, we employed the Spearman rank correlation:

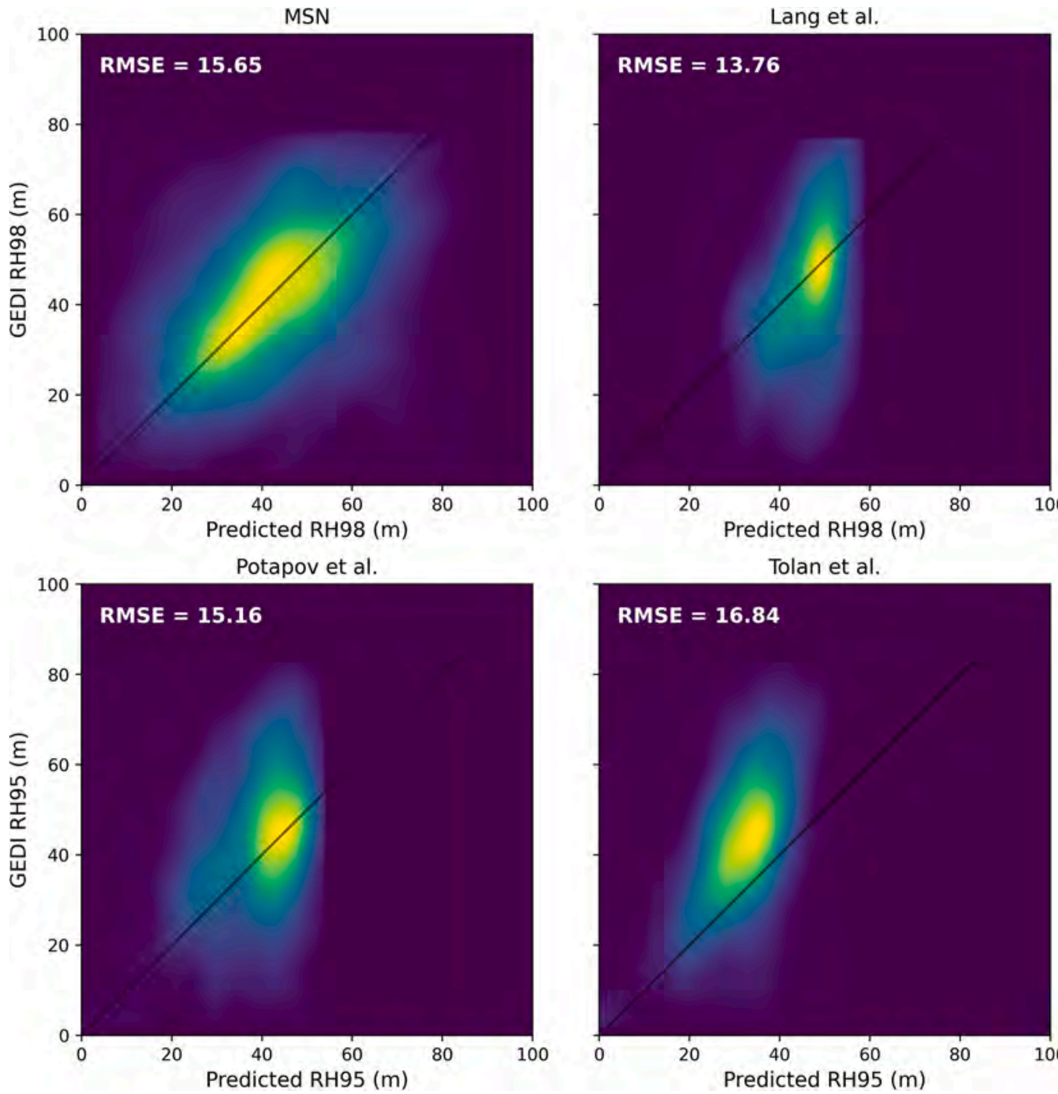

Fig. 3. Density plots comparing observed GEDI RH98 values with predictions from MSN and three global canopy height products in the illustrative study area in Washington. For each product, the selected RH metric corresponds to the reference metric used in that study. The diagonal line denotes the 1:1 relationship, and RMSE values (m) are shown for each method. MSN shows higher RMSE than most other products, but its predictions are less compressed toward the mean.

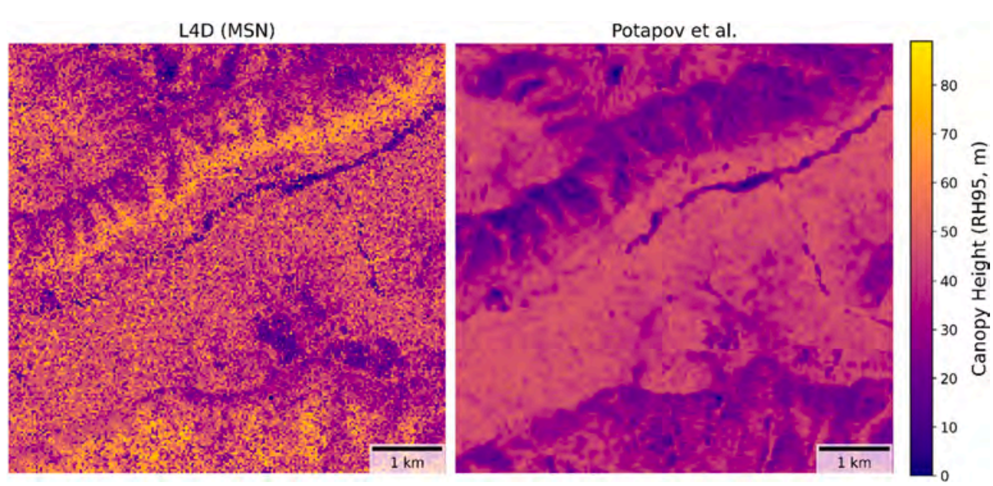

Fig. 4. Large-scale views of two 30 m resolution maps for the study area: L4D generated using MSN (left) and the map from Potapov et al. (right). The MSN-based map shows granular spatial artifacts at the local scale while better preserving both tall and short canopy heights, whereas the Potapov et al. map exhibits more spatial coherence but a compressed range of estimates.

Table 3
RMSE values reported in each study.

MSN at RH98	Lang et al. at RH98	Potapov et al. at RH95	Tolan et al. at RH95
8.5 m	6.0 m	6.6 m	n/a

Table 4
The mean RMSE values and standard deviations calculated on our forest-tile validation set.

MSN at RH98	Lang et al. at RH98	Potapov et al. at RH95	Tolan et al. at RH95
8.4 ± 3.1 m	8.2 ± 2.6 m	7.3 ± 2.6 m	7.7 ± 3.6 m

$$\rho = \frac{cov(R(RH_{98}), R(AGBD))}{\sigma_{R(RH_{98})}\sigma_{R(AGBD)}} \tag{8}$$

where $R(x)$ denotes the rank-transformed values of the variable x, and σ represents the standard deviation of the ranks. This statistic was derived at the 10-km tile level, globally, and compared to the correlation observed over the validation set of waveforms. The difference in the AGBD–RH98 correlation was computed as the predicted correlation minus the observed correlation:

$$\rho_{diff} = \rho^{pred} - \rho^{obs} \tag{9}$$

2.3. Evaluation design and comparison framework

Our validation activities focused on 10 km tiles with at least 50% forest coverage, as identified by the European Space Agency (ESA) WorldCover 10 m 2021 product (Zanaga et al., 2022), and at least 100 validation waveforms. Under these criteria, approximately 30% of 1,085,943 L4D tiles remained for evaluation. The evaluation scores were computed separately for each tile using a randomly selected 20% of the reference GEDI shots within each tile as a validation set. This tile-specific validation set was used only to evaluate performance because MSN has no parameters to tune. We did not use the entire 20% hold-out dataset for calculation of RMSE and KS statistics; each map represents only a single year, and disturbances and growth throughout the GEDI period of operation may compromise evaluation of conditions in a single year. Instead, we matched each map to the appropriate RH value (Table 1) measured in the subset of validation shots measured for the tile during the one-year period closest to the target year of each map (Table 2). Across global forested tiles, matching validation shots only in the nearest year left a mean of 400–480 shots per 10 km tile. For products with spatial resolutions finer than the GEDI footprint (approximately 25 m), we used the maximum value within a 30-m buffer around each GEDI shot to better represent local top-canopy height.

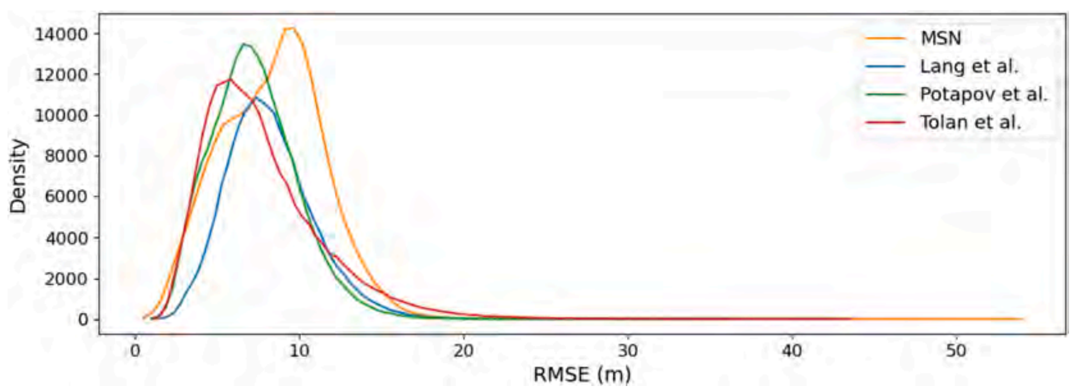

Fig. 5. Distribution of RMSE across forest tiles for global canopy height products and the MSN. The mean RMSE values are listed in Table 4.

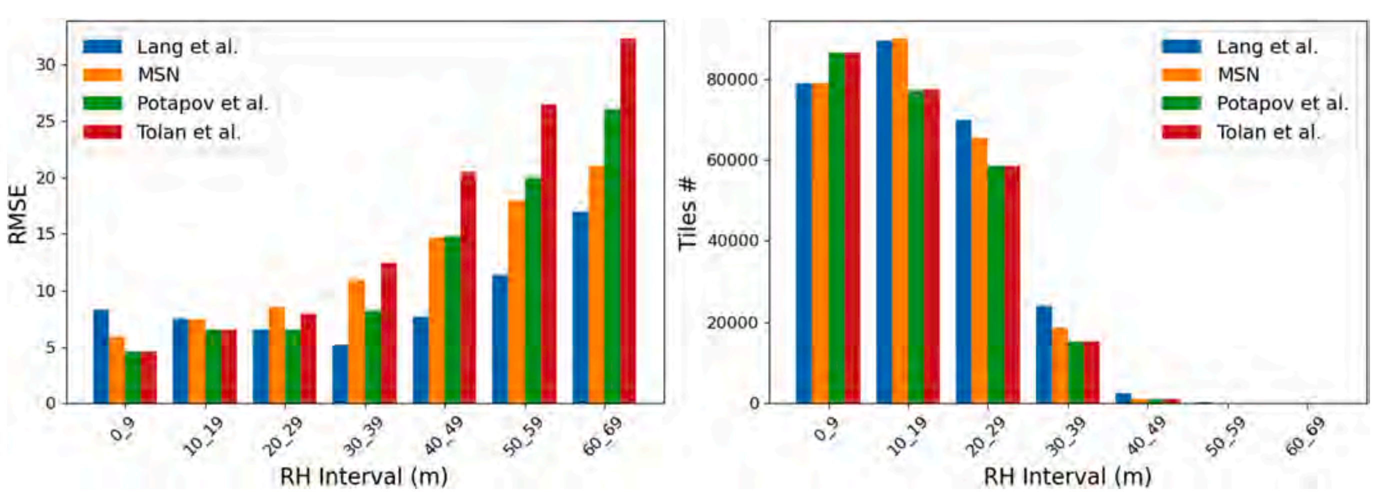

Fig. 6. RMSE across RH interval bins for Lang et al., MSN, Potapov et al., and Tolan et al. products. The left panel shows RMSE by RH interval, and the right panel shows the number of tiles evaluated within each interval. Lang shows lower RMSE at greater canopy heights but higher RMSE at lower canopy heights, consistent with overestimation in shorter forests. Potapov et al. and Tolan et al. show lower RMSE at lower canopy heights but higher RMSE at greater canopy heights, likely reflecting underestimation in taller forests. MSN shows relatively consistent performance across RH intervals and has the second-lowest RMSE in the tallest intervals, behind Lang et al.

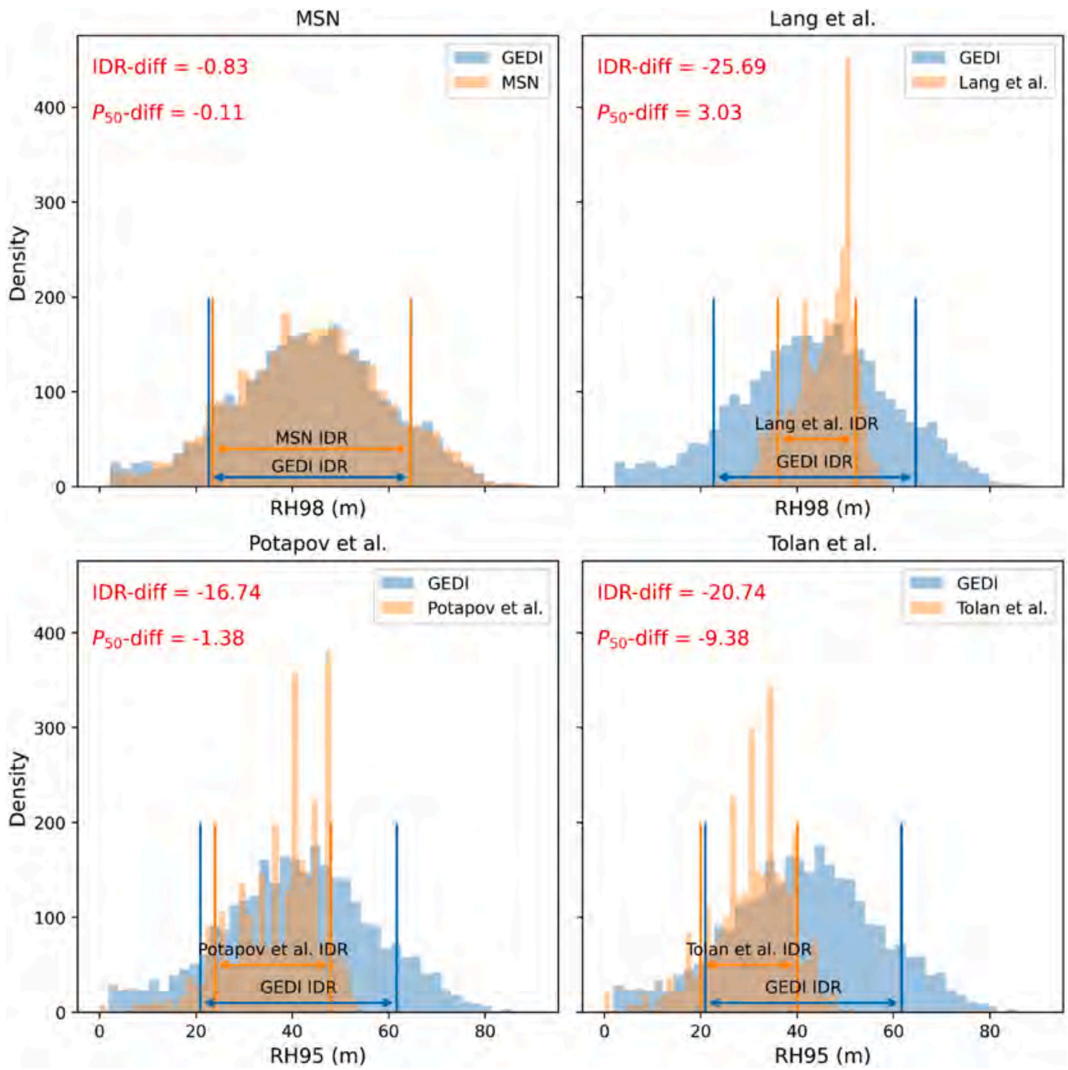

Fig. 7. Comparison of RH98 distributions for predicted and observed GEDI values using MSN and three other global canopy height products over the study area in Washington State shown in Fig. 2. For each method, the IDRs of the predicted values and the GEDI observations are shown, along with the corresponding IDR differences (IDR_{diff}) and median difference ($P_{50\,diff}$). Negative IDR_{diff} values (meters) indicate that the predicted distribution is narrower than the observed one. Negative $P_{50\,diff}$ values indicate systematic underestimation of height, whereas positive values denote overestimation.

The RMSE and KS results for MSN were derived from entirely independent waveforms, whereas we recognize that the other maps we evaluated likely used some of our validation data for training. This non-independence may have positively biased our accuracy measures for those products. IDR predictions for each product were derived from the mapped predictions made for pixels labeled as forest in the ESA WorldCover 10 m 2021 product (Zanaga et al., 2022). The mapped representation of MSN was the publicly available GEDI L4D product, which was produced using all waveforms, including those in our validation set. Predictions at the locations of the validation waveforms, for all map products, were therefore not independent of the waveforms used to determine the tile's reference IDR. For fully forested tiles, this non-independence affected an average of less than 0.005 of the predictions (i.e., about 480 of 110,000 pixels per tile). We mapped the difference between the mapped and reference IDR values at qualifying 10 km tiles for each of the maps.

Covariance and correlation across mapped variables could only be assessed for L4D, as the other products represented only canopy height. By way of context, we compared mapped covariance among RH values and correlation between biomass and RH98 values associated with the L4D MSN map against a nearest-neighbor algorithm using $k = 10$ that we treated as a proxy for algorithms which employ more generalization across predictions for the purpose of minimizing the kind of individual error that RMSE measures (Fig. 1).

It is noted that the GEDI RH metrics and imputed map used here differ slightly from the released GEDI L4D product. Here, RH metrics were retrieved from GEDI Release 2 (V002) using the default waveform processing algorithm parameter set selected for its performance in identifying the ground and canopy. This is the set currently available on Google Earth Engine (Healey et al., 2020). The released GEDI L4D product was imputed using a different algorithm selection scheme to ensure consistency with GEDI L4D AGBD predictions. The only discrepancy was that algorithm setting groups 5 and 10 were excluded for the Evergreen Broadleaf tree stratum (land_cover_data/pft_class == 2) in South America (land_cover_data/region_class == 6) (Seo et al., 2025).

Table 2 summarizes the validation exercises described herein, specifying: the reference dataset; whether or not that reference dataset was independent of the model-building process; the sample of the map considered; the set of tiles to which the validation exercise was applied; and cases where no validation was possible. Prediction distributions were evaluated at the level of the 10-km tiles used to build the MSN models. Coincidence of the modeling and validation domains for L4D presented a potential risk of artificially good agreement for metrics such

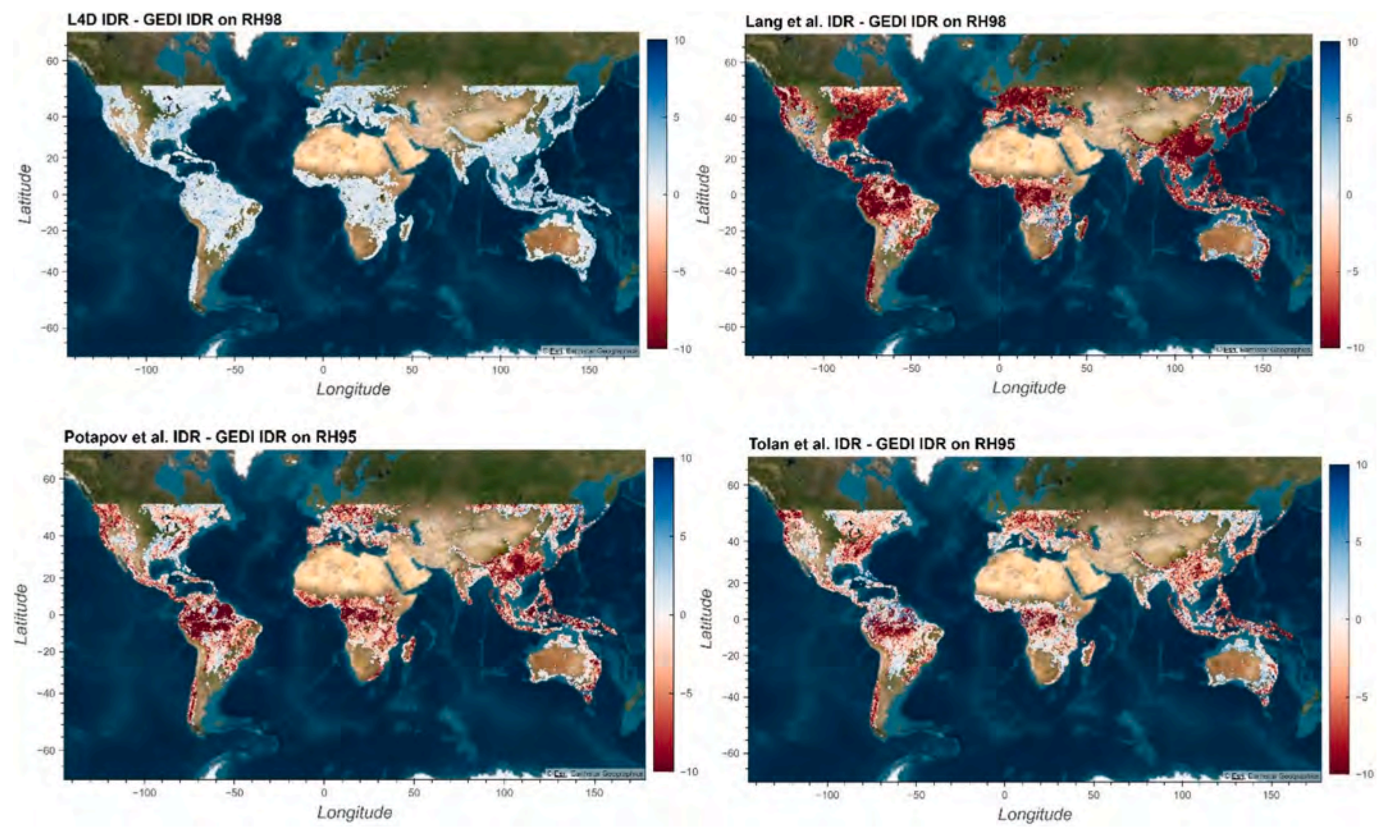

Fig. 8. Degree to which the predicted RH98 (or RH95) height range in each global product is higher (blue) or lower (red) than the range measured from GEDI shots within each 10 × 10 km forest tile. The color bar is capped at −10 and 10 m. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

as IDR – a risk not present for the existing maps. A global evaluation of L4D IDR against GEDI waveform IDR at multiple scales and on grids offset from the modeling tile system found that estimates of IDR_{diff} were robust against these concerns (Appendix A: L4D IDR difference across scales for a subset of tiles). For each error metric, we included an illustrative local-scale validation at a tile selected in Olympic National Park, Washington State, USA, an area characterized by abundant tall trees (Fig. 2). Analysis was then expanded to broader spatial scales, up to the global level.

3. Results

3.1. RMSE

3.1.1. RMSE for a 10-km illustration tile

We directly compared the performance of MSN and the other three global canopy height maps in predicting GEDI RH98 (or RH95) using the 20% validation set from the corresponding one-year period. In the illustrative tile in Washington, there were 1473 validation shots from 04/2019–03/2020, 286 available in the 01/2020–12/2020 period, and 802 available from 04/2022–03/2023. This number varied considerably across tiles and years based upon GEDI's sampling variability. MSN yielded a relatively higher RMSE for the illustration tile (Fig. 3). The range of predictions in the non-MSN products was notably compressed, which would minimize large predictive errors and benefit RMSE values. For example, the Lang et al. product showed the narrowest range of predicted values for this area and achieved the lowest RMSE. Avoidance of predictions at the extremes of the reference distribution also smooths the appearance of non-MSN maps. MSN maps often exhibit speckle noise that occurs when spatial context highlights extreme predictions likely made in error (Fig. 4).

3.1.2. RMSE at the global scale

Although published RMSE estimates (Table 3) among maps were derived using different GEDI data configurations and validation periods, they offer a general perspective on global-scale RMSE performance. Published validations generally focused on vegetated samples, and we likewise report RMSE for MSN only for tiles classified as forest (using the ESA map as described above). Tolan et al. used GEDI samples solely for post-processing their predicted canopy height derived from remote sensing imagery and airborne lidar and did not report a globally summarized RMSE.

The RMSE of MSN RH98 (8.5 m) was substantially greater than RMSEs reported for other products. This difference diminished when RMSE was evaluated for forest tiles using the methods described above (Table 4). Global accuracy of the Tolan et al. product fell between Potapov et al. and the other maps. This comparison provided a broadly comparable measure of canopy height agreement across a shared set of validation samples. The global distribution of RMSE at the 10-km tile level is shown in Fig. 5. We also summarized RMSE across specific tree-height intervals globally (Fig. 6), similar to the analysis conducted for forested regions in China by Chen et al. (2025), and found that accuracy varied differentially by product across trees of different heights. The products of Potapov et al. and Tolan et al. had lower errors than the other products for short trees, for example, those below 20 m, whereas Lang et al. had the lowest errors for tall trees above 30 m, with MSN exhibiting the second-lowest errors from 40 m onward. Additional RMSE results by continent and plant functional type (PFT) are provided in Appendix Table A1.

3.2. IDR difference

3.2.1. IDR difference for a 10-km illustration tile

As defined in Section 2.2.2, IDR provides a measure of distributional

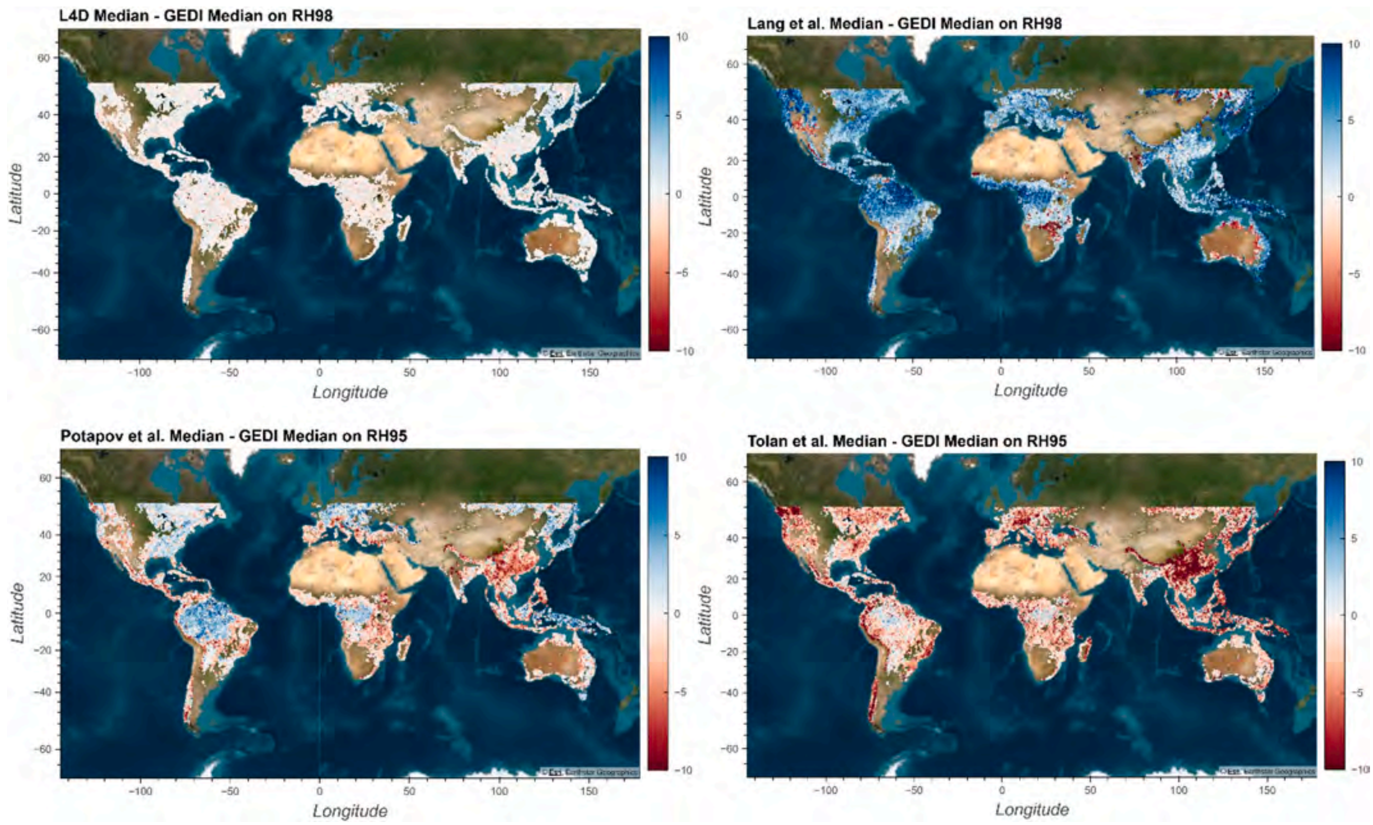

Fig. 9. Amount (meters) by which predicted median RH98 (or RH95) height in each global product is higher (blue) or lower (red) than the median measured from GEDI shots within each 10 × 10 km forest tile. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

spread. Here, we compared the IDR of predicted and observed RH98 values using the results from MSN and three other global canopy height products. Fig. 7 illustrates the distribution of their predictions relative to the validation data, along with the corresponding IDR_{diff} and P_{50diff} values. The distribution of MSN predictions closely matches that of the GEDI observations, whereas the other products are compressed toward the center, failing to reflect the full range of the observed values. Furthermore, the Tolan et al. and Potapov et al. estimates tended to underestimate canopy height in this tile, while the Lang et al. map exhibits systematic overprediction.

3.2.2. IDR difference at the global scale

We extended the IDR evaluation to the global scale. Deviations between GEDI-sampled and mapped IDR values were mapped globally across all forest tiles between latitudes −51.6° and 51.6° (Fig. 8). Values near zero indicate strong concordance with the GEDI-based range; negative values (red) indicate a compressed spread, whereas positive values (blue) reflect a mapped distribution broader than the one measured by GEDI. Across the mapped domain, the L4D product generated using MSN exhibits the closest alignment with the GEDI-derived IDR, while the other products tend to produce much narrower ranges, particularly in forests dominated by tall trees.

Because similar IDR values do not necessarily imply similar distributions—especially when distributions are shifted—we additionally evaluated the median difference (Fig. 9). The median value mapped by Lang et al. shows pronounced overestimation across most forest tiles, whereas Tolan et al. was generally below the median observed at the reference GEDI shots. Potapov et al. displays a mixture of over- and underestimation globally, and L4D shows little systematic bias.

Results for these population-level metrics, stratified by geographic region and PFT, are provided in Appendix Tables A.2 and A.3.

3.3. KS test

3.3.1. KS test for a 10-km illustration tile

Progressing from the IDR, which relies on two specific quantiles of the distribution (the 10th and 90th), we also performed the classical two-sample KS test to compare the full distributions using their CDFs. Fig. 10 shows the CDFs of the GEDI data alongside the predictions from MSN and other products, together with the corresponding KS statistics and *p*-values. The MSN estimates align closely with the GEDI distribution, as evidenced by a small KS statistic and a high p-value. By comparison, the other global canopy-height maps display a significant departure from the GEDI distribution, reflected in a larger KS statistic and a near-zero p-value.

3.3.2. KS test at the global scale

We also applied the KS test to each global product using our forest-tile validation data from the corresponding one-year period. Table 5 summarizes the results by reporting the percentage of tiles rejected at the 0.05 significance level—that is, the proportion of tiles whose predicted distributions are statistically distinguishable from the GEDI distribution. All approaches except MSN yielded predictions that deviated from the GEDI distribution in most cases, and this pattern was also observed when the results were stratified by geographic region and PFT (Appendix Table A4).

3.4. Covariance among RH metrics across the Amazon

A distinguishing feature of GEDI L4D, compared to existing map products, is that it provides multi-dimensional imputed waveforms rather than only canopy top height estimates. When $k = 1$ (i.e., imputing the most similar neighbor), nearest neighbor models are expected to approximately

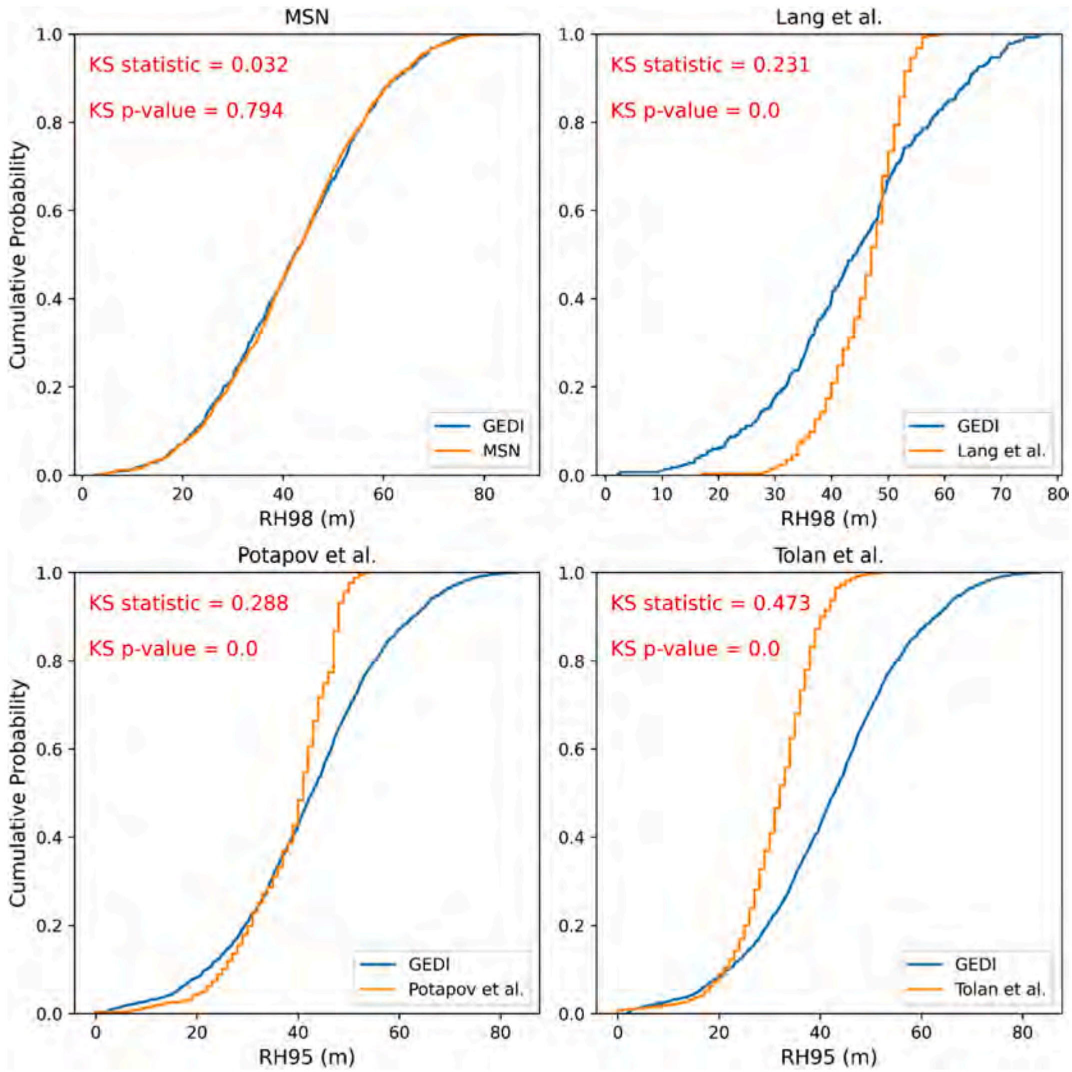

Fig. 10. Comparison of CDFs for predicted and observed GEDI RH98 values using MSN and three global canopy-height products. KS statistics and p-values from the two-sample KS test are reported in each panel. MSN shows a close match to the empirical distribution, whereas the three other maps exhibit marked deviation. Distributions reflect map values in the case study area in Washington.

Table 5
Proportion of tiles whose predicted distributions differ statistically from the GEDI observations.

MSN at RH98	Lang et al. at RH98	Potapov et al. at RH95	Tolan et al. at RH95
26.12%	99.68%	98.24%	97.06%

retain the covariance structure among imputed variables (Eskelson et al., 2009). Fig. 11 affirms this effect when covariance was measured among RH values across the Amazon Basin. For the sake of comparison, an imputation algorithm using 10 neighbors was tested as a proxy for algorithms producing more generalized predictions at the expense of some homogenization around a mean. Covariances among RH values of all levels, especially those near the top of the canopy, were reduced in the 10-NN map relative to the relationships observed in the validation set across the region. This finding aligns with the reduced waveform variability observed in the example tile in Washington State (Fig. 1).

3.5. Global correlation between RH98 and AGBD

Lastly, we examine how the RH98-AGBD relationship observed in the GEDI data compares with that produced by the two k-NN approaches at the global scale (Fig. 12). As with the IDR-difference maps (Fig. 8), values close to 0 indicate little difference relative to the GEDI-derived correlation. Negative values (red) denote that the predicted correlation is lower than the observed one, whereas positive values (blue) indicate an inflated correlation. Across much of the domain, the 10-NN predictions of biomass and canopy height exhibited higher correlation than that measured by GEDI at the footprint level. In contrast, MSN showed minimal difference in the RH98-AGBD correlation overall, with the exception of certain regions (e.g., south of the Amazon) where the mission applied multiple GEDI L4A AGBD models within the same 10-km tiles. Results stratified by continental region and PFT are presented in Appendix Table A5.

4. Discussion

4.1. Tradeoffs among current canopy height maps

The results highlight the differing strengths and limitations of currently available global forest structure maps, underscoring the importance of aligning map selection with specific scientific objectives. Signal saturation is common with optical and SAR sensors, whose sensitivity to canopy structure diminishes in tall, dense forests (Healey et al., 2020; Potapov et al., 2021). Canopy height maps generated

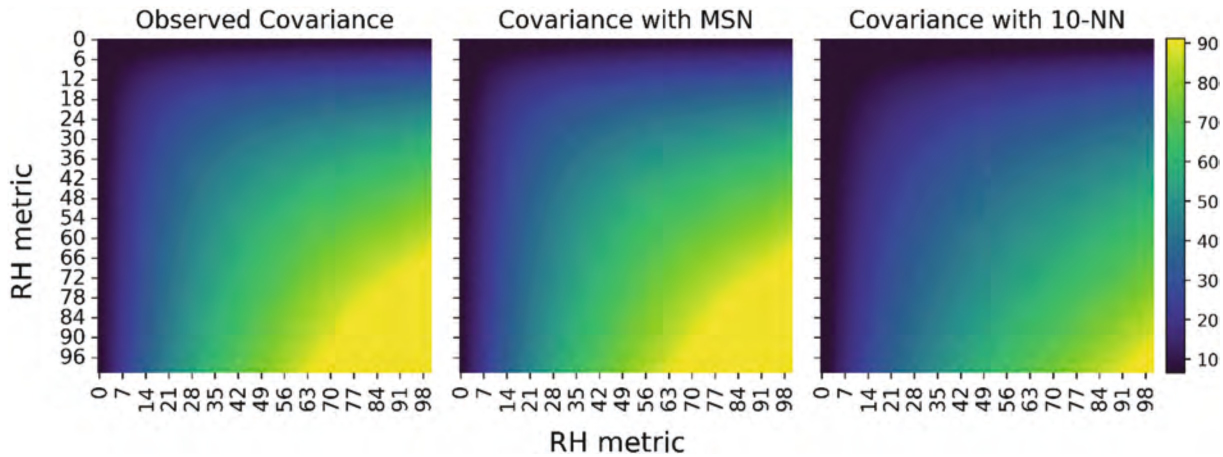

Fig. 11. Covariance matrices of RH metrics derived from observed GEDI shots within the Amazon Forest (left), MSN predictions (middle), and 10-NN predictions (right). The color bar is capped at the 5th and 95th percentiles. The MSN output closely reproduces the covariance structure present in the observations shown in the left panel, whereas the 10-NN predictions yield a diminished covariance pattern.

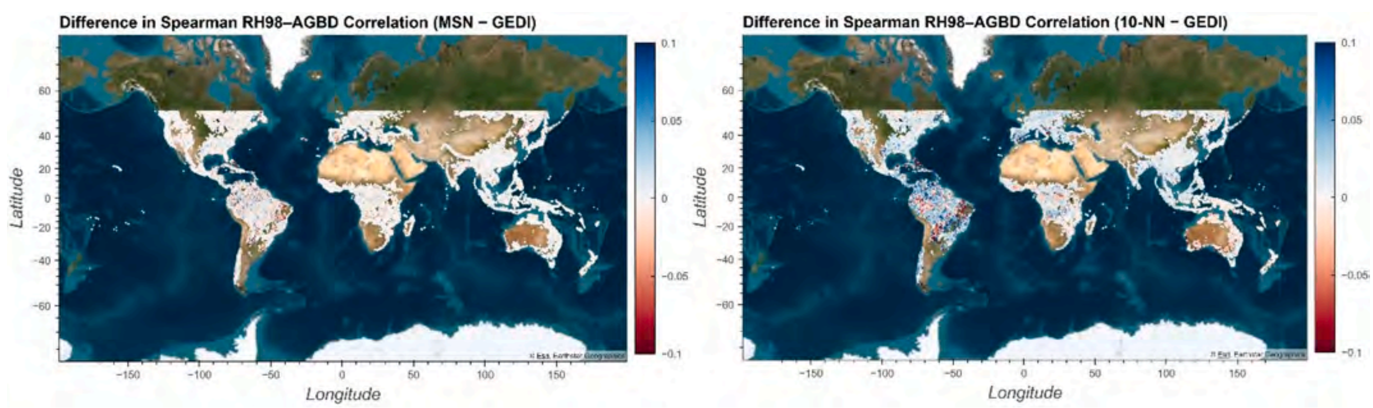

Fig. 12. Degree to which the predicted RH98-AGBD Spearman correlation is higher (blue) or lower (red) than the correlation measured from GEDI shots within each 10 × 10 km forest tile under both MSN and 10-NN imputation. The color bar is capped at −0.1 and 0.1. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

using approaches that minimize mean residual error typically achieve lower RMSE but systematically distort the underlying canopy height distribution, often biasing predictions toward mean values and attenuating extremes. As a result, fine-scale variability and structural heterogeneity—key characteristics of forest canopies—may be underrepresented. In contrast, the canopy height map derived from GEDI L4D prioritizes preservation of the empirical height distribution observed in GEDI measurements. By design, this approach more faithfully reproduces realistic canopy structural heterogeneity across landscapes, albeit at the cost of reduced per-pixel (or point-level) precision. A further limitation is that MSN approaches can yield pronounced local discontinuity (e.g., abrupt pixel-to-pixel changes) when that value is spatially misallocated. In practice, L4D maps often display speckle artifacts at the local level, as spatial context highlights occasional large prediction errors. While low-pass filtering may minimize this effect, the impact of such post-processing on measures of variability such as IDR is currently unknown.

When highly accurate estimates at fine spatial resolution are required, or when the goal is to recover generalized spatial patterns with smoother, more spatially coherent structure, canopy height maps optimized for minimizing mean error may be most appropriate. For example, tree or forest segmentation tasks (Gibril et al., 2023; Ball et al., 2023) are likely to benefit from products such as Tolan's, which are less likely than GEDI L4D to disrupt patch integrity with a large error in predicted height. The same dynamic is likely in applications related to fine-scale species distribution models (Burns et al., 2020; Betts et al., 2022; Denan et al., 2023) or in mapping landscape-scale features that emphasize connectivity.

However, the homogenization we document here of tree heights in previous global maps (at least with respect to the distributions measured by GEDI) can severely compromise decision support in many other practical applications. For instance, prediction of fire spread and effects over intermediate scales benefits from a realistic depiction of height variation (Agee, 1996; Conard and Hilbruner, 2003); overly uniform representations of forest height may degrade our ability to assess fire risk and predict fire effects. Likewise, forest structural diversity is a strong predictor of the biodiversity of animals using the forest (Goetz et al., 2007) and multi-scale tree height variability is often related to the presence of individual species (Shirley et al., 2013; Huang et al., 2014; Ducci et al., 2015; Burns et al., 2020; Seo et al., 2021; Chiaverini et al., 2023; Gerstner et al., 2024; Moudrý et al., 2024). Maps that fail to adequately represent the range of tree heights may therefore lead to underestimation of animal biodiversity and mistaken expectations about which areas are vital for conservation. Lastly, maps that distort the range of tree sizes within a woodshed may introduce errors in GIS-based analysis to support the timber processing industry. Specifically, better matching of mapped procurement demand (Anderson et al., 2011) with the distribution of tree sizes across woodsheds of various sizes may lead to better scoping of transportation costs and future product classes. In the context of these latter applications, the distributional accuracy we document for the L4D product provides a concrete benefit over products optimized only for minimization of RMSE.

4.2. Reconciliation with model-based inference

The GEDI mission was designed to accommodate model-based inference (Dubayah et al., 2022a). Under that paradigm, population units are considered to be random variables, and the uncertainty of inferences such as the local mean biomass specified in the GEDI L4B product is drawn from properties of the models that GEDI enables (Dubayah et al., 2022b). Fitting of the mission's biomass models accordingly preserves information that allows covariance of prediction errors to be propagated across areas of interest to support this mode of inference (Kellner et al., 2023). Linear models may be accommodated in this process through analytical expressions of covariance (Patterson et al., 2019), and machine learning may also be accommodated through bootstrapping approaches that involve either resampling the model-fitting data or varying the individual model-building points through manipulations related to observed residuals (Saarela et al., 2025).

Since the L4D imputed waveform maps are GEDI mission products, it is of interest to explore how they correspond to the model-based paradigm. Nearest neighbor models have been used in a variety of model-based applications, including through a process where multiple maps were built by resampling the model-building “neighbors” with replacement (McRoberts, 2012), and another where alternative maps were produced using synthetic “neighbors” that were iteratively simulated using covariances among dependent and independent variables evident in the model-building data (Ene et al., 2013). If future versions of L4D were to include support for model-based estimation, similar methods might be used to distribute alternative imputations capable of supporting representation of the covarying effects of model error. Such a product might efficiently take the form of a list of imputed shot numbers, indexed by bootstrap, for each pixel. Any subsequent applications (for instance, to predict species distribution from imputed waveforms) would require additional bootstrapping (e.g., Saarela et al., 2025) to leverage the alternative imputations into model-based variance estimates for the population of interest.

4.3. Toward multi-objective canopy height models

As illustrated, no single approach or data product achieves optimal performance across all dimensions of forest structure map quality. However, hybrid approaches may help alleviate apparent tradeoffs between RMSE and distributional error. In particular, optimization frameworks that explicitly balance multiple objectives offer a promising direction for future research. A wide range of tradeoffs has been studied in mathematics and machine learning, including bias–variance (Geman et al., 1992), accuracy–model complexity (Tibshirani, 1996), precision–recall (Powers, 2011), accuracy–fairness (Agarwal et al., 2018), performance–computational cost (Tan et al., 2019), and exploration–exploitation (Sutton and Barto, 2018). These challenges are commonly addressed through approaches such as scalarization, Pareto optimization (Miettinen, 1999; Deb, 2001), constraint-based optimization (Boyd and Vandenberghe, 2004), and multi-task or multi-head learning (Caruana, 1997). Most existing methods, however, primarily emphasize point-wise optimization for computational efficiency, making it non-trivial to integrate per-sample accuracy with richer population-level characteristics beyond simple summary statistics (e.g., the mean).

Vision transformer algorithms (Dosovitskiy et al., 2021) may also hold promise for mitigating the speckle error obvious in MSN maps (Fig. 4), either through establishment of objectives related to prediction coherence or more simply through reformulation of input features to include multi-scalar information. Developing alternative modeling frameworks that better reconcile these perspectives therefore represents a worthwhile avenue for advancing large-scale forest structure mapping using dense GEDI observations.

5. Conclusion

NASA's GEDI instrument was conceived as a sampling tool, providing dense coverage of the range of conditions that occur within Earth's forests (Dubayah et al., 2022a). This paper shows that choice of modeling strategy in using GEDI to calibrate wall-to-wall forest structure maps has implications for the dimensions in which those maps are accurate. Existing global canopy height maps minimize pixel-level errors, but they do so while clustering predictions around a mean value. The GEDI L4D Imputed Waveform product represents a design decision that sacrifices precision to achieve predictions that much better match GEDI's population-level measurements. The result is better characterization of the variability, range, and distribution of canopy heights. The emerging potential of global satellite laser altimetry for quantifying population-level errors may catalyze development of model objective functions that reconcile map strengths with application needs.

CRediT authorship contribution statement

Eugene Seo: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Sean P. Healey:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Zhiqiang Yang:** Writing – review & editing, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **John D. Armston:** Writing – review & editing, Resources, Methodology, Investigation, Data curation. **Ralph O. Dubayah:** Writing – review & editing, Resources, Investigation, Funding acquisition, Data curation, Conceptualization. **Simon Ilyushchenko:** Writing – review & editing, Resources, Data curation. **Tiago De Conto:** Writing – review & editing, Resources, Investigation, Data curation. **Matthew G. Betts:** Writing – review & editing, Resources, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was partially funded by NASA Competed GEDI Science Team grants NNH24OB110 to the USDA Forest Service and 80NSSC26K0147 to Oregon State University. The University of Maryland contributors were funded by NASA Contract NNL15AA03C for the development and execution of the GEDI mission (Dubayah, Principal Investigator). The findings and conclusions in this publication are those of the authors and should not be construed to represent any official USDA or U.S. Government determination or policy. The authors acknowledge the Advanced Research Computing Services (ARCS @OSU) supercomputing resources (https://arcs.oregonstate.edu) made available for conducting the research reported in this paper.

Appendix A. IDR difference at different domain scales

The main analysis was conducted at the 10-km tile scale used for L4D modeling and mapping. To assess the sensitivity of the distributional agreement between the L4D map and GEDI observation to domain size, we examined how the L4D-GEDI IDR difference varies with spatial scale beyond 10 km using 100 randomly selected forest locations (Fig. A.1a). Fig. A.1b summarizes IDR difference across the tested scales, which remain generally low throughout.

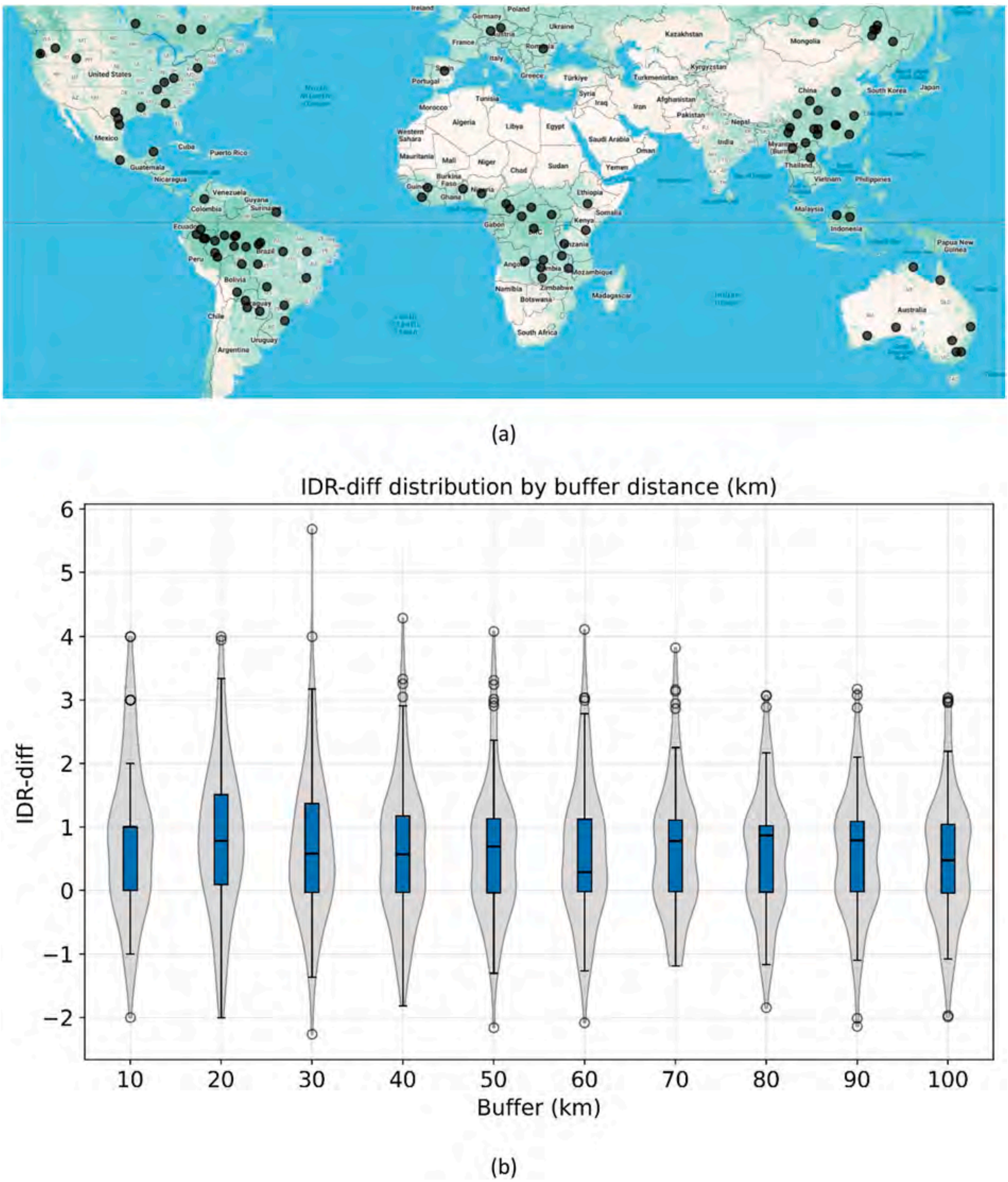

Fig. A.1. Sensitivity of L4D-GEDI IDR difference to domain (buffer) size. (a) One hundred randomly selected locations with at least 1000 forest pixels within the 10-km buffer and few missing GEDI L4D pixels. (b) Distribution of IDR difference between imputed RH98 values from the GEDI L4D map and the corresponding high-quality GEDI observations within each buffer boundary. The violins show the distributions; the box plots show the median and interquartile range (IQR; Q1-Q3); whiskers indicate the non-outlier range; and dots represent individual samples beyond the whiskers.

Table A1
RMSE for the validation set, summarized by continent and by PFT type. Values in each cell are ordered as MSN, Lang et al., Potapov et al., and Tolan et al. Generally, Potapov et al. and Tolan et al. produced lower RMSE across continents and PFT types.

	MSN at RH98	Lang et al. at RH98	Potapov et al. at RH95	Tolan et al. at RH95
Europe	7.9 ± 2.3	7.2 ± 1.6	7.3 ± 2.0	**7.0 ± 2.6**
N. Asia	7.7 ± 2.8	8.1 ± 2.3	**7.4 ± 2.9**	8.0 ± 3.8
Australasia	7.8 ± 3.9	8.7 ± 3.2	**6.5 ± 2.9**	7.2 ± 3.8
Africa	8.3 ± 3.3	8.1 ± 3.0	**6.7 ± 2.5**	6.9 ± 2.9
S. Asia	9.9 ± 2.7	**8.5 ± 2.1**	8.8 ± 2.8	10.2 ± 4.1
S. America	8.7 ± 2.8	8.6 ± 2.9	**7.4 ± 2.2**	7.8 ± 3.4
N. America	7.1 ± 2.9	7.5 ± 2.3	6.4 ± 2.4	**6.3 ± 2.8**

	MSN at RH98	Lang et al. at RH98	Potapov et al. at RH95	Tolan et al. at RH95
ENT	7.4 ± 3.1	7.6 ± 2.4	**6.8 ± 3.0**	7.2 ± 3.6
EBT	9.4 ± 2.8	8.8 ± 2.7	**7.8 ± 2.5**	8.5 ± 3.7
DNT	6.3 ± 1.2	6.8 ± 1.8	**5.7 ± 1.3**	5.9 ± 1.6
DBT	7.3 ± 2.5	7.3 ± 2.2	6.6 ± 2.1	**6.4 ± 2.7**
Mixed Forests	**3.3 ± 1.3**	5.7 ± 1.3	4.1 ± 1.1	3.5 ± 1.2
Closed Shrublands	**4.9 ± 2.0**	6.7 ± 1.9	5.1 ± 2.2	5.0 ± 2.2
Open Shrublands	6.0 ± 1.7	6.9 ± 1.5	5.6 ± 1.3	**5.3 ± 1.5**
Woody Savannas	5.4 ± 1.8	6.4 ± 1.6	5.3 ± 1.5	**4.8 ± 1.4**
Savannas	8.5 ± 1.9	8.8 ± 1.9	7.4 ± 1.5	**6.8 ± 1.5**

Table A2
IDR-diff summarized by continent and PFT type. Values in each cell are ordered as MSN, Lang et al., Potapov et al., and Tolan et al. Negative values indicate that the predicted RH98 or RH95 height range is lower than the GEDI-observed range within each 10 × 10 km forest tile, while positive values indicate that the predicted range is higher. All products except MSN showed a narrower range than the actual range (i.e., negative value) across almost all continental regions and PFT types.

	MSN at RH98	Lang et al. at RH98	Potapov et al. at RH95	Tolan et al. at RH95
Europe	**1.8 ± 1.5**	−8.0 ± 4.2	−2.6 ± 4.3	−3.4 ± 3.9
N. Asia	**1.5 ± 2.0**	−7.0 ± 6.3	−4.9 ± 5.1	−3.0 ± 4.3
Australasia	**1.6 ± 2.0**	−4.9 ± 7.6	−5.4 ± 6.8	−3.4 ± 6.1
Africa	**1.0 ± 1.7**	−6.5 ± 6.9	−7.1 ± 5.8	−4.1 ± 5.6
S. Asia	**1.7 ± 2.1**	−9.5 ± 5.5	−7.6 ± 5.3	−4.3 ± 5.3
S. America	**1.2 ± 1.9**	−8.7 ± 6.0	−8.0 ± 6.4	−3.9 ± 6.0
N. America	**1.6 ± 1.7**	−7.2 ± 5.2	−3.0 ± 4.1	−3.4 ± 3.9

	MSN at RH98	Lang et al. at RH98	Potapov et al. at RH95	Tolan et al. at RH95
ENT	**1.4 ± 1.4**	−7.3 ± 4.7	−4.2 ± 4.3	−3.3 ± 4.2
EBT	**1.4 ± 2.0**	−9.1 ± 5.7	−8.0 ± 6.2	−4.5 ± 5.8
DNT	**1.0 ± 1.6**	−5.4 ± 4.7	−5.3 ± 3.5	−1.7 ± 3.0
DBT	**1.5 ± 2.0**	−6.1 ± 6.5	−3.8 ± 4.3	−3.1 ± 4.0
Mixed Forests	1.0 ± 1.0	1.7 ± 4.4	**−0.7 ± 2.8**	1.5 ± 2.7
Closed Shrublands	1.3 ± 1.7	−0.7 ± 6.2	−1.4 ± 3.7	**−0.3 ± 3.5**
Open Shrublands	2.0 ± 2.0	−6.3 ± 5.9	**−1.3 ± 3.9**	−2.5 ± 4.1
Woody Savannas	1.9 ± 2.2	−3.2 ± 5.8	−1.8 ± 3.8	**−1.4 ± 4.0**
Savannas	**1.9 ± 1.6**	−9.6 ± 4.8	−2.6 ± 4.1	−4.0 ± 3.8

Table A3
Median-diff summarized by continent and PFT type. Values in each cell are ordered as MSN, Lang et al., Potapov et al., and Tolan et al. Negative values indicate underestimation of the GEDI-observed RH metric, while positive values indicate overestimation. MSN and Potapov et al. has very low bias, while Lang et al. has overestimations across all continental regions and PFT types (except Mixed Forests) and Tolan et al. has substantial underestimations across all cases.

	MSN at RH98	Lang et al. at RH98	Potapov et al. at RH95	Tolan et al. at RH95
Europe	−0.2 ± 1.4	4.1 ± 2.6	**0.0 ± 3.8**	−7.6 ± 63.8
N. Asia	**0.2 ± 1.7**	4.4 ± 4.8	−0.9 ± 3.8	−11.4 ± 75.1
Australasia	**0.0 ± 1.6**	5.0 ± 6.0	1.1 ± 4.2	−14.5 ± 104.7
Africa	**−0.4 ± 1.3**	4.4 ± 4.6	−0.4 ± 3.4	−6.7 ± 67.3
S. Asia	**−0.2 ± 1.6**	3.1 ± 3.6	−3.1 ± 3.7	−44.3 ± 192.9
S. America	**−0.2 ± 1.5**	6.0 ± 3.5	1.9 ± 4.1	−11.7 ± 97.9

(continued on next page)

Table A3 (*continued*)

	MSN at RH98	Lang et al. at RH98	Potapov et al. at RH95	Tolan et al. at RH95
N. America	−0.3 ± 1.2	4.4 ± 3.4	**0.0 ± 2.8**	−8.1 ± 66.9

	MSN at RH98	**Lang et al. at RH98**	**Potapov et al. at RH95**	**Tolan et al. at RH95**
ENT	**−0.4 ± 1.2**	4.7 ± 2.8	−0.8 ± 3.1	−12.8 ± 87.4
EBT	**−0.2 ± 1.6**	5.4 ± 3.8	0.4 ± 4.5	−19.1 ± 125.0
DNT	**−0.2 ± 1.1**	2.5 ± 3.4	−0.3 ± 2.3	−4.1 ± 2.7
DBT	**−0.2 ± 1.2**	2.9 ± 4.3	−0.8 ± 3.1	−8.8 ± 71.1
Mixed Forests	**0.2 ± 0.8**	−3.1 ± 4.8	−2.8 ± 2.6	−41.2 ± 195.3
Closed Shrublands	**0.3 ± 1.5**	2.1 ± 6.2	−0.9 ± 3.1	−11.7 ± 97.4
Open Shrublands	**0.9 ± 2.1**	6.2 ± 4.6	1.1 ± 3.8	−10.6 ± 91.3
Woody Savannas	0.8 ± 2.0	4.1 ± 5.5	**−0.3 ± 3.6**	−10.7 ± 96.7
Savannas	**0.6 ± 2.0**	6.9 ± 3.3	0.9 ± 3.5	−2.1 ± 3.3

Table A4
Proportion of tiles for which predicted distributions differed significantly from GEDI observations, summarized by continent and by PFT type. Values in each cell are ordered as MSN, Lang et al., Potapov et al., and Tolan et al. MSN had substantially fewer tiles where the predicted distributions statistically differed from the GEDI distribution compared with all other products, across all continental regions and PFT types.

	MSN at RH98	**Lang et al. at RH98**	**Potapov et al. at RH95**	**Tolan et al. at RH95**
Europe	**19.7**	99.9	99.0	99.1
N. Asia	**20.0**	99.3	99.1	98.0
Australasia	**27.8**	99.9	97.4	97.3
Africa	**23.9**	99.7	96.6	94.0
S. Asia	**23.1**	99.4	98.1	97.1
S. America	**33.6**	99.8	98.9	97.5
N. America	**22.1**	99.9	98.4	97.4

	MSN at RH98	**Lang et al. at RH98**	**Potapov et al. at RH95**	**Tolan et al. at RH95**
ENT	**23.2**	99.8	99.0	98.7
EBT	**29.2**	99.7	97.9	96.6
DNT	**12.7**	98.8	99.0	94.5
DBT	**16.1**	99.4	97.9	95.4
Mixed Forests	**31.2**	100.0	100.0	99.0
Closed Shrublands	**23.2**	99.9	99.8	99.0
Open Shrublands	**24.4**	100.0	99.8	99.6
Woody Savannas	**23.3**	99.8	99.8	99.2
Savannas	**17.0**	100.0	98.7	98.6

Table A5
Difference between the predicted and GEDI-observed RH98–AGBD Spearman correlations, summarized by continent and PFT type for both MSN and 10-NN imputation. Positive values indicate higher predicted correlations than those observed in GEDI data, while negative values indicate lower predicted correlations. Values in each cell are ordered as MSN and 10-NN. Because the mean values are small, standard deviations are also reported. MSN closely preserved the GEDI-observed RH98–AGBD correlations, whereas 10-NN showed larger variability across all cases.

	MSN	10-NN
Europe	**0.00 ± 0.02**	0.01 ± 0.03
N. Asia	**0.00 ± 0.01**	0.01 ± 0.02
Australasia	**0.00 ± 0.01**	0.01 ± 0.02
Africa	**0.00 ± 0.02**	0.00 ± 0.03
S. Asia	**0.00 ± 0.01**	0.01 ± 0.02
S. America	**0.00 ± 0.03**	0.01 ± 0.05
N. America	**0.00 ± 0.01**	0.01 ± 0.02

(*continued on next page*)

Table A5 (*continued*)

	MSN	10-NN
	MSN	**10-NN**
ENT	**0.00 ± 0.01**	0.01 ± 0.02
EBT	**0.00 ± 0.02**	0.01 ± 0.04
DNT	**0.00 ± 0.01**	0.02 ± 0.02
DBT	**0.00 ± 0.02**	0.00 ± 0.03
Mixed Forests	**0.00 ± 0.01**	0.01 ± 0.03
Closed Shrublands	**0.00 ± 0.02**	0.01 ± 0.04
Open Shrublands	**0.00 ± 0.02**	0.01 ± 0.03
Woody Savannas	**0.00 ± 0.02**	0.01 ± 0.04
Savannas	**0.00 ± 0.01**	0.00 ± 0.02

Data availability

I have shared the link to my data at the Attach File step.

ORNL-DAAC

GEDI L4D Imputed Waveforms, Version 2 (Original data), https://developers.google.com/earth-engine/datasets/catalog/LARSE_GEDI_GEDI04_D_002.

References

Agarwal, A., Beygelzimer, A., Dudik, M., Langford, J., Wallach, H., 2018. A reductions approach to fair classification. In: Proceedings of the 35th International Conference on Machine Learning. PMLR, pp. 60–69.

Agee, J.K., 1996. The influence of forest structure on fire behavior. In: Proceedings of the 17th Annual Forest Vegetation Management Conference, pp. 52–68.

Anderson, N.M., Germain, R.H., Bevilacqua, E., 2011. Geographic information system-based spatial analysis of sawmill wood procurement. J. For. 109, 34–42. https://doi.org/10.1093/jof/109.1.34.

Ball, J.G.C., Hickman, S.H.M., Jackson, T.D., Koay, X.J., Hirst, J., Jay, W., Archer, M., Aubry-Kientz, M., Vincent, G., Coomes, D.A., 2023. Accurate delineation of individual tree crowns in tropical forests from aerial RGB imagery using mask R-CNN. Remote Sens. Ecol. Conserv. 9, 641–655. https://doi.org/10.1002/rse2.332.

Betts, M.G., Yang, Z., Hadley, A.S., Smith, A.C., Rousseau, J.S., Northrup, J.M., Nocera, J. J., Gorelick, N., Gerber, B.D., 2022. Forest degradation drives widespread avian habitat and population declines. Nat. Ecol. Evol. 6, 709–719. https://doi.org/10.1038/s41559-022-01737-8.

Bishop, C.M., 2006. Pattern Recognition and Machine Learning. Springer, Berlin.

Boyd, S., Vandenberghe, L., 2004. Convex Optimization. Cambridge University Press.

Bray, J.R., Curtis, J.T., 1957. An ordination of the upland forest communities of southern Wisconsin. Ecol. Monogr. 27, 326–349. https://doi.org/10.2307/1942268.

Burns, P., Clark, M., Salas, L., Hancock, S., Leland, D., Jantz, P., Dubayah, R., Goetz, S.J., 2020. Incorporating canopy structure from simulated GEDI lidar into bird species distribution models. Environ. Res. Lett. 15, 095002. https://doi.org/10.1088/1748-9326/ab80ee.

Caruana, R., 1997. Multitask learning. Mach. Learn. 28, 41–75. https://doi.org/10.1023/A:1007379606734.

Chen, A., Cheng, K., Chen, Y., Qi, Z., Yang, H., Ren, Y., Yang, Z., Chen, M., Xu, J., Zhang, Y., Huang, G., Xiang, T., Zhang, J., Liu, W., Lin, D., Xu, G., Guo, Q., 2025. Validating recent global canopy height maps over China's forests based on UAV lidar data. Remote Sens. Environ. 329, 114957. https://doi.org/10.1016/j.rse.2025.114957.

Chiaverini, L., Macdonald, D.W., Hearn, A.J., Kaszta, Ż., Ash, E., Bothwell, H.M., Can, Ö. E., Channa, P., Clements, G.R., Haidir, I.A., Kyaw, P.P., Moore, J.H., Rasphone, A., Tan, C.K.W., Cushman, S.A., 2023. Not seeing the forest for the trees: generalised linear model out-performs random forest in species distribution modelling for southeast Asian felids. Ecol. Inform. 75, 102026. https://doi.org/10.1016/j.ecoinf.2023.102026.

Conard, S.G., Hilbruner, M., 2003. Influence of Forest Structure on Wildfire Behavior and the Severity of its Effects. USDA For. Serv., Washington, DC.

Deb, K., 2001. Multi-Objective Optimization Using Evolutionary Algorithms. Wiley.

Denan, N., Norhisham, A.R., Sanusi, R., Stone, J., Azhar, B., 2023. Stand-level habitat characteristics and edge habitats drive biological pest control services in the understory of oil palm plantations. Biol. Control 183, 105261. https://doi.org/10.1016/j.biocontrol.2023.105261.

Dhargay, S., Lyell, C.S., Brown, T.P., Inbar, A., Sheridan, G.J., Lane, P.N.J., 2022. Performance of GEDI space-borne lidar for quantifying structural variation in the temperate forests of South-Eastern Australia. Remote Sens. 14, 3615. https://doi.org/10.3390/rs14153615.

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., Houlsby, N., 2021. An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale. International Conference on Learning Representations. OpenReview.

Dubayah, R., Blair, J.B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S., Armston, J., Tang, H., Duncanson, L., Hancock, S., Jantz, P., Marselis, S., Patterson, P.L., Qi, W., Silva, C., 2020. The global ecosystem dynamics investigation: high-resolution laser ranging of the earth's forests and topography. Sci. Remote Sens. 1, 100002. https://doi.org/10.1016/j.srs.2020.100002.

Dubayah, R., Armston, J., Healey, S.P., Bruening, J.M., Patterson, P.L., Kellner, J.R., Duncanson, L., Saarela, S., Ståhl, G., Yang, Z., Tang, H., Blair, J.B., Fatoyinbo, L., Goetz, S., Hancock, S., Hansen, M., Hofton, M., Hurtt, G., Luthcke, S., 2022a. GEDI launches a new era of biomass inference from space. Environ. Res. Lett. 17, 095001. https://doi.org/10.1088/1748-9326/ac8694.

Dubayah, R.O., Armston, J., Healey, S.P., Yang, Z., Patterson, P.L., Saarela, S., Stahl, G., Duncanson, L., Kellner, J.R., 2022b. GEDI L4B gridded aboveground biomass density, version 2. ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/2017.

Dubayah, R.O., Armston, J., Healey, S.P., Yang, Z., Patterson, P.L., Saarela, S., Stahl, G., Duncanson, L., Kellner, J.R., Bruening, J.M., Pascual, A., 2023. GEDI L4B Gridded Aboveground Biomass Density, Version 2.1. ORNL Distributed Active Archive Center. https://doi.org/10.3334/ORNLDAAC/2299.

Ducci, L., Agnelli, P., Di Febbraro, M., Frate, L., Russo, D., Loy, A., Carranza, M.L., Santini, G., Roscioni, F., 2015. Different bat guilds perceive their habitat in different ways: a multiscale landscape approach for variable selection in species distribution modelling. Landsc. Ecol. 30, 2147–2159. https://doi.org/10.1007/s10980-015-0237-x.

Ene, L.T., Næsset, E., Gobakken, T., 2013. Model-based inference for k-nearest neighbours predictions using a canonical vine copula. Scand. J. For. Res. 28, 266–281. https://doi.org/10.1080/02827581.2012.723743.

Eskelson, B.N.I., Temesgen, H., Lemay, V., Barrett, T.M., Crookston, N.L., Hudak, A.T., 2009. The roles of nearest neighbor methods in imputing missing data in forest inventory and monitoring databases. Scand. J. For. Res. 24, 235–246. https://doi.org/10.1080/02827580902870490.

Geman, S., Bienenstock, E., Doursat, R., 1992. Neural networks and the bias/variance dilemma. Neural Comput. 4, 1–58. https://doi.org/10.1162/neco.1992.4.1.1.

Gerstner, B.E., Blair, M.E., Bills, P., Cruz-Rodriguez, C.A., Zarnetske, P.L., 2024. The influence of scale-dependent geodiversity on species distribution models in a biodiversity hotspot. Philos. Trans. Series A Math. Phys. Eng. Sci. 382. https://doi.org/10.1098/rsta.2023.0057.

Gibril, M.B.A., Mohd Shafri, H.Z., Al-Ruzouq, R., Shanableh, A., Nahas, F., Al Mansoori, S., 2023. Large-scale date palm tree segmentation from multiscale UAV-based and aerial images using deep vision transformers. Drones 7, 93. https://doi.org/10.3390/drones7020093.

Goetz, S., Steinberg, D., Dubayah, R., Blair, B., 2007. Laser remote sensing of canopy habitat heterogeneity as a predictor of bird species richness in an eastern temperate forest, USA. Remote Sens. Environ. 108, 254–263. https://doi.org/10.1016/j.rse.2006.11.016.

Gorelick, N., Yang, Z., Arévalo, P., Bullock, E.L., Insfrán, K.P., Healey, S.P., 2023. A global time series dataset to facilitate forest greenhouse gas reporting. Environ. Res. Lett. 18, 084001. https://doi.org/10.1088/1748-9326/ace2da.

Healey, S.P., Yang, Z., Cohen, W.B., Pierce, D.J., 2006. Application of two regression-based methods to estimate the effects of partial harvest on forest structure using Landsat data. Remote Sens. Environ. 101, 115–126. https://doi.org/10.1016/j.rse.2005.12.006.

Healey, S.P., Yang, Z., Gorelick, N., Ilyushchenko, S., 2020. Highly local model calibration with a new GEDI LiDAR asset on Google earth engine reduces Landsat forest height signal saturation. Remote Sens. 12, 2840. https://doi.org/10.3390/rs12172840.

Heisig, J., Milenković, M., Pebesma, E., 2025. High-resolution canopy fuel maps based on GEDI: a foundation for wildfire modeling in Germany. Environ. Res. Ecol. 4, 015003. https://doi.org/10.1088/2752-664X/adaaf9.

Huang, Q., Swatantran, A., Dubayah, R., Goetz, S.J., 2014. The influence of vegetation height heterogeneity on forest and woodland bird species richness across the United States. PLoS One 9, e103236. https://doi.org/10.1371/journal.pone.0103236.

Hudak, A.T., Crookston, N.L., Evans, J.S., Hall, D.E., Falkowski, M.J., 2008. Nearest neighbor imputation of species-level, plot-scale forest structure attributes from

LiDAR data. Remote Sens. Environ. 112, 2232–2245. https://doi.org/10.1016/j.rse.2007.10.009.

Kellner, J.R., Armston, J., Duncanson, L., 2023. Algorithm theoretical basis document for GEDI footprint aboveground biomass density. Earth Space Sci. 10. https://doi.org/10.1029/2022EA002516 e2022EA002516.

Lang, N., Jetz, W., Schindler, K., Wegner, J.D., 2023. A high-resolution canopy height model of the earth. Nat. Ecol. Evol. 7, 1778–1789. https://doi.org/10.1038/s41559-023-02206-6.

Ma, L., Hurtt, G., Tang, H., Lamb, R., Lister, A., Chini, L., Dubayah, R., Armston, J., Campbell, E., Duncanson, L., Healey, S., O'Neil-Dunne, J., Ott, L., Poulter, B., Shen, Q., 2023. Spatial heterogeneity of global forest aboveground carbon stocks and fluxes constrained by spaceborne lidar data and mechanistic modeling. Glob. Chang. Biol. 29, 3378–3394. https://doi.org/10.1111/gcb.16682.

Mandl, L., Stritih, A., Seidl, R., Ginzler, C., Senf, C., 2023. Spaceborne LiDAR for characterizing forest structure across scales in the European Alps. Remote Sens. Ecol. Conserv. 9, 599–614. https://doi.org/10.1002/rse2.330.

McRoberts, R.E., 2012. Estimating forest attribute parameters for small areas using nearest neighbors techniques. For. Ecol. Manag. 272, 3–12. https://doi.org/10.1016/j.foreco.2011.06.039.

Miettinen, K., 1999. Nonlinear Multiobjective Optimization. Springer.

Moeur, M., Stage, A.R., 1995. Most similar neighbor: an improved sampling inference procedure for natural resource planning. For. Sci. 41, 337–359. https://doi.org/10.1093/forestscience/41.2.337.

Moudrý, V., Gábor, L., Marselis, S., Pracná, P., Barták, V., Prošek, J., Navrátilová, B., Novotný, J., Potůčková, M., Gdulová, K., Crespo-Peremarch, P., Komárek, J., Malavasi, M., Rocchini, D., Ruiz, L.A., Torralba, J., Torresani, M., Cazzolla Gatti, R., Wild, J., 2024. Comparison of three global canopy height maps and their applicability to biodiversity modeling: accuracy issues revealed. Ecosphere 15, e70026. https://doi.org/10.1002/ecs2.70026.

Patterson, P.L., Healey, S.P., Ståhl, G., Saarela, S., Holm, S., Andersen, H.-E., Dubayah, R. O., Duncanson, L., Hancock, S., Armston, J., Kellner, J.R., Cohen, W.B., Yang, Z., 2019. Statistical properties of hybrid estimators proposed for GEDI—NASA'S global ecosystem dynamics investigation. Environ. Res. Lett. 14, 065007. https://doi.org/10.1088/1748-9326/ab18df.

Picotte, J.J., Dockter, D., Long, J., Tolk, B., Davidson, A., Peterson, B., 2019. LANDFIRE remap prototype mapping effort: developing a new framework for mapping vegetation classification, change, and structure. Fire 2, 35. https://doi.org/10.3390/fire2020035.

Potapov, P., Li, X., Hernandez-Serna, A., Tyukavina, A., Hansen, M.C., Kommareddy, A., Pickens, A., Turubanova, S., Tang, H., Silva, C.E., Armston, J., Dubayah, R., Blair, J. B., Hofton, M., 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.

Powers, D.M.W., 2011. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. J. Mach. Learn. Technol. 2, 37–63.

Roy, D.P., Kashongwe, H.B., Armston, J., 2021. The impact of geolocation uncertainty on GEDI tropical forest canopy height estimation and change monitoring. Sci. Remote Sens. 4, 100024. https://doi.org/10.1016/j.srs.2021.100024.

Saarela, S., Healey, S.P., Yang, Z., Roald, B.-E., Patterson, P.L., Gobakken, T., Næsset, E., Hou, Z., McRoberts, R.E., Ståhl, G., 2025. A separable bootstrap variance estimation algorithm for hierarchical model-based inference of forest aboveground biomass using data from NASA'S GEDI and Landsat missions. Environmetrics 36, e2883. https://doi.org/10.1002/env.2883.

Seo, E., Hutchinson, R.A., Fu, X., Li, C., Hallman, T.A., Kilbride, J., Robinson, W.D., 2021. StatEcoNet: statistical ecology neural networks for species distribution Modeling. Proc. AAAI Conf. Artif. Intell. 35, 513–521. https://doi.org/10.1609/aaai.v35i1.16129.

Seo, E., Healey, S.P., Yang, Z., Dubayah, R.O., De Conto, T., Armston, J., 2025. GEDI L4D imputed waveforms, version 2. ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/2455.

Shirley, S.M., Yang, Z., Hutchinson, R.A., Alexander, J.D., McGarigal, K., Betts, M.G., 2013. Species distribution modelling for the people: unclassified Landsat TM imagery predicts bird occurrence at fine resolutions. Divers. Distrib. 19, 855–866. https://doi.org/10.1111/ddi.12093.

Simard, M., Pinto, N., Fisher, J.B., Baccini, A., 2011. Mapping forest canopy height globally with spaceborne lidar. J. Geophys. Res. Biogeosci. 116. https://doi.org/10.1029/2011JG001708.

Sutton, R.S., Barto, A.G., 2018. Reinforcement Learning: An Introduction, second ed. MIT Press.

Tan, M., Chen, B., Pang, R., Vasudevan, V., Sandler, M., Howard, A., Le, Q.V., 2019. MnasNet: platform-aware neural architecture search for mobile. In: 2019 IEEE/CVF conf. Comput. Vis. Pattern Recognit (CVPR) 2815–2823. https://doi.org/10.1109/CVPR.2019.00293.

Theobald, D.M., Harrison-Atlas, D., Monahan, W.B., Albano, C.M., 2015. Ecologically-relevant maps of landforms and physiographic diversity for climate adaptation planning. PLoS One 10, e0143619. https://doi.org/10.1371/journal.pone.0143619.

Tibshirani, R., 1996. Regression shrinkage and selection via the lasso. J. R. Stat. Soc. Ser. B Stat. Methodol. 58, 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x.

Tolan, J., Yang, H.-I., Nosarzewski, B., Couairon, G., Vo, H.V., Brandt, J., Spore, J., Majumdar, S., Haziza, D., Vamaraju, J., Moutakanni, T., Bojanowski, P., Johns, T., White, B., Tiecke, T., Couprie, C., 2024. Very high resolution canopy height maps from RGB imagery using self-supervised vision transformer and convolutional decoder trained on aerial lidar. Remote Sens. Environ. 300, 113888. https://doi.org/10.1016/j.rse.2023.113888.

Valbuena, R., Hernando, A., Manzanera, J.A., Martínez-Falero, E., García-Abril, A., Mola-Yudego, B., 2018. Most similar neighbor imputation of forest attributes using metrics derived from combined airborne LIDAR and multispectral sensors. Int. J. Digit. Earth 11, 1205–1218. https://doi.org/10.1080/17538947.2017.1387183.

Wagner, F.H., Dalagnol, R., Carter, G., Hirye, M.C.M., Gill, S., Takougoum, L.B.S., Favrichon, S., Keller, M., Ometto, J.P.H.B., Alves, L., Creze, C., George-Chacon, S.P., Li, S., Liu, Z., Mullissa, A., Yang, Y., Santos, E.G., Worden, S.R., Brandt, M., Ciais, P., Hagen, S.C., Saatchi, S., 2025. Wall-to-wall Amazon forest height mapping with Planet NICFI, aerial LiDAR, and a U-Net regression model. Remote Sens. Ecol. Conserv. https://doi.org/10.1002/rse2.70041.

Zanaga, D., Van De Kerchove, R., Daems, D., De Keersmaecker, W., Brockmann, C., Kirches, G., Wevers, J., Cartus, O., Santoro, M., Fritz, S., Lesiv, M., Herold, M., Tsendbazar, N.E., Xu, P., Ramoino, F., Arino, O., 2022. ESA WorldCover 10 m 2021 v200. https://doi.org/10.5281/zenodo.7254221.

Zhang, X., 2024. VIIRS/NPP land surface phenology yearly L3 global 0.05Deg CMG V002. NASA land processes distributed active archive Center. https://doi.org/10.5067/VIIRS/VNP22C2.002.

Zhu, Z., Woodcock, C.E., 2014. Continuous change detection and classification of land cover using all available Landsat data. Remote Sens. Environ. 144, 152–171. https://doi.org/10.1016/j.rse.2014.01.011.

Zhu, Z., Woodcock, C.E., Holden, C., Yang, Z., 2015. Generating synthetic Landsat images based on all available Landsat data: predicting Landsat surface reflectance at any given time. Remote Sens. Environ. 162, 67–83. https://doi.org/10.1016/j.rse.2015.02.009.