Remote Sensing of Environment 280 (2022) 113174

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

How to consider the effects of time of day, beam strength, and snow cover in ICESat-2 based estimation of boreal forest biomass?

P. Varvia [a,*], L. Korhonen [a], A. Bruguière [a,b], J. Toivonen [a], P. Packalen [a], M. Maltamo [a], S. Saarela [c,d], S.C. Popescu [e]

[a] School of Forest Sciences, University of Eastern Finland, P.O. Box 111, Joensuu, FI-80101, Finland
[b] National Land Survey of Finland, Torikatu 36 A, Joensuu, FI-80100, Finland
[c] Faculty of Forest Sciences, Swedish University of Agricultural Sciences, SLU Skogsmarksgrand 17, Umeå, SE-90183, Sweden
[d] Faculty of Environmental Sciences and Natural Resource Management, Norwegian University of Life Sciences, P.O. Box 5003, NMBU, Ås, NO-1432, Norway
[e] Department of Ecology and Conservation Biology, Texas A&M University, 2258 TAMU, College Station, 77843-2258, TX, United States

ARTICLE INFO

Edited by Marie Weiss

Keywords:
ICESat-2
Above-ground biomass
Boreal forest
Mixed-effect models
Lidar

ABSTRACT

The objective of this study was to explore the effects of (1) the presence/absence of snow and snow depth, (2) solar noise, i.e., day/night and sun angle observations, and (3) strong/weak beam differences on ICESat-2 data in the context of data utility for forest AGB estimation. The framework of the study is multiphase modeling, where AGB field data and wall-to-wall airborne laser scanning (ALS) and Sentinel-2 data are used to produce proxy ALS plots on ICESat-2 track positions. Models between the predicted proxy AGB and the ICESat-2 photon data are then formulated and evaluated by subsets, such as only strong beam data captured in snowy conditions.

Our results indicate that, if possible, strong beam night data from snowless conditions should be used in AGB estimation, because our models showed clearly smallest RMSE (26.9%) for this data subset. If more data are needed, we recommend using only strong beam data and constructing separate models for the different data subsets. In the order of increasing RMSE%, the next best options were snow/night/strong (30.4%), snow/day/strong (33.5%), and snowless/day/strong (34.1%). Weak beam data from snowy night conditions could also be used if necessary (31.0%).

1. Introduction

Forest above-ground biomass (AGB) has a crucial role in the global carbon cycle (Herold et al., 2019), but maps describing its global distribution have large uncertainties (Mitchard et al., 2013; Zhang et al., 2019). The key to improving the accuracy of global AGB maps is the development of sensors that can obtain direct measurements of 3D structure of forests (Duncanson et al., 2019), which passive satellite sensors can only observe indirectly.

Satellite lidar sensors are one such sensor type. Utilization of spaceborne lidar measurements for AGB estimation started with the original ICESat (Ice, Cloud and land Elevation Satellite), launched in 2003, which included a full waveform lidar instrument GLAS (Geoscience Laser Altimeter System) (Zwally et al., 2002). Spaceborne lidars do not provide wall-to-wall coverage similar to, for example, optical satellites, which complicates field plot collection. In early studies (e.g. Lefsky et al., 2005; Nelson et al., 2009b), field plot locations were chosen to coincide with the spaceborne lidar footprint locations, however in remote forests such approach is usually not feasible. Later studies have therefore often used proxy field plots acquired using airborne laser scanning (ALS) (Wulder et al., 2012) to produce biomass models for spaceborne lidar. Both profiling airborne lidar (e.g. Boudreau et al., 2008; Nelson et al., 2009a; Margolis et al., 2015) and small footprint ALS data (e.g. Nelson et al., 2017; Holm et al., 2017) have been utilized.

Back in 2018, NASA (National Air and Space Administration) launched two new spaceborne lidar sensors that are capable of direct tree height measurements. The GEDI (Global Ecosystem Dynamic Investigation) sensor was especially designed for tree height and AGB estimation (Dubayah et al., 2020). However, due to its mounting on the International Space Station, it does not acquire data above 52°N, which leaves unmeasured the northern tier of the circumpolar boreal forest. On the other hand, the ICESat −2 (Ice, Cloud and land Elevation Satellite 2) was designed for snow and ice monitoring (Markus et al., 2017). It is in a polar orbit that provides excellent coverage of the entire boreal forest zone.

* Corresponding author.
E-mail address: petri.varvia@uef.fi (P. Varvia).

https://doi.org/10.1016/j.rse.2022.113174
Received 10 January 2022; Received in revised form 10 July 2022; Accepted 14 July 2022
Available online 25 July 2022

The ICESat −2 ATLAS (Advanced Topographic Laser Altimeter System) is a profiling lidar sensor that operates at visible green wavelength (532 nm). Instead of providing wall-to-wall coverage, profiling sensors record strip samples of terrain height measurements using multiple beams. The 532 nm wavelength provides strong backscattering from snow and ice, but the reflectance from vegetation is considerably weaker compared to GEDI that operates at 1064 nm wavelength. In addition, ATLAS is a photon counting lidar, and therefore subject to solar noise photons reflected from the atmosphere, which must be filtered from the data before it can be used (Popescu et al., 2018). Each ATLAS ground track consists of three pairs of strong and weak lidar beams. The weak beams, one-fourth the power of the strong beams, permit measurement of local across-track slope on glaciers and snowfields (Markus et al., 2017; Neumann et al., 2019). Thus, the capability of weak beams to observe canopy heights can be poor (Neuenschwander et al., 2020). Regardless of these limitations, initial studies have shown that ICESat −2 strong beams can provide relatively accurate estimates of forest canopy height (Neuenschwander et al., 2020) and AGB (Narine et al., 2020). The spatial resolution of estimation is however considerably poorer compared to GEDI or airborne lidar sensors, because the number of photons per meter of ground track is low (c. 0–20 photons, depending on the surface reflectance) (Neumann et al., 2021).

Spaceborne lidar sensors have an advantage in that new data are recorded throughout the year whenever the sky is clear. However, most studies on forest attribute estimation with spaceborne lidars have been conducted with data captured in snowless leaf-on conditions, often due to having a significant deciduous component in the study area. Furthermore, there is little information on how the commonly used airborne lidars perform in the estimation of forest attributes when there is snow. In boreal forests, the snowy season can last roughly from October to May. Snow accumulates both on the forest floor and the trees, which increases the forest reflectance at the 532 nm wavelength compared to summer conditions. This could be beneficial to estimation of forest variables using ATLAS data, especially using weak beams. However, if there is plenty of snow, the allometric relationships between AGB and observed canopy heights and densities could also be affected, especially if the ground elevation is estimated directly from the ATLAS data. One research question, therefore, is if the ATLAS data obtained from snowy boreal forests can be used in combination with snow-free data from the same area.

Another topic to consider is the performance of noise filtering when using day and night data sets in a same model. The ATLAS photon coordinate data is delivered in the ATL03 Global Geolocated Photon product. Each ATL03 photon is classified as noise, ground, or canopy, and this classification is stored in the ATL08 Land and Vegetation Height product. Day data has more solar noise photons, thus night data is expected to perform better in the estimation of forest parameters (Narine et al., 2019). The performance of the default noise classification with data from different times-of-day and beam strengths should also be clarified when the objective is to estimate boreal forest AGB.

Our objective in this study is to investigate and clarify how the effects of time of day, beam strength and snow cover should be considered when constructing models for evergreen boreal forest AGB estimation using ICESat −2 data. Furthermore, we evaluate if ICESat −2 data obtained in different acquisition conditions (snowy or snow-free, day or night) and with different beam strengths can be fused into a single AGB prediction model without losing accuracy. While ICESat −2 canopy height product has been previously validated in snowy conditions (Neuenschwander et al., 2020), to our knowledge, this is the first study where satellite lidar AGB modeling is tested in the presence of snow.

2. Materials & methods

Because ATLAS is a profiling lidar sensor, the ATLAS footprints are unlikely to overlap with the available field plots. Thus, we employ multi-phase modeling similar to e.g. Boudreau et al. (2008), Nelson et al. (2017), Saarela et al. (2018): (1) The AGB of field plots is calculated based on allometric models. (2) The field-measured AGB is linked with ALS data to produce a model to predict AGB from ALS data using the area-based approach. (3) The ALS model is used to predict AGB at ICESat −2 track locations. (4) An AGB model is constructed using ALS-derived AGB and ICESat −2 photon metrics. The workflow of the study is charted in Fig. 1.

2.1. Study area and field measurements

The study site is an approximately 60 × 50 km area located near Valtimo, Finland (N 63°46′ E 28°13′, see Fig. 2), consisting of boreal forest. The area is dominated by Scots pine (*Pinus sylvestris* L.), Norway spruce (*Picea abies* (L.) Karst) and birches (*Betula* spp.). The snowy season typically lasts from late November to late April. ALS-based forest management inventory was conducted in the area by the Finnish Forest Centre during the summer 2019, which included acquisition of ALS data and aerial images in cooperation with the National Land Survey. The distribution of the field plots in the study area is shown in Fig. 3. The field plot data and ALS data are openly available at https://www.metsaan.fi/karttapalvelut (in Finnish) and https://tiedostopalvelu.maanmittauslaitos.fi/tp/kartta, respectively.

The field data contained 797 circular plots with radius of either 5.64 m, 9.00 m, or 12.62 m depending on the forest maturity. Diameter at breast height (DBH) was measured for each tree with DBH≥ 5 cm. The height of a sample tree of each tree species was recorded on each plot and a calibrated height model (Eerikäinen, 2009) was then used to predict the height for the rest of the trees. Using the measured tree diameters and predicted tree heights, the plot level AGB (Mg/ha) was calculated by summing up the individual tree biomass estimates produced by Repola's species-specific biomass models (Repola, 2008, 2009).

2.2. ALS data

The ALS data were collected 7 June–9 July 2019 using a Leica ALS 80HP scanner at 1700 m above ground level, which resulted in a nominal pulse density of 5 pts/m^2 and a footprint diameter of 39 cm. However, the publicly distributed data was only available at resampled 0.5 pts/m^2 pulse density, which was still sufficient for AGB modeling using the area-based approach (ABA) (Kotivuori et al., 2016).

The ALS echoes were height normalized using LAStools (rapidlasso GmbH) and erroneous echoes with canopy heights higher than 40 m above-ground height were discarded. Onwards ALS height refers to height at above-ground level. Canopy metrics were then computed for each plot using two sets of echoes: first of many + only echoes, and last-of-many + only echoes. The metrics included average and maximum heights, standard deviation of heights, height percentiles $p_5, p_{10}, p_{20}, \dots, p_{90}, p_{95}, p_{99}$, canopy density percentiles $b_5, b_{10}, b_{20}, \dots, b_{90}, b_{95}$, canopy cover, and the average and standard deviation of intensities.

2.3. Sentinel-2 data

We used a Sentinel-2 image captured on June 14th 2019. Atmospheric correction of the Sentinel-2 image was performed using Sen2Cor (Main-Knorn et al., 2017). The atmospheric bands (bands 1, 9, and 10) were omitted after this phase. The pixel values extracted from the rest of bands were used as predictors in the AGB models. In addition, several common spectral vegetation indices were calculated.

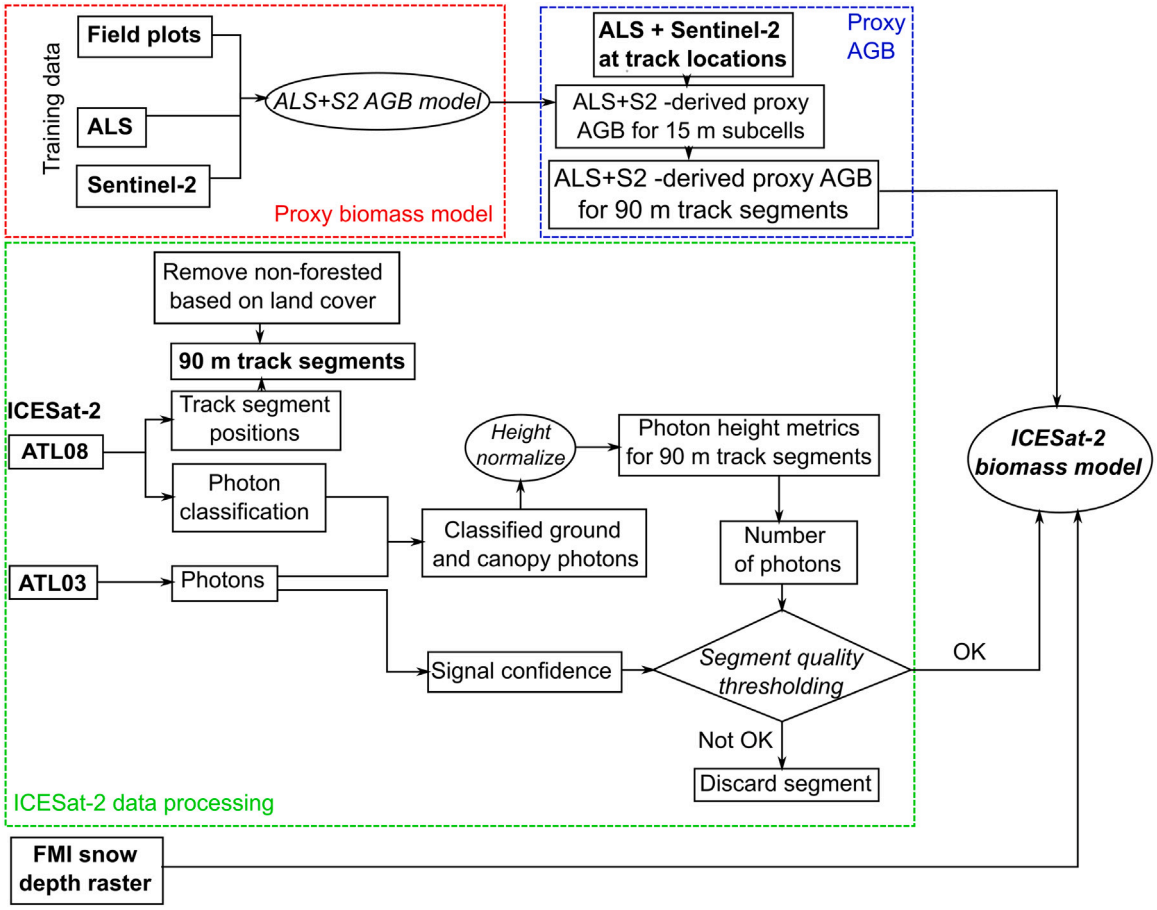

Fig. 1. Workflow for the ICESat −2 AGB modeling.

2.4. ALS-based biomass model

The field plot AGB measurements, ALS metrics and Sentinel-2 bands were used to produce a biomass model to predict AGB from ALS and Sentinel-2 data. We formulated a quadratic model between AGB and four predictors x_i:

$$\mathrm{AGB}_{\mathrm{proxy}} = \left(b_0 + \sum_{i=1}^{4} b_i x_i \right)^2, \quad (1)$$

where b_i are the model coefficients. We used a simulated annealing based variable selection (Packalen et al., 2012) to choose the four predictors x_i freely from the ALS and Sentinel-2 metrics. Simulated annealing is a heuristic optimization algorithm that uses a stochastic process to find an approximate global optimum. In each iteration of the simulated annealing, model (1) was fitted to the training data using nonlinear least squares (function `gnls` in the R package `nlme` (Pinheiro et al., 2020)) and the predictors selected at that iteration. The goodness of fit was evaluated using root mean square error (RMSE). Simulated annealing was run for 10 000 iterations to find the predictors that produced the lowest RMSE.

2.5. ICESat −2 data

All ATL03 (Neumann et al., 2021) and ATL08 (Neuenschwander et al., 2021) data (version 4) captured from the study area during the years 2018 and 2019 were used in this study, covering the period from October 2018 to December 2019. These ICESat −2 tracks are shown in Fig. 3. Data was available on 29 discrete dates, resulting in a good temporal and spatial coverage of the study area.

The snow cover flag from ATL08 was replaced by a 10 km resolution national daily snow depth raster produced by the Finnish meteorological institute (Aalto et al., 2016). While the snow cover info from the national raster corresponded exactly with the ATL08 snow cover flag during the core winter season, there was a significant divergence in autumn during the first snowfalls, and in spring during the snow melt season (Fig. 4).

2.6. ICESat −2 track segmentation

ICESat −2 tracks were split into 90 × 15 m track segments. Each segment consisted of six 15 × 15 m cells (Fig. 5), for which the AGBs were predicted using the ALS model. The centers of the 90 m segments were placed on locations of the ATL08 100 m segments, to facilitate the concurrent use of ATL03 and ATL08 data.

To discard 90 m track segments that were not fully forested, a forest mask was created by combining the CORINE land cover product (European Environment Agency (EEA), 2018) and a canopy height mask derived from the wall-to-wall ALS data. The ALS height mask was constructed using the maximum ALS echo heights with the primary purpose of removing clear cut areas classified as forest in CORINE. If the maximum height was less than 2 m, the CORINE pixel was reclassified as not forested.

After the non-forested 90 m track segments were removed, the ALS data were intersected with the 15 m segment subcells and processed as in Section 2.2. Sentinel-2 reflectances were then extracted for the 15 m subcells. Next, the ALS-based biomass model (Eq. (5)) was used to predict $\mathrm{AGB}_{\mathrm{ALS}}$ on the subcells. The final reference AGB for the 90 m track segments was obtained by averaging the predictions of the six subcells.

Fig. 2. Location of Valtimo study area in Finland.

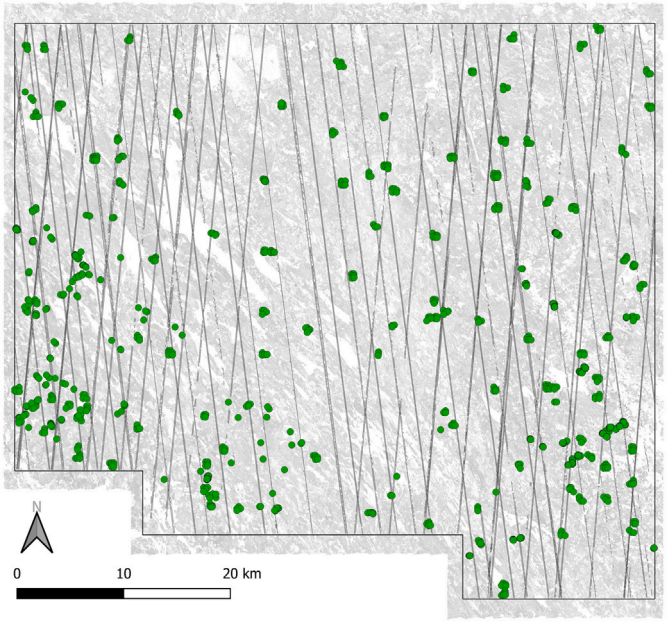

Fig. 3. Locations of field plots (green dots) and ICESat −2 tracks in the study area, superimposed on canopy height model.

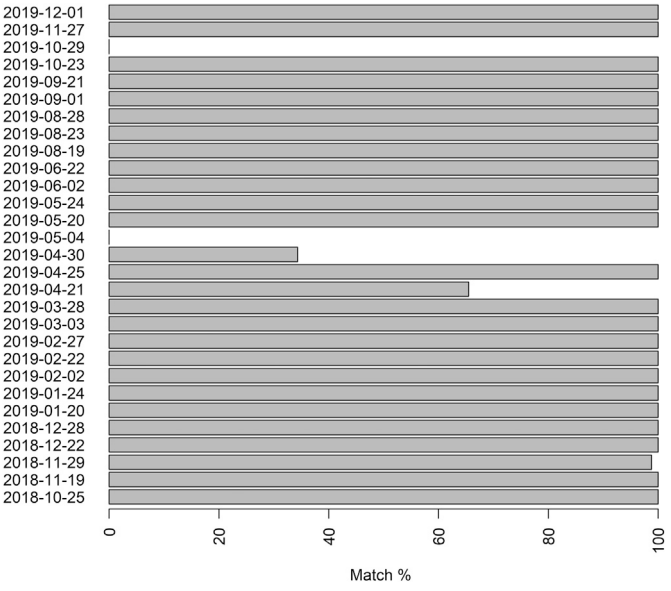

Fig. 4. Comparison of ATL08 and FMI snow cover data by imaging date. Match % is the percentage of track segments with a similar snow status in ATL08 and the FMI snow rasters by date.

2.7. ICESat −2 biomass modeling

We used the ATL03 Global Geolocated Photon Data (Neumann et al., 2021) and ATL08 Land and Vegetation Height (Neuenschwander et al., 2021) data products (version 4). The ATL03 product contains the geolocated photons, i.e. location and height of each measured photon. The ATL08 product includes, in addition to the vegetation height products at 100 m resolution, photon classifications to noise, ground, canopy or top of canopy for each ATL03 photon. Thus, the first step in the ICESat −2 data processing was to link the individual photon classifications from ATL08 with ATL03 photon locations.

After linking the ATL08 photon classifications with the corresponding ATL03 photons, the ATL03 photons unclassified in ATL08 were omitted. The resulting photon clouds containing only the ATL08 classified photons were then clipped to the 90 m track segments. Using the ATL08 classified ground photons, ellipsoidal photon heights relative to the WGS84 ellipsoid were normalized to heights above ground level. A set of ABA metrics were finally calculated without a height threshold. These included the number of photons (canopy only n_c and total n_{all}), average photon height, standard deviation, maximum, height percentiles $p_5, p_{10}, p_{20}, \ldots, p_{90}, p_{95}, p_{99}$, canopy density percentiles $b_5, b_{10}, b_{20}, \ldots, b_{90}, b_{95}$, and average square height (qav).

To omit segments with either poor quality data or too few photons, we first calculated the fraction of photons with a high ATL03 photon signal confidence flag (`signal_conf_ph`) for each segment. This quality measure was then applied with the number of ATL08 classified photons on each segment. Segments with a fraction of high confidence photons less than 0.6 or less than 100 classified photons were omitted.

For modeling, the data were further subdivided based on the presence of snow on the ground, time of day (night/day, based on nautical dawn i.e. sun elevation less than −12°), and weak or strong ATLAS beam. The combination snowless/night/weak ($n = 13$) was omitted due to the low number of segments satisfying the quality threshold.

Linear mixed effect models between square-root transformed ALS and Sentinel-2 -predicted proxy AGB and ICESat −2 predictors were

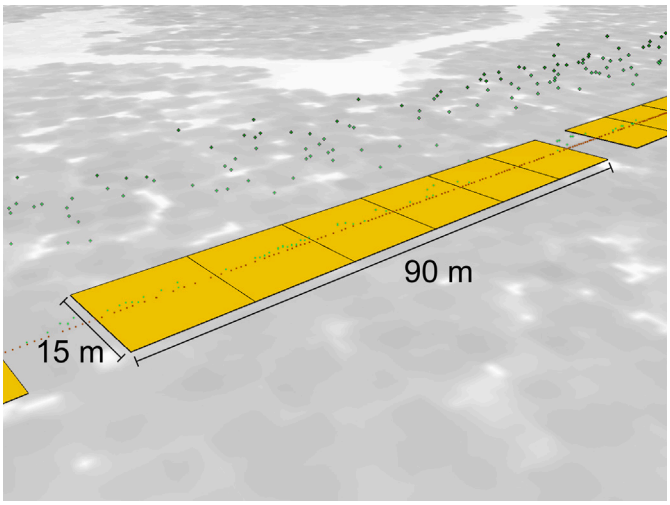

Fig. 5. ICESat −2 track segmentation scheme: one 90 m × 15 m track segment consists of six 15 m × 15 m cells for which the reference AGBs were predicted using ALS and Sentinel-2 data. ATL03 photons are colored by classification.

then formulated:

$$\sqrt{\widehat{\mathrm{AGB}}_{\mathrm{proxy}}} = b_0 + \sum_{i=1}^{5} b_i x_i + b_6 \alpha_s + b_7 d_{\mathrm{snow}} + b_8 \mathrm{strong} + \mu_j, \quad (2)$$

The fixed part of each model included intercept, five ICESat −2 ABA metrics x_i (either untransformed or square-root transformed), sun elevation angle $\alpha_s \in [-90°, 90°]$, and when applicable, snow depth d_{snow} in centimeters from the FMI snow depth raster. Strong/weak beam was included as a dummy variable (strong) in models that combined strong and weak beam data. Random intercept μ_j in the model was grouped by acquisition date j to account for varying weather and seasonal effects. The five predictor variables of each model were selected using simulated annealing (Packalen et al., 2012). In the annealing, optimization was first performed using only the fixed part for 10 000 iterations, which was repeated 10 times. The best solution was then used as a starting point for the variable selection of the full mixed model, which was further annealed for 3 repetitions of 10 000 iterations. In each iteration, the model (2) was fit using restricted maximum likelihood with the function `lme` from the R package `nlme` and the goodness of fit was evaluated using RMSE. Finally, the predicted $\sqrt{\widehat{\mathrm{AGB}}_{\mathrm{proxy}}}$ were back-transformed and bias corrected (Gregoire et al., 2008).

2.7.1. Model evaluation

The performance of the ICESat −2 biomass models was evaluated by calculating RMSE and the relative RMSE (RMSE%) of the fitted model values to the proxy AGB. RMSE and RMSE% are defined as:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\widehat{\mathrm{AGB}} - \widehat{\mathrm{AGB}}_{\mathrm{proxy}}\right)^2} \quad (3)$$

$$\mathrm{RMSE\%} = \frac{\mathrm{RMSE}}{\frac{1}{n}\sum_{i=1}^{n}\widehat{\mathrm{AGB}}_{\mathrm{proxy}}} \quad (4)$$

3. Results and discussion

3.1. ALS-based biomass model

First, the proxy biomass model was fitted to the training data as described in Section 2.4. The resultant model was

$$\widehat{\mathrm{AGB}}_{\mathrm{proxy}} = (0.44\ \mathrm{avg}_f + 0.34\ \mathrm{avg}_l - 0.0013\ \mathrm{NIR} - 0.0012\ \mathrm{SWIR1} + 8.52)^2, \quad (5)$$

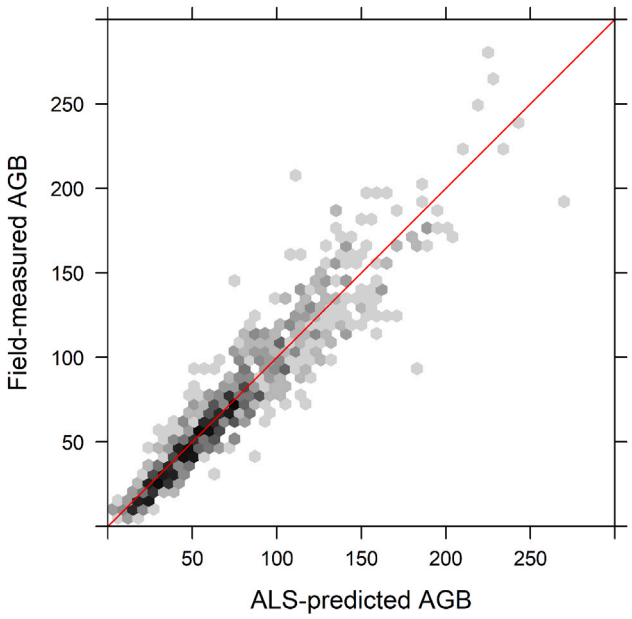

Fig. 6. Scatter density plot of the field-plot AGB vs. AGB predicted from ALS using model (5).

where avg_f and avg_l are the average ALS heights calculated from first of many and only, and from last of many and only echoes, respectively. NIR and SWIR1 are the Sentinel-2 reflectances from the near-infrared and first shortwave-infrared bands. The resultant model had an RMSE of 15.1 Mg/ha for AGB (RMSE% 20.2%). The scatter plot of the fitted ALS biomass model is shown in Fig. 6.

3.2. ICESat −2 biomass model

The RMSEs of the submodels are listed in Table 1 in a descending order of accuracy. The table also includes the number of 90 m track segments, the number of different imaging dates and the standard deviation of the random intercept. Due to the omittance of snowless/night/weak class, some general combinations are missing, for example the snowless/weak class that is equivalent to the snowless/day/weak class. Snowless/night/strong submodel had the best RMSE% performance at 26.9%. In general, the specific submodels had lower RMSE, with, for example, the model using all the data having relatively poor RMSE% of 36.7%. Snowless (31.4%) produced considerably better results than snowy data (36.9%). Furthermore, strong beam data (34.0%) was better than weak beam data (39.0%), and night data (34.3%) had a slightly better performance than day data (36.4%). RMSE% of the combined models when applied to the specific subsets are shown in Table 2 and density scatter plots of the submodels are shown in Fig. 7. The significance of including the random intercept was tested using likelihood ratio test (e.g. Mehtätalo and Lappi, 2020). The inclusion of random effect was significant in all cases, except snow/night/weak and snowless/day/weak.

Based on the performance of the submodels, the best results are obtained in snowless conditions using strong beam data captured during night. The merged models showed worse performance than specific models across the line when applied to the specific subsets. For example, the RMSE% of the best merged model, snowless/strong (30.2%), was better than the RMSE% of subset model snowless/day/strong (34.1%). However, when this merged model was applied to the snowless/day/strong data (Table 2), its performance was slightly worse at 35.3%.

The selected variables and the fixed part weights are in Table 3. Most commonly chosen predictor variables included the average (avg)

Table 1
RMSE and relative RMSE of the submodels in a descending order of RMSE%. The number of 90 m track segments and imaging dates per subset are shown. The column $\sqrt{\mathrm{Var}(\mu_j)}$ shows the absolute magnitude of the standard deviation of the random intercept, the standard deviation relative to the fixed intercept is shown in parentheses. Significance of the random intercept was tested using likelihood ratio test, shown below the $\sqrt{\mathrm{Var}(\mu_j)}$ values: ns = non-significant, * = $p \leq 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$.

Data	RMSE [Mg/ha] (RMSE%)	n	n_{date}	$\sqrt{\mathrm{Var}(\mu_j)}$
Snowless/night/strong	16.2 (26.9%)	3589	4	0.1 (0.03) ^{ns}
Snowless/strong	17.7 (30.2%)	5091	14	1.0 (0.4) ***
Snow/night/strong	18.6 (30.4%)	2699	5	1.2 (0.1) ***
Night/strong	18.6 (30.7%)	6288	9	2.5 (0.4) *
Snow/night/weak	19.2 (31.0%)	2567	3	0.0 (0.0) ***
Snowless	18.4 (31.4%)	5412	14	0.6 (0.1) ^{ns}
Snow/strong	19.4 (32.8%)	8922	16	0.7 (0.1) ***
Snow/day/strong	19.5 (33.5%)	6223	11	0.6 (0.1) ***
Strong	20.0 (34.0%)	14013	29	1.2 (0.2) ***
Snowless/day/strong	18.8 (34.1%)	1502	10	0.9 (1.2) ***
Night	20.9 (34.3%)	8868	9	3.5 (1.4) ***
Day/strong	20.1 (34.9%)	7725	20	0.9 (0.1) ***
Snow/day	20.4 (35.6%)	8025	11	0.8 (0.1) ***
Snow/night	22.0 (35.8%)	5266	5	15.5 (0.4) ***
Day	20.8 (36.4%)	9835	20	1.0 (0.2) ***
Snow/weak	21.5 (36.5%)	4369	12	1.1 (0.1) ***
All	21.6 (36.7%)	18703	29	1.4 (0.2) ***
Snow	21.8 (36.9%)	13291	16	0.9 (0.1) ***
Snowless/day	20.7 (37.2%)	1810	10	0.5 (0.1) ***
Weak	23.0 (39.0%)	4690	23	1.2 (0.1) ***
Snow/day/weak	22.1 (40.3%)	1802	9	1.5 (0.2) ***
Day/weak	23.3 (42.2%)	2110	16	1.0 (0.1) ***
Snowless/day/weak	25.6 (44.3%)	308	8	0.0 (0.0) ^{ns}

Table 2
RMSE% of the combined models when applied to each subset. For example, snowless/strong model has RMSE% of 28.2% when applied only to snowless/night/strong data.

	RMSE%				
Data subset	Specific	Combined models			All
		Night/strong	Snowless/strong	Snowless/night	
Snowless/night/strong	26.9%	29.1%	28.1%	26.9%	33.1%
		Night/strong	Snow/strong	Snow/night	
Snow/night/strong	30.4%	32.5%	30.8%	34.1%	33.3%
		Night/weak	Snow/weak	Snow/night	
Snow/night/weak	31.0%	–	32.9%	37.6%	39.3%
		Day/strong	Snow/strong	Snow/day	
Snow/day/strong	33.5%	34.2%	33.7%	33.6%	36.6%
		Day/strong	Snowless/strong	Snowless/day	
Snowless/day/strong	34.1%	37.7%	35.3%	34.5%	38.7%
		Day/weak	Snow/weak	Snow/day	
Snow/day/weak	40.3%	41.1%	42.0%	42.3%	43.7%
		Day/weak	Snowless/weak	Snowless/day	
Snowless/day/weak	44.3%	47.9%	–	47.3%	76.7%

and average square height (qav), and their transformations. Furthermore, every submodel except snowless/night/strong and snow/day/weak included some form of a photon count metric, n_c or n_{all}, in several cases both.

The standard deviations of the date-specific random intercepts $\sqrt{\mathrm{Var}(\mu_j)}$ (Table 1) showed some variation when compared relatively to the fixed intercepts (Table 3). In most cases, the standard deviations were relatively small compared to the fixed intercepts, but on the submodels snowless/day/strong and night, the standard deviation was larger than the fixed intercept. As the random intercept was grouped by the acquisition date, large standard deviation of the random intercept compared to the fixed intercept means that there is more variation between the data captured on different dates. On the other hand, the two cases (snowless/night/strong and snow/day/weak) where the random intercept was found to be non-significant and having a standard deviation close to zero, implies that temporal differences in the data were smaller.

Table 4 presents the predicted proxy AGBs of the subclasses, ICESat −2 p_{99} calculated from the ATL08-classified ATL03 photons, ALS-derived (first of many and only echoes) p_{99}, used as a proxy for predicted dominant height, and canopy cover from ALS. The 10th and 90th percentiles of the variables are shown in parentheses. AGB_{ALS} and canopy cover showed slight variation between the subclasses due to uneven geographic distribution, but did not show any clear relationship with the model performance. Comparison between the ICESat −2 and ALS-derived p_{99} showed that good performance is likely, when these distributions are close. In the worst performing submodels there was a noticeable tendency that ICESat −2 p_{99} is circa 4 meters lower than its ALS equivalent. The effect was pronouced in the 10th percentiles: the ICESat −2 derived p_{99} were close to 3 meters for the majority

Table 3
Fixed parts of the fitted submodels. In the same order as Table 1. Statistical significance of each model parameter is shown under the parameter: ns = non-significant, * = $p \leq 0.05$, ** = $p \leq 0.01$, *** = $p \leq 0.001$.

Snowless/night/strong	$2.9_{***} + 2.5\text{avg}_{***} - 0.06\text{qav}_{***} + 1.1\text{std}_{***} + 0.2\sqrt{n_c}_{***} - 5.6\sqrt{\text{avg}}_{***} - 0.05\alpha_{s\,***}$
Snowless/strong	$2.6_{***} + 0.002n_{\text{all}\,***} - 0.3p_{20\,***} + 0.1\sqrt{n_c}_{***} + 0.8\sqrt{\text{qav}}_{***} - 0.3\sqrt{p_{90}}_{***} + 0.006\alpha_{s\,ns}$
Snow/night/strong	$11.6_{***} - 0.006n_{\text{all}\,***} - 0.02\text{qav}_{***} + 0.4\sqrt{n_c}_{***} - 8.1\sqrt{\text{avg}}_{***} + 2.4\sqrt{\text{qav}}_{***} - 0.04d_{\text{snow}\,*} + 0.1\alpha_{s\,ns}$
Night/strong	$6.7_{***} - 0.005n_{\text{all}\,***} - 0.02b_{30\,***} + 0.4\sqrt{n_c}_{***} - 4.0\sqrt{\text{avg}}_{***} + 1.3\sqrt{\text{qav}}_{***} - 0.03d_{\text{snow}\,*} - 0.003\alpha_{s\,ns}$
Snow/night/weak	$-2.7_{ns} - 0.1n_{\text{all}\,***} + 0.6\text{std}_{***} - 0.02b_{10\,***} + 2.4\sqrt{n_{\text{all}}}_{***} + 0.5\sqrt{\text{max}}_{***} + 0.02d_{\text{snow}\,***} + 0.08\alpha_{s\,***}$
Snowless	$6.5_{***} + 1.3\text{avg}_{***} - 0.07\text{qav}_{***} + 0.2\sqrt{n_c}_{***} - 6.0\sqrt{\text{avg}}_{***} + 1.9\sqrt{\text{qav}}_{***} - 2.3\text{strong}_{***} + 0.006\alpha_{s\,ns}$
Snow/strong	$6.5_{***} - 0.005n_{\text{all}\,***} - 0.03\text{qav}_{***} + 0.4\sqrt{n_c}_{***} - 7.0\sqrt{\text{avg}}_{***} + 2.3\sqrt{\text{qav}}_{***} + 0.001d_{\text{snow}\,ns} + 0.01\alpha_{s\,ns}$
Snow/day/strong	$8.2_{***} - 0.04\text{qav}_{***} + 0.4\sqrt{n_c}_{***} - 0.2\sqrt{n_{\text{all}}}_{***} - 6.8\sqrt{\text{avg}}_{***} + 2.3\sqrt{\text{qav}}_{***} + 0.01d_{\text{snow}\,ns} + 0.08\alpha_{s\,*}$
Strong	$6.4_{***} - 0.004n_{\text{all}\,***} - 0.01b_{30\,***} + 0.3\sqrt{n_c}_{***} - 4.5\sqrt{\text{avg}}_{***} + 1.5\sqrt{\text{qav}}_{***} - 0.002d_{\text{snow}\,ns} + 0.01\alpha_{s\,ns}$
Snowless/day/strong	$0.8_{ns} - 0.003n_{\text{all}\,**} - 0.4p_{20\,***} + 0.04p_{99\,**} + 0.3\sqrt{n_{\text{all}}}_{***} + 0.7\sqrt{\text{qav}}_{***} - 0.03\alpha_{s\,ns}$
Night	$-2.6_{ns} + 1.1\text{std}_{***} + 0.2\sqrt{n_c}_{***} - 0.2\sqrt{n_{\text{all}}}_{***} + 0.4\sqrt{p_{50}}_{***} - 0.7\sqrt{p_{95}}_{***} - 0.04d_{\text{snow}\,**} + 2.1\text{strong}_{***} - 0.3\alpha_{s\,**}$
Day/strong	$7.4_{***} - 0.03\text{qav}_{***} + 0.3\sqrt{n_c}_{***} - 0.2\sqrt{n_{\text{all}}}_{***} - 5.8\sqrt{\text{avg}}_{***} + 2.0\sqrt{\text{qav}}_{***} - 0.0002d_{\text{snow}\,ns} - 0.05\alpha_{s\,ns}$
Snow/day	$7.7_{***} - 0.04\text{qav}_{***} + 0.4\sqrt{n_c}_{***} - 0.2\sqrt{n_{\text{all}}}_{***} - 6.3\sqrt{\text{avg}}_{***} + 2.2\sqrt{\text{qav}}_{***} + 3e-04d_{\text{snow}\,ns} + 0.7\text{strong}_{***} + 0.09\alpha_{s\,*}$
Snow/night	$-40.7_{***} + 1.7\text{std}_{***} + 0.2\sqrt{n_c}_{***} - 0.3\sqrt{n_{\text{all}}}_{***} + 0.6\sqrt{\text{max}}_{***} - 4.1\sqrt{\text{std}}_{***} - 0.009d_{\text{snow}\,ns} + 2.5\text{strong}_{***} - 1.3\alpha_{s\,***}$
Day	$5.8_{***} - 0.004n_{\text{all}\,***} - 0.03\text{qav}_{***} + 0.3\sqrt{n_c}_{***} - 5.6\sqrt{\text{avg}}_{***} + 2.0\sqrt{\text{qav}}_{***} - 0.02d_{\text{snow}\,ns} + 0.1\text{strong}_{*} - 0.009\alpha_{s\,ns}$
Snow/weak	$9.1_{***} - 0.02n_{\text{all}\,***} + 1.5\text{std}_{***} + 0.2\sqrt{n_c}_{***} + 0.8\sqrt{\text{max}}_{***} - 3.5\sqrt{\text{std}}_{***} + 0.002d_{\text{snow}\,ns} + 0.02\alpha_{s\,ns}$
All	$6.5_{***} + 1.8\text{std}_{***} + 0.2\sqrt{n_c}_{***} - 0.1\sqrt{n_{\text{all}}}_{***} + 0.5\sqrt{\text{max}}_{***} - 4.3\sqrt{\text{std}}_{***} - 0.01d_{\text{snow}\,*} + 0.8\text{strong}_{***} + 0.0004\alpha_{s\,ns}$
Snow	$7.6_{***} + 1.7\text{std}_{***} + 0.01b_{05\,***} + 0.3\sqrt{n_c}_{***} - 0.2\sqrt{n_{\text{all}}}_{***} - 3.4\sqrt{\text{std}}_{***} - 0.005d_{\text{snow}\,ns} + 1.0\text{strong}_{***} + 0.01\alpha_{s\,ns}$
Snowless/day	$5.7_{***} - 0.04\text{qav}_{***} + 0.08\sqrt{n_{\text{all}}}_{***} - 3.1\sqrt{\text{avg}}_{***} + 1.9\sqrt{\text{qav}}_{***} + 0.3\sqrt{p_{60}}_{***} - 2.6\text{strong}_{***} - 0.01\alpha_{s\,ns}$
Weak	$9.7_{***} - 0.03n_{c\,***} + 0.7\text{std}_{***} + 0.6\sqrt{n_c}_{***} - 0.3\sqrt{n_{\text{all}}}_{***} - 0.2\sqrt{b_{20}}_{***} - 0.03d_{\text{snow}\,**} + 0.03\alpha_{s\,*}$
Snow/day/weak	$7.2_{***} - 0.01n_{\text{all}\,***} + 0.8\text{std}_{***} + 0.1b_{05\,***} + 0.3\sqrt{n_c}_{***} - 1.0\sqrt{b_{05}}_{***} - 0.04d_{\text{snow}\,*} + 0.2\alpha_{s\,ns}$
Day/weak	$8.6_{***} + 0.008n_{c\,***} + 1.8\text{std}_{***} - 0.2\sqrt{n_{\text{all}}}_{***} + 1.4\sqrt{\text{max}}_{***} - 5.3\sqrt{\text{std}}_{***} - 0.04d_{\text{snow}\,**} + 0.08\alpha_{s\,*}$
Snowless/day/weak	$6.9_{***} + 0.1\text{max}_{***} - 4.1\text{avg}_{***} + 0.3\text{qav}_{***} + 3.3\sqrt{p_{30}}_{**} + 0.8\sqrt{p_{80}}_{*} + 0.01\alpha_{s\,ns}$

Table 4
The averages of proxy AGB, ICESat-2 p_{99}, ALS p_{99}, and ALS-derived canopy cover (CC) for the submodels. 10th and 90th percentiles of the parameters are in parentheses. Order of the submodels is the same as in Table 1.

Data	RMSE%	AGB_{ALS} [Mg/ha]	I2 p99 [m]	ALS p99 [m]	ALS CC [%]
Snowless/night/strong	26.9%	60.1 (20.1,105.3)	14.6 (8.4,19.6)	16.1 (10.6,20.9)	63.6 (36.9,86.9)
Snowless/strong	30.2%	58.6 (19.4,103.7)	14.0 (7.6,19.2)	15.9 (10.4,20.7)	63.3 (36.1,86.3)
Snow/night/strong	30.4%	61.2 (21.9,103.2)	10.4 (1.2,18.0)	16.3 (11.1,21.0)	65.4 (40.3,86.1)
Night/strong	30.7%	60.6 (21.1,104.6)	12.8 (2.2,19.1)	16.2 (10.8,20.9)	64.4 (38.2,86.5)
Snow/night/weak	31.0%	61.8 (23.3,103.5)	13.4 (6.0,19.1)	16.5 (11.3,21.0)	65.3 (41.0,86.2)
Snowless	31.4%	58.6 (19.4,103.7)	14.0 (7.6,19.2)	15.9 (10.4,20.7)	63.3 (36.1,86.3)
Snow/strong	32.8%	59.1 (21.7,98.2)	11.3 (1.7,17.9)	16.1 (10.9,20.5)	64.3 (39.0,85.9)
Snow/day/strong	33.5%	58.1 (21.7,96.2)	11.7 (2.6,17.8)	16.0 (10.9,20.4)	63.8 (38.7,85.7)
Strong	34.0%	58.9 (21.1,100.1)	12.3 (2.6,18.4)	16.0 (10.7,20.6)	63.9 (38.2,86.0)
Snowless/day/strong	34.1%	55.1 (17.5,95.8)	12.5 (5.7,18.0)	15.3 (9.9,20.2)	62.6 (34.8,85.3)
Night	34.3%	61.0 (21.5,104.6)	13.0 (3.4,19.1)	16.3 (11.0,20.9)	64.7 (39.1,86.4)
Day/strong	34.9%	57.5 (21.0,96.1)	11.9 (2.9,17.9)	15.8 (10.6,20.4)	63.6 (38.1,85.6)
Snow/day	35.6%	57.4 (20.6,95.8)	11.6 (2.6,17.9)	15.9 (10.7,20.4)	63.3 (37.5,85.7)
Snow/night	35.8%	61.5 (22.4,103.3)	11.9 (1.9,18.6)	16.4 (11.2,21.0)	65.4 (40.6,86.2)
Day	36.4%	57.0 (20.3,95.8)	11.7 (2.9,17.9)	15.8 (10.5,20.4)	63.2 (37.2,85.6)
Snow/weak	36.5%	58.9 (20.5,99.6)	12.5 (4.7,18.7)	16.1 (10.8,20.8)	63.7 (38.1,85.9)
All	36.7%	58.9 (20.9,100.0)	12.3 (3.0,18.5)	16.0 (10.7,20.6)	63.9 (38.2,86.0)
Snow	36.9%	59.0 (21.3,98.7)	11.7 (2.1,18.1)	16.1 (10.9,20.6)	64.1 (38.7,85.9)
Snowless/day	37.2%	55.1 (17.5,95.8)	12.5 (5.7,18.0)	15.3 (9.9,20.2)	62.6 (34.8,85.3)
Weak	39.0%	58.9 (20.5,99.6)	12.5 (4.7,18.7)	16.1 (10.8,20.8)	63.7 (38.1,85.9)
Snow/day/weak	40.3%	54.8 (18.0,93.8)	11.1 (2.7,17.9)	15.6 (9.9,20.4)	61.4 (34.0,85.6)
Day/weak	42.2%	54.8 (18.0,93.8)	11.1 (2.7,17.9)	15.6 (9.9,20.4)	61.4 (34.0,85.6)
Snowless/day/weak	44.3%	57.7 (22.3,90.7)	4.3 (0.0,11.9)	15.7 (11.5,19.6)	64.3 (40.5,86.2)

of the submodels, while the 10th percentiles p_{99} were close to 10 meters in the ALS data. The underestimation of canopy heights in a similar boreal setting was previously reported by Neuenschwander et al. (2020), where data captured during snowy conditions similarly showed larger underestimation than snowless data.

Night data being generally better than day data was expected, as the increased solar noise during daytime will make noise filtering more difficult. However, as recently reported by Neuenschwander et al. (2022), day data generally has fewer signal photons than night data in snowless forest conditions, with the most pronounced discrepancy in weak beam data. Our results indicate that similar behavior likely happens also in snowy conditions: while snow/night/weak data work quite well with RMSE of 31.0%, snow/day/weak is one of the worst performing subsets at RMSE of 40.3%.

Due to ATLAS instrument operating in the visible green spectrum (532 nm), strong backscattering from snow can be expected, which

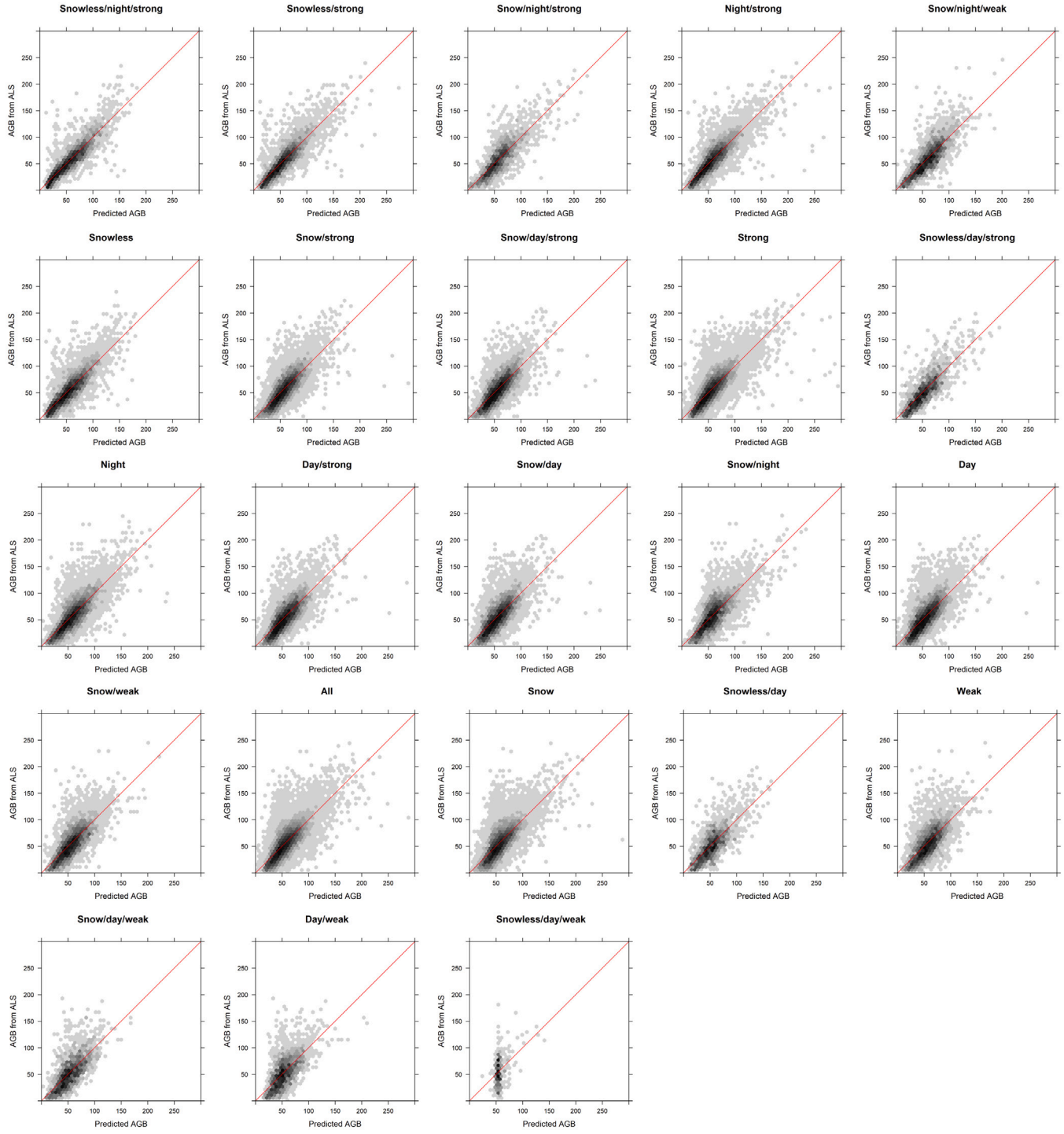

Fig. 7. Density scatter plots of the submodels, in the same order as in Table 1.

makes data captured from snowy forests different from data captured on snowless conditions. The ratio of canopy and ground reflectance is thus expected to vary strongly between snowless and snowy conditions (e.g. Duong et al., 2008), which in a photon counting setting like ICESat −2 can strongly affect canopy metrics, such as canopy cover (Neuenschwander et al., 2022). The average number of ground and canopy photons per segment versus the day of the year is shown in Fig. 8. While there is variation in the photon rates due to weather effects and different spatial distributions of the segments, the winter data have significantly more photons than snowless data. Yet, submodels using data from snowy season have worse performance. Possible explanation may be that the reflectance difference is further affected by presence of frost or snow on the trees, which usually is much less permanent than snow on the ground. Possible evidence of this is the larger canopy photon counts seen in the beginning of the year in Fig. 8. The varying canopy frost and snow conditions could thus make winter data more heterogeneous and harder to model. The presence of heavy snow load in the canopy could also affect tree allometry and canopy cover, which can further obfuscate the relationship between forest AGB and ICESat −2 data. While the study area is dominated by evergreen conifers, there is some deciduous admixture. In winter, the deciduous trees can

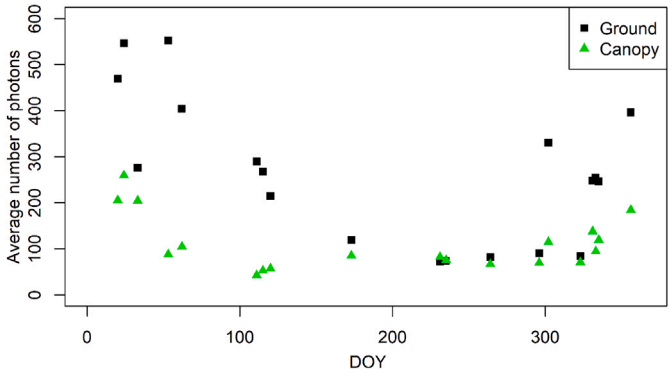

Fig. 8. Average number of ground (black square) and canopy (green triangle) photons per segment versus day of the year. Dates with less than 100 segments were omitted.

be expected to scatter fewer photons, due to the smaller surface area caused by the absence of leaves.

Another possible aspect is the increased cloud cover during the winter months. In general, the quality thresholding used in this study seemed to remove most of the cloud-contaminated data based on the ATLAS cloud confidence flag of the remaining track segments, but some potentially clouded segments remained. During the course of research, we tested discarding also the remaining segments with a cloud confidence flag higher than 1, but it did not improve the performance clearly.

The application of quality criteria to clean the data of unreliable observations is crucial for obtaining reliable AGB estimates from ICESat −2 data. More stringent photon number or canopy photon number thresholding can improve the model performance by discarding track segments with a poor canopy signal, but as a drawback the number of segments is also reduced. Furthermore, given the different radiometric properties of e.g. sapling stands and mature forests, too stringent quality thresholding may result in an unbalanced representation of forest types.

4. Conclusions

In this study, we investigated the effects of acquisition time of day (day/night), beam strength and snow cover on boreal forest AGB estimation using ICESat −2 data. We also examined how merging data acquired in different conditions affected the model performance. This study used a multi-phase modeling framework, where ALS and Sentinel-2 based AGB model was used to produce proxy AGB values on the ICESat −2 track locations. Mixed effect models were then constructed for each data subset and evaluated.

The results indicated that strong beam night data from snow-free conditions performed best, with the clearly smallest RMSE of 26.9%. However, in the boreal forest zone the availability of such data could be limited, because in the midsummer the sun will stay the above the horizon or just barely under it throughout the night. Only four out of thirty passes from our study area represented this type of data. The next best data subsets were snow/night/strong (30.4%), snow/night/weak (31.0%), snow/day/strong (33.5%), and snowless/day/strong (34.2%).

In general, we found that models built for the specific data subset are superior to the merged models. While some of the merged models, such as snowless/strong (30.2%), represented a middle ground between the respective subset models, they performed slightly worse than the specific models when applied to the respective data subset. Furthermore, we found that data captured in snowless conditions performed better than snowy data, night data was slightly better than day data, and strong beam data was superior to weak beam data. While snowy data did not produce the best model performance, the results support that ICESat −2 data captured in snowy conditions could be utilized for AGB estimation in evergreen boreal forests.

A typical spaceborne lidar application scenario would involve prediction of AGB outside of the area where the model was trained, i.e. a model transfer scenario, which will be our objective for the next study. In the next study, we will also explore inference and uncertainty quantification of the predicted AGB.

CRediT authorship contribution statement

P. Varvia: Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization. **L. Korhonen:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition. **A. Bruguière:** Software, Investigation. **J. Toivonen:** Software, Validation, Investigation. **P. Packalen:** Conceptualization, Methodology, Writing – review & editing. **M. Maltamo:** Conceptualization, Writing – review & editing. **S. Saarela:** Conceptualization, Writing – review & editing. **S.C. Popescu:** Conceptualization, Writing – review & editing.

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.

Acknowledgments

This study was supported by the Academy of Finland (project number 332707). The authors also wish to thank the reviewers for their helpful and constructive comments.

References

Aalto, J., Pirinen, P., Jylhä, K., 2016. New gridded daily climatology of Finland: permutation-based uncertainty estimates and temporal trends in climate. J. Geophys. Res.: Atmos. 121 (8), 3807–3823.

Boudreau, J., Nelson, R.F., Margolis, H.A., Beaudoin, A., Guindon, L., Kimes, D.S., 2008. Regional aboveground forest biomass using airborne and spaceborne LiDAR in Québec. Remote Sens. Environ. 112 (10), 3876–3890.

Dubayah, R., Blair, J.B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S., et al., 2020. The global ecosystem dynamics investigation: High-resolution laser ranging of the Earth's forests and topography. Sci. Remote Sens. 1, 100002.

Duncanson, L., Armston, J., Disney, M., Avitabile, V., Barbier, N., Calders, K., Carter, S., Chave, J., Herold, M., Crowther, T.W., et al., 2019. The importance of consistent global forest aboveground biomass product validation. Surv. Geophys. 40 (4), 979–999.

Duong, V., Lindenbergh, R., Pfeifer, N., Vosselman, G., 2008. Single and two epoch analysis of ICESat full waveform data over forested areas. Int. J. Remote Sens. 29 (5), 1453–1473.

Eerikäinen, K., 2009. A multivariate linear mixed-effects model for the generalization of sample tree heights and crown ratios in the Finnish National Forest Inventory. Forest Sci. 55 (6), 480–493.

European Environment Agency (EEA), 2018. Copernicus Land Monitoring Service. European Union.

Gregoire, T.G., Lin, Q.F., Boudreau, J., Nelson, R., 2008. Regression estimation following the square-root transformation of the response. Forest Sci. 54 (6), 597–606.

Herold, M., Carter, S., Avitabile, V., Espejo, A.B., Jonckheere, I., Lucas, R., McRoberts, R.E., Næsset, E., Nightingale, J., Petersen, R., et al., 2019. The role and need for space-based forest biomass-related measurements in environmental management and policy. Surv. Geophys. 40 (4), 757–778.

Holm, S., Nelson, R., Ståhl, G., 2017. Hybrid three-phase estimators for large-area forest inventory using ground plots, airborne lidar, and space lidar. Remote Sens. Environ. 197, 85–97.

Kotivuori, E., Korhonen, L., Packalen, P., 2016. Nationwide airborne laser scanning based models for volume, biomass and dominant height in Finland. Silva Fennica 50 (4), 1567.

Lefsky, M.A., Harding, D.J., Keller, M., Cohen, W.B., Carabajal, C.C., Del Bom Espirito-Santo, F., Hunter, M.O., de Oliveira Jr., R., 2005. Estimates of forest canopy height and aboveground biomass using ICESat. Geophys. Res. Lett. 32 (22).

Main-Knorn, M., Pflug, B., Louis, J., Debaecker, V., Müller-Wilm, U., Gascon, F., 2017. Sen2Cor For Sentinel-2. In: Image and Signal Processing for Remote Sensing XXIII. Vol. 10427. International Society for Optics and Photonics, 1042704.

Margolis, H.A., Nelson, R.F., Montesano, P.M., Beaudoin, A., Sun, G., Andersen, H.-E., Wulder, M.A., 2015. Combining satellite lidar, airborne lidar, and ground plots to estimate the amount and distribution of aboveground biomass in the boreal forest of North America. Can. J. Forest Res. 45 (7), 838–855.

Markus, T., Neumann, T., Martino, A., Abdalati, W., Brunt, K., Csatho, B., Farrell, S., Fricker, H., Gardner, A., Harding, D., et al., 2017. The Ice, Cloud, and land Elevation Satellite-2 (ICESat-2): science requirements, concept, and implementation. Remote Sens. Environ. 190, 260–273.

Mehtätalo, L., Lappi, J., 2020. Biometry for Forestry and Environmental Data: With Examples in R. Chapman and Hall/CRC.

Mitchard, E.T., Saatchi, S.S., Baccini, A., Asner, G.P., Goetz, S.J., Harris, N.L., Brown, S., 2013. Uncertainty in the spatial distribution of tropical forest biomass: a comparison of pan-tropical maps. Carbon Balance Manage. 8 (1), 1–13.

Narine, L.L., Popescu, S.C., Malambo, L., 2019. Synergy of ICESat-2 and Landsat for mapping forest aboveground biomass with deep learning. Remote Sens. 11 (12), 1503.

Narine, L.L., Popescu, S.C., Malambo, L., 2020. Using ICESat-2 to estimate and map forest aboveground biomass: A first example. Remote Sens. 12 (11), 1824.

Nelson, R., Boudreau, J., Gregoire, T.G., Margolis, H., Næsset, E., Gobakken, T., Ståhl, G., 2009a. Estimating Quebec provincial forest resources using ICESat/GLAS. Can. J. Forest Res. 39 (4), 862–881.

Nelson, R., Margolis, H., Montesano, P., Sun, G., Cook, B., Corp, L., Andersen, H.-E., deJong, B., Pellat, F.P., Fickel, T., et al., 2017. Lidar-based estimates of aboveground biomass in the continental US and Mexico using ground, airborne, and satellite observations. Remote Sens. Environ. 188, 127–140.

Nelson, R., Ranson, K.J., Sun, G., Kimes, D.S., Kharuk, V., Montesano, P., 2009b. Estimating Siberian timber volume using MODIS and ICESat/GLAS. Remote Sens. Environ. 113 (3), 691–701.

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., Magruder, L., Guenther, E., Hancock, S., Purslow, M., 2022. Radiometric assessment of ICESat-2 over vegetated surfaces. Remote Sens. 14 (3), 787.

Neuenschwander, A., Pitts, K., Jelley, B., Robbins, J., Klotz, B., Popescu, S., Nelson, R., Harding, D., Pederson, D., Sheridan, R., 2021. ATLAS/ICESat-2 L3A Land and Vegetation Height, Version 4. NASA National Snow and Ice Data Center Distributed Active Archive Center, Boulder, Colorado USA, http://dx.doi.org/10.5067/ATLAS/ATL08.004, (Accessed April 19th 2021).

Neumann, T., Brenner, A., Hancock, D., Robbins, J., Saba, J., Harbeck, K., Gibbons, A., Lee, J., Luthcke, S., Rebold, T., et al., 2021. ATLAS/ICESat-2 L2A Global Geolocated Photon Data, Version 4. NASA National Snow and Ice Data Center Distributed Active Archive Center, Boulder, Colorado USA, http://dx.doi.org/10.5067/ATLAS/ATL03.004, (Accessed April 19th 2021).

Neumann, T.A., Martino, A.J., Markus, T., Bae, S., Bock, M.R., Brenner, A.C., Brunt, K.M., Cavanaugh, J., Fernandes, S.T., Hancock, D.W., et al., 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.

Packalen, P., Temesgen, H., Maltamo, M., 2012. Variable selection strategies for nearest neighbor imputation methods used in remote sensing based forest inventory. Can. J. Remote Sens. 38 (5), 557–569. http://dx.doi.org/10.5589/m12-046.

Pinheiro, J., Bates, D., et al., 2020. nlme: Linear and nonlinear mixed effects models. R package version 3.1-148 URL https://CRAN.R-project.org/package=nlme.

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.

Repola, J., 2008. Biomass equations for birch in Finland. Silva Fennica 42 (4), 605–624.

Repola, J., 2009. Biomass equations for scots pine and Norway spruce in Finland. Silva Fennica 43 (4), 625–647.

Saarela, S., Holm, S., Healey, S.P., Andersen, H.-E., Petersson, H., Prentius, W., Patterson, P.L., Næsset, E., Gregoire, T.G., Ståhl, G., 2018. Generalized hierarchical model-based estimation for aboveground biomass assessment using GEDI and Landsat data. Remote Sens. 10 (11), 1832.

Wulder, M.A., White, J.C., Bater, C.W., Coops, N.C., Hopkinson, C., Chen, G., 2012. Lidar plots—A new large-area data collection option: Context, concepts, and case study. Can. J. Remote Sens. 38 (5), 600–618.

Zhang, Y., Liang, S., Yang, L., 2019. A review of regional and global gridded forest biomass datasets. Remote Sens. 11 (23), 2744.

Zwally, H.J., Schutz, B., Abdalati, W., Abshire, J., Bentley, C., Brenner, A., Bufton, J., Dezio, J., Hancock, D., Harding, D., et al., 2002. ICESat's laser measurements of polar ice, atmosphere, ocean, and land. J. Geodyn. 34 (3–4), 405–445.