Ng et al. *New Zealand Journal of Forestry Science* (2025) 55:14
https://doi.org/10.33494/nzjfs552025x446x
E-ISSN: 1179-5395
published on-line: 05/09/2025

RESEARCH ARTICLE **Open Access**

New Zealand Journal of Forestry Science

Performance of global canopy height models across varied New Zealand vegetation types

Sue Kee Ng[1,*], Pascal Sirguey[1] and Todd Redpath[1,2]

[1]*National School of Surveying, University of Otago, PO Box 56, Dunedin, New Zealand*
[2]*Interpine Group Ltd, PO Box 1209, Rotorua, New Zealand*

*Corresponding author: j.nsk.1408@gmail.com

(Received for publication 6 February 2025; accepted in revised form 24 July 2025)
Editor: Horacio. E. Bown

Abstract

Background: Global canopy height models are becoming prolific yet require evaluation across New Zealand's diverse vegetation types to assess their accuracy and applicability. Accurate measurement of canopy height is crucial for estimating above-ground woody biomass, which is essential for modelling carbon emissions and sequestration in the context of climate change. These models generally rely on remote sensing data and machine learning techniques, with Light Detection and Ranging (LiDAR) technology commonly employed for precise measurement.

Methods: This study validated the three latest global canopy height models, each provided at a different resolution: 30-metre, 10-metre, and 1-metre. We assessed the accuracy of the selected models by comparing them against canopy height estimates derived from local Airborne Laser Scanning (ALS) datasets, which served as our reference data. Eleven regions across New Zealand were selected based on ALS data availability, encompassing five vegetation and land cover types. Our methodology involved utilising and automating the processing of large New Zealand ALS datasets. To align resolutions for comparison, the reference canopy height was calculated by aggregating average or maximum heights at 10 and 30 m spatial resolution. Model performances were assessed using statistical metrics, including root-mean-square error (RMSE), bias, and R^2.

Results: Overall, all models exhibited relatively low R^2 values, indicating limited capture of canopy height variability. The Potapov 30-metre model performed best with average aggregation in shorter vegetation. In contrast, the Lang 10-metre model showed improved accuracy with maximum aggregation, particularly in taller vegetation, but visual boundaries between different vegetation types were not as distinct. The Tolan 1-metre model provided a balanced approach, minimising biases in lower heights but underestimating taller canopies. Results highlight model-specific strengths for varying vegetation structures and the sensitivity of performances to aggregation methods applied to high-resolution reference ALS data.

Conclusions: All three global canopy height models exhibit varied performance across New Zealand's vegetation types. The findings highlight the importance of vegetation-specific applications to optimise each global model's accuracy. Currently, these models are suitable for carbon accounting efforts as supplementary tools rather than replacements for existing methodologies.

Keywords: LiDAR, Canopy height, GEDI, Machine learning, New Zealand vegetation

Introduction

Forests play a pivotal role as one of the most significant natural carbon sinks in the global effort to mitigate the escalating impacts of climate change (Hunt 2009). Carbon sequestration in forests is critical for offsetting anthropogenic emissions and stabilising global temperatures, particularly in the context of international climate goals such as the Paris Agreement (Lorenz & Lal 2010; IPCC 2021). Consequently, there is growing international interest in accurately quantifying forest carbon stocks and monitoring their changes over time (Brown 2002). This interest is especially relevant in New

Zealand, where forest ecosystems are central to national carbon accounting efforts. The country's Emissions Trading Scheme (ETS) emphasises the importance of accurately quantifying forest carbon stocks as growers strive for equitable compensation for carbon sequestration (Ministry for the Environment 2024).

One of the main approaches for assessing forest carbon stocks is measuring canopy height, a critical variable that correlates strongly with aboveground biomass and carbon storage (Coomes et al. 2018). LiDAR (Light Detection and Ranging) technology has been commonly used to measure canopy height with high precision, providing detailed three-dimensional data on forest structures (Wallace et al. 2012; Peterson et al. 2007). LiDAR's ability to capture forest attributes, including wood volume and leaf area index, makes it invaluable for ecological and forest management applications (Lim et al. 2003). However, the widespread application of airborne LiDAR is often limited by high operational costs and logistical challenges, making it impractical for large-scale, frequent surveys. For example, New Zealand's national LiDAR programme, initiated by Land Information New Zealand (LINZ) in 2016, had aimed to cover 80% of the country by 2024 (Lee et al. 2023).

While this represents a major achievement, the ALS datasets are regionally acquired, vary in acquisition dates, and are not regularly updated at a national scale. This temporal inconsistency limits their use for long-term change detection compared to models derived from consistent satellite observations. Global CHMs, with more frequent satellite-based updates, offer potential for monitoring canopy height changes over time at broad scales. Furthermore, the comprehensive ALS archive in New Zealand provides a desirable testing ground to benchmark global CHMs. By leveraging this high-quality archive, this study contributes to both national and international efforts to evaluate the reliability and limitations of emerging global forest monitoring tools.

Global canopy height models (CHMs) provide pre-processed, accessible datasets for assessing forest canopy height and biomass at varying spatial resolutions. Unlike custom-built models that require technical expertise (Fogel et al. 2024) or paid products from providers such as Planet (Planet Labs PBC 2024), global CHMs are freely available and more accessible to implement for large-scale forest assessments. These models are commonly developed using machine learning methods and integrating multiple remote sensing sources (Li et al. 2020). They either rely on spatial or airborne LiDAR measurements, including data from the Global Ecosystem Dynamics Investigation (GEDI) or Airborne Laser Scanning (ALS) (Lang et al. 2023; Potapov et al. 2021; Tolan et al. 2024). These LiDAR measurements are often supplemented by multispectral data (Sentinel-2, Landsat) and occasionally radar data (Sentinel-1, Alos) to predict canopy height. This approach addresses their respective limitations, such as using denser optical imagery to complement the sparse spaceborne LiDAR coverage (Lang et al. 2023). Machine learning techniques have further enhanced the accuracy of these models, allowing for the fusion of various data sources to generate high-resolution CHMs with more extensive coverage (Alvites et al. 2024).

Several notable global CHMs have emerged in recent years. Potapov et al. (2021) created a global canopy height model at a 30-metre resolution using regression tree ensembles trained on Landsat spectral data. Building on this work, Lang et al. (2023) developed a 10-metre resolution CHM using a window-based deep learning approach, incorporating textural and spectral data from Sentinel-2 imagery. Most recently, Tolan et al. (2024) introduced a 1-metre resolution global CHM, utilising advanced machine learning techniques such as self-supervised vision transformers and convolutional decoders on Maxar satellite imagery and aerial LiDAR data for fine-scale accuracy. These advancements in global CHM development hold great potential for forest monitoring at both national and global scales.

While these global CHMs show promise, their applicability at finer spatial scales and across diverse vegetation types remains to be determined. In New Zealand, it is essential to assess the accuracy of these models across different vegetation types with unique structural characteristics. For instance, different vegetation types have distinct relationships between canopy height and biomass, making it crucial to assess the accuracy of global CHMs to avoid misestimating carbon stocks, which directly affects carbon accounting and Emissions Trading Scheme (ETS) compensation (Köhler et al. 2010; Li et al. 2024). Additionally, global models are often trained on datasets that primarily represent broad, global-scale vegetation patterns, potentially overlooking local characteristics such as New Zealand's indigenous forests.

Pearse et al. (2025) demonstrated a deep learning-based forest mapping using high-resolution ALS data to generate detailed descriptions of forests, including canopy height, specifically focusing on exotic plantation forests dominated by *Pinus radiata* in New Zealand. While this represents state-of-the-art progress for exotic forestry in New Zealand, it has yet to be deployed on a national scale and does not address the unique structural attributes of indigenous forests or other vegetation types. This emphasises the need to evaluate global CHMs, which can account for a broader range of forest types, as complementary tools to enhance these local monitoring efforts. Validating the global CHMs using a robust statistical approach in New Zealand can establish a national-scale accuracy baseline and improve the understanding of the model's performance in diverse local ecological settings.

Therefore, this study aims to address these knowledge gaps by evaluating the performance of the three global high-resolution CHMs—developed by Potapov et al. (2021), Lang et al. (2023), and Tolan et al. (2024)—across New Zealand's varied vegetation types. These models were selected based on their recency, representing the latest documented developments in global canopy height modelling using machine learning and satellite imagery, as well as their open availability and global coverage. Specifically, the objectives are to: (1) Assess the accuracy of the three above-mentioned global CHMs by comparing them with local airborne LiDAR datasets as reference data; and (2) Evaluate their performance across a variety of New Zealand vegetation types, including non-vegetation, grasslands, shrublands, indigenous forests, and exotic plantations. This study focused on vegetation-related variation in CHM performance. Although topographic variation can affect canopy height estimates, particularly where models rely on optical imagery, its specific influence was not characterised here. Future work should further explore this aspect to better understand how terrain complexity interacts with global CHM accuracy.

Methods

Study area

The study area comprises eleven regions across New Zealand, selected based on the availability of New Zealand airborne LiDAR datasets that overlapped with the observation dates used in the three selected CHMs: 2019 and 2020 (Table 1). These collective regions—Bay of Plenty, Waikato Reporoa and Upper Piako River, Waikato Hamilton, Gisborne, Canterbury, Tasman, West Coast, Northland, Wellington City, Otago-Balclutha, and Marlborough—encompass approximately 84,000 km^2 (Figure 1). These sites represent a range of ecosystems, from dense temperate rainforests on the West Coast to dry grasslands and shrublands in the eastern regions like Canterbury. The North Island, with warmer temperatures, supports primarily evergreen forests, while the South Island's varied climate is influenced by the Southern Alps, creating wetter western regions and drier eastern ones (Allen et al. 2013; Rogers et al. 2005). Predominant land covers in these regions include high-producing exotic grasslands, indigenous forests, shrubland and exotic plantations (LRNZ 2020). Elevations average 269 metres, and slopes vary from flat to steep (>35°). These regions feature minimal permanent snow and ice coverage, allowing for year-round vegetation analysis. The regions considered span approximately 34°S to 47°S latitude and 169°E to 179°E longitude.

TABLE 1: The date of ALS capture for each region, along with the corresponding LiDAR sensors used, dominant vegetation types, mean ± standard deviation of slope and elevation above mean sea level (a.s.l.). The layer identifier (Layer ID) on LINZ data services repository (https://data.linz.govt.nz/) for each LiDAR dataset is also indicated, including Digital Elevation Model (DEM) and Digital Surface Model (DSM).

Region	Date of ALS	LiDAR sensors	Dominant vegetation types	Slope (degree)	Elevation (m a.s.l)	LINZ Layer ID
Bay of Plenty	27/10/2019 - 24/10/2022	Optech Galaxy PRIME	Indigenous forest	10±10	350±244	105690, 105691
Waikato - Reporoa and Upper Piako River	16/4/2019 - 17/4/2019	Optech Galaxy PRIME	High-producing exotic grassland	5±5	361±178	104108, 104113
Waikato - Hamilton	3/11/2019 - 5/11/2019	Optech Galaxy PRIME	High-producing exotic grassland	1±1	38±12	104772, 104773
Gisborne	31/12/2018 - 9/10/2020	Optech Orion H300	High-producing exotic grassland	12±7	377±269	105614, 105396
Canterbury	14/3/2018 - 1/5/2019	Optech Orion H300	High-producing exotic grassland	2±4	208±143	104931, 104936
	1/5/2020 - 4/2/2023	Optech Galaxy PRIME	High-producing exotic grassland	14±11	737±472	111133, 111135
Tasman	28/1/2020 - 30/1/2022	Optech Galaxy PRIME	Indigenous forest	19±10	746±406	112854, 112856
West Coast	16/5/2020 - 14/2/2022	Optech Galaxy PRIME	Indigenous forest	13±12	360±375	110163, 110164
Northland	1/12/2018 - 1/2/2020	Trimble AX6oi	High-producing exotic grassland	6±5	104±112	110757, 110911
Marlborough	10/2/2020 - 15/2/2022	Optech Galaxy PRIME	Indigenous forest	18±11	676±534	105911, 105912
Wellington City	20/3/2019 - 14/3/2020	Optech Galaxy PRIME	Broadleaved Indigenous hardwoods, high-producing exotic grassland	8±7	115±99	105023, 105024
Otago - Balclutha	16/1/2020 - 18/1/2020	Optech Galaxy Prime 397	High-producing exotic grassland	3±3	46±43	104763, 104764

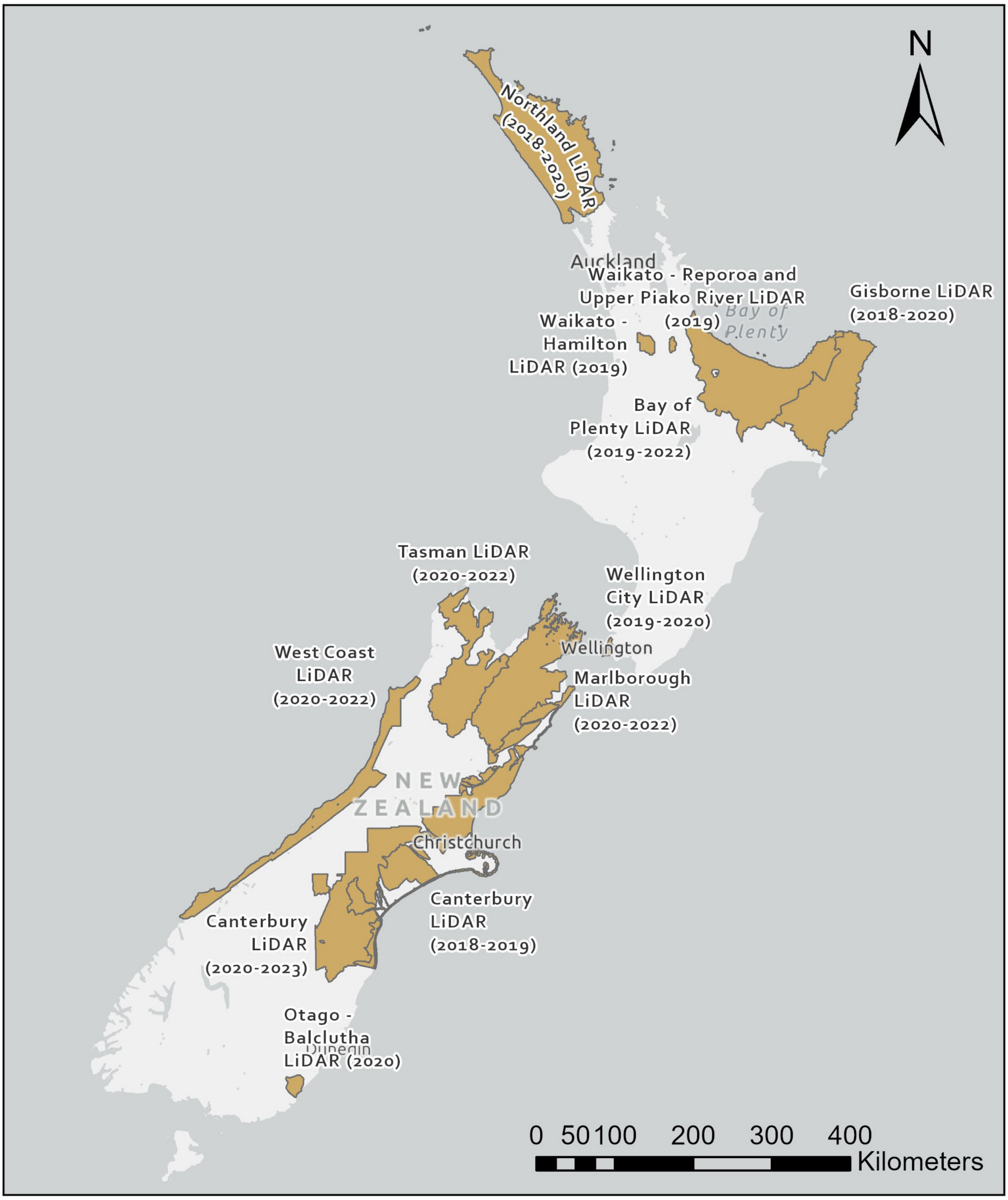


FIGURE 1: The study sites, along with the year of LiDAR acquisition, cover areas from the North Island to the South Island of New Zealand.

Reference data

The New Zealand airborne LiDAR datasets were used as a benchmark in this study to validate the three selected global CHMs. All regional LiDAR acquisitions adhere to the New Zealand National Aerial Lidar Base Specification, ensuring a minimum vertical accuracy of ±0.2 metres, a horizontal accuracy of ±1.0 metres, a pulse density of ≥ 4 pulses per square metre, and gridded at a resolution of 1-metre (LINZ 2022). These accuracy metrics establish the LiDAR data as a reliable benchmark for assessing the global CHMs across New Zealand.

The study used 1-metre spatial resolution airborne LiDAR datasets from the Land Information New Zealand (LINZ) via the Registry of Open Data on Amazon Web Services (AWS) STAC catalogue (LINZ n.d.). Due to large data sizes (40-60GB) for each region, a tile-based processing approach based on LINZ's tiling scheme was implemented, which was found to enhance efficiency.

Once the DEM (Digital Elevation Model) and DSM (Digital Surface Model) tiles were downloaded, canopy height was calculated by subtracting the corresponding DEM values from the DSM within each tile. This process involved calculating both the average and maximum canopy height values within 10 x 10 metre and 30 x 30 metre windows (corresponding to cell factors of 10 and 30, respectively) across each tile. This upscaling was performed to align the resolution with the 30-metre CHM from Potapov et al. (2021) and the 10-metre CHM from Lang et al. (2023). Both average and maximum aggregation methods were considered for evaluating the CHMs to understand the underlying height representation derived from the selected model's machine learning method and see which height estimate is most accurate in New Zealand's environment.

Subsequently, a continuous virtual raster file encompassing the entire study area was generated to provide a mosaic of all the processed CHM tiles. The entire workflow, including data retrieval, processing, and virtual raster creation, was automated using Python scripting with the GDAL library (GDAL/OGR contributors 2024). A virtual raster (VRT) is a GDAL technique in Python for merging multiple raster tiles into a single file (GDAL/OGR contributors 2024). It virtually merges the tiles but references them in a lightweight XML (Extensible Markup Language) file, allowing for efficient management of large datasets without duplicating the files or requiring extensive computational resources.

Global canopy height models

Potapov et al. (2021)

Potapov et al. (2021) developed a global CHM by integrating GEDI LiDAR data with Landsat optical data at a 30-metre spatial resolution. They processed multitemporal Landsat data spanning 1997 to 2019 to create consistent metrics, including surface reflectance and phenology indicators, while using only the 2019 subset of Landsat data to represent forest conditions for that year. A median-based regression tree model was calibrated using the GEDI's footprint-based relative height metric (RH95) to predict forest height based on Landsat metrics. The relative height (RH) represents the height above the ground corresponding to the n-th percentile of the LiDAR energy returned, spanning from the top of the canopy to the signal end (Li et al. 2023).

Lang et al. (2023)

The Lang et al. (2023) 10-metre model utilised Sentinel 2 optical images in 2020 as input and trained with global reference height derived from GEDI raw waveforms (collected between 2019 and 2020). They utilised Level 1B/L1B GEDI waveforms to derive the RH98 height metric as canopy-top height reference data. Their deep learning approach – convolutional neural network (CNN) specifically learned to extract patterns and features of the raw satellite images that are predictive of vegetation structure by training on GEDI data, which provides precise measurements of canopy height at a 10-metre resolution. This model also integrates geographic coordinates, enhancing its performance by allowing it to learn spatial priors based on location-specific vegetation characteristics.

Tolan et al. (2024)

Tolan et al. (2024) produced a 1-metre resolution global CHM with self-supervised training of a Vision Transformer on Maxar imagery from 2018–2020. A dense vision transformer decoder was trained using aerial LiDAR-derived canopy height maps from the National Ecological Observatory Network (NEON) as labels, linking satellite imagery features to canopy height predictions. Furthermore, the aerial LiDAR model's outputs were calibrated using a separate convolutional model trained on GEDI spaceborne LiDAR data. The GEDI model produced scaling factors to adjust the aerial LiDAR-based CHM predictions, thus improving global accuracy.

All three of the global CHMs by Potapov et al. (2021), Lang et al. (2023), and Tolan et al. (2024) used in this study are accessible through Awesome Google Earth Engine's (GEE) Community Catalogue (Roy et al. 2024). GEE is widely used in remote sensing analyses as it offers a user-friendly platform for retrieving and processing geospatial datasets. The analysis for Tolan et al. (2024) was done in GEE due to the substantial data size at a 1-metre resolution, which made local processing impractical. Only datasets from Potapov et al. (2021) and Lang et al. (2023) were retrieved for local processing. The retrieved raster data was reprojected to match the study area's New Zealand Transverse Mercator coordinate system and reference LiDAR datasets (NZTM, EPSG:2193). Bicubic interpolation was employed for resampling, and optimal boundaries were defined based on the study area to optimise file size and computational requirements.

Analysis and evaluation metrics

A random sample of 20,000 points was used to evaluate the five selected land covers across the CHMs and reference ALS datasets. This sample size was selected to ensure statistical robustness while maintaining computational efficiency for large-area national datasets. This random sampling approach serves as an initial assessment by reducing the risk of over- or under-representing any particular class, while laying the groundwork for future stratified sampling approaches that could target specific vegetation categories or regions to refine the evaluation further. The land cover data was obtained from the Land Cover Database (v5.0) to represent varied vegetation and land cover height across the study area: sand or gravel, high-producing exotic grassland, mixed exotic shrubland, indigenous forest and exotic forest (LRNZ 2020). Non-vegetated areas like sand or gravel surfaces were first assessed to understand the model's baseline performance. In the case of Tolan et al. (2024), canopy height values were extracted directly from GEE with the sample points for the analysis.

To comprehensively assess the three selected global CHM's performance against the reference ALS data, this study employed a combination of statistical metrics, including calculating the residuals:

1. Bias measures the trueness of the CHM as the average discrepancy between the CHM map and the actual values measured by the reference ALS data. The bias reflects the overall tendency of the CHM map to either overestimate or underestimate canopy height compared to the ALS data. It is calculated by averaging the difference between predicted and actual values across all data points. A positive bias indicates a systematic overestimation by the model (i.e., the model consistently predicts higher canopy heights than the ALS measurements). Conversely, a negative bias suggests underestimation. Ideally, the bias should be close to zero, signifying minimal systematic bias in the model's predictions.

$$Bias = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)$$

Where y_i are the true canopy height measurements from ALS data and $\hat{y}_i$ are the canopy height predictions from the global CHM.

2. Root Mean Squared Error (RMSE) compounds trueness with precision as a more comprehensive measure of the model's accuracy. It calculates the square root of the average squared difference between corresponding data points. Lower RMSE values indicate better agreement between the two datasets.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}$$

where y_i are the true canopy height measurements from ALS data and $\hat{y}_i$ are the canopy height predictions from the global CHM.

3. Finally, the coefficient of determination R^2 is a statistical measure that reflects the proportion of variance in the ALS data represented by the CHM's estimates. It assesses how well the model captures the linear relationship between the predicted value from CHM and actual canopy heights from the ALS data. R^2 ranges from 0 to 1. A value of 1 indicates a perfect fit, where the model reproduces all the variability observed in the ALS data perfectly. Conversely, a value of 0 suggests that the model has no explanatory power.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

where y_i are the true canopy height measurements from ALS data and $\hat{y}_i$ are the canopy height predictions from the global CHM and are the mean of the true canopy height measurements from ALS data.

Results

The analysis indicated that the accuracy of the CHMs varied depending on the type of vegetation and the model resolution. All models generally did not capture a full data variability with relatively low R^2 values. When viewed as maps, it was revealed that the Lang 10-metre model displayed higher height signals with greater blurriness and showed less distinct boundaries. This effect is likely due to sparse supervision during model training, which limited the model's ability to learn high-frequency canopy height variation between adjacent pixels, as noted by Lang et al. (2023). In contrast, the Potapov 30-metre map exhibited more pixelation, and the Tolan 1-metre map is almost on par with the ALS map (Figure 2). Some height discrepancies were found when comparing the maps because the retrieval dates were different.

Comparison across aggregation types

The average aggregation method provided more consistent results in the Potapov 30-metre model, with lower RMSE and more stable estimates than the maximum aggregation (Figures 3 and 4). This approach was better suited for representing overall canopy heights across all vegetation for this particular model. Conversely, for the Lang 10-metre model, the maximum aggregation method generally outperformed the average aggregation in terms of RMSE across all vegetation types, often resulting in lower values (Figures 3 and 4). Notably, there is a big difference in the bias values. For instance, in indigenous forests, the maximum aggregation achieved a bias of 2.82 m. In comparison, the average aggregation had a higher bias of 8.10 m, similar to the exotic forests where the maximum shows 3.18 m and the average shows 10.85 m.

Comparison across models and vegetation types

The performance of global CHMs varied across models, resolutions, and vegetation types, with relatively low R^2 values ranging from 0.15 to 0.38 and RMSE values between 1.21 m and 13.70 m (Figures 3 and 4). All models fall short of the 1:1 line. The exotic forest class always yields significantly higher R^2, although models tend to exhibit a marked non-linear relationship to the reference ALS data (Figures 3 & 4). In non-vegetated areas, the Potapov 30-metre average model and Tolan 1-metre model were the most effective, with RMSE values below 2 m and minimal bias despite a relatively low R^2 value of <0.35. Lang's 10-metre model consistently displayed the largest positive bias, in particular systematically suggesting vegetation in bare areas. Both the Potapov and Tolan models performed better in shorter vegetation than the Lang 10-metre model. However, the Tolan model consistently displayed negative biases, whereas the Potapov model showed positive biases. The Potapov 30-metre model's approach of setting observations below a 3-metre threshold to zero is also reflected in the plots (Figure 3).

For taller vegetation, such as indigenous and exotic forests, the performance of these two models generally declined, with higher RMSE and bias. The Lang 10-metre

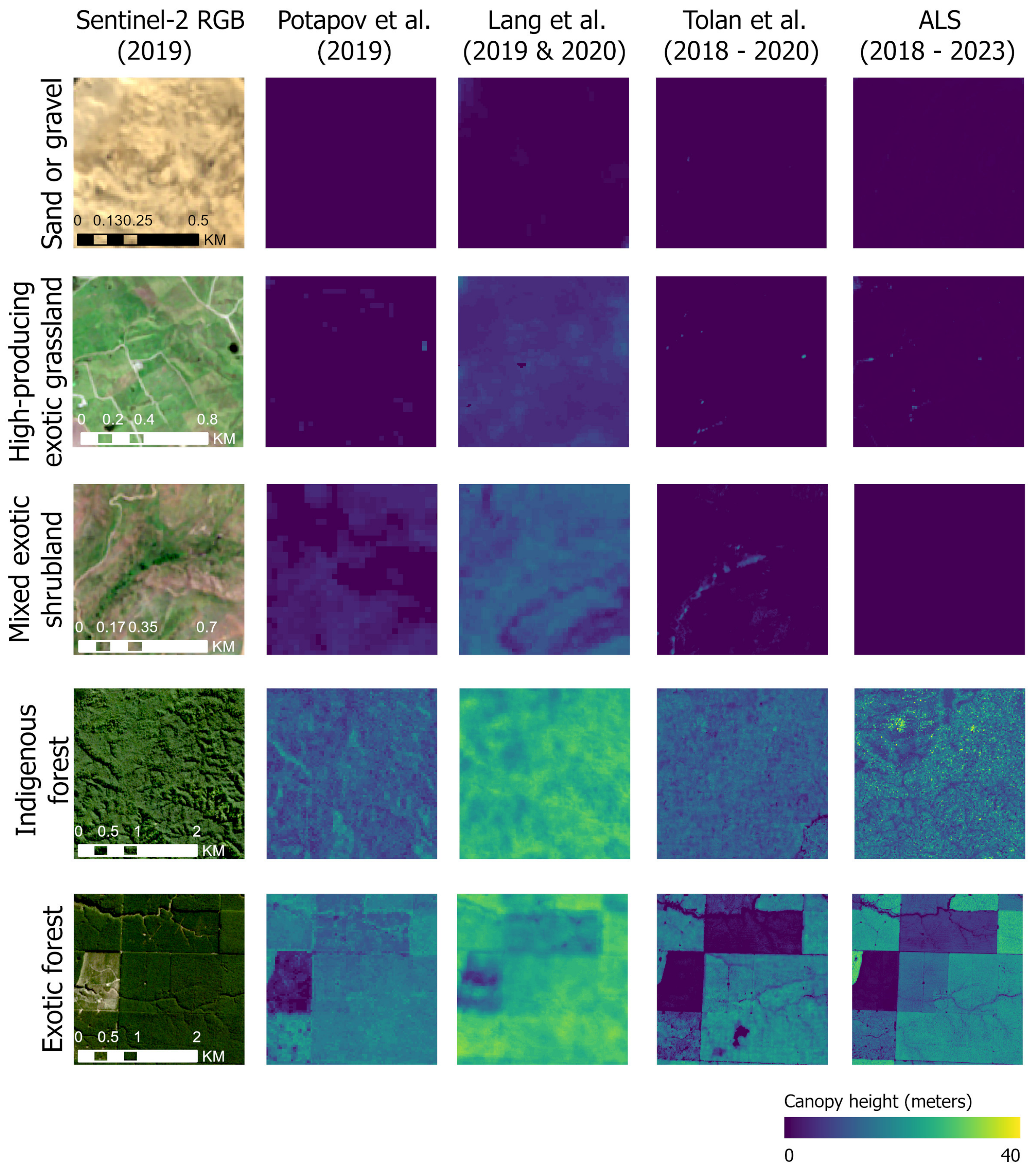


FIGURE 2: Comparison of Potapov 30-metre CHM, Lang 10-metre CHM, Tolan 1-metre CHM and the ALS-derived canopy height map (fifth column), with Sentinel-2 RGB imagery as reference. Each land cover is represented by sand or gravel, high-producing exotic grassland, mixed exotic shrubland, indigenous forest and exotic forest. The year shows the retrieval time of each model.

maximum aggregation model generally outperformed the Potapov 30-metre and Tolan 1-metre models for these vegetation types in terms of R^2 and bias, although the RMSE is slightly higher in exotic forests. When the aggregation approach is comparable, the Lang model also displayed greater heights in exotic forests than other models, saturating around 35-40 metres. However, the indigenous forest outperformed the exotic forest in terms of bias and RMSE for both the Lang 10-metre and Tolan 1-metre models.

Residual analysis was limited to exotic forests, which exhibited the widest canopy height range and the most linear relationship between ALS and CHM values, allowing for clearer and more statistically meaningful residual patterns than other land cover types. The residuals reveal a common trend among models: both

FIGURE 3: Density scatter plot comparing canopy height estimates across various models (30-metre, 10-metre and 1-metre) with ALS-derived canopy height measurements over non-vegetated areas (sand or gravel) and short vegetation (high-producing exotic grassland and mixed exotic shrubs), along with their statistical metric results. Point colours represent data density, transitioning from blue (high density) to green and yellow (lower density). The black dotted line represents the 1:1 relationship in the scatter plots, while the red solid line shows the linear regression fit.

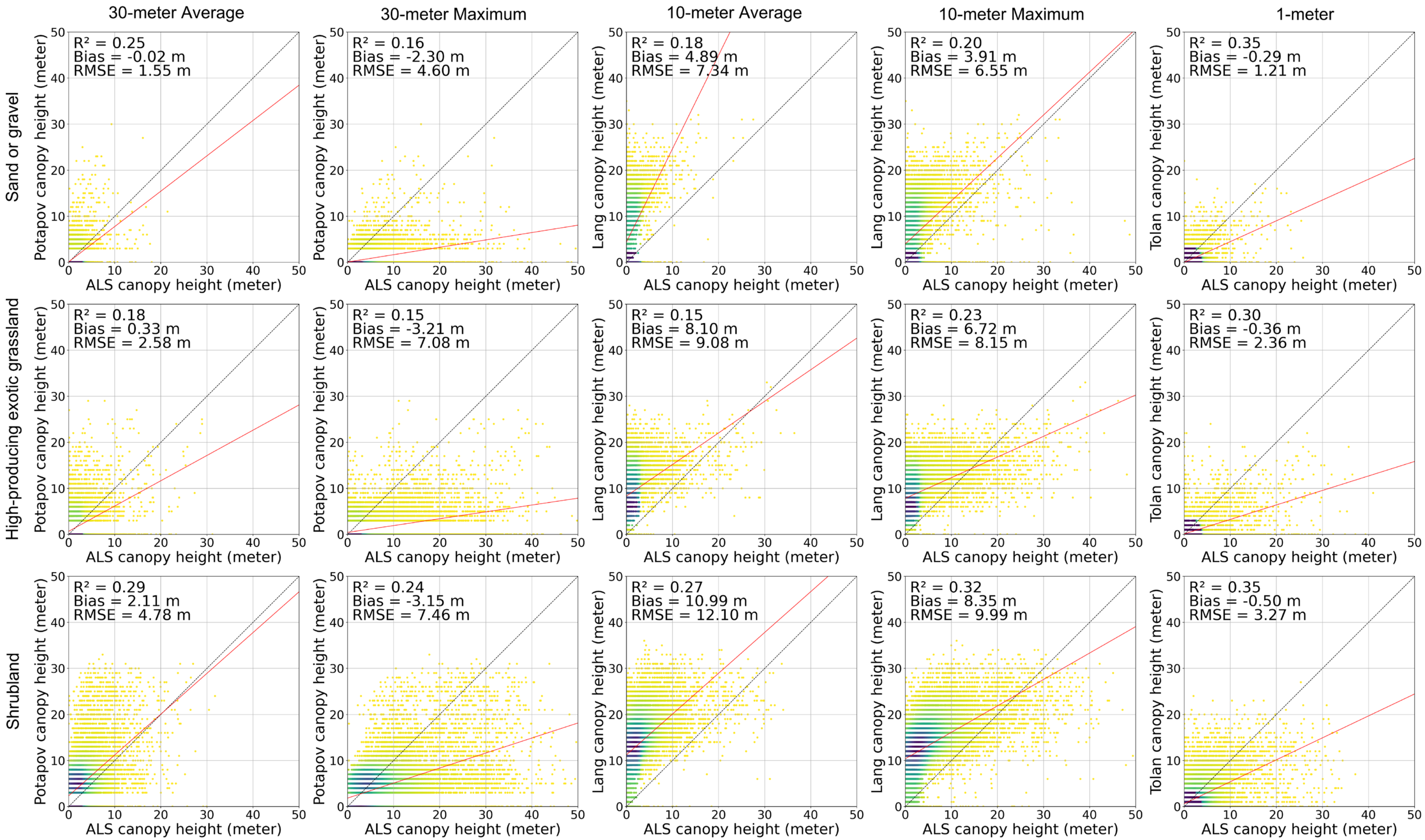
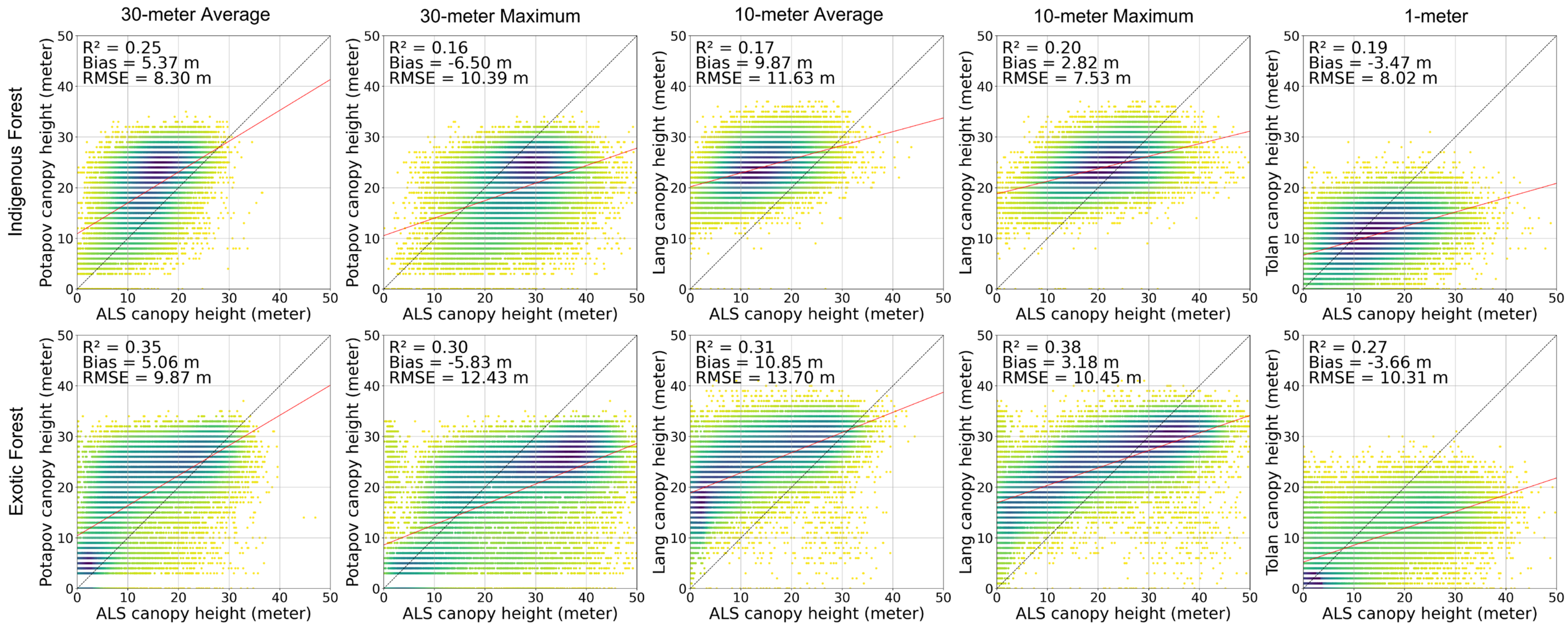


FIGURE 4: Density scatter plot comparing canopy height estimates across various models (30-metre, 10-metre and 1-metre) with ALS-derived canopy height measurements over tall vegetation (indigenous forest and exotic forest), along with their statistical metric results. Point colours represent data density, transitioning from blue (high density) to green and yellow (lower density). The black dotted line represents the 1:1 relationship in the scatter plots, while the red solid line shows the linear regression fit.



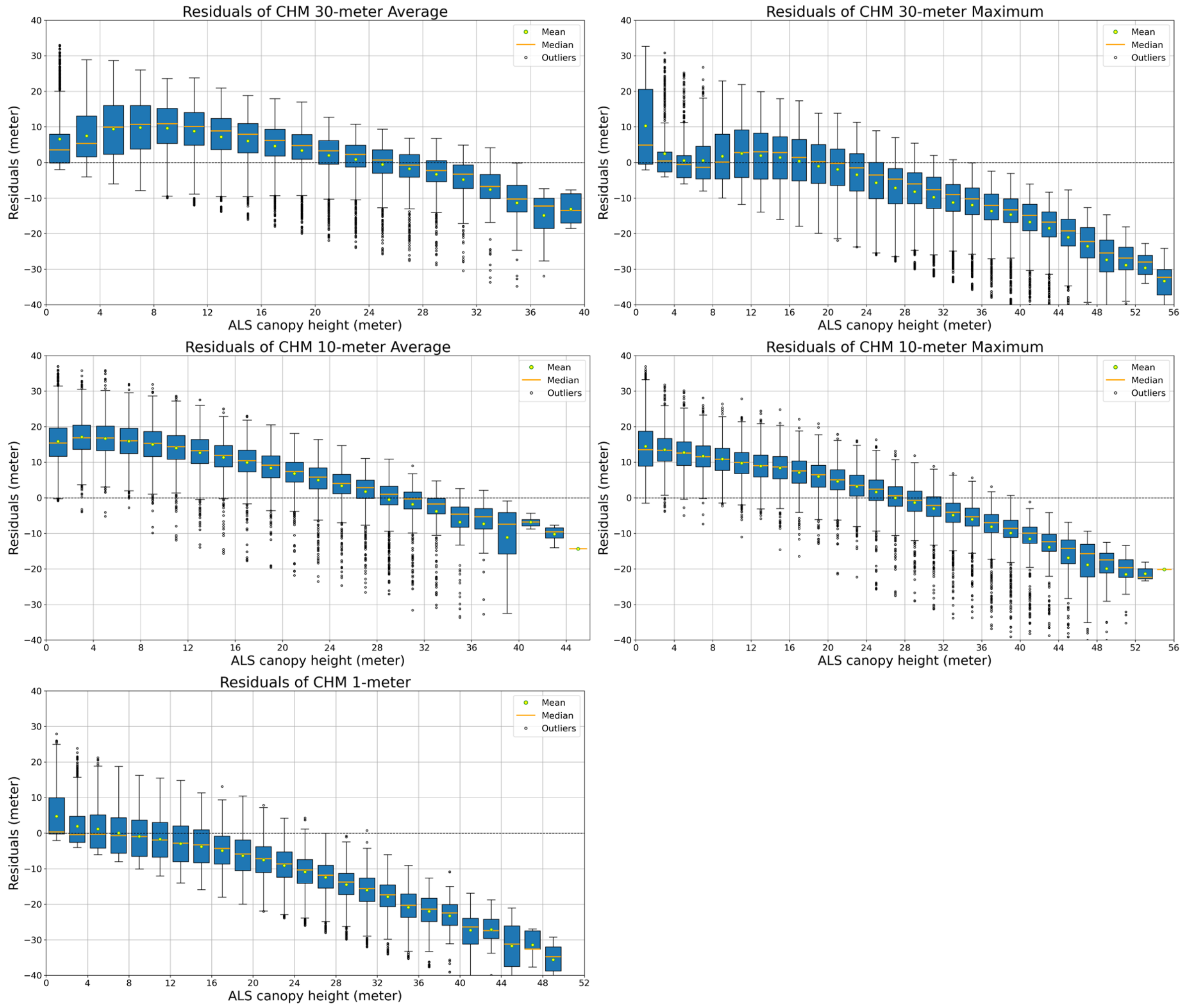


FIGURE 5: Residuals of canopy height estimates for exotic forests across various models at 2-metre intervals.

the Potapov 30-metre average and maximum models tend to overestimate canopy heights in the lower to mid ranges (until 20-24 m) and underestimate them as heights increase beyond that (Figure 5). Similarly, the Lang 10-metre models show overestimation at shorter canopy heights, transitioning to underestimation for taller canopies around 24-32 m. In contrast, the Tolan 1-metre model performed differently, minimising overestimation at lower heights (<8 m) but increasing underestimation for medium to tall canopies.

Discussion

This study evaluated the accuracy of three global canopy height models (CHMs)—Potapov et al. (30-metre), Lang et al. (10-metre), and Tolan et al. (1-metre)—across varied New Zealand vegetation types, using ALS data as a benchmark. The findings reveal distinct performance patterns for each model and provide insight into their suitability for forest carbon accounting and management applications.

Potapov et al. (2021) model

Potapov et al. (2021) developed a 30-metre global CHM using regression tree ensembles trained on Landsat spectral data and GEDI RH95 height metrics. Our results revealed that the average aggregation approach produced more stable estimates and lower RMSE values than maximum aggregation. The results aligned with the model's reliance on a regression tree ensemble that outputs median values. However, this design limits the model's ability to accurately represent the shortest and tallest forest structures, a limitation further exacerbated by the medium resolution of Landsat data (Potapov et al. 2021; Hansen et al. 2016).

In examining non-vegetated areas, the model exhibited negligible bias (-0.02 m) and low RMSE values (below 2 m), suggesting that the model is well-calibrated to detect and correctly classify near-zero canopy heights. Their integration of phenology-based differentiation and temporal metrics allows accurate classification, with non-vegetated areas most likely identified through static reflectance and low NDVI (Normalized Difference Vegetation Index); while dynamic seasonal changes and peak productivity characterise vegetated areas (Potapov et al. 2021). The model's baseline performance implies that deviations observed in vegetated regions may indicate specific model limitations related to canopy complexity rather than systemic calibration issues.

Inherent challenges associated with using GEDI RH95 height metrics exist in sparse and multilayered vegetation areas. In our study, the Potapov 30-metre model exhibited small positive biases (<3 m) in the average aggregation method for high-producing exotic grasslands and mixed exotic shrublands. This result is consistent with the model's reliance on GEDI's RH95 metric, which can overestimate canopy heights in sparsely vegetated areas due to the tendency to include non-canopy returns (Zhu et al. 2022; Yu et al. 2024). The higher pulse density (≥4 points/m^2) and fine resolution (1 metre) of New Zealand ALS data likely exaggerated this overestimation by capturing detailed canopy structures and delineating non-canopy elements that RH95 may mistakenly include. This positive bias also aligns with Potapov's visual analysis findings of overestimation in grasslands, particularly in New Zealand and Lesotho, though their statistical evaluation suggested an overall underestimation due to the disproportional sampling of tall tropical forests used in their ALS validation and their regression tree approach.

For taller vegetation, such as indigenous and exotic forests, the Potapov model struggled to accurately represent canopy heights, with higher RMSE and positive bias than shorter vegetation types shown in our study. The maximum aggregation method amplified these discrepancies by showing underestimation and higher RMSE, revealing the model's limitations in capturing peak canopy heights. Dorado-Roda et al. (2021) observed underestimation biases when comparing RH95-derived heights to ALS data for Mediterranean oak and pine forests. Their approach differed from ours, as they normalised ALS data to 1 point/m^2 and applied a 12.5-metre buffer around GEDI footprints. These differences underscore how ALS data resolution and processing methods can shape the interpretation of RH95's performance, with finer ALS resolutions exposing overestimation in tall vegetation and normalised ALS datasets highlighting underestimation under similar vegetation height.

Despite Potapov's relative advantage in shorter vegetation compared to the Lang 10-metre model, it was still outperformed by the Tolan 1-metre model. Moreover, its limitations in capturing taller canopy heights and vegetation complexity underscore the need for nuanced application and further refinement. The Potapov 30-metre model's overall strengths lie in its stability and utility for large-scale monitoring, particularly in non-vegetated and shorter vegetation types. By leveraging its capability for temporal monitoring, the model holds promise for advancing consistent global forest structural analysis and carbon stock monitoring. Currently, the model is only suited for a generalised assessment of the New Zealand ecosystem and is not recommended for a detailed level carbon stock assessment.

Lang et al. (2023) model

The Lang et al. (2023) 10-metre canopy height model employs a window-based deep learning approach that leverages Sentinel-2 spectral and textural data alongside GEDI RH98 metrics. By training CNN with sparse supervision from GEDI LIDAR-derived canopy top height data, the model effectively incorporates global spectral-textural information to predict canopy heights. However, like other global models, Lang-CHM's performance is influenced by the structural diversity of vegetation and the limitations inherent in its design and training data.

To evaluate the Lang 10-metre model's baseline performance, the model exhibited higher bias and RMSE for bare areas. Their window-based approach could inadvertently incorporate signals from adjacent vegetation into the estimation in areas that suggest minimal or zero vegetation. For example, even small patches of low-lying vegetation or shrubs near bare surfaces might influence the overall signal within the window, leading to inflated height estimates. These challenges might be compounded by the reliance on GEDI RH98 metrics. The wide pulse width of GEDI's laser signals (approximately 4.5 metres) can inflate RH values, misrepresenting surface height (Dubayah et al. 2020). While Lang et al. (2023) mitigated this limitation by using RH98 instead of RH100, residual inaccuracies persist, introducing baseline height signals even in non-vegetated regions (Li et al. 2023). As seen from our results, these inflated values likely propagate through short and tall vegetation estimates, creating systematic biases across short and tall vegetation.

Interestingly, the Lang model performed relatively better for New Zealand's indigenous forests than exotic plantations despite the indigenous forests' unique vegetation structure. This finding contrasts with expectations as managed exotic forests, especially the dominant species radiata pine (*P. radiata)*, are widely distributed across other continents (North America) and are likely well-represented in the training data used for global models (Lang et al. 2022). This observation aligns with the broader understanding of global models' reliance on generalised training data and GEDI's characteristics. As Schwartz et al. (2024) noted, GEDI-based models tend to perform better in regions with dense, diverse canopy layers because RH98 metrics effectively capture upper canopy features. In indigenous forests, the diverse vertical layering might inadvertently align with the Lang model's emphasis on spectral-textural inputs, explaining its improved performance. For exotic forests in the Lang model, the relatively higher bias and RMSE might have resulted from spectral similarities between canopy and understorey elements in plantation settings, which the model cannot distinguish due to its reliance on training datasets generalised to different biomes and forest types. This challenge reflects findings in similar studies, such as Fayad et al. (2024), where the uniformity of plantation forests posed difficulties for global models trained on diverse vegetation types.

This study also confirms that the Lang 10-metre model's performance improves significantly when maximum aggregation methods are applied, particularly for taller canopies. Their use of GEDI RH98—a height metric particularly suited for peak height approximation—enhances alignment with maximum height measurements (Besic et al. 2024). This height metric contrasts with GEDI RH95-based methods, such as those by Potapov et al. (2021), which tend to yield lower canopy height estimates and introduce slight underestimation biases for tall forests (Kacic et al. 2023). For example, in exotic forests, the bias from the Lang 10-metre model was notably lower with maximum aggregation (3.18 m) compared to average aggregation (10.85 m). This observation is consistent with Lang's approach as they had already explicitly addressed the saturation effect on tall canopies, thus reflecting the expected results of our findings.

However, despite these improvements, Lang et al. (2023) revealed error and low bias in non-vegetated and low-height areas, such as deserts and temperate grasslands. They reported a performance decline in denser or taller vegetation types, such as mangroves, tundra, and tropical coniferous forests, where it exhibits a consistent positive bias, with overestimations averaging approximately 2.5 metres. Our findings support these observations, with the Lang 10-metre model overestimating both short and tall vegetation but better estimates in tall vegetation. Similar overestimations have been observed in other studies—Moudrý et al. (2024), Torresani et al. (2023), Tsao et al. (2023), and Alvites et al. (2024). Moudrý et al. (2024) found that the transition between forest and non-forest remains unclear. Lang et al. (2023) acknowledged that the trade-off in their model's design for improving the performance of tall canopies caused a slight overestimation of low canopy heights. Our findings suggest that this trade-off is quite pronounced and may have also included the systematic bias observed in the GEDI system, as previously discussed.

Moudrý et al. (2024) evaluated Lang, Potapov and Tolan CHMs in Mount Richmond Forest stratified by height bins. For the 20–30 m class, they found a Lang bias of +1 m (RMSE 5 m), Potapov bias of –5 m (RMSE 9 m) and Tolan bias of –11 m (RMSE 12 m). In our national-scale indigenous forest analysis, the Lang 10 m (maximum) model exhibited a bias of +2.82 m (RMSE 7.53 m), the Potapov 30 m (average) model +5.37 m (RMSE 8.30 m), and the Tolan 1 m model –3.47 m (RMSE 8.02 m). Although absolute biases are somewhat larger in our national-scale analysis, the direction of errors is consistent: Lang and Potapov tend to overestimate mid-height canopies, while Tolan underestimates. Likewise, all three models increasingly misestimate heights beyond ~30 m. This agreement between site-specific and nationwide assessments reinforces the need for caution when applying these CHMs to forest stands within New Zealand.

Overall, the Lang et al. (2023) model demonstrates considerable strengths in estimating taller canopies due to its reliance on GEDI RH98 metrics and advanced deep-learning techniques. However, its systematic overestimation of low canopy heights, compounded by inherent limitations in the GEDI system and spectral-textural ambiguities, suggests that the model could be more suited for applications focused on taller canopies. Localised adjustments or alternative models may be required for regions with significant short vegetation cover to address the inherent biases observed.

Tolan et al. (2024) model

Tolan et al. (2024) developed a global canopy height map at 1-metre resolution using high-resolution Maxar satellite imagery and airborne LiDAR-derived canopy height models. They employed a self-supervised vision transformer and a deep-learning approach to estimate canopy height from RGB satellite imagery. These techniques enable the model to provide unattainable detailed spatial insights with coarser-resolution models like those of Potapov et al. (2021) and Lang et al. (2023).

The Tolan model consistently outperformed the other CHMs in non-vegetated and shorter vegetation areas, achieving lower RMSE and bias values in our study. Their results indicate a strong baseline performance comparable to that of Potapov et al. (2021). Moudrý et al. (2024) noted its high sensitivity in capturing the transitions between forest and grassland, emphasising its suitability for applications that demand precision. This capability aligns with findings from Wagner et al. (2024), demonstrating the model's robustness in submetre-scale canopy height mapping for California. The model's ability to minimise overestimation in shorter canopies compared to other CHM underscores its suitability for accurately estimating biomass in regions where precision in capturing minor height variations is critical.

However, challenges remain for the Tolan 1-metre model in estimating taller canopies despite displaying a better performance than Potapov's model. Our analysis and observations by Fogel et al. (2024) found that the model tended to underestimate canopy heights in taller forests despite its high resolution. Moreover, Bermudez et al. (2024) found underestimation across the entire data range for Tolan's model in their study, and Wagner et al. (2024) reported underestimation for tall trees >45 m. Such underestimation suggests that further refinement in training data to include wider biomes or adjustments in the model's loss functions could improve its accuracy for tall canopy environments (Wagner et al. 2024; Lang et al. 2023).

Despite these limitations, the Tolan et al. (2024) model represents a global high-resolution canopy height mapping benchmark. Its fine spatial detail and accuracy for shorter canopies position it as a transformative tool for vegetation monitoring, with significant implications for national carbon accounting frameworks and biodiversity assessments. In New Zealand, where substantial expanses of native and exotic grasslands and shrublands dominate the landscape, the model provides a promising pathway to advance both policy and practice in managing and conserving these critical ecosystems.

Limitations and opportunities

This study did not investigate the impact of GEDI geolocation errors, slope, and terrain variations on global CHM measurements. Previous studies (Quirós et al. 2021; Moudrý et al. 2022; Lang et al. 2023) have highlighted these factors as significant sources of error in CHMs. For instance, GEDI geolocation uncertainties could result in canopy height estimates being off by approximately 2 metres (Li et al. 2023). Additionally, GEDI's larger footprint (25 metres) is prone to bias on steep slopes and underperforms in fragmented landscapes like mountainous regions, reducing data accuracy and spatial representation (Mandl et al. 2023). Developing geolocation optimisation processes with high-precision ALS data could reduce these systematic errors (Tang et al. 2023). While Lang et al. (2023) considered geolocation uncertainty unlikely to hinder their model's utility, its implications for local-scale applications like New Zealand's forests remain underexplored.

A significant limitation in this study stems from temporal discrepancies between ALS acquisition and global CHM retrieval times, with differences of up to ±2-3 years across the 11 regions analysed. Such temporal mismatches can lead to inconsistencies in canopy height comparisons, particularly in fast-growing or structurally dynamic ecosystems like exotic forests. Incorporating species-specific growth models could help adjust ALS canopy heights to match CHM acquisition dates. For instance, height-age curves tailored to *P. radiata* or other dominant species could account for expected growth over the temporal gap, enabling more accurate comparisons (van der Colff & Kimberley 2013). Integrating species-specific growth models with regional exotic forest descriptions could support more targeted and locally calibrated temporal adjustment methods. Another fundamental limitation is the restricted spatial distribution of training data used in the Lang et al. (2023) model. As global models often rely on training datasets that emphasise a few specific biomes, they may not fully capture unique structural attributes of local vegetation, such as those found in New Zealand. Spatial misalignments between the global CHMs and ALS-derived canopy height maps could also contribute to sampling uncertainty. Although visual inspections indicated overall alignment, small discrepancies, particularly for the GEE-processed Tolan model, may affect point-based canopy height comparisons.

Future efforts could incorporate terrain and environmental factors into the analysis to address these limitations and assess these models more comprehensively. For instance, incorporating slope, elevation, and other topographical variables could enhance the study to account for terrain-induced variations in canopy height accuracy assessment (Besic et al. 2024; Li et al. 2024b). The random sampling approach in this study did not adequately capture the height variation in indigenous forests, where tree growth tends to be more stable than exotic plantations. A stratified sampling design targeting specific growth stages and vegetation characteristics could improve model validation, providing deeper insights into its application to areas where ALS data is unavailable in New Zealand.

Our analysis revealed non-linear relationships between CHMs and ALS data for all land cover types, suggesting potential improvements through mathematical transformations. For exotic forests, where the relationship shows a curved trend, applying non-linear transformations to Lang-CHM values could better capture height variations, reduce bias, and improve model fit. Testing various transformations could help refine the regression model, particularly for vegetation types with significant height variability.

New Zealand's extensive, high-resolution ALS datasets provide an excellent opportunity to enhance global CHMs. Among all global CHMs, Lang et al. (2023) stand out by providing open access to their code and model, offering valuable opportunities to improve its performance by reducing bias and capturing the unique structural characteristics of New Zealand's forests. This approach could facilitate the development of a national-level canopy height map at a 10-metre resolution, which could subsequently be translated into a carbon stock map similar to those produced by Lang et al. (2021).

While our results indicate that all global CHMs show notable biases and relatively low R^2 values, reflecting limited ability to fully capture canopy height variability, the models still exhibit distinct strengths depending on vegetation structure and data aggregation methods. Therefore, these CHMs could be useful as complementary tools for New Zealand's forest monitoring and carbon accounting efforts, especially where large-scale updates are needed. However, users should be aware of the potential for substantial biases, and consider these uncertainties when applying the models in policy or compliance contexts. Importantly, our findings reinforce the need for locally developed and calibrated models that can better reflect New Zealand's various vegetation structures and provide the accuracy required for regional-scale applications such as ETS reporting.

An additional avenue for exploration involves harmonising spatial resolutions for a more equitable comparison between CHMs. Aggregating Lang's 10-metre and Tolan's 1-metre CHMs to the coarser 30-metre resolution of Potapov's model could offer valuable insights. This harmonised comparison may reveal statistical benefits that highlight the relative strengths of higher-resolution CHMs when evaluated on an equivalent spatial scale. Such an approach could help disentangle the effects of resolution from the underlying model design, providing a clearer picture of their relative performance across varied vegetation types.

In summary, the analysis highlighted distinct strengths across the three models. The Potapov 30-metre model provided reliable general canopy height estimates at broad scales, while the Lang 10-metre model offered improved performance in taller canopies but exhibited challenges in low-vegetation areas. The Tolan 1-metre model aligned well with ALS data for shorter vegetation, presenting potential for fine-scale applications. Recognising these complementary strengths is essential when selecting a model for a given spatial scale or management objective.

Conclusions

The analysis revealed that the Potapov 30-metre model is well-suited for a broad-scale general canopy height assessment due to its reliance on a median-based regression tree ensemble and coarser resolution that limits its applicability for detailed assessments. The Lang 10-metre model, while excelling in estimating taller canopies through its advanced deep learning approach, struggled with non-vegetated areas and short vegetation, reflecting the trade-offs inherent in its design. The Tolan 1-metre model demonstrated strong alignment with ALS data for shorter vegetation, highlighting its potential for fine-scale vegetation mapping, though its underestimation of taller canopies warrants further caution.

These findings emphasise the strengths and limitations of each CHM, illustrating the importance of tailored approaches for vegetation-specific applications. At present, global CHMs are best suited as complementary tools for carbon accounting rather than replacements for existing methodologies. While they provide valuable large-scale insights, their current limitations require careful consideration when applied to local-scale assessments. By addressing the limitations and capitalising on opportunities for refinement, insights on each global CHM can improve in New Zealand's context and play a critical role in advancing forest monitoring, enhancing carbon accounting, and supporting sustainable forest management within the broader context of climate change mitigation efforts.

List of abbreviations

ALS – Airborne Laser Scanning
AWS – Amazon Web Services
CHM – Canopy Height Model
CNN – Convolutional Neural Network
DEM – Digital Elevation Model
DSM – Digital Surface Model
EPSG – European Petroleum Survey Group
ETS – Emissions Trading Scheme
GEE – Google Earth Engine
GEDI – Global Ecosystem Dynamics Investigation
IPCC – Intergovernmental Panel on Climate Change
LiDAR – Light Detection and Ranging
LINZ – Land Information New Zealand
L1B – Level 1B (GEDI Data Processing Level)
NDVI – Normalized Difference Vegetation Index
NEON – National Ecological Observatory Network
NZTM – New Zealand Transverse Mercator
RH – Relative Height
RMSE – Root Mean Square Error
R^2 – Coefficient of Determination
SAR – Synthetic Aperture Radar
STAC – SpatioTemporal Asset Catalog
TIFF – Tagged Image File Format
VRT – Virtual Raster
XML – Extensible Markup Language

Authors' contributions

Sue Kee participated in collecting and processing the ALS data and global CHMs, producing maps, conducting statistical analysis, and writing the manuscript. Todd and Pascal contributed to the study design, assisted in interpreting the data, and reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Acknowledgements

The authors acknowledge the National School of Surveying, University of Otago, for supporting this research through a publishing bursary.

References

Allen, R.B., Bellingham, P.J., Holdaway, R.J., & Wiser, S.K. (2013). New Zealand's indigenous forests and shrublands. *Ecosystem services in New Zealand-condition and trends.* Lincoln, New Zealand: Manaaki Whenua Press, pp. 34-48.

Alvites, C., O'Sullivan, H., Francini, S., Marchetti, M., Santopuoli, G., Chirici, G., & Bazzato, E. (2024). High-resolution canopy height mapping: Integrating NASA's Global Ecosystem Dynamics Investigation (GEDI) with multi-source remote sensing data. *Remote Sensing, 16*(7), 1281. https://doi.org/10.3390/rs16071281

Besic, N., Picard, N., Vega, C., Hertzog, L., Renaud, J.P., Fogel, F., Schwartz, M., Pellissier-Tanon, A., Destouet, G., Mortier, F., Planells-Rodriguez, M., & Ciais, P. (2024). Remote sensing-based high-resolution mapping of the forest canopy height: some models are useful, but might they be even more if combined? *Geoscientific Model Development Discussions*, *2024*, 1-26. https://doi.org/10.5194/gmd-2024-95

Bermudez, J., Rogers, C., Sothe, C., Cyr, D., & Gonsamo, A. (2024). A deep learning approach to estimate canopy height and uncertainty by integrating seasonal optical, SAR and limited GEDI LiDAR data over northern forests. *arXiv preprint arXiv:2410.18108*

Brown, S. (2002). Measuring carbon in forests: current status and future challenges. *Environmental pollution, 116*(3), 363-372. https://doi.org/10.1016/S0269-7491(01)00212-3

Coomes, D.A., Šafka, D., Shepherd, J., Dalponte, M., & Holdaway, R. (2018). Airborne laser scanning of natural forests in New Zealand reveals the influences of wind on forest carbon. *Forest Ecosystems*, *5*, 1-14. https://doi.org/10.1186/s40663-017-0119-6

Dorado-Roda, I., Pascual, A., Godinho, S., Silva, C.A., Botequim, B., Rodríguez-Gonzálvez, P., González-

Ferreiro, E., & Guerra-Hernández, J. (2021). Assessing the accuracy of GEDI data for canopy height and aboveground biomass estimates in Mediterranean forests. *Remote Sensing*, *13*(12): 2279. https://doi.org/10.3390/rs13122279

Dubayah, R., Blair, J.B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., & Silva, C. (2020). The global ecosystem dynamics investigation: High-resolution laser ranging of the Earth's forests and topography. *Science of Remote Sensing*, *1*: 100002. https://doi.org/10.1016/j.srs.2020.100002

Fayad, I., Ciais, P., Schwartz, M., Wigneron, J.P., Baghdadi, N., de Truchis, A., d'Aspremont, A., Frappart, F., Saatchi, S., Sean, E., Pellissier-Tanon, A., & Bazzi, H. (2024). Hy-TeC: a hybrid vision transformer model for high-resolution and large-scale mapping of canopy height. *Remote Sensing of Environment*, *302*, 113945. https://doi.org/10.1016/j.rse.2023.113945

Fogel, F., Perron, Y., Besic, N., Saint-André, L., Pellissier-Tanon, A., Schwartz, M., & Ciais, P. (2024). Open-canopy: A country-scale benchmark for canopy height estimation at very high resolution. *arXiv preprint arXiv:2407.09392*

GDAL/OGR contributors. (2024). *GDAL/OGR geospatial data abstraction software library*. Open Source Geospatial Foundation. Retrieved February 14, 2024, from https://gdal.org

Hansen, M.C., Potapov, P.V., Goetz, S.J., Turubanova, S., Tyukavina, A., Krylov, A., Kommareddy, A., & Egorov, A. (2016). Mapping tree height distributions in Sub-Saharan Africa using Landsat 7 and 8 data. *Remote Sensing of Environment*, *185*, 221-232. https://doi.org/10.1016/j.rse.2016.02.023

Hunt, C.A. (2009). *Carbon sinks and climate change: forests in the fight against global warming*. Edward Elgar Publishing. https://doi.org/10.4337/9781849802109

IPCC. (2021). *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (V. Masson-Delmotte, P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B. R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, & B. Zhou, Eds.). Cambridge University Press.

Kacic, P., Thonfeld, F., Gessner, U., & Kuenzer, C. (2023). Forest structure characterization in Germany: novel products and analysis based on GEDI, sentinel-1 and sentinel-2 data. *Remote Sensing*, *15*(8): 1969. https://doi.org/10.3390/rs15081969

Köhler, P., & Huth, A. (2010). Towards ground-truthing of spaceborne estimates of above-ground life biomass and leaf area index in tropical rain forests. *Biogeosciences*, *7*(8), 2531-2543. https://doi.org/10.5194/bg-7-2531-2010

Land Resource Information System (LRIS). (n.d.). *LCDB v5.0 - Land Cover Database version 5.0, Mainland New Zealand*. LRIS Portal. Retrieved February 2, 2024, from https://lris.scinfo.org.nz/layer/104400-lcdb-v50-land-cover-database-version-50-mainland-new-zealand/

Lang, N., Schindler, K., & Wegner, J.D. (2021). High carbon stock mapping at large scale with optical satellite imagery and spaceborne LIDAR. *arXiv preprint arXiv:2107.07431*

Lang, N., Kalischek, N., Armston, J., Schindler, K., Dubayah, R., & Wegner, J.D. (2022). Global canopy height regression and uncertainty estimation from GEDI LIDAR waveforms with deep ensembles. *Remote Sensing of Environment*, *268*: 112760. https://doi.org/10.1016/j.rse.2021.112760

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. https://doi.org/10.1038/s41559-023-02206-6

Lee, Y., Beck, M., & Philips, R. (2023). *About the elevation programme.* Toitū Te Whenua Land Information New Zealand. Retrieved March 4, 2024, from https://storymaps.arcgis.com/stories/59af2967e2c64664aebce8ca154c5c49

Li, W., Niu, Z., Shang, R., Qin, Y., Wang, L., & Chen, H. (2020). High-resolution mapping of forest canopy height using machine learning by coupling ICESat-2 LiDAR with Sentinel-1, Sentinel-2 and Landsat-8 data. *International Journal of Applied Earth Observation and Geoinformation*, *92*: 102163. https://doi.org/10.1016/j.jag.2020.102163

Li, X., Wessels, K., Armston, J., Hancock, S., Mathieu, R., Main, R., and Scholes, R. (2023). First validation of GEDI canopy heights in African savannas. *Remote Sensing of Environment*, *285*: 113402. https://doi.org/10.1016/j.rse.2022.113402

Li, S., Zhu, Z., Deng, W., Zhu, Q., Xu, Z., Peng, B., Guo, F., Zhang, Y., & Yang, Z. (2024). Estimation of aboveground biomass of different vegetation types in mangrove forests based on UAV remote sensing. *Sustainable Horizons*, *11*: 100100. https://doi.org/10.1016/j.horiz.2024.100100

Li, Y., Lu, D., Lu, Y., & Li, G. (2024b). Examining the impact of topography and vegetation on existing forest canopy height products from ICESat-2 ATLAS/GEDI Data. *Remote Sensing, 16*(19): 3650. https://doi.org/10.3390/rs16193650

Lim, K., Treitz, P., Wulder, M., St-Onge, B., & Flood, M. (2003). LiDAR remote sensing of forest structure. *Progress in Physical Geography*, *27*(1), 88-106. https://doi.org/10.1191/0309133303pp360ra

Lorenz, K., & Lal, R. (2010). The importance of carbon sequestration in forest ecosystems. In *Carbon sequestration in forest ecosystems* (pp. 45-79). Springer, Dordrecht. https://doi.org/10.1007/978-90-481-3266-9_6

Mandl, L., Stritih, A., Seidl, R., Ginzler, C., & Senf, C. (2023). Spaceborne LiDAR for characterizing forest structure across scales in the European Alps. *Remote Sensing in Ecology and Conservation*, *9*(5), 599-614. https://doi.org/10.1002/rse2.330

Ministry for the Environment. (2024). *About the New Zealand emissions trading scheme.* Retrieved July 2, 2024, from https://environment.govt.nz/what-government-is-doing/areas-of-work/climate-change/ets/about-nz-ets/

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*(10): e70026. https://doi.org/10.1002/ecs2.70026

Moudrý, V., Gdulová, K., Gábor, L., Šárovcová, E., Barták, V., Leroy, F., Špatenková, O. Rocchini, D., & Prošek, J. (2022). Effects of Environmental Conditions on ICESat-2 Terrain and Canopy Heights Retrievals in Central European Mountains. *Remote Sensing of Environment 279*: 113112. https://doi.org/10.1016/j.rse.2022.113112

Pearse, G. D., Jayathunga, S., Camarretta, N., Palmer, M.E., Steer, B.S., Watt, M.S., Watt, P., & Holdaway, A. (2025). Developing a forest description from remote sensing: Insights from New Zealand. *Science of Remote Sensing*, *11*: 100183. https://doi.org/10.1016/j.srs.2024.100183

Peterson, B., Dubayah, R., Hyde, P., Hofton, M., Blair, J.B., & Fites-Kaufman, J. (2007). Use of LIDAR for forest inventory and forest management application. In R.E. McRoberts, G.A. Reams, P.C. Van Deusen, & W.H. McWilliams (Eds.), *Proceedings of the seventh annual forest inventory and analysis symposium* (pp. 193-202). US Department of Agriculture, Forest Service.

Planet Labs PBC. (2024). *Forest Carbon: Reliable, scalable aboveground carbon estimates*. Retrieved March 11, 2024, from https://www.planet.com/products/forest-carbon/

Potapov, P., Li, X., Hernandez-Serna, A., Tyukavina, A., Hansen, M.C., Kommareddy, A., Pickens, A., Turubanova, S., Tang, H., Edibaldo Silva, C., Armston, J., Dubayah, R., Blair, J.B., & Hofton, M. (2021). Mapping global forest canopy height through integration of GEDI and Landsat data. *Remote Sensing of Environment*, *253*, 112165. https://doi.org/10.1016/j.rse.2020.112165

Quirós, E., Polo, M.E., & Fragoso-Campón, L. (2021). GEDI Elevation Accuracy Assessment: A case study of Southwest Spain. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 14,* 5285-5299. https://doi.org/10.1109/JSTARS.2021.3080711

Rogers, G.M., Walker, S., & Lee, W.G. (2005). *The role of disturbance in dryland New Zealand: past and present* (No. 258, pp. 122-pp). Wellington: Department of Conservation.

Roy, S., Swetnam, T., Shen, C., Somasundaram, D., & Saah, A. (2024). *samapriya/awesome-gee-community-datasets: Community Catalog (Version 3.0.0)* [Data set]. Zenodo.

Schwartz, M., Ciais, P., Ottlé, C., De Truchis, A., Vega, C., Fayad, I., Brandt, M., Fensholt, R., Baghdadi, N., Morneau, F., Morin, D., Guyon, D., Dayau, S., & Wigneron, J.P. (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. https://doi.org/10.1016/j.jag.2024.103711

Tang, H., Stoker, J., Luthcke, S., Armston, J., Lee, K., Blair, B., & Hofton, M. (2023). Evaluating and mitigating the impact of systematic geolocation error on canopy height measurement performance of GEDI. *Remote Sensing of Environment, 291*: 113571. https://doi.org/10.1016/j.rse.2023.113571

Toitū Te Whenua Land Information New Zealand (LINZ). (2022). *New Zealand national aerial LiDAR base specification (v1.2).* Retrieved April 18, 2024, from https://www.linz.govt.nz/system/files/2022-10/New%20Zealand%20National%20Aerial%20LiDAR%20Base%20Specification%20v1.2.pdf

Toitū Te Whenua Land Information New Zealand (LINZ). (n.d.). *New Zealand Elevation*. Toitū Te Whenua Land Information New Zealand. Retrieved March 1, 2024, from https://registry.opendata.aws/nz-elevation

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 Sensing of Environment*, *300*, 113888. https://doi.org/10.1016/j.rse.2023.113888

Torresani, M., Rocchini, D., Alberti, A., Moudrý, V., Heym, M., Thouverai, E., Kacic, P., & Tomelleri, E. (2023). LiDAR GEDI derived tree canopy height heterogeneity reveals patterns of biodiversity in forest ecosystems. *Ecological Informatics, 76*, 102082. https://doi.org/10.1016/j.ecoinf.2023.102082

Tsao, A., Nzewi, I., Jayeoba, A., Ayogu, U., & Lobell, D.B. (2023). Canopy height mapping for plantations in Nigeria using GEDI, Landsat, and Sentinel-2. *Remote Sensing*, *15*(21): 5162. https://doi.org/10.3390/rs15215162

van der Colff, M. & Kimberley, M.O. (2013). A National height-age model for *Pinus radiata* in New Zealand. *New Zealand Journal of Forestry Science, 43*: 4. https://doi.org/10.1186/1179-5395-43-4

Wagner, F.H., Roberts, S., Ritz, A.L., Carter, G., Dalagnol, R., Favrichon, S., Hirye, M.C.M., Brandt, M., Ciais, P., & Saatchi, S. (2024). Sub-meter tree height mapping of California using aerial images and LiDAR-informed U-Net model. *Remote Sensing of Environment, 305*: 114099. https://doi.org/10.1016/j.rse.2024.114099

Wallace, L., Lucieer, A., Watson, C., & Turner, D. (2012). Development of a UAV-LiDAR system with application to forest inventory. *Remote Sensing, 4*(6), 1519-1543. https://doi.org/10.3390/rs4061519

Yu, Q., Ryan, M. G., Ji, W., Prihodko, L., Anchang, J. Y., Kahiu, N., Nazir, A., Dai, J., & Hanan, N.P. (2024). Assessing canopy height measurements from ICESat-2 and GEDI orbiting LiDAR across six different biomes with G-LiHT LiDAR. *Environmental Research: Ecology*, *3*(2): 025001. https://doi.org/10.1088/2752-664X/ad39f2

Zhu, X., Nie, S., Wang, C., Xi, X., Lao, J., & Li, D. (2022). Consistency analysis of forest height retrievals between GEDI and ICESat-2. *Remote Sensing of Environment*, *281*: 113244. https://doi.org/10.1016/j.rse.2022.113244