Remote Sensing of Environment 279 (2022) 113112

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Effects of environmental conditions on ICESat-2 terrain and canopy heights retrievals in Central European mountains

Vítězslav Moudrý [a,b,c,*], Kateřina Gdulová [a], Lukáš Gábor [a], Eliška Šárovcová [a], Vojtěch Barták [a], François Leroy [a], Olga Špatenková [a], Duccio Rocchini [a,d], Jiří Prošek [a,c]

[a] Department of Spatial Sciences, Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Kamýcká 129, Praha-Suchdol, 165 00, Czech Republic
[b] Institute for Environmental Studies, Faculty of Science, Charles University, Benátská 2, 12801 Prague 2, Czech Republic
[c] Institute of Botany of the Czech Academy of Sciences, Zámek 1, CZ-252 43 Průhonice, Czech Republic
[d] BIOME Lab, Department of Biological, Geological and Environmental Sciences, Alma Mater Studiorum University of Bologna, Bologna, Italy

ARTICLE INFO

Edited by Jing M. Chen

Keywords:
ATL08
Evaluation
Filtering
Forest
Ground
Laser altimeter
Radiometry
Scattering

ABSTRACT

The ICESat-2 ATL08 land and vegetation product includes several flags that can be used for the assessment of LiDAR-environment interactions and can help select data of the highest quality. However, the usability of these flags has not been sufficiently studied to date. Here, we aimed to evaluate the effects of atmospheric scattering, the presence of snow, canopy cover, terrain slope, beam strength, and solar angle on the accuracy of terrain and canopy height of the ATL08 product as well as on providing recommendations on how to filter data in order to minimize errors. We evaluated the vertical accuracy of ATL08 terrain and canopy height in European mountains by comparing them with the digital terrain model and canopy height model derived from airborne laser scanning data. Our results indicate that the assessment of atmospheric effects using the cloud confidence flag (*cloud_flag_atm*; i.e. number of cloud layers) is better than the previously used multiple scattering warning flag (*msw_flag*). Day acquisitions with more than one layer of clouds yielded a terrain elevation RMSE of 3.22 m in forests while night acquisitions with no more than a single layer of clouds resulted in RMSE of 1.73 m. The increasing atmospheric scattering effects increased the photons' path length, resulting in terrain height underestimation. The presence of snow had a strong positive effect on the number of identified ground photons, independently of the canopy cover, but resulted in an overestimation of terrain height in higher altitudes. Accordingly, the presence of snow cover resulted in a significant underestimation of canopy height in forests. The canopy height in broadleaf/mixed as well as coniferous forests was in summer underestimated on average by 2.1 m (%ME of −15.3%) and 1.2 m (%ME of −8.2%), respectively; in winter, however, the underestimation increased to 8.5 m (%ME of −56.8%) and 5.7 m (%ME of −38.3%), respectively. Canopy height estimates had better accuracy for the strong beam (RMSE of 5.09 m; %RMSE of 35.4%) than for the weak beam (RMSE of 7.03 m; %RMSE of 51.3%). Our results show that the ATL08 terrain height accuracy decreases with uneven distribution of signal photons within individual segments and further deteriorates with increasing terrain slope. Filtering out segments with poor distribution of photons, more than one layer of clouds during the day, and snow cover in high altitudes is the best approach for minimizing the error while maximizing the number of segments left for subsequent analysis.

1. Introduction

Forests cover >4.1 billion hectares of the Earth's surface and store considerable amounts of carbon, acting as an important global carbon sink (Pan et al., 2011). Forest carbon stock is an essential component of climate action plans increasingly made by many states to implement the Paris Agreement on Climate Change and the 2030 Agenda for Sustainable Development (Hein et al., 2018). However, estimates of forest carbon stocks and rates of change remain uncertain due to data limitations and availability (Goetz and Dubayah, 2011; Pugh et al., 2019). Field inventory campaigns are labor-intensive and airborne laser scanning (ALS) surveys are too expensive; in effect, both provide estimates

* Corresponding author at: Department of Spatial Sciences, Faculty of Environmental Sciences, Czech University of Life Sciences Prague, Kamýcká 129, 16500 Praha-Suchdol, Czech Republic.

E-mail address: moudry@fzp.czu.cz (V. Moudrý).

https://doi.org/10.1016/j.rse.2022.113112
Received 23 December 2021; Received in revised form 6 May 2022; Accepted 4 June 2022
Available online 17 June 2022
0034-4257/© 2022 Elsevier Inc. All rights reserved.

from limited samples in space and time (although many developed countries provide ALS point clouds free of charge, the renewal frequency is usually low). On the other hand, space-based remote sensing methods have a great potential to map forests on a global scale (Herold et al., 2019; Marselis et al., 2022; Mulverhill et al., 2022).

NASA's Ice, Cloud, and land Elevation Satellite-2 (ICESat-2) is a space-based laser altimetry mission launched in September 2018 (Markus et al., 2017). The ICESat-2 mission is primarily designed for mapping changes in Earth's cryosphere (Brunt et al., 2019; Farrell et al., 2020), but it also provides data for other geophysical applications such as mapping forest height and above-ground biomass (Nandy et al., 2021). Onboard, the ICESat-2 satellite carries the Advanced Topographic Laser Altimeter System (ATLAS), a single-photon counting laser altimeter that allows the determination of the time of flight of individual photons (Degnan, 2002; Neumann et al., 2019). The ATLAS instrument splits the output laser pulse into three pairs of beams, each of which consists of a weak and a strong beam with an energy ratio of approximately 1:4 (Neumann et al., 2019). The single-photon approach differs from the conventional discrete laser ranging systems that require the acquisition of thousands of backscattered photons to trigger a detection (Harding et al., 2011). It also differs from full-waveform systems such as GLAS (Geoscience Laser Altimeter System), ATLAS's predecessor on board the first ICESat satellite (2003–2009; Schutz et al., 2005). GLAS produced large footprints (diameter of 65 m) with poor vertical resolution causing problems when disentangling terrain and vegetation signal (i.e. forest canopy height) in mountain environment (Chen, 2010; Hilbert and Schmullius, 2012; Bolton et al., 2013). ICESat-2 allows the detection of both terrain and canopy surface even in high relief areas and can considerably improve our ability to monitor global forest biomass (Nandy et al., 2021).

ICESat-2 mission generates geophysical products over various types of surface, one of which is the land and vegetation product (ATL08). Several studies have already contributed to the validation of ATL08 accuracy. In general, studies reported higher accuracy for terrain than canopy heights. In their evaluation of ATL08 in Finland, Neuenschwander and Magruder (2019) and Neuenschwander et al. (2020) determined the terrain accuracy – in terms of root mean square error (RMSE) – to be <1 m and that of the canopy height to be approximately 3 m. Slightly worse estimates were found for the United States by Liu et al. (2021), with the RMSE of the terrain height of about 2 m and that of canopy height of approximately 7 m. In their validation over several ecozones in the United States, Malambo and Popescu (2021) reported the mean absolute deviations (MAE) between terrain elevations and canopy heights acquired from ATL08 and by ALS to be 1.2 m and 3.5 m, respectively.

The ATL08 accuracy is positively affected by the number of retrieved ground photons as the higher density of photons reflected from a surface facilitates the correct identification of signal photons. A laser beam propagating through the atmosphere is affected by atmospheric layers such as clouds that can cause multiple scattering, thus increasing photon path length (Winker, 2003). Furthermore, the presence of clouds affects the number of photon returns from the surface; in addition, returns from low clouds may be misclassified as ground photons (Smith et al., 2019). Consequently, the precision of ground detected under several layers of clouds may be limited and render ATL08 data unsuitable for terrain and vegetation characterization. Therefore, segments acquired under adverse atmospheric conditions are typically filtered out at the beginning of data processing (Neuenschwander et al., 2020; Queinnec et al., 2021).

Besides the signal photons, ICESat-2 geolocated photon data (ATL03) also contain solar background photons (i.e., noise photons) resulting from sunlight reflected off the Earth's surface (Swatantran et al., 2016; Neumann et al., 2019). It is, therefore, necessary to perform quite extensive processing to filter out the background noise before the production of any higher-level ICESat-2 products. Indeed, it has been shown that night acquisitions (with a lower amount of solar background noise photons) have RMSE of terrain height approximately 0.15 m lower than daytime acquisitions (Liu et al., 2021). The laser pulse energy level has similar effects. The strong beam results in more detected signal photons and hence simplifies the solar background noise filtering (Neuenschwander et al., 2020); however, it has been shown previously that weak beams may provide data of equivalent quality (Malambo and Popescu, 2021).

Furthermore, the accuracy of the retrieved canopy and terrain height may be affected by the presence of snow-cover. The ATLAS laser energy level is set to detect approximately 10 photons per strong beam shot on a snow-covered surface (Neuenschwander et al., 2020). However, for land and vegetation, the number of reflected photons is considerably lower due to the lower surface reflectance. Neuenschwander et al. (2020) reported detection of approximately one photon per shot from terrain for a strong beam under no-snow conditions. Therefore, the presence of snow can considerably improve the number of detected ground photons and, hence, the accuracy of terrain height estimates compared to snow-free segments (Neuenschwander et al., 2020). On the other hand, the presence of snow cover hinders the detection of the actual surface. Depending on its thickness, the snow cover may result in an overestimation of the terrain height in the ATL08 product (Deems et al., 2013). Furthermore, the ATL08 accuracy is also considerably affected by the terrain characteristics (Tian and Shan, 2021; Liu et al., 2021), particularly in combination with canopy cover characteristics (Malambo and Popescu, 2021). Dense and tall canopies reduce the number of photons that reach the ground and are reflected back to the satellite and, consequently, complicate the terrain height retrieval (e.g. Liu et al., 2021). Conversely, sparse canopy poses a challenge for the estimates of vegetation height because relatively few photons are reflected from sparse vegetation and, as a consequence, are difficult to distinguish from the solar background noise (Neuenschwander and Pitts, 2019).

Here, we use ATL08 flags derived from auxiliary data that describe the conditions under which the data were acquired and terrain and vegetation characteristics derived form ALS and Corine land cover data to study the influence of the (i) atmospheric conditions, (ii) snow cover, (iii) the distribution of detected photons along the ground track, (iv) solar background noise, (v) laser pulse energy level, (vi) canopy cover, and (vii) terrain slope, on the accuracy of the ATL08 (version 4) terrain and canopy height retrievals in the mountain environment. In addition, we evaluated the effects of some of these characteristics on the number of detected photons. Besides being useful for the assessment of LiDAR-environment interactions, these flags are expected to provide information on data usability and have great potential to facilitate data filtering. Therefore, we also aim to provide users with the best approach for the identification of problematic measurements. This should, in turn, allow the accurate detection and improve the selection of high accuracy data necessary for the generation of higher-level products such as ATL18 gridded ground surface height, canopy height, and canopy cover estimates.

2. Data and methods

2.1. Study area

The study area covers about 4500 km^2 and consists of three Central European mountain ranges and their surrounding areas situated in Germany, Czech Republic, and Poland (Fig. 1): the Ore mountains, the Giant mountains, and the Bohemian Forest. The altitudes range between 300 and 1600 m a.m.s.l. and the terrain slopes are highly variable (0°-50°). The tree line traverses the altitudinal range of 1200–1350 m a.m.s.l. and the vegetation consists of croplands and pastures, natural grasslands, spruce monocultures, and remnants of original deciduous and mixed mountain forests.

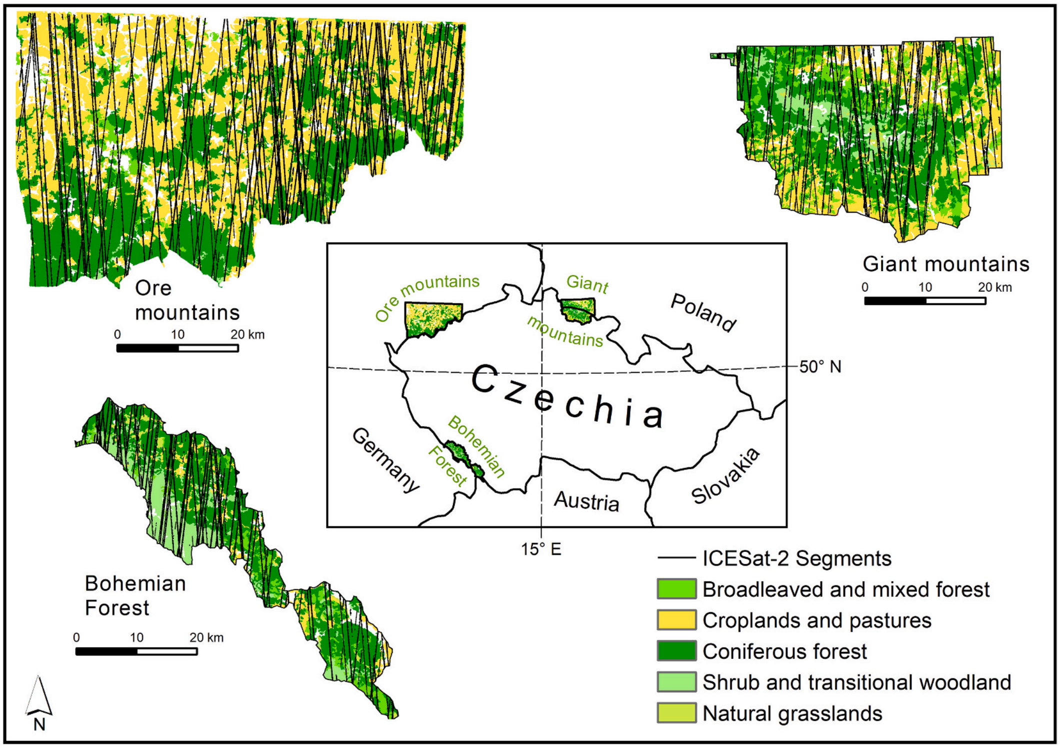

Fig. 1. Land cover derived from pan-European Corine Landcover 2018 data (Feranec et al., 2016) and location of the study areas: Ore mountains, Giant mountains, Bohemian Forest.

2.2. Reference airborne laser scanning data

The airborne LiDAR data for the study area were collected between 2012 and 2017. The data in all study areas were acquired during the leaf-on period. The LiDAR point cloud density is between 4 and 55 points per m^2 (Table 1). For the complete description of airborne laser scanning reference data processing, see Gdulová et al. (2021). The classified point clouds were processed in LAStools (version 210720). Digital terrain models (DTM) and canopy height models (CHM) for the Giant mountains and the Bohemian Forest were generated at a 1 m resolution, and the DTM and CHM for the Ore mountains were provided by "Staatsbetrieb Geobasisinformation und Vermessung Sachsen (GeoSN)" at a 2 m resolution. In order to match the horizontal datum, we used the bilinear resampling method to transform the reference ALS data into WGS84 UTM33N.

2.3. ICESat-2 ATL08 data

ICESat-2 operates in a 91-day exact repeat orbit, with 1387 orbits per cycle. The ATLAS instrument onboard the ICESat-2 satellite is a single-photon counting laser altimeter that uses green laser beams (wavelength of 532 nm) to measure the distance to the surface (Markus et al., 2017). The benefit of using a single-photon approach is that the laser can operate at a much higher repetition rate. In the case of ATLAS, the pulse repetition frequency is 10 kHz, which allows along-track sampling at 0.7 m intervals, with overlapping footprints of ~12 m. The ATLAS instrument splits the output laser pulse into three pairs of beams that are arranged to produce ground tracks with the distance of 3.3 km between the pairs and of 90 m between the tracks of each pair. The pairs consist of one weak and one strong beam with an energy ratio of approximately 1:4 (Neumann et al., 2019). Theoretically, ATLAS can detect up to 16 photons (4 × 4 detector array) per outgoing shot (a negligible fraction of all emitted photons; Markus et al., 2017; Neumann et al., 2019).

Here, we evaluated the accuracy of the terrain and canopy heights that are provided as a part of the land and vegetation height product (ATL08; Neuenschwander et al., 2021). ATL08 is derived from the global geolocated photon data product (ATL03) that serves as a single source of all photon data and ancillary information needed by higher-level products. The critical step for the development of the ATL08 product lies in the correct detection of the signal photons (i.e., photons emitted by the instrument) and filtering out the solar background noise (i.e., the photons originating from sunlight reflected off the Earth's surface; Neumann et al., 2019; Neuenschwander et al., 2021).

The filtering utilizes two signal-finding methods (i.e., histogramming approach and Differential, Regressive, and Gaussian Adaptive Nearest Neighbor) that aim to discriminate signal photons from the noise photons originating from the solar background. Subsequently, a series of iterative filters of the signal photons is adopted to capture the ground and top of canopy surface. The individual photons are then classified as ground, canopy, or noise based on their distance from the estimated canopy and ground surfaces and metrics for canopy and terrain surfaces are then provided at a segments of 100 m × 12 m (along the ground track

Table 1
LiDAR data characteristics.

Study area	Area (km^2)	Year	Reference system (EPSG)	Point cloud density
Bohemian Forest	680	2017	31468	55 p/m^2
Ore mountains	2640	2015–2017	25833	4 p/m^2
Giant mountains (Czechia)	480	2012	5514	5 p/m^2
Giant mountains (Poland)	720	2012	2180	4 p/m^2

× cross the ground track). The segment size was chosen so that a sufficient number of photons reflecting from both the terrain and canopy surfaces were available for terrain and canopy metrics estimation (Popescu et al., 2018; Neuenschwander and Pitts, 2019; Neuenschwander et al., 2021).

The terrain metrics within each segment include mean (*h_te_mean*) and median (*h_te_median*) heights of all ground photons within a segment, and also two estimates of terrain height at the location of the mid-point of a segment (*h_te_interp* and *h_te_bestfit*). The canopy metrics within each segment include minimum, mean, maximum height, and several percentile heights. Previous studies mostly concentrated on evaluating the relative canopy height at the 98th percentile (Neuenschwander et al., 2020; Liu et al., 2021); mean vegetation height, however, can provide important supplementary information characterizing vegetation structure (Chen, 2010; Malambo and Popescu, 2021). Therefore, in this study, we concentrated on the mean values (*h_te_mean* and *h_mean_canopy*).

We downloaded ATL08 version 4 containing data acquired between October 2018 and September 2021 using the Earthdata portal (accessed September 2021). The ICESat-2 mean terrain height is given as ellipsoidal height while the vertical datum of ALS data is given as normal height. Therefore, to match the heights of ICESat-2 and the reference ALS data, we used a quasigeoid of Czechia. We first interpolated the height of the quasigeoid of Czechia and surrounding areas from 53,550 positions (latitude, longitude) at a resolution of 1′ x 1.5′, and, subsequently, subtracted it from the ICESat-2 mean terrain height. Note that we used canopy heights represented as relative heights above the terrain and, thus, there was no need to match their vertical datum. In order to match the horizontal datum with ALS data, we projected the centroids of ICESat-2 ATL08 segments into WGS84 UTM33N. We did not apply any horizontal geolocation offset corrections as we are interested in the accuracy from the perspective of users who typically do not have necessary data (e.g. high-resolution ALS-DTM) to assess horizontal offsets and perform the corrections. Besides, studies that performed such assessment reported that horizontal geolocation offset tended to be relatively low (between 0 and 3 m, which is much lower than the mission requirement of 6.5 m; Neuenschwander et al., 2020; Malambo and Popescu, 2021).

The ATL08 product includes several flags that describe target (e.g. surface reflectance) and acquisition (e.g. atmospheric scattering) characteristics and inform the user on segment data quality and usability. We combined these flags with external data sources describing the target characteristics (e.g. landcover) to identify the usability of these flags as well as the causes of bias in terrain and canopy height retrievals.

2.3.1. Atmospheric effects and surface reflectance

The atmospheric layers and cloud interferences are identified using the density-dimension algorithm, and besides providing data for atmospheric sciences (as part of the ATL09 product; Herzfeld et al., 2021), they are also delivered as a part of the ATL08 product (Neuenschwander et al., 2021). There are three flags that represent atmospheric conditions: Cloud confidence flag (*cloud_flag_atm*), Multiple scattering warning flag (*msw_flag*), and layer flag (*layer_flag*). The cloud confidence flag indicates the number of aerosols or cloud layers identified in the atmosphere (i.e., 0 means no clouds, 1 corresponds to the presence of one layer of clouds, 2 corresponds to two layers, etc.). Note that the cloud confidence flag (*cloud_flag_atm*) replaced the original parameter (*cloud_flag_asr*), which was found unsuitable for cloud cover detection over dry land due to varying surface reflectance (Neuenschwander et al., 2021). The multiple scattering warning flag is estimated based on the height, thickness, and optical depth of the layer and can range from 0 to 5 where zero means no multiple scattering (i.e. no layers were detected) and 5 the greatest scattering (i.e., an atmospheric layer that touches the ground, such as fog, blowing snow, or dust storm, was detected). Finally, the layer flag simply indicates the presence or absence of clouds. Note that all above flags are observed only for strong beams, assuming the cloud condition for the corresponding weak beams is the same.

The presence of snow can significantly increase the number of detected ground photons compared to snow-free segments due to the high reflectance of the snow surface at the wavelength of 532 nm. Therefore, we also used the snow cover mask (*segment_snowcover*) that indicates a likely presence of snow within a segment (Neuenschwander et al., 2020). This flag is extracted from the National Oceanic and Atmospheric Administration (NOAA) daily snow cover product. According to the snow cover flag, snow cover was present from November through April in the study area and the median number of ground photons per 100 m segment for snow-covered surface increased more than three times compared to the snow-free surface; in non-forested areas, the values increased from 83 to 254 and in forests from 31 to 113 when snow was present.

2.3.2. Distribution of photons within a segment

Almost all negative effects affecting signal detection are bound to finally manifest themselves in the distribution of signal photons within a segment. The variability in ground photons distribution can be assessed using the subset terrain flag (*subset_te_flag*) reflecting the distribution of ground photons within each 100 m segment (each of these is divided into five 20 m sub-segments). This flag provides the user with information on whether the signal photons used to estimate the terrain height within the segment are evenly distributed or not (Neuenschwander et al., 2021).

2.3.3. Laser pulse energy, solar background noise, terrain slope, and canopy cover

We also evaluated the effect of signal strength (i.e. strong/weak beams), solar background noise (i.e., day/night acquisitions), slope, and canopy cover on terrain and canopy height retrievals. Canopy cover was estimated from ALS-CHM within each 100 m segment as the proportion of cells with canopy height greater than two meters. Note that in some analyses, we use a simple binary representation of canopy cover – forests and non-forested areas. We derived the terrain slope from an ALS DTM. We used Horn's algorithm with a 3 × 3 cell neighborhood implemented in the Slope tool of ArcGIS (version 10.8.1). To distinguish landcover categories, we used the Corine Landcover 2018 data rather than the segment landcover flag available as a part of ATL08 as the latter is based on relatively coarse MODIS land cover data.

2.4. Data pre-processing and sample selection

We limited our evaluation to the following categories of landcover: Croplands and pastures, Broadleaf and mixed forest, Coniferous forest, Transitional woodlands and shrub, and Natural grasslands (Fig. 1). We used pan-European Corine Landcover 2018 data at a 100 × 100 m resolution (Feranec et al., 2016). We used the uncertainty of ground height estimates (*h_te_uncertainty)* flag to remove invalid segments from the evaluation (Neuenschwander et al., 2021). If the number of ground photons within a segment is below the limit (50 photons), this flag shows an invalid value (3.4028E+38) for the particular segment. In addition, we noticed that some of the differences between the ATL08 terrain height and ALS-derived DTM were extremely large (i.e. outliers with height errors of hundreds of meters that may be, for example, caused by the background noise or some instrumental errors that have been incorrectly classified as ground) and should not be used for a reliable assessment. Such extremely incorrect values were often consecutive observations within a single beam, which only supports the notion of the instrumentation error being the likely cause. We limited the effect of such data by removing outliers (i.e., 1% of segments with the greatest error, segments that are in the first 0.5% and the last 0.5% of differences between ATL08 terrain height and ALS-derived DTM). This resulted in a total of 69,624 segments retained for the assessment of the terrain height (Fig. 2a). There was no further removal of segments for any of the ATL08 terrain height accuracy analyses (i.e. all analyses use all segments,

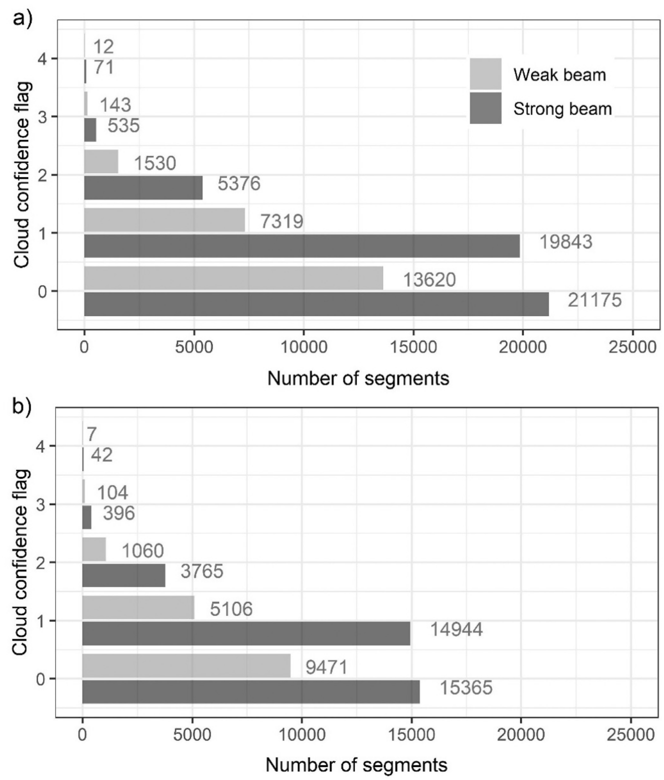

Fig. 2. Numbers of segments containing (a) terrain and (b) canopy height information grouped by beam strength (light grey – weak; dark grey – strong) and cloud confidence flag.

which were stratified according to evaluated flags). For assessment of the canopy height, only segments associated with forest landcover types Broadleaf and mixed forest, and Coniferous forest were used. Besides, we only used segments where >5% of classified photons were classified as canopy (Neuenschwander et al., 2021). This reduced the number to 28,658 segments retained for the assessment of canopy height (Fig. 2b).

2.5. Assessment of accuracy and relative importance

We calculated vertical differences between the mean terrain height and mean canopy height derived from the ATL08 (*h_te_mean* and *h_mean_canopy*) and those derived from ALS-based DTM and CHM, respectively, using pairwise combinations on a segment basis. To generate segments of 100 m × 12 m, the centroids of the first and last segments were used to calculate the track inclination, and subsequently, a rectangular buffer was generated. We used the height differences to calculate several error metrics; namely, we calculated the mean error (ME) and root mean square error (RMSE), expressed as:

$$ME = \frac{1}{n}\sum_{i=1}^{n}(h_{ATL08i} - h_{REFi})$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(h_{ATL08i} - h_{REFi})^2}$$

where h_{ATL08i} is the i_{th} elevation from ATL08, h_{REFi} is the corresponding "true" elevation from ALS DTM or ALS CHM, and n is the number of samples. In addition, for the canopy height estimates, we calculated percent mean error (%ME) and percent root mean square error (% RMSE) as follows:

$$\%ME = \frac{ME}{\overline{h}_{REF}} \times 100$$

$$\%RMSE = \frac{RMSE}{\overline{h}_{REF}} \times 100$$

where $\overline{h}_{REF}$ is the mean of the reference values. We stratified the error metrics with respect to the aforementioned flags (see the previous section).

In addition, we assessed the overall importance of the flags on both the number of retrieved photons and the ATL08 height accuracy. This assessment was performed both for the ground and the canopy photons. We used flags described in Section 2.3 with binary variables (the presence of snow, beam strength, and solar angle), categorical (landcover, photons distribution in sub-segments, and cloud flag), and continuous (canopy cover, slope, and the number of terrain and canopy photons, respectively) variables. We investigated the relative contribution of each of the flags to the ME of terrain (*h_te_mean*) and canopy height (*h_mean_canopy*), respectively, and to the number of photons retrieved using Random Forest (Breiman, 2001). The hyperparameters that need to be specified for Random Forest are the minimal size for allowing node splitting (nodesize), the structure and size of the forest (i.e., the number of trees), and the number of variables considered as candidate splitting variables at each split (mtry). We used models with minimal size of a node of 5, 150 trees, and mtry set to one-third of the total number of predictors (flags). The performance of the models was evaluated by calculating the OOB coefficient of determination (R^2). To show the effect of each variable, we assessed their relative importance and produced partial dependency plots. Random Forest model provides the importance of predictors by calculating their increases in the predictive error by randomly permuting each predictor through out-of-bag observations of each tree and calculating the subsequent decrease in out-of-bag (OOB) accuracy (we scaled the relative importance of each variable so that the sum adds to 100). Random forest models were conducted in R programming language version 4.1.1. using the ranger package version 0.13.1 (Wright et al., 2020).

3. Results

We observed excellent agreement between the ATL08 mean terrain height and ALS-DTM. The overall ME of ATL08 mean terrain height with respect to ALS-DTM was −0.27 m. The RMSE of ATL08 with respect to ALS-DTM was estimated to be 1.84 m. The fact that ATL08 mean terrain height is very close to ALS-DTM is also evident from the scatterplot showing a nearly one-to-one relationship (Fig. 3a). The agreement between the ATL08 mean canopy height and ALS-CHM was considerably weaker compared to the terrain (Fig. 3b). Although most points cumulated in the vicinity of the one-to-one line, many others tend to overestimate or underestimate the mean canopy height. Overall, ATL08 underestimated ALS canopy height by an average of 2.30 m representing %ME of −16.2%. The overall RMSE was 5.67 m representing %RMSE of 39.8%.

3.1. Terrain height

3.1.1. Clouds and day/night

From the distributions, we observed a trend showing that an increasing number of cloud or aerosol layers (cloud confidence flag; *cloud_flag_atm*) resulted in larger terrain elevation errors. RMSE for daytime acquisitions increased from 1.82 m (cloud-free acquisitions) to 2.74 m (cloud confidence flag equal to two). In addition, our results show a slight increase in the underestimation of terrain with the growing cloud confidence flag. ME for daytime acquisitions increased from −0.33 m (no clouds) to −0.56 m (cloud confidence flag equal to two), likely due to the increased photon path length caused by multiple scattering in clouds. However, the number of segments available for cloud confidence flag three and, especially, four was very low. Therefore, it is difficult to infer whether such high number of cloud layers

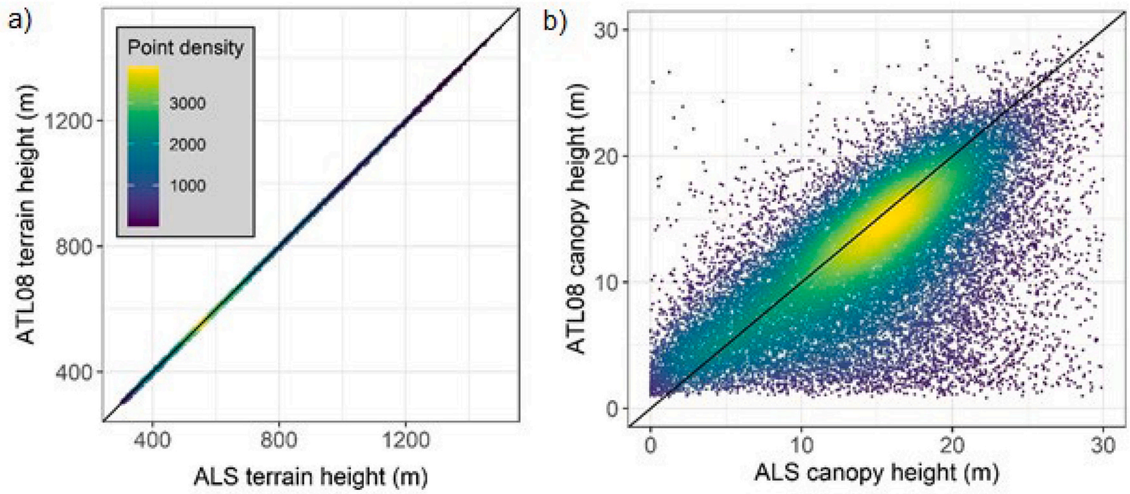

Fig. 3. Scatterplots of (a) ATL08 mean terrain height versus ALS-DTM mean terrain height and (b) ATL08 mean canopy height versus ALS-CHM mean canopy height. The solid line indicates y = x.

Table 2
The terrain height accuracy measures for increasing cloud cover (Cloud confidence flag) in relation to the solar angle (day/night).

Cloud confidence flag	Day			Night		
	Number of segments	ME (m)	RMSE (m)	Number of segments	ME (m)	RMSE (m)
0	22,555	−0.33	1.82	12,240	−0.10	1.46
1	12,274	−0.46	2.22	14,888	−0.16	1.75
2	2002	−0.56	2.74	4904	−0.17	1.59
3	216	−0.24	2.05	462	−0.22	1.82
4	4	0.33	0.51	79	−0.44	1.31

further decreases the accuracy (Table 2). The density plots of the ATL08 terrain mean error show a unimodal distribution symmetric around zero for all levels of the cloud confidence flag and both daytime and nighttime acquisitions (Fig. 4). The nighttime acquisitions have a higher accuracy than daytime acquisitions and this positive effect is evident for all levels of cloud confidence flag except for the highest level with very few segments available (Table 2).

3.1.2. Snow cover and land cover

The presence of snow cover hinders the detection of the actual surface, resulting in terrain height overestimation and, consequently, in canopy height underestimation. Our results show that snow-free segments of ATL08 tend to underestimate the terrain height by lower tens of centimeters, but snow-covered segments in high altitudes overestimate it (Table 3). In low altitudes, the density plots of the ATL08 terrain mean error show a unimodal distribution symmetric around zero for both snow-covered and snow-free segments. In higher altitudes, the distribution of height differences for snow-covered segments is shifted into positive values, while for snow-free segments, it remains symmetric around zero (Fig. 5). At altitudes above 1000 m, the presence of snow resulted in an average overestimation of the terrain by almost 50 cm (Table 3). The largest terrain overestimation due to the presence of snow

Table 3
Terrain accuracy measures for increasing altitudes in relation to the presence/absence of snow cover.

Elevation (m a.m.s.l.)	Snow-free segments			Snow-covered segments		
	Number of segments	ME (m)	RMSE (m)	Number of segments	ME (m)	RMSE (m)
200–500	13,560	−0.37	1.94	2424	−0.35	1.85
500–800	27,253	−0.31	1.78	6738	−0.23	1.61
800–1000	8550	−0.28	1.95	1884	0.03	2.04
1000–1200	5196	−0.24	1.90	2086	0.45	1.66
1200–1600	1334	−0.24	2.22	599	0.43	2.10

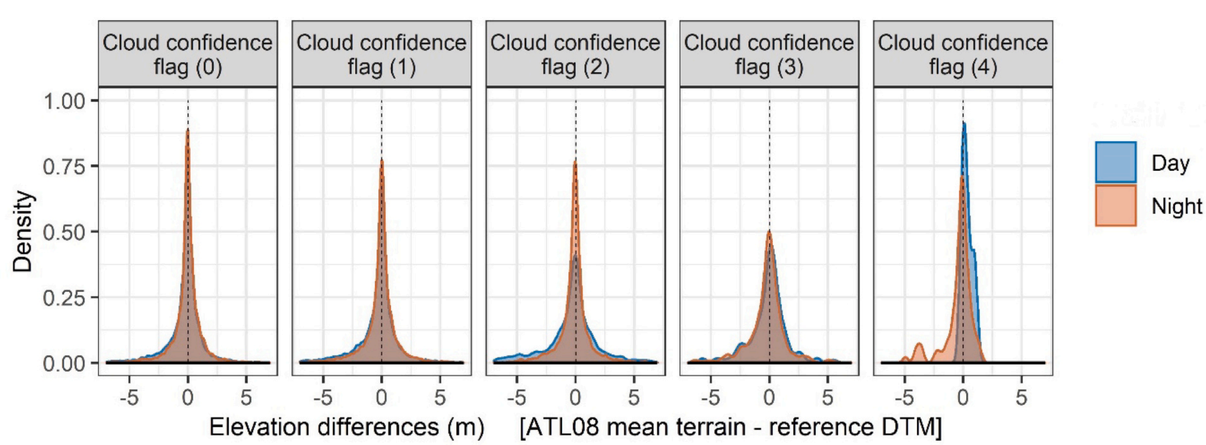

Fig. 4. Density plots illustrating the distribution of mean terrain height differences (ATL08 mean terrain height - mean DTM height), in meters, according to the cloud confidence flag grouped by the solar angle (day/night).

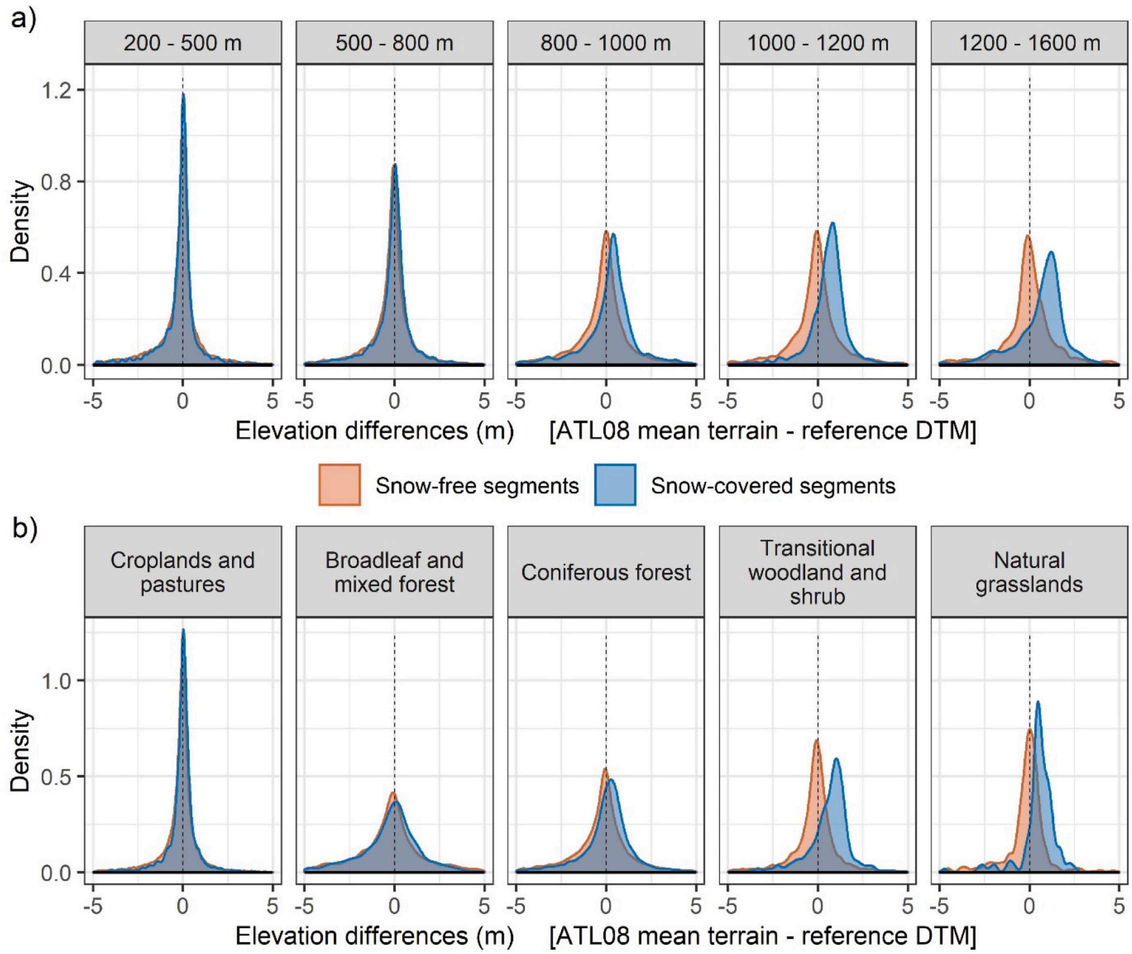

Fig. 5. Density plots illustrating the distribution of mean terrain height differences (ATL08 mean terrain height - mean DTM height), in meters, according to a) altitude and b) landcover type grouped by the snow cover flag (−covered, −free).

Table 4
Terrain height accuracy measures for main landcover types in relation to the presence/absence of snow cover.

Landcover	Snow-free segments			Snow-covered segments		
	Number of segments	ME (m)	RMSE (m)	Number of segments	ME (m)	RMSE (m)
Croplands and pastures	27,464	−0.24	1.43	5480	−0.10	1.18
Broadleaf and mixed forest	3186	−0.55	2.76	948	−0.48	2.40
Coniferous forest	20,520	−0.40	2.22	5837	−0.15	2.07
Transitional woodland and shrub	4261	−0.27	1.69	1319	0.51	1.68
Natural grasslands	462	−0.15	1.39	147	0.38	1.40

was observed for transitional woodlands and shrubs as well as for grasslands, which is a landcover predominating in the study area at high altitudes (Table 4; Fig. 5). On the other hand, the RMSEs for snow-free and snow-covered segments are relatively similar (Table 3), but considerably higher for forests than for grasslands and pastures (Table 4).

3.1.3. Interaction of the effect of clouds, snow cover, day/night, and land cover type

We examined the distribution of the vertical accuracy of ATL08 terrain height measurements grouped by cloud confidence flag for snow-covered and snow-free segments, further grouped by the solar angle (daytime vs nighttime acquisitions) and land cover type (forest/non-forest). From the distributions, we observed a trend showing that the increasing number of cloud layers resulted in considerably larger terrain elevation errors, which were particularly evident for segments with the cloud confidence flag values ≥ 2 (Table 5; Fig. 6); the error variability further increased for daytime acquisitions and snow- and/or forest-covered segments. In addition, weak beams appeared to be slightly more sensitive to atmospheric attenuation (Fig. 6). This shows that the cloud confidence flag (*cloud_flag_atm*) is a useful indicator of ATL08 terrain height accuracy.

3.1.4. Distribution of photons along ground track and slope

Almost all parameters influencing signal detection, both positive (snow cover, beam strength) and negative (signal attenuation by canopy or cloud cover), inevitably affect the distribution of ground returns within a segment. This, in turn, can greatly affect the estimated terrain accuracy, especially in rough terrain such as the mountain environment. Indeed, our results show that ATL08 terrain height accuracy considerably deteriorates with the lower number of 20 m sub-segments containing signal photons (Table 6; Fig. 7). In addition, the accuracy deteriorates as the slope of the terrain increases.

Table 5
Root mean squared error (RMSE) of ATL08 terrain height in meters with respect to the cloud cover flag, presence of snow, landcover, and solar angle (green – lowest error; purple – highest error).

		Forest		Non-forest	
		Day	Night	Day	Night
Cloud confidence flag ≤ 1	Snow cover	2.16	1.73	1.20	1.09
Cloud confidence flag ≤ 1	Snow free	2.57	2.05	1.65	1.18
Cloud confidence flag > 1	Snow cover	2.95	2.04	2.59	2.08
Cloud confidence flag > 1	Snow free	3.22	1.96	2.09	1.14

3.1.5. Optimal selection of segments

Based on the above analysis, we propose an optimal approach for the selection of ATL08 segments with the highest terrain accuracy as follows: (i) only data with photons in all five sub-segments should be kept (*subset_te_flag*); (ii) daytime acquisitions (*night_flag*) that have two or more layers of clouds (*cloud_flag_atm*) should be removed; and snow-covered segments (*segment_snowcover*) in high altitudes should also be removed (here, we removed snow-covered segments at altitudes above 1000 m). Such filtering removed 28% of available segments in our study area and resulted in a considerable improvement in accuracy (Fig. 8). The RMSE of removed segments was 3.02 m while the RMSE of segments left for the analysis was 1.1 m. In contrast, removal of all segments with snow cover or having cloud cover flag higher than zero (which is the commonly used approach; Neuenschwander et al., 2020; Queinnec et al., 2021) would result in the removal of 61% of data and only minimal improvement in accuracy (RMSE of 1.89 m and 1.78 of removed segments and segments left for the analysis, respectively). See Fig. 8 for density plots of the height differences between the ATL08 terrain and LiDAR ALS DTM for removed segments and segments left for the analysis.

3.2. Canopy height estimates

The error in the canopy height estimates is generally larger than that of terrain height estimates. From the distributions, we observed a trend showing that an increasing number of cloud or aerosol layers (cloud confidence flag; *cloud_flag_atm*) resulted in larger errors in canopy height estimates, especially for daytime acquisitions (Table 7). For nighttime acquisitions, the effect of cloud cover became evident for cloud confidence flag higher than two. The accuracy of nighttime acquisitions was higher than that of daytime acquisitions and this positive effect is evident for all levels of cloud confidence flag, for both strong and weak beams, and regardless of the presence/absence of snow (Table 8).

The most notable error was observed in the underestimation of the mean forest canopy height, resulting from the presence of snow on the terrain (Fig. 9). The canopy height of snow-free segments in broadleaf/mixed and in coniferous forests was on average underestimated by 2.1 m (%ME of −15.3%) and 1.2 m (%ME of −8.2%), respectively. The underestimation of snow-covered segments increased to 8.5 m (%ME of −56.8%) and 5.7 m (%ME of −38.3%), respectively; see Fig. 10 for density plots showing the distribution of the height error of ATL08 canopy height in relation to the presence/absence of snow in forests. The accuracy of the ATL08 mean canopy height also deteriorated with decreasing canopy cover and with increasing terrain slope (Fig. 11).

An examination of the agreement between ATL08 and ALS-CHM mean canopy heights by beam strength showed a lower accuracy for the weak than for the strong beam. For the weak beam, the canopy height was on average underestimated by 3.5 m representing %ME of −25.7%, while for the strong beam, the underestimation was 1.86 m, representing %ME of −12.9%. The overall RMSEs were 7.03 m (%RMSE of 51.3%) and 5.09 m (%RMSE of 35.4%) for the weak and strong beams, respectively.

3.3. Random forest variable importance

The RF models of the number of ground and canopy photons showed R^2 of 0.67 and 0.54, respectively (Fig. 12). The overall number of ground photons was strongly affected by the snow cover, beam strength, and canopy cover. The photons distribution in sub-segments, slope, cloud cover, and landcover had a moderate effect, and the effect of the solar angle was weak. Where canopy photons are concerned, the effect of slope and cloud flags on the number of photons was slightly stronger than on ground photons while the effect of canopy cover and photons distribution in sub-segments on this number was lower (Fig. 12; see Fig. S8-S9 for partial dependency plots). The RF models of ATL08 terrain and canopy height accuracy showed R^2 of 0.12 and 0.42, respectively (Fig. 12). The accuracy of ATL08 mean terrain height was particularly affected by the number of ground photons, their distribution within sub-segments, and canopy cover. The accuracy of the ATL08 mean canopy height was strongly affected by the canopy cover, presence of snow, the number of canopy photons, terrain slope, and beam strength. The effects of the remaining variables were considerably smaller (Fig. 12; see Fig. S10-S11 for partial dependency plots).

4. Discussion

4.1. Effect of atmospheric scattering on terrain and canopy height retrievals

One of the goals of our study was to examine the effect of photons attenuation in the atmosphere on the vertical accuracy of ATL08 terrain and canopy height estimates and to determine, which of the available flags is the best suited for this purpose. Our results show that cloud-free acquisitions have the best accuracy (Table 5; Fig. 6). However, the increasing number of cloud or aerosol layers alone has only a relatively low effect on the terrain and canopy height estimates accuracy. This, however, assumes that the surface was successfully detected even when cloud layers are present, which is rarely true, especially for ground photons. The increasing number of cloud or aerosol layers was associated with a decrease in the number of ground photons within individual segments (Fig. S1) as well as of the number of segments containing terrain information (Fig. 2). The number of ground photons steeply declined with cloud confidence flag values higher than one (Fig. S1).

Nighttime acquisitions over non-forested areas on snow-covered surfaces are ideal for the illustration of multiple scattering in clouds as there is no interference of other effects such as the solar background noise or canopy cover (Fig. 6). While segments acquired under the clear sky (cloud confidence flag ≤1) generally tend to overestimate the terrain height due to the presence of snow, segments acquired under a cloud cover (cloud confidence flag >1) tend, on average, to underestimate the terrain despite the presence of snow. Multiple scattering of photons in the dense clouds increases the photon path length, making the surface

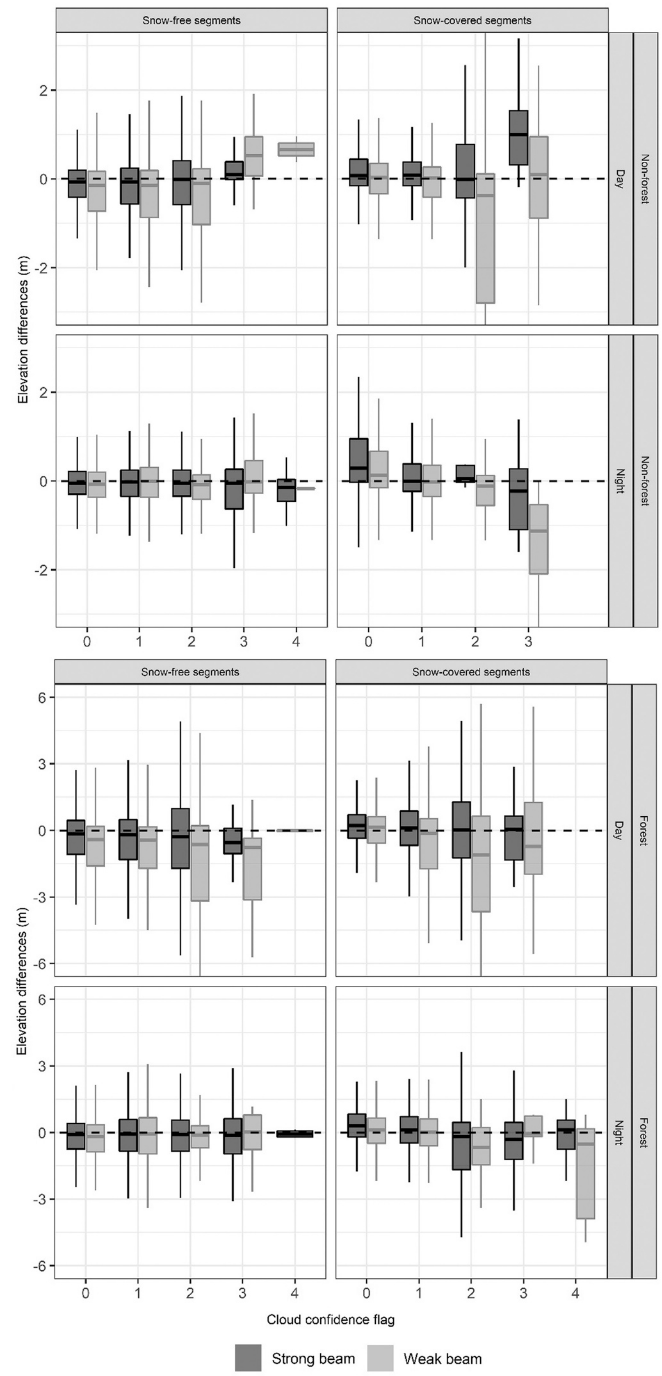

Fig. 6. Box plots showing the accuracy of ATL08 terrain height estimates (i.e., the elevation difference between ATL08 and reference DTM) considering also the cloud confidence flag (i.e. number of cloud layers), the solar angle (i.e., day and night), presence of snow (i.e., summer and winter season), landcover (i.e., forests and non-forested areas) and beam strength (strong beam; weak beam). The central horizontal line in the box indicates the median, boxes interquartile range (25th to 75th percentile), and the whiskers 1.5 times the interquartile range.

appear lower than it actually is. The magnitude of the error caused by the multiple-scattered signal depends on the height, thickness, and optical depth of the scattering layer (Winker, 2003).

Such effects are the reason why prior studies typically filtered out the segments acquired under adverse atmospheric conditions and did not evaluate their accuracy at all. For example, Neuenschwander et al. (2020) used only cloud-free data in their analysis. Similarly, Queinnec et al. (2021) removed all segments with multiple scattering warning flag (*msw_flag*) higher than zero. However, this can lead to a significant reduction in the number of segments available for further analysis (Carabajal and Boy, 2020). For example, in our study area, such brute-force filtering would result in a 50% reduction in the number of available segments (Fig. 2), which could significantly affect the generation of higher-level products such as ATL18 gridded ground-surface height, canopy height, and canopy cover estimates. Note, however, that in other situations, where accuracy is much more important than the number of available segments, the removal of any potentially inaccurate segments might be a legitimate approach. In addition, it is not clear, which flag is the most useful for such filtering. Some studies use the multiple scattering warning flag (Queinnec et al., 2021) while others the cloud confidence flag (Li et al., 2021). Our results show that from three available flags (cloud confidence flag, multiple scattering warning flag, and layer flag), the cloud confidence flag is the most useful for the assessment of atmospheric effects (Fig. 6, Fig. S1 – S7). It was previously recommended to use the cloud confidence flag (*cloud_flag_atm*) only for daytime acquisitions (Palm et al., 2020), which corresponds with our results as daytime acquisitions with a high confidence flag were associated with greater error than nighttime ones (Fig. 6). For nighttime acquisitions, it was recommended to use the multiple scattering warning flag (*msw_flag*; Palm et al., 2020). However, we did not observe any patterns that would suggest that the multiple scattering warning flag is more appropriate for nighttime acquisitions; in fact, we did not observe any relationship between the number of ground photons or terrain accuracy and the multiple scattering warning flag or layer flag (*layer_flag*) at all (Fig. S1 – S7). Note, however, that we have relatively few multiple scattering warning flag records with values of 4 and, especially, 5 (i.e., the atmospheric layer that touches the ground), which, therefore, requires further exploration.

Table 6
Root mean squared error (RMSE) of ATL08 mean terrain height in meters in relation to the number of subsegments with ground points and terrain slope (green – lowest error; purple – highest error).

Number of subsegments with ground points	Slope (0°-10°)	Slope (10°-5°)	Slope (25°-50°)
5	0.69	1.91	3.49
4	1.73	3.32	5.11
3	2.43	4.39	6.68
2	2.81	5.56	6.22
1	3.18	6.41	8.79

4.2. Effects of the evaluated factors on terrain accuracy

Besides atmospheric attenuation, the density of detected photons can be affected by landcover (canopy cover) and surface reflectance. Indeed, our results show that the decrease in the number of ground photons caused by attenuation in the atmosphere further deteriorated in forests due to the capture of photons by vegetation canopy (Fig. S1). The same effect of reduced photon penetration through the canopy cover was recently observed by Liu et al. (2021) and Malambo and Popescu (2021). In contrast, the presence of snow significantly increased the number of detected ground photons (Fig. S1). Similar seasonal variation in the average number of detected ground photons due to the presence of snow with an approximately 50% reduction in photon count when there was no snow cover was recently observed by Tian and Shan (2021). The high density of signal photons is especially important from the perspective of algorithms used for filtering as the higher density of photons reflected from a surface makes the detection of ground photons easier. Noise filtering is a critical step for accurate terrain and canopy height estimates.

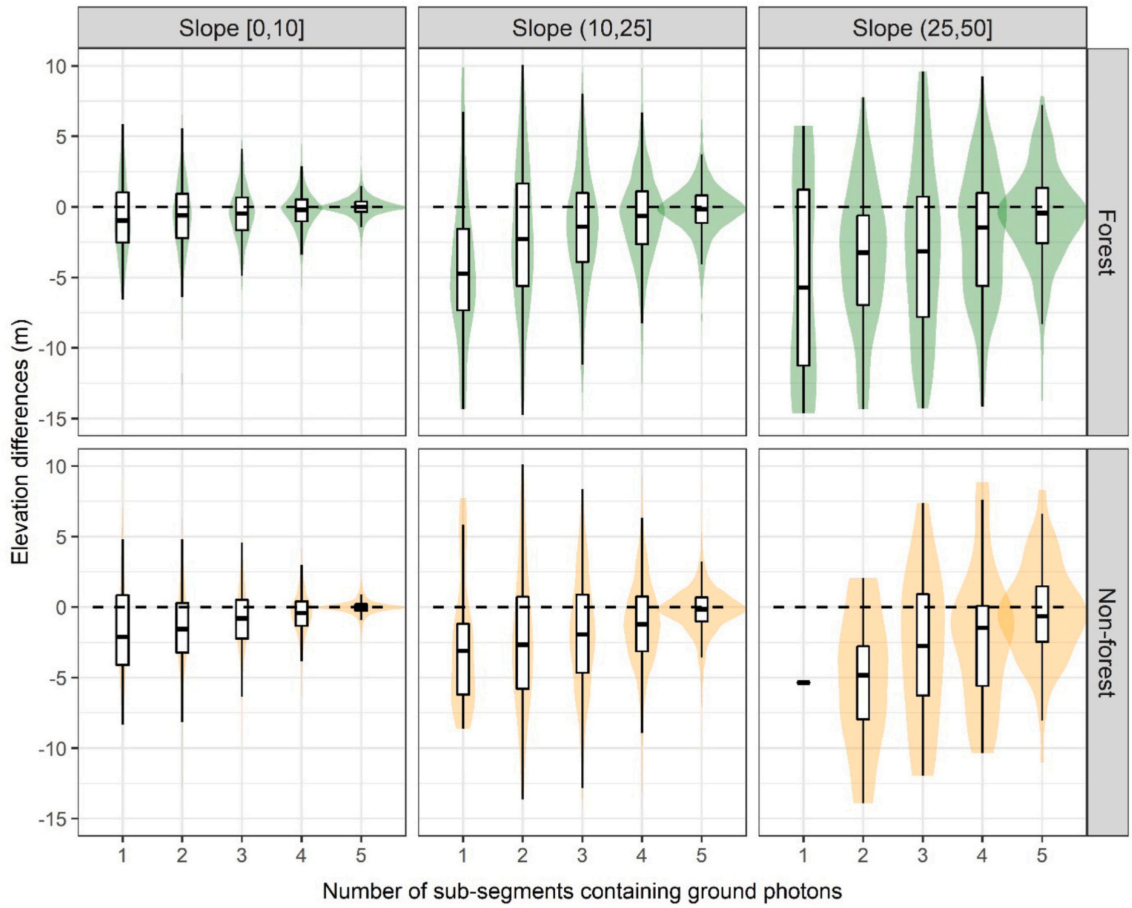

Fig. 7. Violin plots showing the ATL08 terrain height error in relation to the distribution of photons within segments (i.e. number of sub-segments containing ground photons; *subset_te_flag*), landcover (forest vs non-forest), and slope in degrees. The central horizontal line in the box indicates the median, boxes interquartile range (25th to 75th percentile), and the whiskers 1.5 times the interquartile range.

It is expected that strong beams will provide products of better accuracy – given the higher density of photons, which facilitates filtering – than weak beams (Neuenschwander et al., 2020). In general, strong beams indeed resulted in more segments and more signal photons within a segment than weak beams (Fig. 2). However, we did not observe any major differences in the accuracy of segments acquired using strong and weak beams. This is in agreement with a recent study by Malambo and Popescu (2021) who have shown that weak beams may provide data of similar quality as strong beams.

The distribution of points within a segment is a better indicator of the accuracy of terrain estimates than beam strength. Almost all effects affecting signal detection, including beam strength, are reflected in the distribution of ground returns within a segment. According to Random Forest model, the accuracy of ATL08 mean terrain height was particularly affected by the number of ground photons and their distribution within sub-segments, terrain slope, and canopy cover (Fig. 12). It is, however, important to note that R^2 of this particular model was only 12%, which is relatively low and the relative importance of flags from this model has to be interpreted with caution.

Our results show that ATL08 terrain height accuracy considerably deteriorates with the decreasing number of sub-segments containing signal (ground) photons and that the magnitude of this deterioration increases with the terrain slope (Fig. 7). In flat terrain, the effect of missing ground points in sub-segments is less severe than in steep terrain as the effect of missing sub-segments on the mean terrain height estimates in the entire segment is minimal. This corresponds with the results of previous research showing that slope is the parameter with the greatest impact on the accuracy of terrain retrieval (Liu et al., 2021). The observed effect of slope may be due to the horizontal displacements of the ICESat-2 ATL08 segments relative to the reference data as the effect of horizontal displacement on vertical accuracy increases with steep slope terrain (horizontal error reported for ICESAT-2 tends to be 0–3 m; Neuenschwander et al., 2020; Malambo and Popescu, 2021).

Ground filtering is particularly complicated by solar background noise, which is especially high in daytime acquisitions (Magruder et al., 2012; Popescu et al., 2018). Indeed, our results as well as results of previous studies show that the accuracy of daytime acquisitions is worse than that of nighttime acquisitions (Neuenschwander et al., 2020; Tian and Shan, 2021). In view of this, the presence of snow cover (causing a higher density of photons reflected from the ground), can, in a way, be considered beneficial. For example, Neuenschwander et al. (2020) suggested that the presence of snow had a greater impact on terrain accuracy than beam strength or acquisition time. In addition, they suggested that a high density of ground photons from snow-covered ground can in turn improve the results of the ground filtering algorithm as it becomes less vulnerable to background noise photons. However, our results show this is true only for data acquired under favorable weather conditions. Note that Neuenschwander et al. (2020) used only cloud-free data in their analysis; in our study, the greatest error variance in the terrain quality was observed in snow-covered segments acquired during the day and under clouds (Fig. 6). It is likely that the combination of the overestimation of terrain due to the snow layer, and its underestimation due to the increased photon path length (caused by multiple scattering of photons in clouds) made the ground photon filtering even more complicated and led to the observed decrease in terrain accuracy.

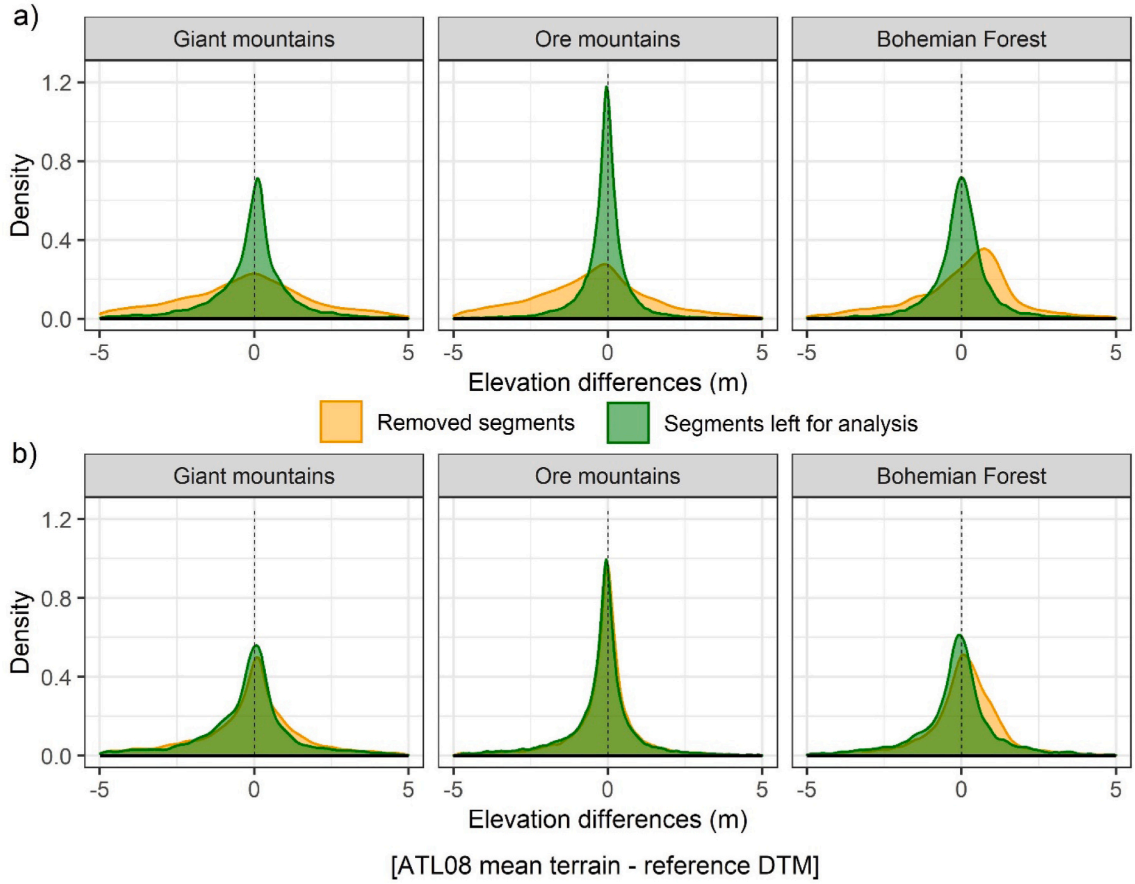

Fig. 8. Density plots illustrating the accuracy of ATL08 segments suggested for removal (orange) and segments with the highest terrain accuracy (green) left for analysis in three study areas. Two approaches for the removal of low accuracy segments are compared. The top figure (a) represents the removal of segments using the criteria suggested in this study (we removed segments with photons in less than five sub-segments; daytime acquisitions that have two or more layers of clouds; and snow-covered segments at altitudes above 1000 m ASL), while the bottom figure (b) is based on the removal of all segments with snow cover or having a cloud cover flag higher than zero, which is a common approach used in existing studies. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 7
Canopy height accuracy measures for increasing cloud cover (Cloud confidence flag) in relation to solar angle (day/night).

Cloud confidence flag	Day			Night		
	Number of segments	%ME (%)	% RMSE (%)	Number of segments	%ME (%)	% RMSE (%)
0	8025	−21.8	43.2	5233	−9.8	33.9
1	4992	−23.7	49.9	6963	−11.6	31.8
2	880	−23.0	60.9	2194	−6.3	30.9
3	97	−10.0	74.9	239	−24.2	44.1
4	0			35	−3.6	17.7

On the other hand, in high altitudes, the presence of snow cover resulted in terrain overestimation and, in turn, canopy height underestimation. We observed an approximately 45 cm overestimation of the terrain height caused by the presence of snow in altitudes higher than 1000 m. For example, Neuenschwander et al. (2020) observed an average terrain overestimation of 33 cm caused by snow cover in Finland. Again, this is a reason why some studies remove snow-covered segments from their analysis (e.g. Queinnec et al., 2021). This would, however, result in a 20% reduction in the number of segments in our study area. Still, as explained above, not all snow-covered segments are necessarily erroneous and we recommend combining the flag (*segment_snowcover*) with some others (e.g., altitude, time of the year), which can serve as a proxy for the likely presence of thick snow cover.

Table 8
Percent root mean square error (%RMSE) of ATL08 canopy height with respect to the cloud cover flag, presence of snow, beam strength, and solar angle (green – lowest error; purple – highest error).

		Strong beam		Weak beam	
		Day	Night	Day	Night
Cloud confidence flag ≤ 1	Snow cover	51.5	45.7	66.1	47.6
Cloud confidence flag ≤ 1	Snow free	34.6	27.4	45.6	40.4
Cloud confidence flag > 1	Snow cover	76.5	60.0	81.5	31.5
Cloud confidence flag > 1	Snow free	40.7	29.6	68.3	32.5

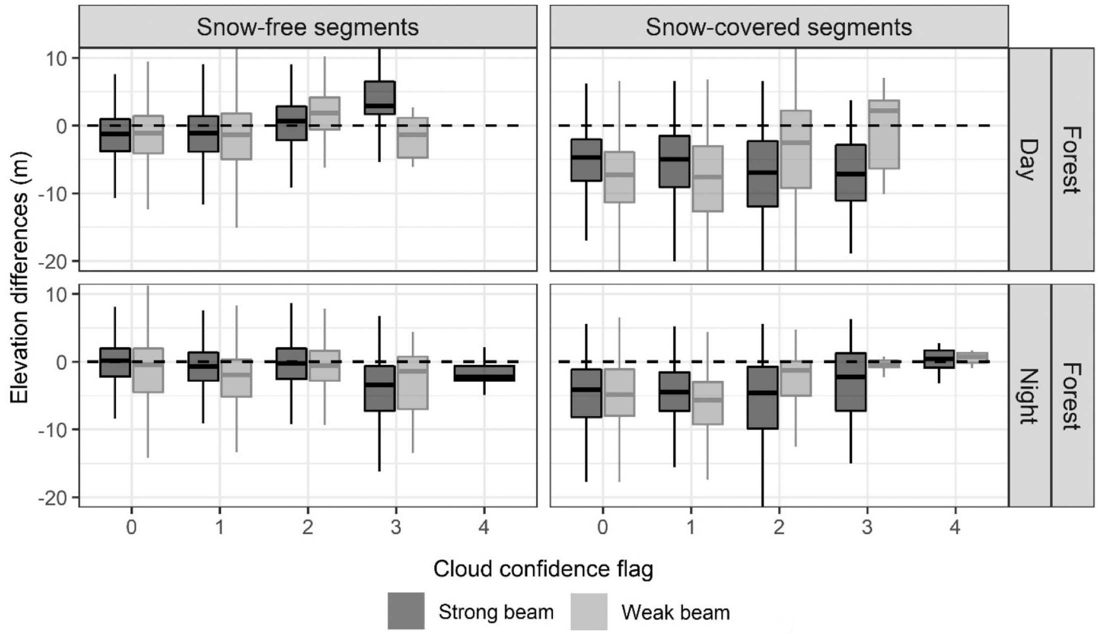

Fig. 9. Box plots showing the relationship between the accuracy of the ATL08 forest canopy height (i.e., the elevation difference between ATL08 and reference CHM) and the cloud confidence flag (i.e., number of cloud layers) with respect to the solar angle (i.e., day and night), presence of snow (i.e., summer and winter season), and beam strength (strong vs weak) beam. The central horizontal line in the box indicates the median, boxes interquartile range (25th to 75th percentile), and the whiskers 1.5 times the interquartile range.

4.3. Effects of evaluated factors on canopy accuracy

The error of canopy height estimates is considerably larger than that of terrain height and in summer, the ATL08 mean canopy height was in general underestimated by a few meters. The mean error for snow-free conditions in broadleaf/mixed forests and in coniferous forests was −2.1 m and − 1.2 m, respectively, which corresponds to the underestimation reported by prior studies (e.g., Liu et al., 2021; Neuenschwander et al., 2020). In winter, however, the underestimation considerably increased due to the presence of snow and seasonal changes in tree foliage (Fig. 10; i.e., the reference ALS data were acquired in summer). Besides, errors in canopy height estimation were larger for the weak beam than for the strong beam. Neuenschwander et al. (2020) recommended not to use the weak beam for canopy height estimation. Our results are in line with this recommendation, particularly where daytime acquisitions are concerned (Table 8). According to the Random Forest models, the accuracy of the ATL08 mean canopy height was considerably affected by the number of retrieved canopy photons, presence of snow, and the beam strength (Fig. 12). It is, therefore, evident that the number of retrieved photons is important in estimating the canopy height. In addition, canopy cover and terrain slope also play an important role. Our results show that the accuracy of ATL08 mean canopy height depends on the extent of canopy cover and deteriorates with the increasing terrain slope (Fig. 11). This finding corresponds to the results by Liu et al. (2021) who observed deterioration in the accuracy of canopy height estimates with the extent of canopy cover and, similarly to us, also reported slope to be an important error-forming factor. The best accuracy of the mean canopy height was

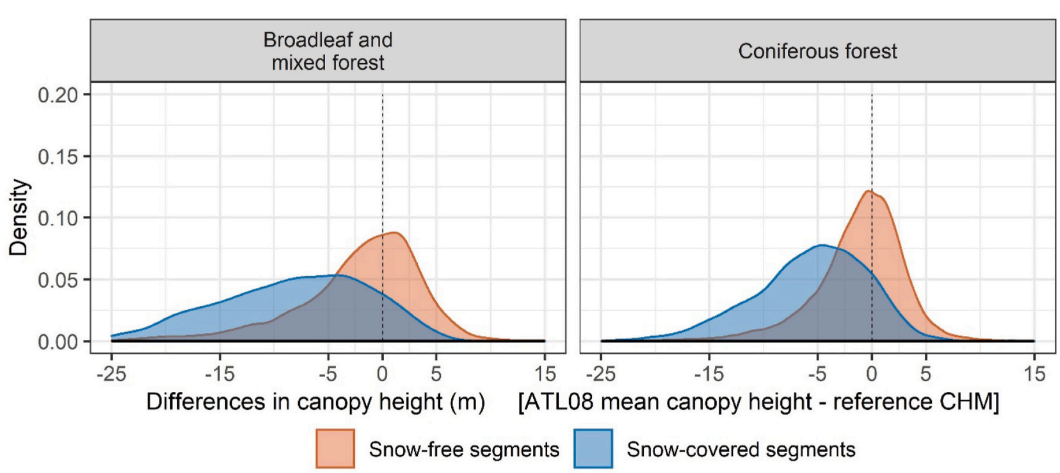

Fig. 10. Density plots showing the distribution of the absolute height error of the ATL08 mean canopy height in relation to the presence of snow according to the snow cover flag (blue and brown colors represent segments with snow and without snow, respectively) and forest type (broadleaf/mixed forest vs coniferous forest). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

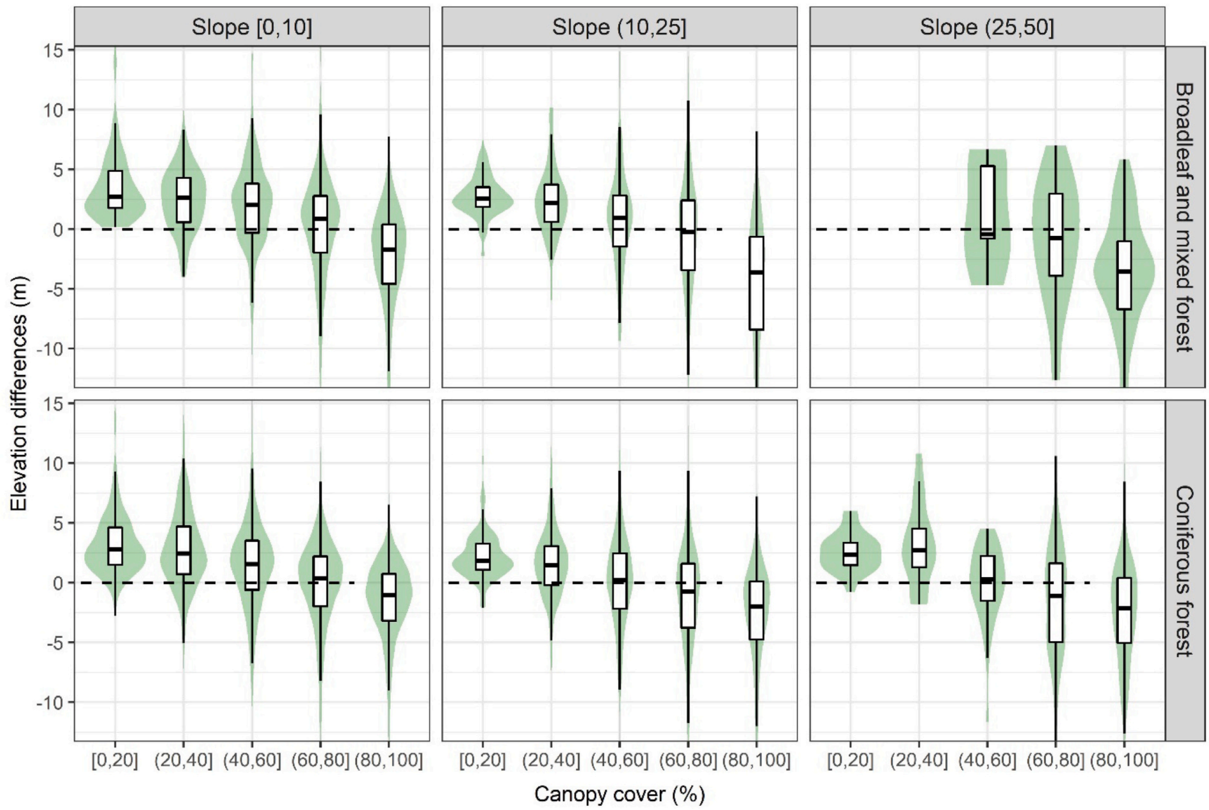

Fig. 11. Violin plots showing the ATL08 canopy height error in association with canopy cover and slope in degrees. Only segments without snow cover were used for this evaluation. The central horizontal line in the box indicates the median, boxes interquartile range (25th to 75th percentile), and the whiskers 1.5 times the interquartile range.

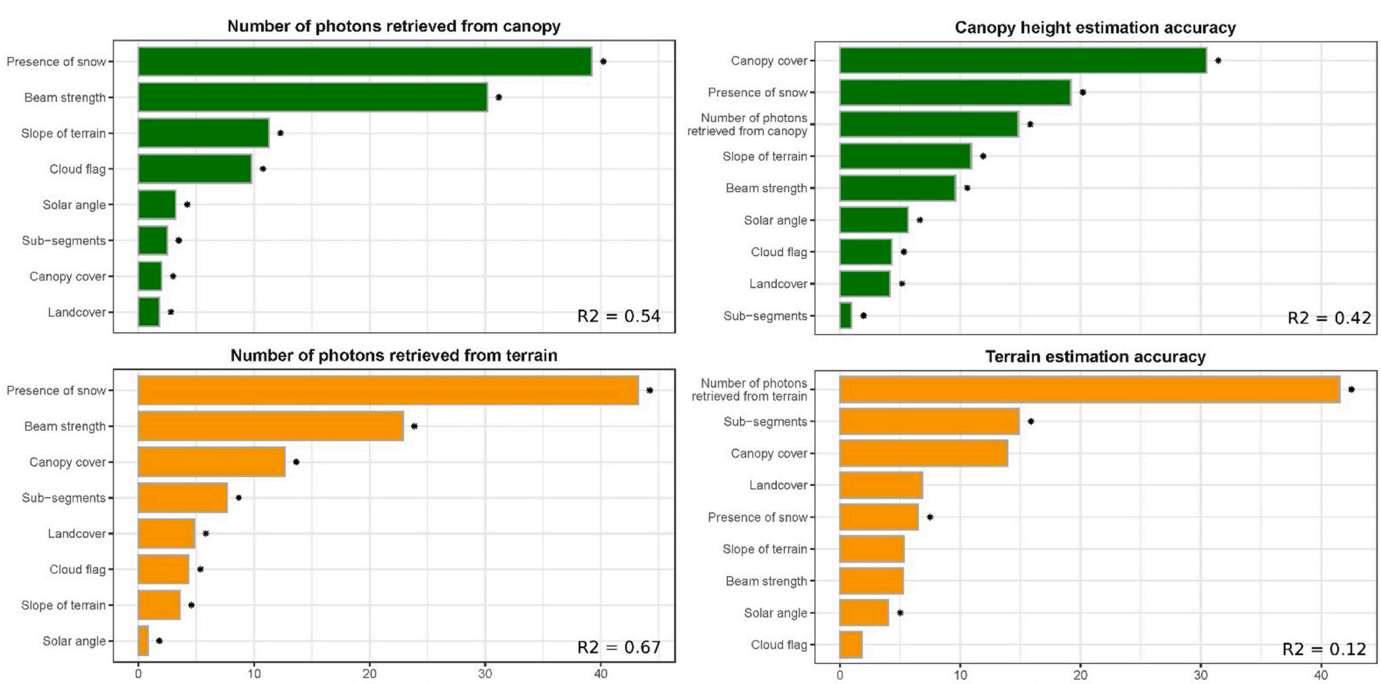

Fig. 12. Relative importance (estimated using permutation importance) of individual flags on ATL08 mean terrain and canopy height accuracy. The relative importance of each variable is scaled so that the sum adds to 100. Flags marked with an asterisk show a significant effect on the response variable. The *p*-values were computed using the Altmann et al. (2010) method with 100 permutations and the variable is considered significant if the permutation p-value is <0.05.

observed for the canopy cover ranging from 40% to 60%. Below that range, the mean height is overestimated but the error variance is relatively low. Above that range, on the other hand, canopy height is underestimated and the error variance is high, particularly on high slopes (Fig. 11). Similarly, Neuenschwander et al. (2020) reported the highest accuracy of canopy height estimates within the 40–80% canopy cover range. This is likely due to the fact that in areas with low canopy cover, the chance of a photon being reflected from the canopy is low. In contrast, with high canopy cover, fewer photons are returned from the ground, leading to its lower accuracy, especially on steep slopes, and consequently to poor estimates of canopy height (Pang et al., 2022).

4.4. Selection of segments with the highest accuracy

We suggest that, so far, practices for data filtering have been, on the one hand, unnecessarily strict (e.g. removing all segments with non-zero multiple scattering warning flag or all segments with snow cover), but, on the other hand, naïve and potentially resulting in keeping inaccurate records (but see Carabajal and Boy (2020) for selection of data that meet high accuracy requirements). Although this might be reasonable for exploratory studies, we need to develop more rigorous methods to minimize error while maximizing the number of segments left for subsequent analysis, such as the generation of higher-level products or the correction of global DTMs (Magruder et al., 2021). We suggest that ATL08 segments should be first filtered according to the distribution of ground photons in the segment (*subset_te_flag*; i.e., only data with photons in all five sub-segments should be kept). Although Liu et al. (2021) recently proposed that it is unnecessary to filter terrain data according to acquisition time, we suggest that it is reasonable to remove daytime acquisitions that have two or more layers of clouds (*cloud_flag_atm*) due to the synergic effect of solar background noise and increased photon path length due to atmospheric scattering. Finally, to minimize the negative effect of snow cover on terrain height estimation, we suggest removing snow-covered segments (*segment_snowcover*) in high altitudes, the combination of which serves as a substitute for the likely presence of a thick snow cover. As far as canopy height is concerned, in addition to above, we recommend removing all snow-covered segments (or even better would be removal of acquisitions in leaf-off period) and also removing weak beam acquisitions during the daytime (Table 8). This corresponds with a study by Liu et al. (2021) who recommend using strong-beam and nighttime acquisitions for canopy height estimates. Note, however, that most important influence on the accuracy of canopy height estimates has the density of canopy cover itself (Fig. 12) and the accuracy of canopy height estimates using ICESat-2 is limited (see Liu et al. (2021), for a comparison with GEDI).

5. Conclusions

We showed that atmospheric attenuation, surface reflectance, laser pulse energy level, solar background noise, canopy cover, and terrain slope are interlinked effects that affect the number of detected photons as well as the accuracy of ATL08 terrain and canopy height estimates. Generally, the error of canopy height estimates is considerably larger than that of terrain height. Results show that the accuracy of nighttime acquisitions is better than that of daytime acquisitions and that the increasing number of cloud layers causes a lower number of photons in a segment and greater error variability, especially of terrain estimates. This decrease was quite abrupt for the cloud confidence flag higher than one (i.e., data having less than two layers of clouds were of very good quality). Consequently, the accuracy of the ground detected under several layers of clouds may be limited and ATL08 segments unsuitable for terrain characterization. In general, ATL08 mean terrain estimates tend to underestimate the terrain (reference DTM) by a few tens of centimeters in summer (i.e., no snow cover), but overestimate it in winter, particularly in altitudes higher than 1000 m due to the presence of thick snow cover. Congruently, the canopy height is underestimated by a few meters in summer, and this underestimation considerably increased due to the presence of snow and seasonal loss of tree foliage in winter. The accuracy of the canopy height estimates depends on the extent of canopy cover and deteriorates with the increasing slope of the terrain. Almost all parameters affecting signal detection are reflected in the distribution of ground returns within a segment. The ATL08 terrain height accuracy deteriorates with the lower number of sub-segments containing signal photons and the magnitude of this decline increases for steep slopes. The presence of snow was associated with the strongest positive effect on the number of detected ground photons, the number of ground photons detected over snow-covered surfaces increased three times compared to snow-free surfaces. Removing segments with the poor distribution of photons, more than one layer of clouds during the day, and snow cover in high altitudes is the best approach for data filtering that minimizes errors while maximizing the number of segments left for subsequent analysis.

CRediT authorship contribution statement

Vítězslav Moudrý: Writing – original draft, Writing – review & editing, Conceptualization, Methodology, Software, Data curation, Investigation, Visualization, Supervision. **Kateřina Gdulová:** Data curation, Investigation, Writing – review & editing. **Lukáš Gábor:** Software, Investigation, Writing – review & editing. **Eliška Šárovcová:** Software, Investigation, Writing – review & editing. **Vojtěch Barták:** Software, Writing – review & editing, Formal analysis. **François Leroy:** Software, Writing – review & editing, Formal analysis. **Olga Špatenková:** Writing – review & editing. **Duccio Rocchini:** Writing – review & editing. **Jiří Prošek:** Writing – review & editing, Visualization, Conceptualization.

Declaration of Competing Interest

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

Acknowledgements

We are grateful to four anonymous reviewers for their constructive comments that significantly improved the quality of the manuscript. We are also grateful to the administrations of Šumava National Park, Krkonose Mountains National Park, and the Head Office of Geodesy and Cartography in Poland for providing the airborne laser scanning data. We also thank the "Staatsbetrieb Geobasisinformation und Vermessung Sachsen (GeoSN)" for providing access to the DTM and CHM of Saxony as open data under the dl-de/by-2-0 license. This work was supported by the Internal Grant Agency of the Faculty of Environmental Sciences, Czech University of Life Sciences Prague (project No. 2022B0035). D.R. was partially supported by the H2020 Project SHOWCASE (Grant agreement No. 862480). VM and JP were supported by long-term research development project RVO 67985939 (Czech Academy of Sciences).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.rse.2022.113112.

References

Altmann, A., Toloşi, L., Sander, O., Lengauer, T., 2010. Permutation importance: a corrected feature importance measure. Bioinformatics 26 (10), 1340–1347.
Bolton, D.K., Coops, N.C., Wulder, M.A., 2013. Investigating the agreement between global canopy height maps and airborne Lidar derived height estimates over Canada. Can. J. Remote. Sens. 39 (sup1), S139–S151.
Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.

Brunt, K.M., Neumann, T.A., Smith, B.E., 2019. Assessment of ICESat-2 ice sheet surface heights, based on comparisons over the interior of the Antarctic ice sheet. Geophys. Res. Lett. 46 (22), 13072–13078.
Carabajal, C.C., Boy, J.P., 2020. ICESat-2 altimetry as geodetic control. Int. Arch. Photogram. Remote Sens. Spat. Inform. Sci. 43, 1299–1306.
Chen, Q., 2010. Retrieving vegetation height of forests and woodlands over mountainous areas in the Pacific Coast region using satellite laser altimetry. Remote Sens. Environ. 114 (7), 1610–1627.
Deems, J.S., Painter, T.H., Finnegan, D.C., 2013. Lidar measurement of snow depth: a review. J. Glaciol. 59 (215), 467–479.
Degnan, J.J., 2002. Photon-counting multikilohertz microlaser altimeters for airborne and spaceborne topographic measurements. J. Geodyn. 34 (3–4), 503–549.
Farrell, S.L., Duncan, K., Buckley, E.M., Richter-Menge, J., Li, R., 2020. Mapping Sea ice surface topography in high fidelity with ICESat-2. Geophys. Res. Lett. 47 (21) e2020GL090708.
Feranec, J., Soukup, T., Hazeu, G., Jaffrain, G., (Eds.)., 2016. European Landscape Dynamics: CORINE Land Cover Data. CRC Press.
Gdulová, K., Marešová, J., Barták, V., Szostak, M., Červenka, J., Moudrý, V., 2021. Use of TanDEM-X and SRTM-C data for detection of deforestation caused by bark beetle in central European Mountains. Remote Sens. 13 (15), 3042.
Goetz, S., Dubayah, R., 2011. Advances in remote sensing technology and implications for measuring and monitoring forest carbon stocks and change. Carbon Manag. 2 (3), 231–244.
Harding, D.J., Dabney, P.W., Valett, S., 2011. Polarimetric, two-color, photon-counting laser altimeter measurements of forest canopy structure. In: International Symposium on Lidar and Radar Mapping 2011: Technologies and applications. International Society for Optics and Photonics, p. 828629.
Hein, J., Guarin, A., Frommé, E., Pauw, P., 2018. Deforestation and the Paris climate agreement: an assessment of REDD+ in the national climate action plans. Forest Policy Econ. 90, 7–11.
Herold, M., Carter, S., Avitabile, V., Espejo, A.B., Jonckheere, I., Lucas, R., De Sy, V., 2019. The role and need for space-based forest biomass-related measurements in environmental management and policy. Surv. Geophys. 40 (4), 757–778.
Herzfeld, U., Hayes, A., Palm, S., Hancock, D., Vaughan, M., Barbieri, K., 2021. Detection and height measurement of tenuous clouds and blowing snow in ICESat-2 ATLAS data. Geophys. Res. Lett. 48 (17) e2021GL093473.
Hilbert, C., Schmullius, C., 2012. Influence of surface topography on ICESat/GLAS forest height estimation and waveform shape. Remote Sens. 4 (8), 2210–2235.
Li, B., Xie, H., Liu, S., Tong, X., Tang, H., Wang, X., 2021. A method of extracting high-accuracy elevation control points from ICESat-2 altimetry data. Photogramm. Eng. Remote. Sens. 87 (11), 821–830.
Liu, A., Cheng, X., Chen, Z., 2021. Performance evaluation of GEDI and ICESat-2 laser altimeter data for terrain and canopy height retrievals. Remote Sens. Environ. 264, 112571.
Magruder, L.A., Wharton III, M.E., Stout, K.D., Neuenschwander, A.L., 2012. Noise filtering techniques for photon-counting ladar data. In: Laser Radar Technology and Applications XVII, vol. 8379. International Society for Optics and Photonics, p. 83790Q. May.
Magruder, L., Neuenschwander, A., Klotz, B., 2021. Digital terrain model elevation corrections using space-based imagery and ICESat-2 laser altimetry. Remote Sens. Environ. 264, 112621.
Malambo, L., Popescu, S.C., 2021. Assessing the agreement of ICESat-2 terrain and canopy height with airborne lidar over US ecozones. Remote Sens. Environ. 266, 112711.
Markus, T., Neumann, T., Martino, A., Abdalati, W., Brunt, K., Csatho, B., Zwally, J., 2017. The ice, cloud, and land elevation Satellite-2 (ICESat-2): science requirements, concept, and implementation. Remote Sens. Environ. 190, 260–273.
Marselis, S.M., Keil, P., Chase, J.M., Dubayah, R.O., 2022. The use of GEDI canopy structure for explaining variation in tree species richness in natural forests. Environ. Res. Lett. 17 (4), 045003. https://doi.org/10.1088/1748-9326/ac583f.
Mulverhill, C., Coops, N.C., Hermosilla, T., White, J.C., Wulder, M.A., 2022. Evaluating ICESat-2 for monitoring, modeling, and update of large area forest canopy height products. Remote Sens. Environ. 271, 112919. https://doi.org/10.1016/j.rse.2022.112919.
Nandy, S., Srinet, R., Padalia, H., 2021. Mapping forest height and aboveground biomass by integrating ICESat-2, Sentinel-1 and Sentinel-2 data using random forest algorithm in northwest Himalayan foothills of India. Geophys. Res. Lett. 48 (14) e2021GL093799.
Neuenschwander, A.L., Magruder, L.A., 2019. Canopy and terrain height retrievals with ICESat-2: a first look. Remote Sens. 11 (14), 1721.
Neuenschwander, A., Pitts, K., 2019. The ATL08 land and vegetation product for the ICESat-2 Mission. Remote Sens. Environ. 221, 247–259.
Neuenschwander, A., Guenther, E., White, J.C., Duncanson, L., Montesano, P., 2020. Validation of ICESat-2 terrain and canopy heights in boreal forests. Remote Sens. Environ. 251, 112110.
Neuenschwander, A., Popescu, S., Nelson, R., Harding, D., Pitts, K., Robbins, J., Sheridan, R., 2021. Ice, Cloud, and Land Elevation Satellite 2 (ICESat-2) Algorithm Theoretical Basis Document (ATBD) for Land-Vegetation Along-Track Products (ATL08) release 004. National Aeronautics and Space Administration, Washington, DC, USA. Tech. Rep., Accessed: Nov. 18, 2021. [Online]. Available. https://nsidc.org/sites/nsidc.org/files/technical-references/ICESat2_ATL08_ATBD_r004.pdf.
Neumann, T.A., Martino, A.J., Markus, T., Bae, S., Bock, M.R., Brenner, A.C., Thomas, T. C., 2019. The ice, cloud, and land elevation satellite–2 Mission: a global geolocated photon product derived from the advanced topographic laser altimeter system. Remote Sens. Environ. 233, 111325.
Palm, S., Yang, Y., Herzfeld, U., 2020. ICESat-2 Algorithm Theoretical Basis Document for the Atmosphere, Part I: Level 2 and 3 Data Products Version 8.3. National Aeronautics and Space Administration, Washington, DC, USA. Tech. Rep., Accessed: Nov. 20, 2021. [Online]. Available. https://nsidc.org/sites/nsidc.org/files/technical-references/ICESat2_ATL04_ATL09_ATBD%28part%201%29_r003.pdf.
Pan, Y., Birdsey, R.A., Fang, J., Houghton, R., Kauppi, P.E., Kurz, W.A., Hayes, D., 2011. A large and persistent carbon sink in the world's forests. Science 333 (6045), 988–993.
Pang, S., Li, G., Jiang, X., Chen, Y., Lu, Y., Lu, D., 2022. Retrieval of forest canopy height in a mountainous region with ICESat-2 ATLAS. For. Ecosyst. 9, 100046. https://doi.org/10.1016/j.fecs.2022.100046.
Popescu, S.C., Zhou, T., Nelson, R., Neuenschwander, A., Sheridan, R., Narine, L., Walsh, K.M., 2018. Photon counting LiDAR: an adaptive ground and canopy height retrieval algorithm for ICESat-2 data. Remote Sens. Environ. 208, 154–170.
Pugh, T.A., Lindeskog, M., Smith, B., Poulter, B., Arneth, A., Haverd, V., Calle, L., 2019. Role of forest regrowth in global carbon sink dynamics. Proc. Natl. Acad. Sci. 116 (10), 4382–4387.
Queinnec, M., White, J.C., Coops, N.C., 2021. Comparing airborne and spaceborne photon-counting LiDAR canopy structural estimates across different boreal forest types. Remote Sens. Environ. 262, 112510.
Schutz, B.E., Zwally, H.J., Shuman, C.A., Hancock, D., DiMarzio, J.P., 2005. Overview of the ICESat mission. Geophys. Res. Lett. 32 (21).
Smith, B., Fricker, H.A., Holschuh, N., Gardner, A.S., Adusumilli, S., Brunt, K.M., Siegfried, M.R., 2019. Land ice height-retrieval algorithm for NASA's ICESat-2 photon-counting laser altimeter. Remote Sens. Environ. 233, 111352.
Swatantran, A., Tang, H., Barrett, T., DeCola, P., Dubayah, R., 2016. Rapid, high-resolution forest structure and terrain mapping over large areas using single photon lidar. Scientific reports 6 (1), 1–12.
Tian, X., Shan, J., 2021. Comprehensive evaluation of the ICESat-2 ATL08 terrain product. IEEE Trans. Geosci. Remote Sens. 59 (10), 8195–8209. https://doi.org/10.1109/TGRS.2021.3051086.
Winker, D.M., 2003. Accounting for multiple scattering in retrievals from space lidar. In: 12th International Workshop on Lidar Multiple Scattering Experiments, vol. 5059. International Society for Optics and Photonics, pp. 128–139. April.
Wright, M.N., Wager, S., Probst, P., 2020. Ranger: A Fast Implementation of Random Forests. R Package Version 0.13.1.