Retrieval of forest height information using spaceborne LiDAR data
a comparison of GEDI and ICESat-2 missions for Crimean pine (Pinus nigra) stands

Vatandaslar, Can; Narin, Omer Gokberk; Abdikan, Saygin

DOI
10.1007/s00468-022-02378-x

Publication date
2022

Document Version
Final published version

Published in
Trees - Structure and Function

Citation (APA)
Vatandaslar, C., Narin, O. G., & Abdikan, S. (2022). Retrieval of forest height information using spaceborne LiDAR data: a comparison of GEDI and ICESat-2 missions for Crimean pine (Pinus nigra) stands. *Trees - Structure and Function*, *37*(3), 717-731. https://doi.org/10.1007/s00468-022-02378-x

Trees
https://doi.org/10.1007/s00468-022-02378-x

ORIGINAL ARTICLE

Retrieval of forest height information using spaceborne LiDAR data: a comparison of GEDI and ICESat-2 missions for Crimean pine (*Pinus nigra*) stands

Can Vatandaslar[1] · Omer Gokberk Narin[2,3] · Saygin Abdikan[4]

Received: 17 August 2022 / Accepted: 8 December 2022

Abstract

Key message **Despite showing a cost-effective potential for quantifying vertical forest structure, the GEDI and ICESat-2 satellite LiDAR missions fall short of the data accuracy standards required by tree- and stand-level forest inventories.**

Abstract Tree and stand heights are key inventory variables in forestry, but measuring them manually is time-consuming for large forestlands. For that reason, researchers have traditionally used terrestrial and aerial remote sensing systems to retrieve forest height information. Recent developments in sensor technology have made it possible for spaceborne LiDAR systems to collect height data. However, there is still a knowledge gap regarding the utility and reliability of these data in varying forest structures. The present study aims to assess the accuracies of dominant stand heights retrieved by GEDI and ICESat-2 satellites. To that end, we used stand-type maps and field-measured inventory data from forest management plans as references. Additionally, we developed convolutional neural network (CNN) models to improve the data accuracy of raw LiDAR metrics. The results showed that GEDI generally underestimated dominant heights (RMSE = 3.06 m, %RMSE = 21.80%), whereas ICESat-2 overestimated them (RMSE = 4.02 m, %RMSE = 30.76%). Accuracy decreased further as the slope increased, particularly for ICESat-2 data. Nonetheless, using CNN models, we improved estimation accuracies to some extent (%RMSEs = 20.12% and 19.75% for GEDI and ICESat-2). In terms of forest structure, GEDI performed better in fully-covered stands than in sparsely-covered forests. This is attributable to the smaller height differences between canopy tops in dense forest conditions. ICESat-2, on the other hand, performed better in thin forests (DBH < 20 cm) than in large-girth and mature stands of Crimean pine. We conclude that GEDI and ICESat-2 missions, particularly in hilly landscapes, rarely achieve the standards needed in stand-level forest inventories when used alone.

Keywords Light detection and ranging (LiDAR) · Global ecosystem dynamics investigation (GEDI) · Ice cloud and land elevation satellite-2 (ICESat-2) · Height metrics · Canopy height model · Convolutional neural network (CNN)

Communicated by R. Guy.

✉ Can Vatandaslar
canvatandaslar@artvin.edu.tr

[1] Department of Forest Engineering, Artvin Coruh University, 08100 Artvin, Turkey

[2] Geoscience & Remote Sensing Department, Faculty of Civil Engineering, Technische Universiteit Delft, Delft, Netherlands

[3] Department of Geomatics Engineering, Afyon Kocatepe University, Afyon, Turkey

[4] Department of Geomatics Engineering, Hacettepe University, Ankara, Turkey

Abbreviations

LiDAR	Light detection and ranging
LS	Laser scanning
ALS	Airborne laser scanning
SLS	Spaceborne laser scanning
MLS	Mobile laser scanning
TLS	Terrestrial laser scanning
RMSE	Root mean square error
GLAS	Geoscience laser altimeter system
GEDI	Global ecosystem dynamics investigation
ISS	International space station
ICESat-2	Ice, cloud and land elevation satellite-2
NN	Neural network
CNN	Convolutional neural network
WGS84	World geodetic system 1984

Published online: 23 December 2022

ATL08	Land and vegetation height product of ICESat-2
GDF	Turkish general directorate of forest
GIS	Geographic information systems

Introduction

Stand height is one of the essential parameters in forest inventories. It is needed to characterize the vertical structure of forest ecosystems to estimate aboveground biomass, carbon stocks, stand volume, site index, and many other silvicultural or biodiversity-related metrics (Duncanson et al. 2020; Özkal et al. 2021; Potapov et al. 2021; Liu et al. 2022). The future growth of the forest may also be simulated at the stand or landscape level based on the height information when other inventory parameters are known (Laar and Akça 2007). Thus, the long-term sustainability of forests and their ecosystem services can be ensured with more accurate and up-to-date data for informed decision-making (Bettinger et al. 2009; Baskent 2020; Biber et al. 2020).

In conventional forest inventories, stand height is calculated based on in situ tree height measurements performed using mechanical or electronic hypsometers (e.g. Blume-Leiss, Vertex). Many stand height calculation approaches exist in the literature, including the arithmetic mean height ($\hbar$), Lorey's mean height (h_L; mean height weighted by trees' basal area), and dominant height (h_{100}) (Laar and Akça 2007). While the $\hbar$ averages the height values of all trees in a sample plot, the h_{100} averages only dominant trees' heights, taking 100 trees per ha. h_L, on the other hand, calculates the mean height weighting by individual trees' basal area. Regardless of which approach is used, height measurement in forest environments is a daunting task by manual means (Hyyppa et al. 2020a). Moreover, it may have systematic or random errors stemming from the operator, sampling schemes, measurement devices, and/or terrain slope (Sibona et al. 2017; Ganz et al. 2019; Persson and Stahl 2020). Therefore, for practical reasons, tree heights are often estimated by regression models (a.k.a. h–d models or height curves) based on the relationship between diameter-at-breast height (DBH) and height in many countries (Vidal et al. 2016; Bolat et al. 2022). These regression models also have estimation errors due to the embedded statistical assumptions (Laar and Akça 2007; Persson and Stahl 2020) and improper use across different eco-regions (Bolat et al. 2022; Seki and Sakici 2022a; 2022b). That is why the permissible error limit for tree height is more flexible than other forest inventory parameters, such as the area coverage of stands. For example, an error rate of 15% is permissible for tree height measurements, while it is only 5% for the sub-compartment area (Qiu et al. 2018).

Given the limitations mentioned above, researchers try to measure or model the stand height using active remote sensing systems. Light Detection and Ranging (LiDAR, a.k.a. laser scanning – LS) is one of the active systems measuring the distance of objects or surfaces from the sensor using laser pulses. LiDAR systems may be classified in different ways according to the laser altimetry technique (i.e. discrete return, full-waveform, photon-counting) or their platforms (airborne LS/ALS, terrestrial LS/TLS, mobile LS/MLS, drone LS/DLS, spaceborne LS/SLS) (Wang and Fang 2020). Each system has its unique advantages and disadvantages for specific fields of application. Discrete return systems, for example, detect a few returns from each short pulse during the flight of the system. The discrete returned peak points are recorded in the waveform, which may limit the estimation of actual canopy height. By contrast, full-waveform systems record the whole energy returning at equal intervals. Thus, they can record many returns yielding a more detailed representation of forests' vertical structure (Salas 2021). The photon-counting sensors, on the other hand, are able to collect energy with individual photon-sensitive detectors (Marcus et al. 2017).

In the forestry field, there is numerous research conducted with different LiDAR platforms. Kanja et al. (2019), for example, modeled stand height using ALS-derived metrics in a Mediterranean forest dominated by Calabrian pine (*Pinus brutia* (Ten.)). Their multiple regression models explained 73% and 80% of the variations in the dominant and mean heights, respectively. The relative RMSEs of the models were about 15% (ca. 2 m) which were within the permissible error limits for most forest inventory missions (Qiu et al. 2018; Hyyppa et al. 2020a). However, ALS data provision is expensive, especially for undeveloped and developing nations, because it is captured via overlapped flight lines in the areas of interest (Sefercik et al. 2021). Alternatively, transects are used for sampling the population in question (Coops et al. 2021). Therefore its spatial coverage is often limited, like DLS applications (Ganz et al. 2019). In addition, the insufficient sampling density of ALS point clouds might be another limitation in some cases. The sparse point clouds used by Kanja et al. (2019) had a sampling density of 8–9 points/m^2, and they stated that the height models could be improved significantly by increasing point density. In this sense, close-range LS provides a clear advantage over ALS (Sefercik et al. 2021). In a handheld MLS-based study by Vatandaşlar and Zeybek (2021), more than 3000 points fell per m^2 in a mixed forest plot. Using an MLS system, dominant stand heights can be directly captured without modeling with a bias of less than 1 m (3.4%) at the forest level (Vatandaşlar et al. 2022). If static TLS systems are used, the bias decreases even more (Cabo et al. 2018; Liang et al. 2019). Nevertheless, TLS of a forest may not be the best approach due to the occlusion problem and the device's weight (Vatandaşlar and Zeybek 2020, 2021). Thus, one of the most reliable and practical approaches for the dominant height estimation today seems to be the use of Simultaneous Localization and Mapping (SLAM)-based MLS (Vatandaşlar et al. 2022; Hyyppa et al. 2020a, b), sometimes added-on to drones (Liang et al. 2019; Hyyppa et al. 2021).

As airborne (ALS, DLS) and terrestrial platforms (TLS, MLS) provide only local-based data for a limited area (Coops et al. 2021), new generation SLS systems have been launched to fill the need for forest monitoring globally. SLS differs from other LS instruments in the sense of its mission because they are specifically designed for global monitoring purposes. SLS also uses full-waveform or photon-counting techniques, while discrete return and full-waveform are common in conventional ALS. Regarding resolution, footprint size varies between 12 and 25 m in SLS systems. Meanwhile, it ranges from 0.1 m to 3 m for small footprints, and 10 m to 30 m for large footprints of ALS (Beland et al. 2019). NASA's *Global Ecosystem Dynamics Investigation (GEDI)* is the first full-waveform SLS system that specifically collects data on forest structure (Duncanson et al. 2020). Combining the GEDI and Landsat data, a 30-m-resolution forest canopy height map of the world has been generated for 2019. Taking the global validation data and available ALS resources as the reference, the RMSE and R^2 values of the map were found as about 6 m and 0.60, respectively (Potapov et al. 2021). On the other hand, Liu et al. (2021) assessed the accuracies of GEDI and ICESat-2 height metrics using ALS data as a reference. They revealed that GEDI outperformed ICESat-2 for canopy height estimations in forestlands of the US. In another study, Liu et al. (2022) integrated GEDI, ICESat-2, and Sentinel-2 data to produce China's national canopy height map. They applied the neural network (NN)-guided interpolation method and estimated canopy height with RMSEs between 4.88 m and 5.32 m and R^2 values between 0.55 and 0.60 using three validation datasets. Fayad et al. (2021a) also applied a convolutional NN-based approach to GEDI data to estimate the canopy height of *Eucalyptus* stands in Brazil. Different scenarios showed that the RMSE values ranged from 1.54 m to 1.94 m, with R^2 values varying between 0.86 and 0.91.

Another mission of NASA is *Ice, Cloud, and Land Elevation Satellite-2 (ICESat-2),* launched in 2018. It is a photon-counting laser altimeter focusing on Polar Regions, glaciers, and seas (Markus et al. 2017). ICESat-2 is the continuation of its predecessor, the *Geoscience Laser Altimeter System (GLAS)* on the ICESat, which was the first spaceborne laser altimeter operated between 2003 and 2009. It also focused on ice thickness estimation and monitoring of the Arctic Ocean and Antarctic ice sheets (Neuenschwander et al. 2020; Markus et al. 2017). Studies on ICESat and ICESat-2 showed that they could provide consistent forest biomass and canopy height estimates, as well. Using ICESat data, Lefsky et al. (2005) estimated biomass ($R^2 = 73\%$) and canopy height (RMSE < ~ 12.7 m) over Brazil and the US. Helmer et al. (2009), on the other hand, combined ICESat and Landsat data to estimate the accumulation rates of forest biomass in Brazil. Recent studies have also shown the performance of ICESat-2 data in forest research. For instance, Sun et al. (2020) used both Landsat and ICESat missions and monitored the change in canopy height over China from 2005 to 2019. Combining ICESat-2 and Sentinel-1 data, Nandy et al. (2021) modeled canopy height and aboveground biomass with R^2 values of 0.84 and 0.83, respectively. Neuenschwander and Magruder (2019) studied the boreal forests of Finland and estimated canopy heights with an RMSE of 3.2 m using ALS data as a reference. They stated that the 4-year time difference between the ICESat-2 and ALS data might negatively affect the estimation accuracy. Finally, Neuenschwander et al. (2020) compared canopy heights derived from ICESat and ALS data over the boreal forests of Finland. They concluded that RMSE reaches up to 2.7 m during winter when the terrain is covered with snow. They also stated that forest stands whose canopy cover ratios were between 40 and 85% yielded the most accurate estimation in their study.

As seen, previous studies were conducted at the global (Potapov et al. 2021), national (Liu et al. 2022), or regional (Duncanson et al. 2020; Dorado-Roda et al. 2021) levels. That is why their estimation accuracies rarely satisfy the needs of forest-stand inventories, mostly conducted at the landscape level. Qui et al. (2018) state that the maximum permissible error for tree height measurements is 15% in operative forest management. On the other hand, published studies generally use ALS as reference data, which may cause additional errors in height estimates. From a forest management perspective, there is a need to assess the accuracy of satellite-based height data using field-based reference measurements. Minimizing the errors of the height data by advanced modeling techniques, e.g. machine learning, deep learning and artificial intelligence, is also necessary for properly characterizing forest ecosystems. With the increase in new data sources, the amount of data collected has also increased considerably. In this context, systems similar to nerve neurons in the human brain and computer-controlled or unsupervised deep learning methods have been developed. The relationship between the remote sensing data and field-measured forest attributes can be built by methods including conventional ones, such as linear regression and multiple regression; or up-to-date machine learning algorithms, such as random forest, artificial neural network, and convolutional neural network (CNN) (Coops et al. 2021; Fayad et al. 2021b; Özkal et al. 2021; Colkesen et al. 2022; Ercanli et al. 2022). Amongst them, the superior performance of CNN has been shown by researchers for forest height estimates (Narin et al. 2022), forest cover mapping (Ferreira et al. 2021), and canopy modeling (Shah et al. 2020). CNN is a sort of deep learning approach, having four layers (input, convolution, pooling, and fully-connected layers) within its architecture. Thus, it is capable of processing one-, two-, and three-dimensional data in the forms of signal, image, or video (Li et al. 2018). Here, we use one-dimensional CNN architecture, as will be illustrated in the Methodology section in detail.

The main objective of the present study is to estimate stands' dominant heights using SLS systems at the landscape level. Specifically, we focus on (1) assessing the accuracy of GEDI and ICESat-2 data by taking the field-measured heights as a reference, (2) determining which stand types achieve the best estimation results, and (3) developing a CNN model for minimizing error rates in dominant height estimates. This is one of the first studies focusing on a direct comparison of forest stand height information retrieved from the GEDI and ICESat-2 sensors. Moreover, the vertical structure of Crimean pine (*Pinus nigra*) forests will be evaluated through SLS data for the first time. The findings of this study would be useful to forest managers and timber surveyors in their regular forestry operations.

Material and methods

Study area

The study area is located in the central Anatolian region of Turkey (Fig. 1) and consists of two state forest enterprises, i.e. Hocalar and Sinanpaşa. Nearly one-third of it is forested, covering a total area of 145,649 ha. Less than half of the forest (20,906 ha) is characterized by productive stands whose canopy cover ratio is > 10%. The remainder of the forest (28,974 ha) is degraded due to the past anthropogenic pressure and unfavorable climate condition in this region (GDF 2015a, b). Turkey's macro climate map shows the region is in the Central-West Anatolian climate zone (Türkeş 2010). The average annual precipitation total and air temperature are reported as 466 mm and 12 ℃, respectively (TMS 2015). While the number of frost days is 95 in a year, the minimum and maximum air temperatures are − 22 ℃ and 39.8 ℃. That is why the forest is dominated by Crimean pine (*Pinus nigra* J.F. Arnold subsp. *pallasiana* (Lamb.) Holmboe), a common tree species resistant to natural disturbances (Seki and Sakici 2022a). It is followed by oak *(Quercus sp.)*, juniper (*Juniperus comminus, J. oxycedrus*), and Lebanese cedar (*Cedrus libani*) species. Productive forests are mostly covered by pure, even-aged stands of Crimean pine at different developmental stages. Given the forest management plans, the site

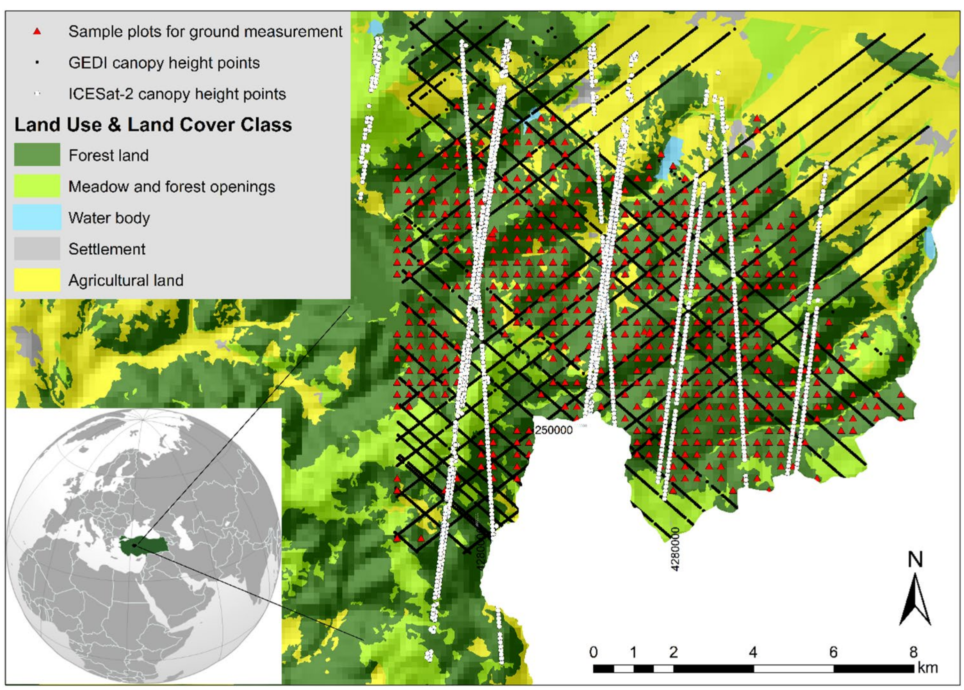

Fig. 1 Location of the study area, field-measured plots, and LiDAR data points with main land use & land covers

quality class changes from moderate (3rd bonitet class) to poor (5th bonitet class). There is no high-quality site in the study area. While the altitude is between 920 and 1940 m asl, the average slope is 25% for the entire area. The average slope increases to 35% in forestlands. The steep forested slopes can be seen in the hillshade map shown in Fig. 1.

Satellite data

Different data sets from two satellite missions were employed in the current investigation. The first mission was GEDI and the second one was ICESat-2. GEDI launched its mission in late 2018 with the main goal of global forest monitoring. It has a full-waveform laser altimeter system mounted on the *International Space Station (ISS)* (Potapov et al. 2021). Between latitudes 51.6° North and 51.6° South, the system scatters laser beams in the range of 1064 nm (the electromagnetic spectrum's near-infrared region). The World Geodetic System 1984 (WGS84) coordinate system is used by GEDI to collect data (Dubayah et al. 2020).

GEDI data were downloaded from http://www.daac.ornl.gov and processed in R (R Core Team 2020) using the *rGEDI* package (Silva et al. 2020). Within the scope of this study, all data coded as "Quality_flag 1" were used because a quality flag value of zero means that such data was of insufficient quality. First, all available data were downloaded for the year 2019 (July 4, July 13, Sep. 7, Sep. 16, Sep. 29, Oct. 11) but the data dated Sep.16 was not used since its quality flag was zero. Thus, a total of 2344 data points were analyzed for GEDI. During the study, only the relative height values at the 95th percentile (RH95) metric were evaluated. We applied no additional filtering because the number of data points was limited for our area of investigation.

The ICESat-2 satellite provided the second data source. It began collecting data in 2018 as a follow-up to the ICESat/GLAS satellite. A photon-counting laser altimeter system with a beam wavelength of 532 nm (the electromagnetic spectrum's green region) is used in the ICESat-2. The system's temporal resolution is approximately three months. It gathers data between the latitudes of 88° North and 88° South (Neuenschwander and Magruder 2019).

The Advanced Topographic Laser Altimeter System (ATLAS) instrument aboard ICESat-2 generates a variety of products for various missions, including global geolocated photon data product (ATL03), land ice elevation (ATL06), sea ice elevation (ATL07), land and vegetation height product (ATL08), atmosphere (ATL09), ocean elevation (ATL12), and inland water height (ATL13) (Marcus et al. 2017). The ATL08 product (Neuenschwander et al. 2020) of the ICESat-2 was downloaded for this study on http://www.openaltimetry.org (Khalse et al. 2020). This product collects elevation data for land and vegetation. Covering the year 2019 (Feb. 28, Aug. 1, Aug. 29, Oct. 29, and Oct. 30), a total of 833 data points were used from ICESat-2. However, only the "h_canopy" metric was evaluated because other metrics were outside the scope of this study.

Methodology

In situ measurements and height calculations

In situ measurements were based on forest inventory data included in the management plans of the two forest enterprises (GDF 2015a, b). Forest inventory surveys were performed by forest professionals from the Turkish General Directorate of Forest (GDF). In order to renew the forest management plans, sample plots were systematically distributed to forestland in GIS. The size of the circular plots were 400 m^2, 600 m^2 and 800 m^2 for fully-covered (canopy cover > 70%), medium-covered (40% – 70%) and loosely-covered (10% – 40%) stands, respectively. The distance between the plot centers was roughly 300 m × 300 m. Each plot was visited during the summer of 2014. All trees with DBH equal to or more than 8 cm were measured and recorded in the plots. Other measured/identified inventory parameters were as follows: tree species, age and heights of dominant trees, canopy cover ratio, timber quality, and forest health.

Tree heights were measured from the dominant trees using an electronic hypsometer having a sensitivity of ± 1 cm. The age of the dominant trees was also

Table 1 The age and site class information for natural Crimean pine forests of Turkey (Kalıpsız 1963)

Age class	Class boundaries	Midpoint of the age class	Site quality class	Dominant height at standard age[a]	Dominant height at the class mid-point
1	0–20 years	10-year-old	1 (very high)	30–34 m	32 m
2	21–40 years	30-year-old	2 (high)	25–29 m	27 m
3	41–60 years	50-year-old	3 (moderate)	20–24 m	22 m
4	61–80 years	70-year-old	4 (low)	15–19 m	17 m
–	–	–	5 (very low)	10–14 m	12 m

[a]The standard (a.k.a. base) age is 100 years for Turkey's growth and yield tables

determined by counting annual rings collected with a Haglöf increment borer. Thus, site quality (a.k.a. bonitet) classes were identified for each plot based on Crimean pine's national site index table (Kalıpsız 1963). Further details on the field methods used for height measurement in these forests may be found in Laar and Akça (2007).

Since forest surveys were carried out in 2014, we updated the inventory data measured from 1242 sample plots across two state forest enterprises. However, only 633 of them overlapped with the LiDAR data points and this subset was used in the analyses. Adding five years to each, stand ages were forwarded to late 2019, the data acquisition year of the satellites. Accordingly, measured tree heights were recalculated given the Crimean pine growth and site index tables generated by Kalıpsız (1963). Stand height and age class information for the Crimean pine forests in Turkey can be seen in Table 1.

Growth and yield models are useful tools in forest management for developing predictions of future states of stands. However, they generally require several assumptions to simplify and track the problem (Bettinger et al. 2009). Thus, we used some assumptions suggesting that forest stands undisturbedly develop in a controlled environment, which is a rare situation in a real forest. Table 2 specifies the assumptions made when updating stand ages based on the site index model (Kalıpsız 1963; Table 1). In Table 2, the readers can also see why those assumptions are necessary and whether they are valid for the case study areas.

GIS analysis

First, the updated stand heights were entered into the attribute tables of the forest cover maps in the GIS environment. Then, the maps of the two forest enterprises were combined into one GIS layer. The new layer was spatially joined to LiDAR data points based on the *intersect* rule using the "Join One to Many" command in ArcGIS 10.2 (ESRI 2012).

In a next step, the "Target Field ID" column was summarized by averaging the heights of the LiDAR data points that fell into each polygon. Here, the polygons represent stand types in the landscape. Each stand type has unique forest inventory parameters, such as species mix, dominant height, age & site classes, canopy cover ratio, and developmental stage.

Finally, the GIS summary table was transferred to an Excel sheet for further data analysis. All analyses were separately performed for GEDI and ICESat-2 data sets. Thus, it could be possible to assess the dominant height estimates of the two satellite systems at the stand level. Table 3 summarizes how many samples we have for different data types.

Table 2 Simplifying assumptions made for using the site index model of Kalıpsız (1963)

Assumption	Justification	Validation
All stands in the study area are (near-) natural	Turkey's growth and site index models have been developed for natural forest conditions	More than 90% of Turkey's forests are (near-) natural in origin (GDF 2020)
Only dominant tree species (i.e. Crimean pine) is taken into account for mixed forest stands	In Turkey, there has yet to be a growth and site index model for mixed forests of Crimean pine	While 90.6% of all stands are pure in our study area, 98.5% of the total growing stock comprises Crimean pine species (GDF 2015a, b)
Each stand is at the midpoint of its age class interval. For example, a young stand in the first-age class would be ten years old (see Table 1)	In Turkish forest plans, stand age information is provided as age classes only	The age class interval used for Crimean pine is 20 years in GDF (2015a, b) because it is not a fast-growing tree species
Each stand continues to grow based on the natural forest development without any management interventions or biotic & abiotic disturbances	"*To be able to use quantitative methods, we may make simplifying assumptions so that problems are tractable (useable)*" (Bettinger et al. 2009, p. 93)	The forests in our area of investigation are mostly set aside for conservation due to soil erosion problems and poor site conditions. No known (a)biotic agent significantly disturbs the forest in these forest enterprises (GDF 2015a, b)
The country-level site index model is suitable for the case study area	There is no local site index model developed for Hocalar and Sinanpaşa sub-regions	The average site condition in our study area generally reflects the typical structure of Crimean pine forest in Turkey (GDF 2015a, b)

Table 3 Summary table for the number of data collected and used in the assessment

	GEDI RH95 metric	ICESat-2 ATL08 (h_canopy metric)	Grand total	GIS data type
The number of LiDAR data	5510 (2344) [a]	1848 (833)	7358 (3177)	Point
The number of sample plots	N/A [b]	N/A [b]	1242 (633)	Point
The number of forest stands	806 (249)	320 (116)	N/A[c]	Polygon

[a]While the values outside the parenthesis are the total number of data that fell in the entire study area, the values in parenthesis show the number of data used in the analyses

[b]Not applicable because GEDI and ICESat-2 points are not overlapping with ground sample plots exactly

[c]Not applicable because some of the forest stands are the same in GEDI and ICESat-2 coverages

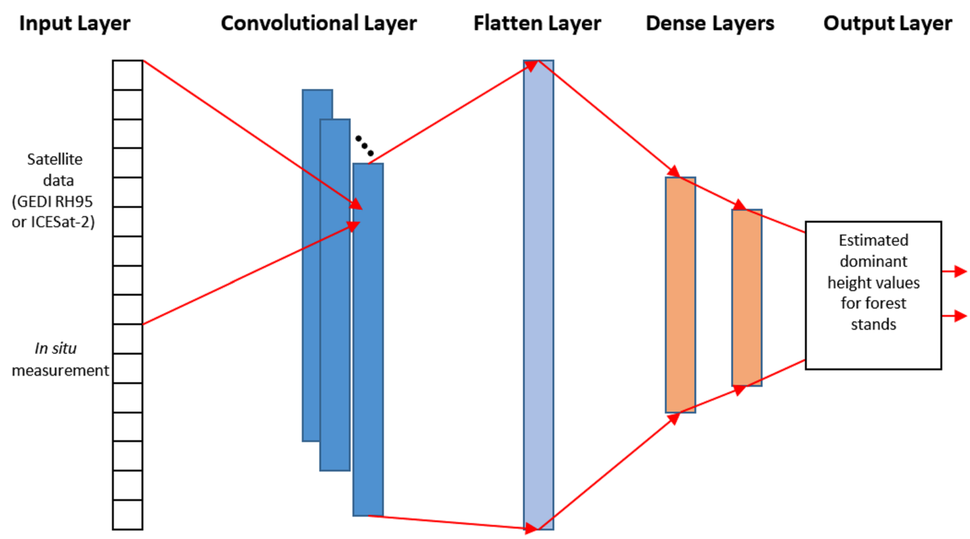

Fig. 2 The model architecture for the CNN developed. Both GEDI and ICESat-2 models are composed of one convolution, one flatten, one output, and two dense layers

Modeling (CNN)

We developed convolutional neural network (CNN) models based on the strong relationship between field-measured and satellite-derived height values to estimate the dominant heights of stands more accurately. To generate CNN models, a one-dimensional network was preferred. Because our data points were limited, they were not classified during the model development stage based on slope, species mix, and canopy cover classes. Namely, the GEDI and ICESat-2 models were developed for all stands (i.e., unclassified data). However, we used the K-fold cross-validation technique repeated four times.

The CNN models were developed using the R programming language (URL1). We used *keras* (Allaire and Chollet 2022) and *caret* (Kuhn 2022) packages for the model development, whose architecture was depicted in Fig. 2. The input layer in the figure included both field-measured and satellite-derived dominant height data. The filter and activation functions in the convolutional layer were set to "256" and "relu", respectively. The dense layer was divided into two levels. In the first level, the units and activation functions were "1024" and "relu". In the second level, the same parameters were "1" and "linear". Finally, we ran the model with a large number of epochs and batches of 250 and 16, respectively.

Accuracy assessment

We assumed the updated in situ measurements were true and used them as the reference dataset for evaluation. Then, the accuracy of GEDI- and ICESat-derived dominant heights estimated by models were assessed based on the root mean squared error (RMSE), relative RMSE (%RMSE), bias, and Pearson's correlation coefficient (r) statistics. They were calculated using the following formulas:

$$RMSE = \sqrt{\frac{1}{n}\sum\nolimits_{i=1}^{n}\left(y_i^{act} - y_i^{cal}\right)} \quad (1)$$

$$\%RMSE = \sqrt{100 * \frac{1}{n}\sum\nolimits_{i=1}^{n}\left(y_i^{act} - y_i^{cal}\right)} \quad (2)$$

$$bias = \left(y_i^{cal} - y_i^{act}\right) \quad (3)$$

Table 4 The definitions for stand characteristics used for data classification

Stand variable	Class name	Definition	# of plots
Canopy cover	Fully-covered forest	Canopy cover is more than 70%	459
	Sparsely-covered forest	Canopy cover is between 10 and 70%[a]	174
Development stage	Medium-tree stage	Tree DBH is between 20 and 52 cm	360
	Thin-tree stage	Tree DBH is equal or less than 20 cm	273
Species mixture	Mixed forest	Stands with more than one tree species	47
	Pure forest	Stands with only one tree species[b]	586
Terrain slope	Forests in flatter areas	The slope gradient is less than 35%[c]	371
	Forests on a steep slope	The slope gradient is equal or higher than 35%	262

[a]Degraded forests (canopy cover < 10%) are excluded from the analysis because they have no reliable field-measured height data

[b]As long as the second tree species does not exceed 10% of the total number of trees, it is considered a pure forest

[c]35% was used as a threshold because it was the mean slope rate of the forestland in the study area

$$r = \frac{\sum\left(y_i^{act} - \overline{y_i^{act}}\right)\left(y_i^{cal} - \overline{y_i^{cal}}\right)}{\sqrt{\sum\left(y_i^{act} - \overline{y_i^{act}}\right)^2\left(y_i^{cal} - \overline{y_i^{cal}}\right)^2}} \quad (4)$$

where n is the number of samples; y_i^{act} is the field-measured height; y_i^{cal} is the estimated height by satellite missions or the CNN models; $\overline{y_i^{act}}$ and $\overline{y_i^{cal}}$ are the average field-measured and estimated height values, respectively.

Error statistics were separately calculated for GEDI and ICESat-2 using their whole datasets (all stands). To understand the potential error sources, we also classified the data given several stand types and recalculated the same statistics. The assessed stand types and their characteristics can be seen in Table 4.

Results

GEDI data

The residuals for satellite-based dominant heights were shown as box plots in Fig. 3. Considering all stands together ($N=249$), a bias of − 0.55 m was seen between in situ measurements and GEDI data (Table 5). When only fully covered stands were assessed, the bias decreased to − 0.38 m. In contrast, it increased to − 1.20 m for sparsely covered stands. Regarding the development stage, both thin- (DBH < 20 cm) and large-girth (DBH > 20 cm) stands' heights were underestimated with biases of − 0.58 m and − 0.52 m, respectively.

We also assessed the RH95 metric regarding the terrain slope and species mix classes (Fig. 3c, d, e). Accordingly, a bias of − 0.62 m was found between GEDI and the reference data for pure stands, and it decreased to − 0.23 m for mixed stands. However, it should be noted that the number of samples for mixed stands ($N=25$) was significantly lower than the pure ones ($N=223$). Figure 3e also showed that the residuals were mostly negative in the box plot. While RMSEs were close to each other, the pure stands had smaller residuals, 10% lower %RMSE, and a much higher R^2 value (Table 5). The terrain slope, on the other hand, did not significantly affect the accuracy of dominant heights (Fig. 3c). Even though RMSE and R^2 values were lower for flatter areas than steep slopes, their %RMSE values were close to each other. Table 5 showed that RMSE and %RMSE for unclassified data were 3.06 m and 21.8%, respectively. Other

Table 5 Error statistics for satellite-based dominant heights by all stands and different stand types

Error statistics	GEDI RH95 / ICESat-2 canopy height metrics								
	Unclassified data (all stands)	Canopy cover		Stand development stage		Species mixture		Terrain slope	
		Fully-covered	Sparsely-covered	Medium-tree (> 20 cm)	Thin-tree (< 20 cm)	Mixed stands	Pure stands	Flat areas < 35%	Steep slopes > 35%
RMSE (m)	3.06/4.32[a]	2.78/4.24	3.40/4.64	3.33/4.76	2.61/2.99	3.11/0.99	3.02/4.50	2.70/3.81	3.15/4.87
%RMSE	21.80/30.76	20.20/31.41	23.08/28.50	20.15/28.35	25.15/27.15	30.83/8.65	20.84/31.38	21.34/29.18	20.78/31.96
Bias (m)	− 0.55/1.99	− 0.38/1.89	− 1.21/2.41	− 0.52 / 3.21	− 0.58/0.27	− 0.23/− 0.41	−0.62/2.30	− 0.62/1.11	− 0.56/3.08
Pearson's r	0.78/0.79	0.79/0.76	0.81/0.81	0.60/0.68	0.56/0.69	0.10/0.97	0.78/0.78	0.83/0.86	0.71/0.60

[a]While the values before the forward slash (/) refer to GEDI, the values after the forward slash are for ICESat-2

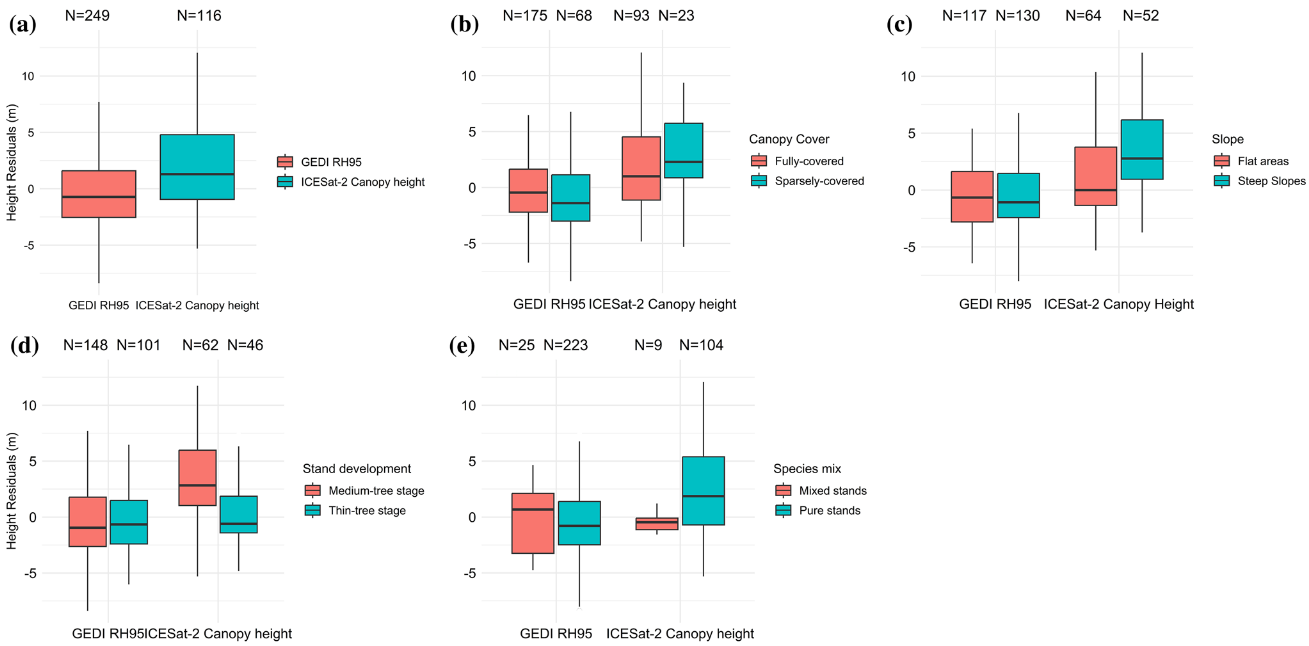

Fig. 3 The box plots demonstrating the residuals of dominant heights retrieved by GEDI and ICESat-2 for **a** unclassified data, **b** canopy cover, **c** terrain slope, **d** developmental stage, and **e** species mix classes. In situ height measurements are taken as the reference

error statistics for different stand types (classified data) can be seen in Table 5.

ICESat-2 data

In the case of analysis for ICESat-2, a total of 116 data points were assessed. Considering all stands, a bias of 1.99 m was found between in situ measurements and ICESat-2 data. Figure 3a indicated that ICESat-2 retrieved greater dominant height values than the field reference.

The accuracy of dominant stand heights stayed mostly the same according to canopy cover classes (Table 5). However, we observed considerable differences between the stands at the thin- and medium-tree developmental stages (Fig. 3d). While the bias for thin stands was 0.27 m, it was 3.21 m for the large-girth stands. The difference can be attributable to large-girth trees generally taller than smaller-diameter trees. Similar %RMSE values for the two development stages (i.e. %RMSE = 28.4%, %27.2) supported this inference.

We also assessed the satellite-based dominant heights according to species mix and slope classes. Although we have limited data points for mixed stands ($N=9$), the errors in the dominant heights of pure stands were noticeably higher than in the mixed stands (Table 5). The box plot for the mixed stands (Fig. 3e) was also very narrow, probably due to the lesser data points. As expected, the dominant height values of the stands on flatter areas were retrieved more accurately than those on steep slopes (Table 5).

Modeling results (CNN)

The dominant stand heights estimated by the CNN models were plotted against the reference data (i.e., in situ measurements) in Figs. 4b, c, d. The error statistics for the estimates can also be seen in Table 6. Except for R^2 values, all the statistics showed that the CNN models provided more accurate estimates compared to the raw height metrics of both GEDI and ICESat-2. The RMSE values decreased from 3.06 m to 2.82 m and 4.32 m to 2.77 m for GEDI and ICESat-2 models, respectively. With a decrease of nearly 11% in RMSE%, the estimation performance of the ICESat-2 model was better than the GEDI model, which had an improvement of 2% in RMSE%.

Regarding the estimation residuals, the GEDI model was almost unbiased. Its bias was only − 0.01 m (Table 6), suggesting the CNN technique significantly improved stand height retrievals by GEDI. On the other hand, the ICESat-2

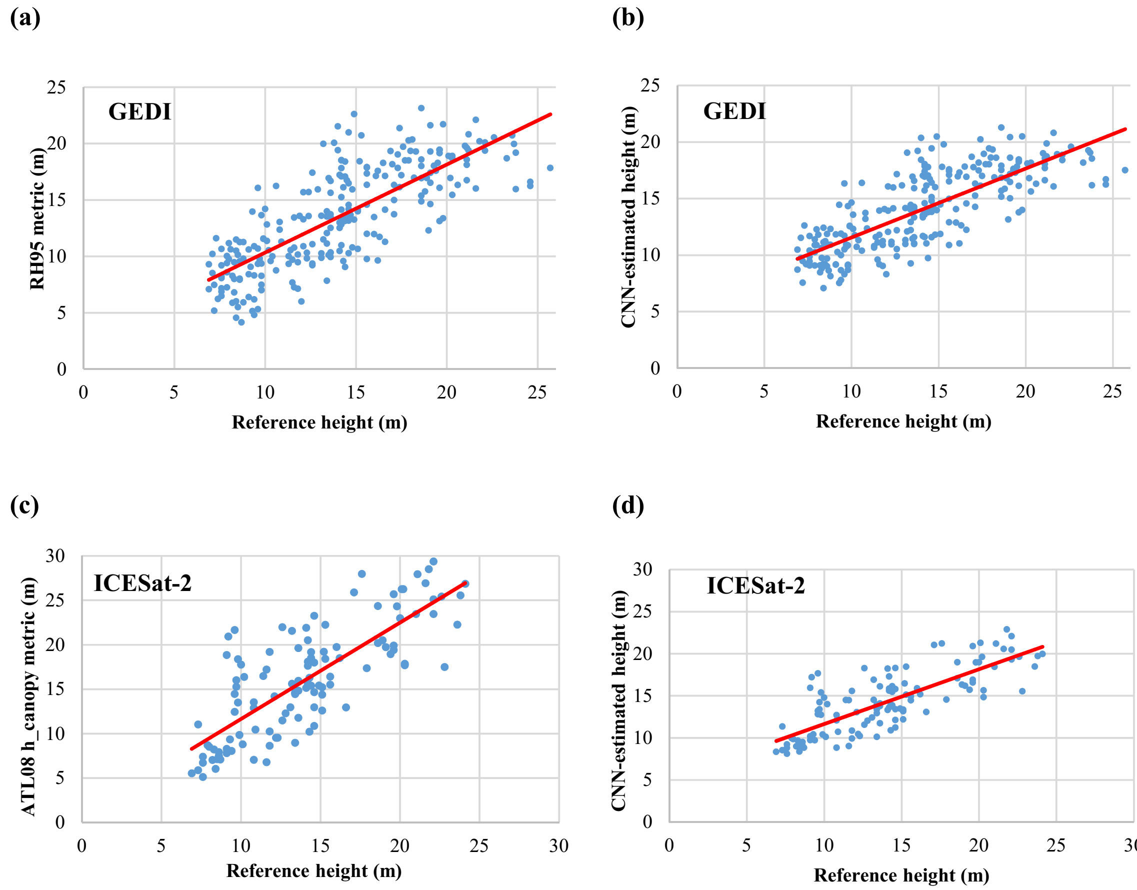

Fig. 4 The reference height values (field data) plotted against **a** raw GEDI metric, **b** modeled GEDI metric, **c** raw ICESat-2 metric, and **d** modeled ICESat-2 metric

Table 6 Error statistics for CNN-modeled dominant heights using GEDI and ICESat-2 data

Error statistics	Estimation results for stand dominant heights		Improvements compared to raw metrics	
	CNN model for GEDI	CNN model for ICESat-2	GEDI RH95	ICESat-2 h_canopy
RMSE (m)	2.82 (3.06)[a]	2.77 (4.32)	0.24	2.77
%RMSE (%)	20.12 (21.80)	19.75 (30.76)	1.68	11.01
Bias (m)	−0.01 (−0.55)	0.24 (1.99)	0.54	1.75
Pearson's r	0.78 (0.78)	0.79 (0.79)	0.00	0.00

[a]The values in the parenthesis indicate the error statistics for raw satellite data before modeling with CNN method

model slightly overestimated the stands' dominant heights and showed a slight positive bias of 0.24 m. However, this is still much lower than ICESat-2's raw height metrics. The improved estimates can also be observed by comparing scatter plots pairwisely (Figs. 4a, b, Figs. 4c, d).

Discussion

In the present study, we compared the dominant stand heights captured by GEDI and ICESat-2 in naturally regenerated, even-aged Crimean pine forests. Taking the in situ measurements as a reference, GEDI underestimated dominant heights for all stand types (bias = − 0.55 m). On the contrary, ICESat-2 generally overestimated the same inventory parameter with a bias of almost 2.00 m. These results are consistent with the recently published works. For instance, Potapov et al. (2021) stated that GEDI's global canopy map underestimated dominant heights with a bias of − 3.80 m compared to ALS-based validation data. In another study, Liu et al. (2021) reported a bias value of 2.43 m for ICESat-2 canopy heights using only strong beams acquired at night.

In general, the raw metrics of GEDI and ICESat-2 (i.e. RH95 and canopy height) provided less accurate data for the stand dominant height variable. Modeled with the CNN technique, the data quality could be improved clearly. Even though R^2 values remained somewhat the same, RMSE, %RMSE, and bias decreased to some extent. While the GEDI RH95 data yielded RMSE = 3.06 m, %RMSE = 21.80%, and bias = − 0.55 m compared to in situ height measurements, the CNN-based GEDI model estimated dominant heights with RMSE = 2.82 m, %RMSE = 20.12%, and bias = 0.01 m. Similarly, the error statistics for the ICESat-2 canopy height data were RMSE = 4.32 m, %RMSE = 30.76%, and bias = 2.00 m. After modeling, they decrease to RMSE = 2.77 m, %RMSE = 19.75%, and bias = 0.24 m. Thus, it can be suggested that CNN is a useful technique in modeling forests' biophysical structures based on active remote sensing data. It is also possible to state that the raw ICESat-2 data benefited more from the CNN modeling than the GEDI data in our case. It may be attributed to the potential differences in the distribution of GEDI, and ICESat-2 data points fell into various forest stands. Namely, the low number of data in certain stand types made our model development effort challenging, particularly during the division of training and test datasets in a statistically robust manner. Based on this experience, we recommend that future studies be conducted in data-rich areas to better train a CNN model. Aside from CNN, our results showed that the accuracy of the raw GEDI data was higher than that of the raw ICESat-2 data without modeling. Liu et al. (2021) also found that GEDI outperformed ICESat-2 in canopy height estimations in their study conducted in the forestland of the US.

The accuracy of dominant heights may change when the data are grouped based on certain stand parameters. The GEDI RH95 metric, for example, provided better results for fully-covered forest stands than sparsely-covered stands. The bias values differed nearly three times in the same direction (Table 5). This can be attributed to the fact that fully-covered forests demonstrate a smoother and more lateral top layer with continuous tree crowns (Fig. 5a). In the sparsely-covered forest, contrastingly, there are small or large gaps among tree crowns, resulting in significant height differences between the canopy and forest floor (Fig. 5b). Using ALS data for coniferous forests of Italy, Sibona et al. (2017) had a similar experience on this sense. Their data accuracy significantly decreased in shorter trees with irregular canopy cover. Thus, they concluded that height predictions with ALS data are more reliable for taller trees with a homogeneous canopy structure. While this is the case for GEDI and ALS data, the canopy cover class had almost no impact on the data accuracy of ICESat-2 in our case. This is an interesting research question that needs to be deeply examined by forest and remote sensing professionals with comparative studies conducted in vast forestlands. Nevertheless, a possible answer could be the difference between the laser altimetry techniques of the two missions. Although photon-counting

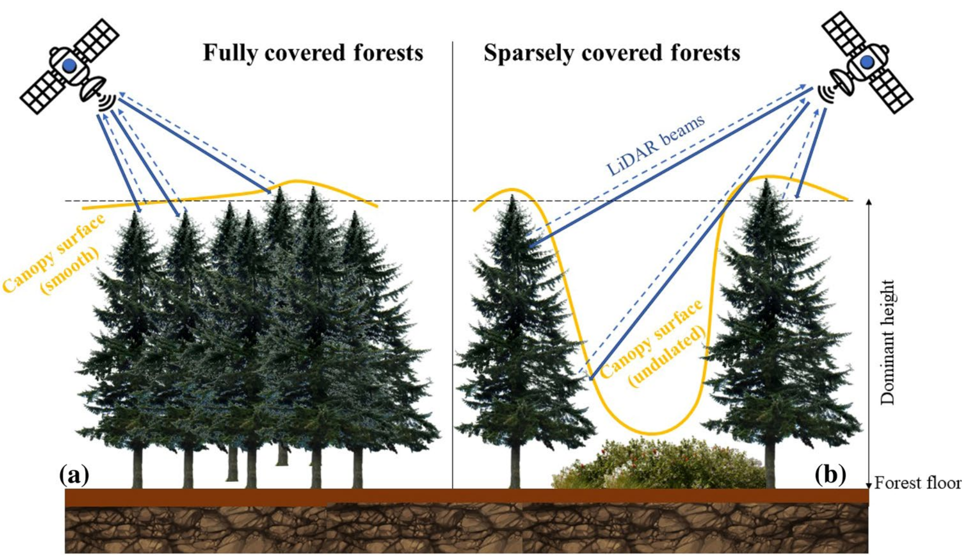

Fig. 5 Illustration of typical canopy structures for **a** fully- and **b** sparsely covered forests. The variation in height metrics (LiDAR beams) is generally smaller in fully covered forests owing to their compact canopy top surfaces. In contrast, it is generally greater in sparsely covered forests because some beams may directly hit low vegetation or ground (forest floor) due to large forest openings

(ICESat-2) and full-waveform (GEDI) are both state-of-the-art SLS techniques, they have their technical advantages and disadvantages depending on the application area. Therefore, Liu et al. (2021) state that it may be unfair to directly compare the terrain or canopy height values retrieved by GEDI and ICESat-2 for a limited area of investigation.

Regarding tree size classes, we observed no considerable differences in the data quality of GEDI (Fig. 3d). However, ICESat-2 yielded better results for thin stands (Table 5). The errors in the dominant heights generally started to increase with stand maturation. While RMSE was 2.99 m for young stands (DBH ≤ 20 cm), it increased to 4.76 m in mature stands with large-girth trees (20 cm < DBH < 52 cm). This is attributable to the larger variations in stand variables of mid-aged and mature forests of Crimean pine, as previously shown by Seki and Sakici (2017). Based on our experience from this limited study, the photon-counting laser altimetry technique of ICESat-2 seems more vulnerable to changes in forests' biophysical structure than the GEDI system. A possible reason for this might be the uninterrupted nature of the GEDI's full-waveform technique recording the total energy (Salas 2021).

As for the slope effect, the data quality of ICESat-2 decreased in forest stands located on steep slopes. The dominant heights of stands on flatter areas were retrieved by ICESat-2 raw data with an RMSE and bias of 3.81 m and 1.11 m, respectively. When the mean slope exceeded 35%, the same statistics increased to 4.87 m and 3.08 m, in the same order. To the best of our knowledge, no study focuses on the effect of slope on ICESat-2's data quality. For GEDI, however, Fayad et al. (2021a) investigated data quality based on different slope ranges in Brazilian plantation forests. They concluded that terrain slope affected volume estimations more than canopy height estimates. While the dominant heights of *Eucalyptus* stands were estimated with an RMSE of 2.06 m on flatter areas (slope < 10%), it increased to 2.11 m on slopes ranging between 10 and 20%. When it exceeded 20%, the RMSE value reached 3.26 m in their case. In the present case, GEDI data provided slightly better results in flatter areas; nonetheless, it was not as pronounced as in ICESat-2 data.

The present study also has several drawbacks; for instance, no filtering was applied to the raw metrics obtained from the satellite data. In other words, we did not separate the datasets based on strong/weak photons and day/night acquisition. Duncanson et al. (2020) investigated which photons are better for biomass estimation using simulated GEDI and ICESat-2 data. They indicated that daytime data of GEDI resulted in a higher error, and the low photon rate of ICESat-2 was more sensitive to uncertainty in biomass estimation. Although the filtered data yields better estimation results statistically, the number of points per area decreases and this situation limits us from filtering. Instead, we examined whether state-of-the-art modeling techniques could increase the accuracy without reducing the number of data points. As seen in Fig. 3, even though we applied no filter on LiDAR data, there were less than 30 data points in certain stand types, which could decrease the statistical quality of the accuracy assessment for some classes. For example, N was 25 for GEDI points fell into the limited mixed stands (Fig. 3e) because the forest complex was dominated by pure stands of Crimean pine. This is another drawback of the present study that can be resolved by future studies conducted in larger and multi-species forest complexes.

As stated earlier, measuring the heights of standing trees in the forest is generally harder than measuring other inventory parameters (e.g. DBH, basal area, and stand density) and it is typically error-prone due to many factors affecting the measurement (Sibona et al. 2017; Ganz et al. 2019; Hyyppa et al. 2020a; Özkal et al. 2021). So-called h–d models may also be erroneous when used improperly (Persson and Stahl 2020; Bolat et al. 2022; Seki and Sakici 2022a; 2022b). Since there was a five-year difference between the dates of field and satellite data, we could not directly use on-the-ground measurements of sample plots. Instead, we simulated ground-measured dominant heights based on yield curves developed by Kalıpsız (1963) for the Crimean pine forests of Turkey. By doing so, we may have introduced additional errors to our reference data due to some assumptions made (see Table 2). Bettinger et al. (2009) mentioned that accepting some simplifying assumptions and uncertainty is inevitable when using growth & yield models because forest dynamics are so complex and natural disturbances may often be unpredictable. The assumptions make the complex problems more tractable and resolvable; however, the uncertainty inherent to models needs to be evaluated carefully relying on biological realism in forestry (Ercanli et al. 2022). In fact, we could use ground measurements just as they are but Crimean pine is a dynamic forest tree species, particularly in its early successional stages. The yield curves of Kalıpsız (1963) show that the five-year height increment of this species could be as high as 2 m (~ 25%) between the ages of 10 and 40, particularly in good sites. Therefore, an accuracy assessment without considering forest dynamics would be more erroneous for LiDAR-derived dominant heights. Ideally, researchers should minimize the temporal difference between ground measurements and satellite data acquisition dates as much as possible.

Finally, it should be noted that naturally regenerated stands of Crimean pine mostly cover our case study areas. Unlike plantation forests, (near-) natural forests often demonstrate a heterogeneous structure regarding tree height diversity and spatial patterns (von Gadow et al. 2012). Besides, the dense understory layer on the forest floor includes a lot of shrubs, irregular trees, and deadwood in even-aged stands of Turkey (GDF 2015a, b) due to lacking vegetation control practices, such as prescribed fire and herbicide treatment, even in intensively managed forests. These elements, along with the broken topography of the study area, might have slightly decreased the estimation accuracies reported in this work.

Conclusion

The objective of the present study was to quantify dominant stand heights at the forest landscape level using products from cutting-edge spaceborne laser altimeters. We assessed the GEDI RH95 and ICESat-2 canopy height metrics in terms of data accuracy through pairwise comparison against in situ height measurements for Crimean pine forests in Turkey. The potential effects of forest structure on data accuracy were also investigated depending on certain stand characteristics, including canopy cover, development stage, species types, and terrain slope. Furthermore, an advanced deep learning technique (i.e. CNN) was employed to develop independent models for GEDI and ICESat-2 missions in order to estimate dominant heights more accurately.

Based on the study's findings, the following conclusions can be drawn: (i) The dominant height data retrieved by GEDI is more accurate (RMSE = 3.06 m) than ICESat-2 (RMSE = 4.02 m) for near-natural forest stands of Crimean pine in Turkey, (ii) in general, GEDI underestimates the stand dominant height parameter (bias = − 0.55 m), while ICESat-2 overestimates it (bias = 1.99 m), (iii) the data accuracy decreases with increasing slope gradient, particularly in the ICESat-2 mission, (iv) while GEDI retrievals perform better in fully-covered stands, ICESat-2 outperforms GEDI in thin and younger stands, (v) the performance of the height retrievals can be enhanced using the CNN modeling technique (e.g. RMSEs decreased to 2.82 m (20.12%) and 2.77 m (19.75%) both for GEDI and ICESat-2), (vi) despite the fact that GEDI and ICESat-2 are useful for acquiring forest height data, their overall accuracies (%RMSEs = ~20%) are still insufficient for tree- and stand-level forest inventories since the permissible error for tree height measurements is maximum 15% (Qiu et al. 2018).

Nevertheless, the data accuracy of GEDI and ICESat-2 may be sufficient for operational forest management and planning in more homogeneous forests than in the current case. Therefore, the focus of future research should be shifted from natural and structurally complex highland forests toward industrial plantations, intensively managed forests, and human-modified woods located in flat lowlands.

Acknowledgements Forest management plan data were purchased from the Turkish General Directorate of Forest (GDF) in official ways. We thank the GDF for the data provision.

Author contribution statement All authors contributed to the study's conception and design. OGN and CV performed data processing, analysis, and mapping. OGN developed convolutional neural network models. CV wrote the first draft of the manuscript and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Data availability The datasets generated and analyzed during the current study are available from the corresponding author upon reasonable request.

Declarations

Competing interests The authors have no relevant financial or non-financial interests to disclose.

References

Allaire J, Chollet F (2022) keras: R Interface to 'Keras'. R package version 2.9.0, https://CRAN.R-project.org/package=keras.

Baskent EZ (2020) A framework for characterizing and regulating ecosystem services in a management planning context. Forests 11(1):102. https://doi.org/10.3390/f11010102

Beland M, Parker G, Sparrow B, Harding D, Chasmer L, Phinn S, Antonarakis A, Strahler A (2019) On promoting the use of lidar systems in forest ecosystem research. For Ecol Manage 450:117484. https://doi.org/10.1016/j.foreco.2019.117484

Bettinger P, Boston K, Siry JP, Grebner DL (2009) Forest management and planning. Academic press, USA

Biber P, Felton A, Nieuwenhuis M, Lindbladh M, Black K, Bahýl' J, et al (2020) Forest biodiversity, carbon sequestration, and wood production: modeling synergies and trade-offs for ten forest landscapes across Europe. Front Ecol Evol 8:291. https://doi.org/10.3389/fevo.2020.547696

Bolat F, Ürker O, Günlü A (2022) Nonlinear height-diameter models for Hungarian oak (Quercus frainetto Ten) in Dumanlı Forest Planning Unit Çanakkale/Turkey. . Aust J Forest Sci 139(3):199–220. https://doi.org/10.2154/aust.5487

Cabo C, Del Pozo S, Rodriguez-Gonzalvez P, Ordonez C, Gonzalez-Aguilera D (2018) Comparing terrestrial laser scanning (TLS) and wearable laser scanning (WLS) for individual tree modeling at plot level. Remote Sens 10:540–556. https://doi.org/10.3390/rs10040540

Colkesen I, Kavzoglu T, Atesoglu A, Tonbul H, Ozturk MY (2022) Multi-seasonal evaluation of hybrid poplar (P. Deltoides) plantations using Worldview-3 imagery and state-of-the-art ensemble learning algorithms. Adv Space Res. https://doi.org/10.1016/j.asr.2022.10.044. ((**in press**))

Coops NC, Tompalski P, Goodbody TRH et al (2021) Modelling lidar-derived estimates of forest attributes over space and time: a review of approaches and future trends. Remote Sens Environ 260:112477. https://doi.org/10.1016/j.rse.2021.112477

Dorado-Roda I, Pascual A, Godinho S et al (2021) Assessing the accuracy of GEDI data for canopy height and aboveground biomass estimates in Mediterranean forests. Remote Sens 13(12):2279. https://doi.org/10.3390/rs13122279

Dubayah R, Blair JB, Goetz S et al (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

Duncanson L, Neuenschwander AL, Hancock S et al (2020) Biomass estimation from simulated GEDI, ICESat-2 and NISAR across environmental gradients in Sonoma County. Calif Remote Sensing Environ 242:111779. https://doi.org/10.1016/j.rse.2020.111779

Eler Ü (1986) Yield studies for Crimean juniper in Turkey. Technical Bulletin, No 192, Forestry Research Institute Press, Ankara (in Turkish with English abstract).

Eraslan İ (1954) Research studies for the management of oak forests in the Thrace. General Directorate of Forest Publishing, Ankara (in Turkish)

Ercanli İ, Muammer Şenyurt, Bolat F (2022) A major challenge to machine learning models: Compatible predictions with biological realism in forestry: A case study of individual tree volume. Proceeding at the 3rd International Conference on Environment and Forest Conservation, Feb. 21–23, Kastamonu, Turkey.

ESRI (2012) ArcGIS 10.2 for Desktop. Environmental Systems Research Institute Inc, California

Evcimen BS (1963) Yield and management studies for the cedar forests in Turkey. General Directorate of Forest Publishing, Ankara (in Turkish)

Fayad I, Baghdadi N, Alvares CA et al (2021a) Terrain slope effect on forest height and wood volume estimation from GEDI data. Remote Sens 13:2136. https://doi.org/10.3390/rs13112136

Fayad I, Ienco D, Baghdadi N, Gaetano R, Alvares CA, Stape JL, Scolforo HF, Le Marire G (2021b) A CNN-based approach for the estimation of canopy heights and wood volume from GEDI waveforms. Remote Sens Environ 265:112652. https://doi.org/10.1016/j.rse.2021.112652

Ferreira MP, Lotte RG, D'Elia FV, Stamatopoulos C, Kim D-H, Benjamin AR (2021) Accurate mapping of Brazil nut trees (*Bertholletia excelsa*) in Amazonian forests using WorldView-3 satellite images and convolutional neural networks. Eco Inform 63:101302. https://doi.org/10.1016/j.ecoinf.2021.101302

Ganz S, Käber Y, Adler P (2019) Measuring tree height with remote sensing—a comparison of photogrammetric and LiDAR data with different field measurements. Forests 10(8):694. https://doi.org/10.3390/f10080694

GDF (2015a) Ecosystem-based multifunctional forest management plan of Hocalar Planning Unit for the period 2015–2034. General Directorate of Forestry Publishing, Ankara (in Turkish)

GDF (2015b) Ecosystem-based multifunctional forest management plan of Sinanpaşa Planning Unit for the period 2015–2034. General Directorate of Forestry Publishing, Ankara (in Turkish)

GDF (2020) 2019 Turkey Report of Sustainable Forest Management Criteria & Indicators. General Directorate of Forestry Publishing, Ankara (in Turkish)

Helmer EH, Lefsky MA, Roberts DA (2009) Biomass accumulation rates of amazonian secondary forest and biomass of old-growth forests from landsat time series and the geoscience laser altimeter system. J Appl Remote Sens 3(1):033505. https://doi.org/10.1117/1.3082116

Hyyppa E, Yu X, Kaartinen H et al (2020) Comparison of backpack, handheld, under-canopy UAV, and above-canopy UAV laser scanning for field reference data collection in boreal forests. Remote Sens 12:3327. https://doi.org/10.3390/rs12203327

Hyyppa J, Yu X, Hakala T et al (2021) Under-canopy UAV laser scanning providing canopy height and stem volume accurately. Forests 12(7):856. https://doi.org/10.3390/f12070856

Hyyppä E, Kukko A, Kaijaluoto R, White JC, Wulder MA, Pyörälä J, Liang X, Yu X, Wang Y, Kaartinen H, Virtanen JP, Hyyppä J (2020) Accurate derivation of stem curve and volume using backpack mobile laser scanning ISPRS. J Photogramm Remote Sens 161:246–262. https://doi.org/10.1016/j.isprsjprs.2020.01.018

Kalıpsız A (1963) Studies on the natural structure and productivity of Crimean pine stands (*Pinus nigra Arnold*) in Turkey. General Directorate of Forest Publishing, İstanbul (in Turkish)

Kanja K, Karahalil U, Çil B (2019) Modeling stand parameters for Pinus brutia (Ten.) using airborne LiDAR data: a case study in Bergama. J App Remote Sens 14(2):022205. https://doi.org/10.1117/1.JRS.14.022205

Khalsa SJS, Borsa A, Nandigam V et al (2020) OpenAltimetry-rapid analysis and visualization of Spaceborne altimeter data. Earth Sci Inf. https://doi.org/10.1007/s12145-020-00520-2

Kuhn M (2022) caret: Classification and Regression Training. R package version 6.0–93 https://CRAN.R-project.org/package=caret.

Laar AV, Akça A (2007) Forest mensuration. Springer, Dordrecht

Lefsky MA, Harding DJ, Keller M et al (2005) Estimates of forest canopy height and aboveground biomass using ICESat. Geophys Res Lett 32:L22S02. https://doi.org/10.1029/2005GL023971

Li Y, Zhang H, Xue X, Jiang Y, Shen Q (2018) Deep learning for remote sensing image classification: a survey. Wi Data Min Knowl Discovery 8(6):e1264. https://doi.org/10.1002/widm.1264

Liang X, Wang Y, Pyorala J et al (2019) Forest in situ observations using unmanned aerial vehicle as an alternative of terrestrial measurements. Forest Ecosyst 6:20. https://doi.org/10.1186/s40663-019-0173-3

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

Liu X, Su Y, Hu T et al (2022) Neural network guided interpolation for mapping canopy height of China's forests by integrating GEDI and ICESat-2 data. Remote Sens Environ 269:112844. https://doi.org/10.1016/j.rse.2021.112844

Markus T, Neumann T, Martino A et al (2017) The ice, cloud, and land elevation satellite-2 (icesat-2): science requirements, concept, and implementation. Remote Sens Environ 190:260–273. https://doi.org/10.1016/j.rse.2016.12.029

Nandy S, Srinet R, Padalia H (2021) Mapping forest height and aboveground biomass by integrating icesat-2 sentinel-1 and sentinel-2 data using random forest algorithm in northwest himalayan foothills of india. Geophys Res Lett. https://doi.org/10.1029/2021GL093799

Narin OG, Vatandaşlar C, Abdikan S (2022) Estimating stand top height using freely distributed ICESat-2 LiDAR data: a case study from multi-species forests in Artvin. Forestist 72(3):294–298. https://doi.org/10.5152/forestist.2022.21044

Neuenschwander AL, Magruder LA (2019) Canopy and terrain height retrievals with ICESat-2: a first look. Remote Sensg 11(14):1721. https://doi.org/10.3390/rs11141721

Neuenschwander AL, Pitts KL, Jelley BP et al (2020) ATLAS/ICESat-2 L3A land and vegetation height, Version 3 [Data Set]. NASA National Snow and Ice Data Center Distributed Active Archive Center, Boulder, Colorado USA

Özkal MK, Atar D, Aydın M, Tunç F (2021) Comparison of traditional regression models and artificial neural network models for height-diameter modeling in uneven-aged fir stands. J Biometry Stud. https://doi.org/10.29329/JofBS.2021.348.01

Persson HJ, Stahl G (2020) Characterizing uncertainty in forest remote sensing studies. Remote Sens 12:505. https://doi.org/10.3390/rs12030505

Potapov P, Li X, Hernandez-Serna A, Tyukavina A et al (2021) Mapping global forest canopy height through integration of GEDI and

Landsat data. Remote Sens Environ 253:112165. https://doi.org/10.3390/rs12030505
Qiu Z, Feng Z, Jiang J, Lin Y, Xue S (2018) Application of a continuous terrestrial photogrammetric measurement system for plot monitoring in the Beijing Songshan national nature reserve. Remote Sens 10(7):1080. https://doi.org/10.3390/rs10071080
R Core Team (2020) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. https://www.R-project.org/
Salas EAL (2021) Waveform LiDAR concepts and applications for potential vegetation phenology monitoring and modeling: a comprehensive review. Geo-Spatial Info Sci 24(2):179–200
Sefercik UG, Ateşoğlu A, Atalay C (2021) Airborne laser scanning performance analysis in the forest stand map production. J Geomat 6(3):179–188
Seki M, Sakıcı OE (2017) Dominant height growth and dynamic site index models for Crimean pine in the Kastamonu-Taşköprü region of Turkey. Can J for Res 47(11):1441–1449. https://doi.org/10.1139/cjfr-2017-0131
Seki M, Sakici OE (2022a) Ecoregion-based height-diameter models for crimean pine. J for Res 27(1):36–44. https://doi.org/10.1080/13416979.2021.1972511
Seki M, Sakici OE (2022b) Ecoregional variation of Crimean pine (Pinus nigra subspecies pallasiana [Lamb.] Holmboe) stand growth. For Sci. https://doi.org/10.1093/forsci/fxac030(inpress)
Shah SAA, Manzoor MA, Bais A (2020) Canopy height estimation at landsat resolution using convolutional neural networks. Mach Learn Knowl Extract 2(1):23–36. https://doi.org/10.3390/make2010003
Sibona E, Vitali A, Meloni F, Caffo L, Dotta A, Lingua E, Motta R, Garbarino M (2017) Direct measurement of tree height provides different results on the assessment of LiDAR accuracy. Forests 8(1):7. https://doi.org/10.3390/f8010007
Silva CA, Hamamura C, Valbuena R, et al. (2020) rGEDI: NASA's Global Ecosystem Dynamics Investigation (GEDI) data visualization and processing. Version 0.1.9. https://CRAN.R-project.org/package=rGEDI
Sun T, Qi J, Huang H (2020) Discovering forest height changes based on spaceborne lidar data of ICESat-1 in 2005 and ICESat-2 in 2019: a case study in the Beijing-Tianjin-Hebei region of China. ForEcosyst. https://doi.org/10.1186/s40663-020-00265-w
TMS (2015) Climate records from the Afyonkarahisar meteorological station (1994–2015). Turkish Meteorology Service, Ankara
Türkeş M (2010) Climatology and Meteorology (1st edt). Kriter Publishing, İstanbul (in Turkish)
URL1 (2022) https://www.datatechnotes.com/2020/01/how-to-fit-regression-data-with-cnn.html Accessed 16 July 2022
Vatandaşlar C, Zeybek M (2020) Application of handheld laser scanning technology for forest inventory purposes in the NE Turkey. Turk J Agric for 44:229–242. https://doi.org/10.3906/tar-1903-40
Vatandaşlar C, Zeybek M (2021) Extraction of forest inventory parameters using handheld mobile laser scanning: a case study from Trabzon. Turkey Measurement 177:109328. https://doi.org/10.1016/j.measurement.2021.109328
Vatandaşlar C, Zeybek M, Çankaya EÇ et al (2022) Using handheld mobile LiDAR technology in forest inventories: Artvin-Şavşat case. Turk J for Res 9(1):81–96
Vidal C, Alberdi I, Hernández L, Redmond JJ (2016) National forest inventories: assessment of wood availability and use. Springer, Switzerland. https://doi.org/10.1007/978-3-319-44015-6_1
von Gadow K, Zhang CY, Wehenkel C et al (2012) Forest structure and diversity. In: Pukkala T, von Gadow K (eds) Continuous cover forestry. Springer, Dordrecht, pp 29–83
Wang Y, Fang H (2020) Estimation of LAI with the LiDAR technology: a review. Remote Sens 12(20):3457. https://doi.org/10.3390/rs12203457