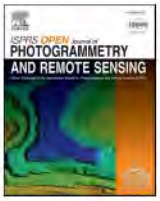
ISPRS Open Journal of Photogrammetry and Remote Sensing 11 (2024) 100053

Contents lists available at ScienceDirect

ISPRS Open Journal of Photogrammetry and Remote Sensing

journal homepage: www.journals.elsevier.com/isprs-open-journal-of-photogrammetry-and-remote-sensing



ICESat-2 noise filtering using a point cloud neural network

Mariya Velikova *, Juan Fernandez-Diaz, Craig Glennie

Department of Civil and Environmental Engineering, University of Houston, Houston, 77004, TX, United States

ARTICLE INFO

Keywords:
ICESat-2
Lidar
Point cloud
Noise filtering

ABSTRACT

The ATLAS sensor onboard the ICESat-2 satellite is a photon-counting lidar (PCL) with a primary mission to map Earth's ice sheets. A secondary goal of the mission is to provide vegetation and terrain elevations, which are essential for calculating the planet's biomass carbon reserves. A drawback of ATLAS is that the sensor does not provide reliable terrain height estimates in dense, high-closure forests because only a few photons reach the ground through the canopy and return to the detector. This low penetration translates into lower accuracy for the resultant terrain model. Tropical forest measurements with ATLAS have an additional problem estimating top of canopy because of frequent atmospheric phenomena such as fog and low clouds that can be misinterpreted as top of the canopy. To alleviate these issues, we propose using a ConvPoint neural network for 3D point clouds and high-density airborne lidar as training data to classify vegetation and terrain returns from ATLAS. The semantic segmentation network provides excellent results and could be used in parallel with the current ATL08 noise filtering algorithms, especially in areas with dense vegetation. We use high-density airborne lidar data acquired along ICESat-2 transects in Central American forests as a ground reference for training the neural network to distinguish between noise photons and photons lying between the terrain and the top of the canopy. Each photon event receives a label (noise or signal) in the test phase, providing automated noise-filtering of the ATL03 data. The terrain and top of canopy elevations are subsequently aggregated in 100 m segments using a series of iterative smoothing filters. We demonstrate improved estimates for both terrain and top of canopy elevations compared to the ATL08 100 m segment estimates. The neural network (NN) noise filtering reliably eliminated outlier top of canopy estimates caused by low clouds, and aggregated root mean square error (RMSE) decreased from 7.7 m for ATL08 to 3.7 m for NN prediction (18 test profiles aggregated). For terrain elevations, RMSE decreased from 5.2 m for ATL08 to 3.3 m for the NN prediction, compared to airborne lidar reference profiles.

1. Introduction

Forests can store a considerable amount of carbon in their living biomass, making them an important carbon sink. Ultimately the carbon gets released back into the atmosphere by mechanisms of respiration, decomposition, or disturbance. Changes in carbon storage are necessary to monitor because they can help mitigate or, conversely - exacerbate climate change. Tropical forests especially have high importance as a carbon sink (Baccini et al., 2012). A substantial part of the storage is concentrated in woody biomass, while the rest of the forest pool is organic forest floor litter and soils. The carbon uptake they provide has been historically very high. Recent studies show that tropical forest carbon sequestration amounts have declined, mainly because of trees mortality and deforestation (Hubau et al., 2020).

Airborne lidar mapping or airborne laser scanning (ALS) can be an essential tool for monitoring forest metrics. Airborne lidar provides accurate wall-to-wall height estimates for both terrain and top of canopy elevations in forested areas, but is cost-prohibitive to provide data on a global scale. Space-borne lidars such as NASA's Ice, Cloud, and Land Elevation Satellite – 2 (ICESat-2) or the Global Ecosystem Dynamics Investigation (GEDI) mission provide global coverage but with only sparse sampling. Nevertheless, various studies have used ICESat-2 data to provide better tree height estimates and, subsequently, better aboveground biomass estimates for forest sites around the world. Initial results using ICESat-2 for vegetation mapping and biomass were obtained using simulated data (Glenn et al., 2016; Gwenzi et al., 2016; Narine et al., 2019; Duncanson et al., 2020). Narine et al. (2020) provided an early assessment of the use of canopy metrics for aboveground biomass (AGB) estimation using ICESat-2 transects over temperate forests in south-east Texas; their results confirmed the utility of ICESat-2 data for characterizing AGB. Neuenschwander et al. (2020) validated terrain and canopy height estimates in boreal forests in southern

* Corresponding author.
E-mail address: mgveliko@cougarnet.uh.edu (M. Velikova).

https://doi.org/10.1016/j.ophoto.2023.100053
Received 22 August 2023; Received in revised form 31 October 2023; Accepted 1 December 2023
Available online 6 December 2023

Finland. They recommended the use of the ATLAS instrument strong beams (vs. using weak beams) for mapping canopy heights. Malambo and Popescu (2021) assessed the agreement between ICESat-2-derived terrain elevations and canopy heights and reference ALS data in various ecozones in the US, reporting that ATL08 terrain heights had better agreement than ATL08 canopy heights with corresponding ALS heights. Contrary to Neuenschwander et al. (2020), they recommended the use of both strong and weak beams for forest height retrievals. Narine et al. (2022) examined the applicability of ICESat-2 as a data source for canopy cover information over forests in south-east Texas and southern Alabama. Their results showed high correlations with ALS canopy cover estimates (R^2 ranging from 0.75 to 0.84), suggesting the possibility of using ICESat-2 data for the development of a gridded canopy cover product. Others have studied the opportunities of integrating ICESat-2 data with GEDI or space-borne imagery and SAR instruments for obtaining aboveground wall-to-wall biomass estimates (Liu et al., 2022; Guerra-Hernández et al., 2022; Luo et al., 2023). Studies using ICESat-2 in tropical forests have not been widely reported in the literature, and previous work has shown that accurate estimation of elevations in tropical forests is challenging (Fernandez-Diaz et al., 2022). Urbazaev et al. (2022) further confirmed that terrain elevations in tropical forests are hard to retrieve and have the lowest accuracy compared to ALS data among a variety of biomes and canopy closures. Musthafa et al. (2023) compared both ICESat-2 and GEDI forest height estimates to field data in tropical and sub-tropical forests in India and found that GEDI achieved lower RMSE values (compared to field reference data) than ICESat-2.

The main instrument onboard the ICESat-2 satellite is the Advanced Topographic Laser Altimeter System (ATLAS). The sensor employs a photon-counting lidar (PCL) with return sensitivity at the single-photon level. The idea behind PCL is to emit low-power pulses and record each returning photon. This mode of operation is in contrast to a discrete mode lidar that needs thousands of photons to reach a predetermined threshold value to record a return (Swatantran et al., 2016; Brown et al., 2020). The number of returning signal photons depends on the outgoing laser energy, surface reflectance of the illuminated area, solar conditions, scattering and attenuation in the atmosphere, and sensitivity of the detector (Neuenschwander et al., 2020). A big challenge for detecting at the single-photon level is the noise introduced by solar photons. The sensor is susceptible to solar noise in the 532 nm operational wavelength, producing point clouds that may contain numerous noise photon events (see Fig. 1). This is because any returned photon, whether from the reflected signal or solar background, can trigger an event within the detector (Neuenschwander et al., 2020).

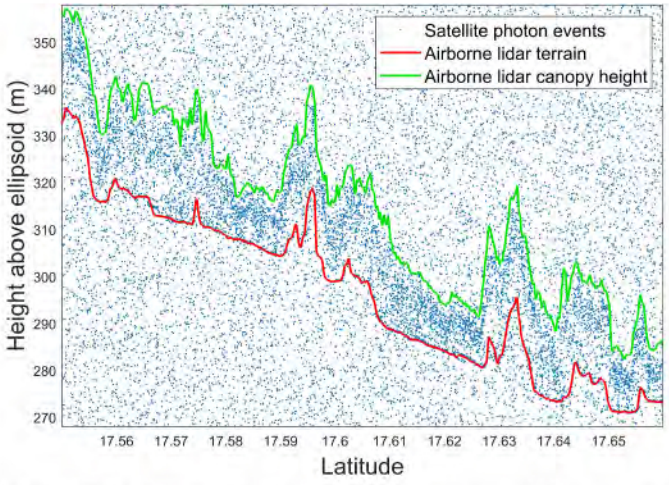


Fig. 1. ATL03 photon events (blue dots) with overlapping ALS estimates of terrain (red line) and top of the canopy (green line). All the photon events above and below the dense band outlined by the airborne lidar data are considered noise. All the events within the outlines are signal, as it contains returns from both terrain and vegetation. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

One of the main challenges for PCL is the reduced canopy penetration compared to linear-mode lidar (Stoker et al., 2016). There is a lower probability for photons to reach the ground and return to the detector through the tree canopy, and as a result, the terrain below the trees often cannot be mapped with high accuracy. This will skew the terrain elevation estimates, which can propagate inaccuracies in the calculation of biomass estimates through vegetation height errors. In the presence of dense vegetation, like in tropical forests, penetration is exacerbated, making it more difficult to estimate terrain elevations under a closed canopy and dense understory. Tropical climates can also result in an increase in incorrect photon event classification and significant outliers. This is primarily caused by the occasional presence of low-lying clouds or fog that can produce photon returns that might be interpreted as top of canopy returns, resulting in significant over-estimates of canopy height, as presented in Fig. 2.

Recent results (Fernandez-Diaz et al., 2022) show that in tropical forests with dense vegetation, terrain elevation estimation errors grow proportionally with the heights of the trees. Outlier values can skew the top of canopy elevation estimates due to atmospheric conditions such as fog or low-lying clouds and the intrinsic sensitivity of the PCL detector. There has been a concerted effort to develop noise filtering algorithms for photon counting lidar, specifically for ICESat-2 data. Most algorithms are based on the idea that signal photons have higher density and are clustered together compared to randomly distributed background noise photons. First derivations of such algorithms were tested on simulated ICESat-2 data collected by airborne photon-counting lidars such as the Multiple Altimeter Beam Experimental Lidar (MABEL) in preparation for the ICESat-2 mission. (Herzfeld et al., 2013; Wang et al., 2016). Popescu et al. (2018) applied multi-level noise filtering to reduce the number of noise photons and then classified the terrain and top of canopy by means of overlapping moving windows and cubic splines. One of the sites where they tested their noise filtering was a tropical forest in Gabon. RMSEs between 3.11 m and 4.48 were reported for night data acquisition and 4.41 m–5.59 m during the day. Since their data was simulated, the performance was not tested against the algorithms used for the ICESat-2 ATL03 or ATL08 products, and we cannot directly compare the obtained in Gabon results to ours. After the launch of ICESat-2, many authors published methods that focused on noise filtering for forest heights retrieval using the newly available ATL03 geolocated photon data. These methods rely on expanding and combining a wide range of traditional approaches for noise filtering such as photon density clustering, neighborhood search within a circle or an ellipse, histogram-based approaches, distance threshold, along with more elaborate combinations of these (Huang et al., 2023; Xie et al., 2022). In Gao et al. (2022), the authors use a succession of large-scale and small-scale search radiuses to improve on the existing DRAGANN (Differential, Regressive, and Gaussian Adaptive Nearest Neighbor) algorithm used for the noise filtering of the released ATL08 data. A further improvement is using an ellipse instead of a circle for the neighborhood search, as seen in Zhu et al. (2020). An even more elaborated search method named local outlier factor algorithm with a rotating search area (LOFR) is presented in He et al. (2023), demonstrating that LOFR algorithm can adaptively adjust the search area and direction. This method works well in complex terrain environments because the search ellipse can follow the topography, thus capturing the correct terrain and top of the canopy. While those methods were reported to be effective and provided good results for their respective testing sites, none of them treated the challenging environment of tropical forest regions and the specific types of noise that could develop in those regions, for example, the presence of low clouds. The only method that was applied exclusively to tropical forests was Li et al. (2020), which proposed an approach based on relative neighbor relationships and locally weighted

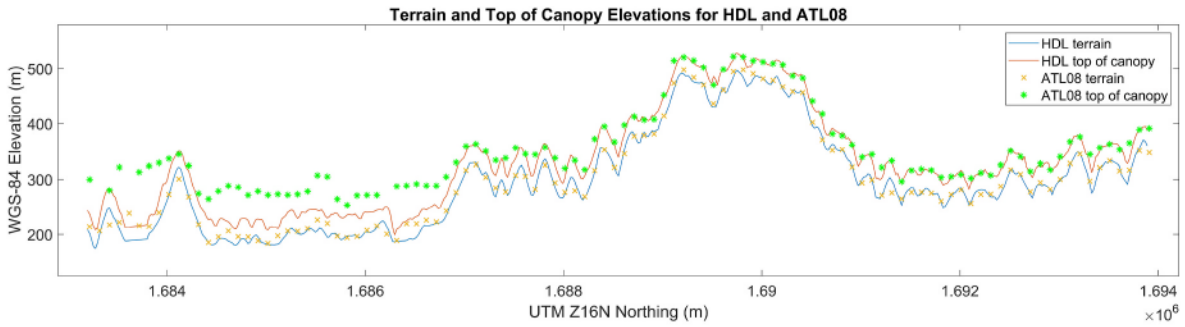


Fig. 2. Example of a profile with an occurrence of both accurate top of canopy elevations (as compared with airborne lidar ground reference) in the right side of the figure, as well as outlier values (probably related to low-lying clouds) in the left side of the figure.

distance statistics. The relative neighboring relationship describes the relative density distribution of the neighboring photon points around two photon points. They use the statistical characteristics of the mean local weighted distance to choose an adaptive threshold that separates the signal from the noise photons. The authors report good results compared to a previously used local distance statistics-based filtering method, but they do not provide any comparison to the standard algorithm used for the data distributed to ICESat-2 ATL08 users. A further drawback is that the location of the study site is not provided, and results are reported for only a single satellite track, which is a very limited test. Currently, two noise filtering methods (Neuenschwander and Pitts, 2019; Neuenschwander et al., 2020; Neumann et al., 2021) are used for the released ATL08 data products to reduce the noise photons before estimating forest heights and terrain elevations, one of which is The Differential, Regressive, and Gaussian Adaptive Nearest Neighbor (DRAGANN). Both methods are described in more detail in section 3.4. While some studies (Liu et al., 2022; Peng et al., 2022) have examined the use of neural networks with ICESat-2, they did not examine noise filtering of ATL03 data or top of canopy (TOC) and terrain estimation in dense forests. Similar to us, Meng et al. (2022) used a neural network for noise filtering, and they reported results only in shallow bathymetry study sites. Unlike the case of a neural network, which allows direct input of 3D points, their method does not directly input the photon events for training the network. They construct a feature vector from a K-nearest neighbors search within a horizontal elliptical region that is used as an input, which results in additional steps. A recent paper by Lin and Knudby (2023) used PointNet++, a neural network for 3D points, to extract bathymetric photons from ICESat-2 data. Their work differs from ours in two fundamental ways: the first is that bathymetric data is inherently easier to classify compared to forested areas, as the photons that delineate the sea floor are in most cases easy to distinguish. The datasets are devoid of high amounts of noise, unlike tropical forests, which do not have any structure that can be easily discerned. Secondly, they use manual classification of the training data, while we employ airborne lidar data for automatic labeling, which significantly speeds and scales the process. In this paper, we propose an approach that, to our knowledge, has yet to be studied: the use of a neural network working directly on the 3D point clouds for automated separation of noise and signal events in dense tropical forests. We are motivated by the large and growing availability of high-density, high-accuracy tropical airborne lidar data that can serve as reference datasets for the training of neural networks. Using airborne lidar data to prepare labeled training datasets is a faster and easier method than fully manual labeling. The neural network method leverages geometric and local signal density information from overlapping airborne lidar measurements to train a deep neural network to separate signal and noise photon events. We use a ConvPoint architecture (Boulch, 2020) for 3D point clouds that is applied to the raw ICESat-2 ATL03 point cloud records, producing reliable and continuous noise filtering results even in densely vegetated areas. The main benefit of the ConvPoint architecture, compared to traditional 2D convolutional neural networks, is the direct application on irregular 3D photon data without conversion to a regularized intermediate representation such as an image. The benefit is two-fold: on the one hand, saving time by directly using readily available ATL03 data product photon events, and on the other, not having data inevitably lost during conversion to a lower dimension format (i.e., from 3D to 2D). We validate our results for ICESat-2 terrain and top of canopy elevations in the dense tropical forests of Mexico, Belize, Guatemala, Honduras, Brazil, and Puerto Rico by comparing the neural network classified data with high-density airborne lidar data from nine study areas. The results show better consistency, removal of outliers, and overall improvement in terrain and top of canopy elevation estimates compared to current ICESat-2 noise filtering algorithms delivered by the ATL08 product. We are examining the most challenging environment for the ATLAS sensor, which is to monitor tropical forests. The method proposed here has the potential to improve terrain and top of canopy estimates in all vegetated environments. The rest of the paper is organized as follows: first, we introduce the study areas and datasets used in this project. We cover in more detail the current noise filtering used on ATL03 data, and next, we present the current research on neural networks for point clouds. We then describe the architecture we use, namely ConvPoint and the datasets preparation, training, and testing. In the final section, we present the results and discuss why the method is well-suited for use in forested areas and some important considerations when adopting the methodology.

2. Materials and methods

2.1. ICESat-2 ATL03 and ATL08 data

ICESat-2 data is freely available and can be downloaded from the National Snow and Ice Data Center (NSIDC). The data product used for the training and testing of the neural network is the ATLAS/ICESat-2 L2A Global Geolocated Photon Data, Version 3 (Neumann et al., 2020), which is referred to as ATL03 data. We have used only data from the strong beams to ensure enough photon events for the training of the model. The data consists of photon events tagged by latitude, longitude, and height (Fig. 1). It should be noted that although the data is in a 3D format, the satellite's footprint is so small (about 13 m), and the transects so regular that it can be represented as a profile. In some profiles, slight variations in the longitude direction along the profile lines are observed, and those come from small variations in the instrument pointing. The ATLAS/ICESat-2 L3A land and vegetation height product, or ATL08, is used during the final comparison of our results to the terrain finding algorithm that is employed by the ICESat-2 team (Neuenschwander and Pitts, 2019). ATL08 data contains heights for both terrain and top of canopy in the along-track direction relative to the WGS-84 ellipsoid. The highest-level terrain and top of canopy elevations are provided within 100-m segments. Depending on the number of classified photons, some segments may not provide elevations if the confidence in

the classification is low. For terrain elevations, users can choose to use the interpolated terrain values (h_te_interp), which are lower confidence 100-m elevations. The high-confidence elevations are denoted as h_canopy_abs and h_te_mean in the ATL08 Land Water Vegetation Elevation data product (Neuenschwander et al., 2020).

2.2. Linear mode high-density airborne lidar (HDL) as ground reference data

The airborne lidar data used for a large part of the project was collected by the National Center for Airborne Laser Mapping (NCALM), based at the University of Houston. The pulse density is higher than 15 pulses/m^2, with canopy measurements of more than 20 points/m^2. The data was automatically segmented into ground and above ground returns using an algorithm similar to that described in Axelsson (2000), and then the classification was manually verified. After segmentation and verification, the data was processed into 20 × 20 m rasters of canopy heights and terrain elevations. The corresponding lidar data was then extracted from the raster overlap locations with the ICESat-2 ground tracks (Fig. 3). The extracted airborne lidar elevation estimates for terrain and top of canopy were then used as ground reference (truth) for the neural network training. All of the NCALM-collected HDL data in this paper has already been used for the validation of ICESat-2 elevation estimations in Fernandez-Diaz et al. (2022). Therefore, more details regarding the airborne HDL used for this project can be found in Fernandez-Diaz et al. (2022). The only other lidar data used was for the Brazilian sites over the Amazon forest. It is openly available (Dos-Santos et al., 2019) and was processed in the same manner as the NCALM data to arrive at both terrain and top of canopy estimates.

2.3. Study sites

We selected HDL data from five sites in Central America, two sites in Brazil and one in Puerto Rico (Fig. 4). We used two of the Central American sites, Maya Biosphere Reserve in Guatemala and Puuc in Mexico, for the training of the neural network. The total area of the two training sites is 2718 km^2, with 756 km of ATL03 transects for the 76 training files. The total length of the ATL03 transects for the test (validation) sites is 169 km in 18 test files. The vegetation type covering the majority of the sites is dense, old-growth tropical forests. A main consideration when using the HDL data as ground reference is the temporal offset between the ICESat-2 and HDL acquisitions. We have selected areas with primarily mature intact forests, such as national forests and nature (biosphere) reserves. Thus, the differences in the tree canopy heights between acquisitions should be minimized. In Table 1 we list the HDL acquisition dates together with the corresponding ATL03 HDF5 files used for the area. Fig. 5 shows the selected sites in Central America with an overlay of the Intact Forest Landscapes map (Potapov et al., 2017). In addition to those considerations, visual inspections were also undertaken for selected profiles to ensure the best possible agreement between ICESat-2 and the reference airborne lidar data, as described in section 3.6.

2.4. ATL03 and DRAGANN noise filtering algorithms

The ICESat-2 ATL03 data product provides a classification for each photon event to be used as input for subsequent higher-level data products such as ATL08. The methodology behind the ATL03 classification algorithm leverages the assumption that background noise photons follow a Poisson distribution. The algorithm searches for outliers to

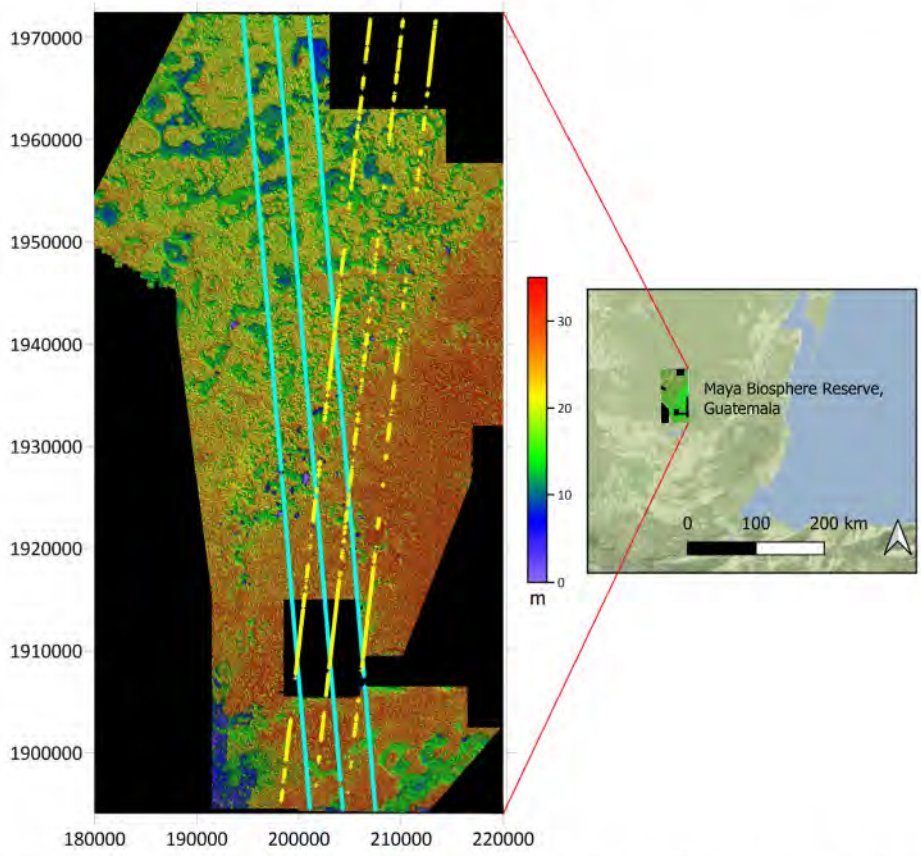


Fig. 3. High-density airborne lidar data for Maya Biospere Reserve and an example of ICESat-ATL08 ground tracks overlap. The yellow and cyan colors of the tracks represent two ATL08 files from different dates. The color of the underlying high-density airborne lidar data raster represents the canopy height model at 20 × 20 m resolution. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

Table 1
HDL data acquisition dates with corresponding date of ATL03 files used for the area.

Area	Year of HDL Acquisition	HDF5 File Acquisition Date
Maya Biosphere Reserve, Guatemala	2019	20181028
		20190326
		20190727
		20200125
		20200323
		20200421
		20210322
Puuc, Mexico	2017	20181122
		20190920
Xpujil, Mexico	2016	20181028
		20190127
Rio Bravo Conservation Area, Belize	2016	20191019
		20200721
Rio Platano Biosphere Reserve, Honduras	2012	20181102
		20220125
Adolpho Ducke National Forest, Brazil	2010	20190802
Saraca-Taquera National Forest, Brazil	2013	20191126
Saraca-Taquera National Forest, Brazil	2014	20210823
El Yunque National Forest, Puerto Rico	2011	20200223
		20220215

the Poisson distribution, which are then labeled as candidate signal photons. Histograms of the photon events are aggregated into along-track and vertical bins. The background noise photons are distributed randomly among the histogram bins while the signal photons will cluster into single or several adjacent bins (Neumann et al., 2021). The approach performs well on surfaces such as ice sheets, but it has been observed that over vegetated areas, the algorithm does not faithfully catch the top of the canopy photons as signal (Neuenschwander and Pitts, 2019). The ATL08 data processing combines the ATL03 classification results with an additional noise filtering method - the Differential, Regressive, and Gaussian Adaptive Nearest Neighbor (DRAGANN), developed by the ATL08 science team, to obtain the best possible signal from noise separation prior to running surface finding algorithms. DRAGANN exploits the assumption that signal photons will be more clustered in space than random noise photons. The algorithm consists of several steps: (1) an adaptive nearest neighbor search to compute the number of nearby photons within a specified search radius for each photon in the point cloud. The radius is established by computing the probability of how many photons will be in a given search area and is a function of the total number of photons in the point cloud. The adaptive search is flexible as different thresholds can be applied, each suitable for varying amounts of background noise and surface reflectance. (2) a histogram of the number of neighbors within a radius is generated, where two distinct peaks are expected to appear - a noise peak and a signal peak, (3) Gaussian curves are fitted to the histograms, one for the signal and one for noise peak - the intersection of the two Gaussian curves is used as a threshold value (optimized search radius) to distinguish between signal and noise photons, and, (4) the threshold classified data is then used as input for final determination of the canopy and ground surfaces.

2.5. Neural networks for 3D point clouds

Deep learning using convolutional neural networks (CNN) is a well-established tool for processing 2D images and structured data such as speech and text, with generally the aim of object or language recognition (LeCun et al., 2015; Deng et al., 2014; Goodfellow et al., 2016). 3D point cloud data is widely used for tasks such as autonomous driving; thus, the need for algorithms that can recognize and automatically classify point cloud data with good accuracy has arisen. Unfortunately, due to the unordered and unstructured nature of point cloud data, a 2D convolution operator cannot be used directly on 3D point clouds without first gridding or regularizing the point cloud data, which may result in a loss of information. Therefore, a CNN that works on unstructured 3D point clouds would be desirable. Fortunately, there are many emerging approaches for the direct processing of 3D point clouds with neural networks, for example, voxelization, multi-view, and graph neural networks (Maturana and Scherer, 2015; Wu et al., 2015; Qi et al., 2016). PointNet is a pioneering neural network that directly processes unordered and unstructured point clouds (Qi et al., 2017a). Geometric deep learning is the generic term for the methods that operate on manifolds, graphs, or directly on point clouds (Bronstein et al., 2017). A good review of the methods for geometric deep learning is provided in Guo et al. (2020), and an overview of semantic segmentation for point clouds using deep learning is provided in Xie et al. (2020) and more recently Zhang et al. (2023). A large portion of the deep learning models applied to point clouds attempts to discriminate objects with a well-defined shape coming from either indoor scenes or data collected by a mobile lidar scanner for the needs of the autonomous driving industry, where classes of objects are, for example, cars, pedestrians, buildings and curb lines. Airborne lidar datasets of outdoor scenes can have some objects with a distinct shape, such as buildings, but they also include many mathematically difficult objects to model, such as clusters of trees. There is a rise in the number of publications that use increasingly elaborate point-based methods of segmentation for airborne lidar datasets, for example (Huang et al., 2021; Lin et al., 2021; Zeng et al., 2023; Yu et al., 2022), among others. Many expand on previous point-based neural networks such as PointNet (Qi et al., 2017b) or KPConv (Thomas et al., 2019). Large-scale airborne lidar datasets have been built specifically with the goal of training and evaluation of deep learning algorithms for applications such as ground filtering or semantic segmentation of airborne point clouds (Varney et al., 2020; Qin et al., 2021; Ye et al., 2020). Semantic3D (Hackel et al., 2017) and the 3D Point Cloud and Modeling (Nuage de Points et Modelisation 3D - NPM3D), Roynard et al. (2018), are two other projects which provide labeled point clouds for outdoor scenes. Because the ATL03 datasets in tropical environments are relatively devoid of regular objects, we selected the ConvPoint architecture as it was implemented in NPM3D and Semantic3D datasets and performed well for outdoor point clouds. An additional reason to choose ConvPoint over, for example, a model designed explicitly for ALS point clouds, is the relative ease of understanding and implementation of the ConvPoint architecture, which is an elegant extension of discrete 2D convolution kernels to continuous 3D space. Additionally, other researchers have confirmed that ConvPoint performs well for semantic segmentation of urban outdoor scenes (Diab et al., 2022) as well as airborne lidar datasets (Turgeon-Pelchat et al., 2021).

2.5.1. ConvPoint architecture and nearest neighbors search

The ConvPoint convolutional network is implemented using the PyTorch library (Boulch, 2020). It employs continuous convolutional kernels, in contrast to the discrete ones used by regular 2D or 3D convolutions on image or voxel data. In discrete convolutions over gridded data, such as an image, each kernel element has an exact overlap with a corresponding image pixel. This is not possible for point clouds, as they are unstructured, and the coordinates of a point are not arranged on a grid but represent a continuum in 3D space. Continuous convolution can be used on unstructured data such as point clouds; continuous, in this case, means that the kernel element locations do not overlap with the point cloud locations but compute the distances between each kernel element and the specific point from the point cloud on which the kernel is applied. The kernel element locations are initialized randomly and are optimized by a multi-layer perceptron (MLP) layer during the training phase of the network (Fig. 6). Learning the kernel element positions is advantageous compared to having fixed-position kernel elements as the network becomes more flexible to adapt to the geometry of the data

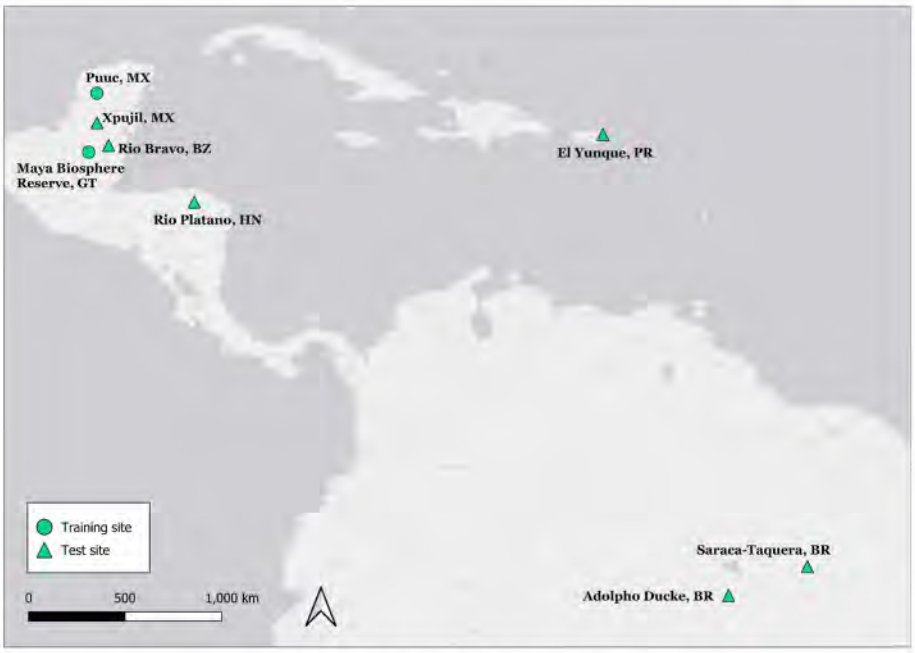


Fig. 4. All high-density airborne lidar data sites.

(Thomas et al., 2019). It is outside the scope of the paper to explain how ConvPoint learns from the distribution of photon events in space to arrive at predictions, and the reader is referred instead to (Thomas et al., 2019) and (Boulch, 2020) for further reading on how the adaptive kernel positions lead to better learning outcomes.

The semantic segmentation part of the network uses a structure similar to U-net with an encoder network followed by a decoder network (Boulch, 2020; Ronneberger et al., 2015). The kernel is applied to the k-nearest neighbors (k-nn) of a set of randomly selected points. In the case of ConvPoint layers, the number of nearest neighbors varies between 16 and 4. As stated in Thomas et al. (2019), a nearest neighbors search is not robust for input point clouds having varying densities. A theoretically better approach would be to use a radius search where the features would be more related to the geometry of the input points rather than the sampling number (Thomas et al., 2019). This is because for a dense point cloud eight nearest neighbors will be much closer to each other than in a sparse cloud case. With a radius search, that problem is eliminated, and the kernel will be applied only to points within a given radius. One drawback will be that the number of points within each radius search will vary. The varying number of points is an issue when building the index matrices for the nearest points within a radius. In the case of k-nn, the number of nn is fixed, and the resulting matrices are easy to compose, and optimization tools such as Cython are easier to implement. We tested the hypothesis that a radius search would be a better-suited choice for the ATL03 point cloud data by implementing a radius search within the existing ConvPoint architecture. A maximum number of neighbors within a radius was selected to circumvent the problem with the unequal number of neighbors for each search point. If the neighboring points were less than the selected maximum, the rest of the indices were set to zero. It was also later necessary to set the calculated distances for those non-existent points to zero before they are passed through the MLP that learns the weights and positions of the kernel elements. Surprisingly, the experiments showed that the original k-nn implementation provided higher accuracy than using a radius search. An unexpected drawback of the radius search was slower training, which resulted in longer run times compared to the k-nn implementation with the same number of epochs.

2.6. Data preparation, training, validation, and testing

For images, semantic segmentation attempts to provide each pixel with a label corresponding to various object categories. In the case of 3D point cloud data, semantic segmentation means that each point from the point cloud will be given a label. For the training data set preparation, we used two classes to divide the ATL03 point clouds: signal and noise. Each point (photon event) was labeled as either zero for noise or one for signal. Within the tree canopy, it would be difficult to discern noise photons because the randomness of returns from vegetation mirrors the randomness of noise event photons. Therefore all photon events lying in the band between the terrain and the top of the canopy, as delineated by HDL, were labeled as signal. In the validation or testing phase, the network output is a signal or noise label prediction for each point. The labeling for the training datasets was performed as follows. The ATL03 photon event data was plotted together with the high-density airborne lidar data (HDL). The MATLAB® polygon function (inpolygon) was used as a boundary to enclose all the points inside the polygon outlined by the HDL terrain and top of canopy elevations. The ATL03 photon events that fall within the polygon were labeled as signal (label 1), while the points outside the polygon were labeled as noise (label 0). The automatic labeling did not always enclose all the correct signal labels because of small differences between the HDL ground truth and the ATL03 data; as an illustration, Fig. 7 shows such a discrepancy around a water body that was incorrectly labeled as noise. In this case, the difference is likely caused by seasonal variations in the water level between acquisitions. Cases like this, therefore, required additional manual labeling using TerraScan software to ensure all the correct signal photons were labeled as such; otherwise, the neural network learning might be hampered. The Results and Discussion section presents more details about the additional manual labeling. An overview of the workflow is presented in Fig. 8.

The x and y coordinates of ATL03 photon events are provided in latitude and longitude. To ensure that the neural network will operate optimally by having all dimensions in the same units, we converted both ATL03 and HDL data into Universal Transverse Mercator (UTM) format using the Point Data Abstraction Library (PDAL) Butler et al. (2021). After separating the ATL03 data into two classes, the data files were used as the input for the training, validation, and testing of the neural

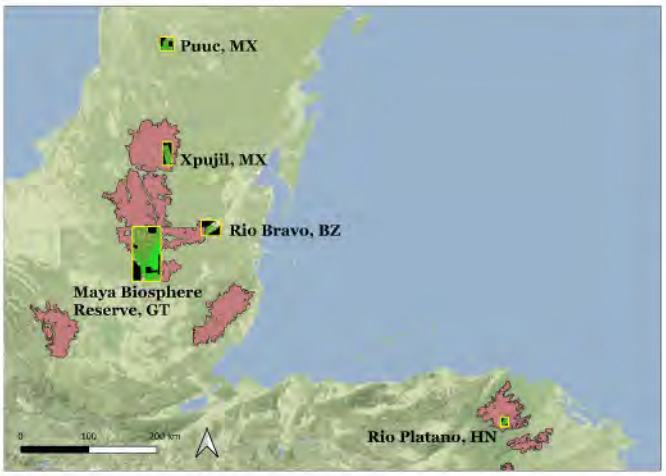


Fig. 5. High-density airborne lidar data coverage areas in Central America overlaid with the intact forest landscapes map for the year 2020.

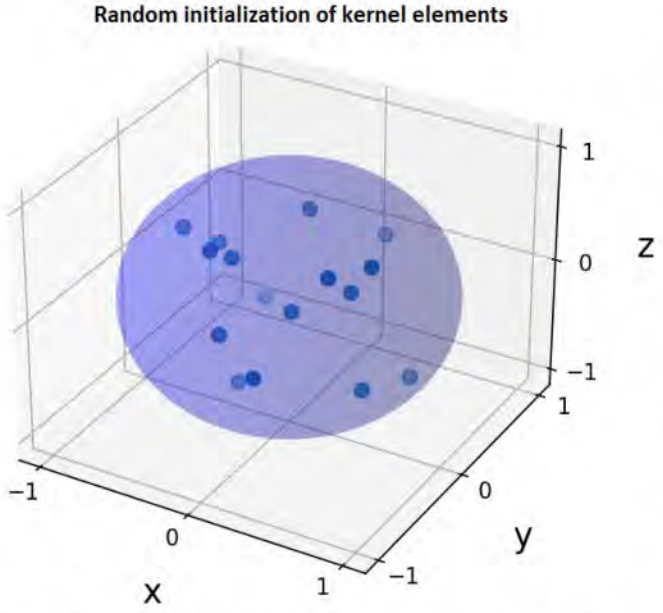


Fig. 6. Random initialization of kernel elements. Each dot represents a kernel element (center), that is initialized within the unit sphere.

network. The 100 training files were divided into 76 files for training, 6 for the validation phase for parameter tuning, and 18 for model testing. We used data from only two sites in Central America for training: the Maya Nature (Biosphere) Reserve in Guatemala and the Puuc site in Mexico. Data from the rest of the sites were only used during the testing phase (13 datasets). This was done to evaluate whether the neural network model was generalized enough to provide good results in areas with different forest biophysical parameters not included in the training. Thus, the number of test datasets in the regions the network was not trained on was primarily constrained by data availability - we tried our best to find as many samples in a diverse set of tropical forest regions that also had coincident HDL data. We chose the 5 test datasets from the two network training sites to represent harder cases with either large outliers or increased solar noise. The validation files were randomly selected at approximately 10% of the training files, and we added a couple more training files later upon the release of newer ICESat-2 datasets in our regions of interest. Since the Maya Biosphere Reserve (MBR) was the largest high-density airborne lidar dataset, the amount of coincident ICESat-2 data was the largest; thus, we used MBR as a training site to get the highest possible number of training datasets. Each Numpy file used for training has five fields for each photon event: (1) northing, (2) easting, (3) height above the ellipsoid, (4) feature, and (5) label. Since the ATL03 point cloud does not contain any features, meaning it has no color or intensity value, for example, the feature column is a vector of ones. What this means is that the convolutional neural network prediction is based purely on geometric features (Boulch, 2020). The input for the testing phase is only the northing, easting, and height fields for each test file. In order to achieve good results, adjustments to the original ConvPoint code were implemented. For instance, the number of required random input points was decreased, as traditional point clouds collected with mobile laser scanners usually contain millions of points. In the case of ATL03 data, we selected the number of random points picked for the first layer to be 1024. This is because the ATL03 points per file varied between ~12,000 and 190,000. Each individual training and testing file was created to be ~0.1° of latitude (~10 km). The reason for this choice is the fact that the ATL08 ground finding algorithm as described in Neuenschwander et al. (2020) is based on 10 km segment lengths and we have used a similar approach to calculate the terrain and top of canopy elevations after the ConvPoint label predictions. We also determined that the normalization of the points to fit the unit ball lowered the accuracy of the results so we removed it.

The kernel points in the convolutional kernel are initialized randomly within the unit ball (Fig. 6) and then the best positions for each kernel element are learned through a multi-layer perceptron (MLP) during the training phase. We experimented with initializing the kernel points using both a flattened sphere or ellipsoid to better represent the different directionality of ATL03 point clouds within a 10 km segment because of the small laser beam footprint. Initializing the kernel

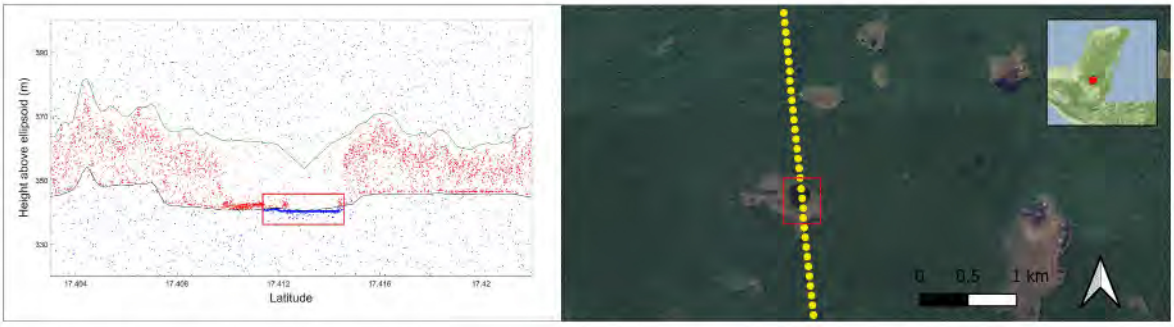


Fig. 7. An example of discrepancy between ICESat-2 ATL03 data and HDL ground reference. The left figure shows the labeled ATL03 photon events from automatic labeling. Red points are photon events labeled signal and blue points are noise. On the right side is a plan view of the ICESat-2 ATL08 data (yellow circles) over the same region. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

elements within an ellipsoid did not result in a significant jump in accuracy. It is worth noting that the training accuracy was already high, at ~98% and the directional kernel initialization resulted in training accuracy that again achieved 98%. We trained the model for 350 epochs, with a batch size of 32, using Adam optimizer and cross-entropy loss. The learning rate was set at 0.001. For the ConvPoint-specific parameters at training, we set the number of input points to the first layer at 1024 and the block size at 110.

2.6.1. Ground and top of canopy finding algorithm

In the ATL08 data product, the top of canopy and terrain elevation estimates are provided at a fixed step size of 100 m in the along-track direction, defined as a segment. For both visual and quantitative interpretation and validation of the neural network predictions, we compare them directly to the ATL08 100 m segment estimates. After obtaining the noise-filtered data, we further process it in 100 m segments for direct comparison with ATL08 data (h_canopy_abs and h_te_interp). We run a similar ground and top of canopy algorithm to the one provided in the Algorithm Theoretical Basis Document (ATBD) for Land - Vegetation Along-Track Products (ATL08) (Neuenschwander and Sheridan, 2020). We use a succession of min/max, Savitsky-Golay, and mean and median filtering to obtain the terrain and top of canopy curves and then use the center latitude from each ATL08 segment to divide our data into corresponding 100 m segments. We then take the 100 m ATL08 and the 100 m segments obtained from the neural network prediction and compare both to the 100 m high-density airborne lidar data (HDL) ground reference, which was segmented in an identical manner.

3. Results and discussion

We demonstrate improved estimates for both terrain and top of canopy elevations compared to ATL08, with mean terrain estimates improved by about 2 m and mean top of canopy by about 4 m (aggregated for all 18 test files given in Tables 3 and 4). For RMSE calculation, we use the 100 m segments of ATL08 data and corresponding 100 m segments from neural network prediction and the HDL reference. Overall RMSE (18 test profiles from all test areas) for the terrain elevations decreased from 5.2 m for ATL08 to 3.3 m for the ConvPoint Neural Network (NN) prediction, compared to the HDL ground reference. Overall RMSE decreased from 7.7 m to 3.7 m for the top of canopy elevations. The overall accuracy in the training, validation, and testing phases is presented in Tables 2 and is computed from the confusion matrix of the predicted point labels as:

(1)

For the terrain elevations RMSE comparison, we used the interpolated terrain values provided by ATL08 (h_te_interp) as the mean terrain elevations h_te_mean are not available for each ATL08 100 m segment. For top of canopy comparison, we used the h_canopy_abs 100 m segments. To ensure that results are not skewed towards the training forest structure, test profiles were selected from areas in Honduras, Brazil, and Puerto Rico (the training data is solely from Guatemala and Mexico). Results suggest that the rules learned by the neural network are not biased towards data from areas it has trained on, but perform well on data from forests with different biophysical parameters (Figs. 9 and 11, Table 4). These findings suggest that the neural network noise filtering method proposed here can be used for other forested environments around the globe. We have used only 76 files for training with a total length of 756 km from two training areas. Extrapolating the results would mean that a network could be trained for global coverage noise filtering in forested areas with a relatively low amount of training with airborne lidar reference data. The applicability of this proposed method is strengthened by the growing global library of freely available high-density airborne lidar data that could serve as a ground reference.

Table 2
Overall accuracy (OA) results for the 18 test profiles. The top five profiles in the table are test datasets from training sites. The lower 13 profiles are test data from sites that were not included in the training dataset. OA is calculated from the confusion matrix and it refers to the accuracy of the neural network in predicting the label of each photon event for each test file (Eq. (1)). GT Guatemala, MX Mexico, BZ Belize, HN Honduras, BR Brazil, PR Puerto Rico, NR Nature (biosphere) reserve, NF National Forest. npoints is the number of ATL03 photon events within each of the test profiles.

File date for test dataset ATL03	Area	Latitude bounds	Beam	OA	npoints
20181028	Maya NR, GT	17.55,17.65	1r	0.9939	17661
20190727	Maya NR, GT	17.455,17.555	1 l	0.9795	82772
20200125	Maya NR, GT	17.5,17.6	3r	0.9983	15221
20200323	Maya NR, GT	17.55,17.65	1 l	0.9770	25358
20190920	Puuc, MX	20.17,20.25	1r	0.9984	12913
20181028	Xpujil, MX	18.74,18.84	1r	0.9990	11729
20190127	Xpujil, MX	18.7,18.84	2 l	0.9898	23270
20191019	Rio Bravo, BZ	17.708,17.793	1r	0.9974	10485
20191019	Rio Bravo, BZ	17.714,17.811	2r	0.9982	9017
20191019	Rio Bravo, BZ	17.731,17.83	3r	0.9978	12840
20200721	Rio Bravo, BZ	17.776,17.855	1 l	0.9923	58016
20181102	Río Platano NR, HN	15.216,15.312	1r	0.9392	16586
20220125	Río Platano NR, HN	15.237,15.277	1 l	0.9924	6303
20191126	Saraca-Taquera NF, BR	1.692, 1.654	3r	0.9827	5783
20210823	Saraca-Taquera NF, BR	1.636, 1.618	1r	0.9817	7809
20190802	Adolpho Ducke NF, BR	2.963, 2.9375	1 l	0.9885	3572
20200223	EL Yunque NF, PR	18.24, 18.286	3r	0.9959	9790
20220215	EL Yunque NF, PR	18.241, 18.30	2 l	0.9870	7514

Table 3
Terrain and top of canopy RMSE results for test datasets from the Maya Biosphere Reserve, Guatemala (GT), and Puuc, Mexico (MX). Both ATL08 and neural network (NN) RMSE are calculated with the HDL data serving as reference for the 100 m segments.

Dataset	Beam	Area	Day/ Night	Top of Canopy		Terrain	
				ATL08 RMSE (m)	NN RMSE (m)	ATL08 RMSE (m)	NN RMSE (m)
20181028	1r	GT	Night	5.6	2.7	6.1	3.7
20190727	1 l	GT	Day	4.5	3.2	6.7	3.7
20200125	3r	GT	Night	37.0	1.8	6.7	2.5
20200323	1 l	GT	Day	2.5	1.9	2.5	2.1
20190920	1r	MX	Night	3.3	2.8	1.9	1.5
Mean				10.6	2.5	4.8	2.7

3.1. Outlier reduction

Neural network noise filtering eliminates large top of canopy outliers in both Figs. 9 and 10. The latter is a test dataset in the Maya Biosphere Reserve NN training site. The test file is for beam 3r; beams 1r and 2r were included in the training dataset. The outliers seem to be caused by fog or low-lying clouds that were mistaken for the top of the canopy because of the density of the returns. ATL08 terrain estimates also show

Table 4
Terrain and Top of canopy RMSE results for the 13 test-only datasets in regions excluded from network training. Both ATL08 and neural network (NN) RMSE are calculated with the HDL data serving as reference for the 100 m segments. GT = Guatemala, MX = Mexico, BZ=Belize, HN=Honduras, BR=Brazil, PR = Puerto Rico.

Dataset	Beam	Area	Day/ Night	Top of Canopy		Terrain	
				ATL08 RMSE (m)	NN RMSE (m)	ATL08 RMSE (m)	NN RMSE (m)
20181028	1r	MX	Night	1.7	0.9	3.3	1.1
20190127	2 l	MX	Night	3.5	2.2	1.2	0.7
20191019	1r	BZ	Night	4.3	3.7	5.0	3.1
20191019	2r	BZ	Night	4.3	2.7	5.2	3.2
20191019	3r	BZ	Night	3.6	2.6	4.9	3.0
20200721	1 l	BZ	Night	4.4	6.3	3.2	1.8
20181102	1r	HN	Night	28.6	7.9	9.9	7.1
20220125	1 l	HN	Night	8.2	6.3	9.1	5.9
20191126	3r	BR	Day	2.3	2.2	3.7	2.7
20210823	1r	BR	Night	4.5	2.9	6.5	4.6
20190802	1 l	BR	Night	4.4	3.0	7.0	4.8
20200223	3r	PR	Night	6.8	6.1	5.1	2.9
20220215	2 l	PR	Night	9.2	7.7	5.5	4.6
Mean				6.6	4.2	5.4	3.5

large inaccuracies compared to the HDL ground reference. The neural network noise filtering eliminates the band of low-lying clouds, leading to higher-accuracy estimates for both the top of canopy and terrain. The former (Fig. 9) is from an area in Honduras (no data from Honduras area was included in the training dataset). Again, low-lying clouds complicate the ATL08 top of canopy predictions between latitudes 15.20 and 15.25. Even though the ConvPoint prediction for signal photons is not entirely continuous, it correctly identified the clouds/fog as noise. The NN results in a reduction of top of canopy RMSE from 28.6 to 7.9 m, while the improvement in terrain is from 9.9 to 7.1 m (Table 4). For this test profile, there is a gap of signal photons incorrectly identified as noise between latitudes 15.23 and 15.24. The discontinuity is the reason for the one incidence of lower overall accuracy of 0.9392 shown in red in Table 2. Another example of noise caused by fog or very low clouds is shown in Fig. 11; the clouds are noticeable between latitudes 18.80 and 18.815. While the neural network incorrectly labels some of the fog layer as signal between latitudes 18.80 and 18.805, it correctly identifies it as noise between 18.805 and 18.815 latitude. The 100 m NN segments are closer to the HDL reference, while the ATL08 prediction shows larger errors.

3.2. 3D vs. 2D neural network input data

Using a 3D point-based neural network offers an advantage in terms of flexibility both for training and test file dimensions when compared to 2D image CNNs. Using direct 3D input requires no data aggregation that may result in loss of resolution. Input files for 2D convolutional architectures such as U-net must have the same rows and columns size. This constraint can decrease the training set size because coincident ground reference data might not always be available for a whole segment length or else has to be cropped. While there is flexibility in the length of the files, if the number of points is too low compared to the mean number of points, the network might predict with lower accuracy.

3.3. Automatic labeling considerations

The automatic labeling described in section 3.6 presented challenges for several ICESat-2 profiles. In some instances, some signal points were outside the HDL bounds while in other instances noise points were labeled as signals; this was confirmed by visual examination. One reason for the discrepancies between the satellite-collected data and the HDL reference might be seasonal variations and the time difference between acquisitions. Some discrepancies also existed around water bodies where ICESat-2 detected shallow bathymetry and associated "ringing". For example, in Fig. 7, it is visually evident that the water surface automatically labeled as noise (red rectangle on the left plot) needs to be re-labeled as signal before the neural network training. The mislabeled water surface is likely a pond that has seasonal water level variations, which caused the discrepancy between ATL03 and HDL elevations. Those clustered signal points that were mislabeled as noise would have

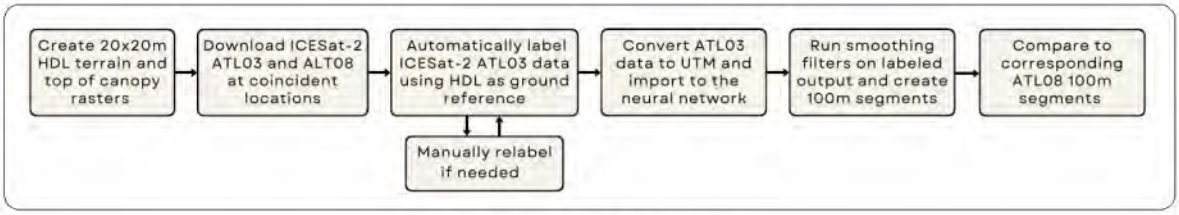


Fig. 8. Overview of the workflow.

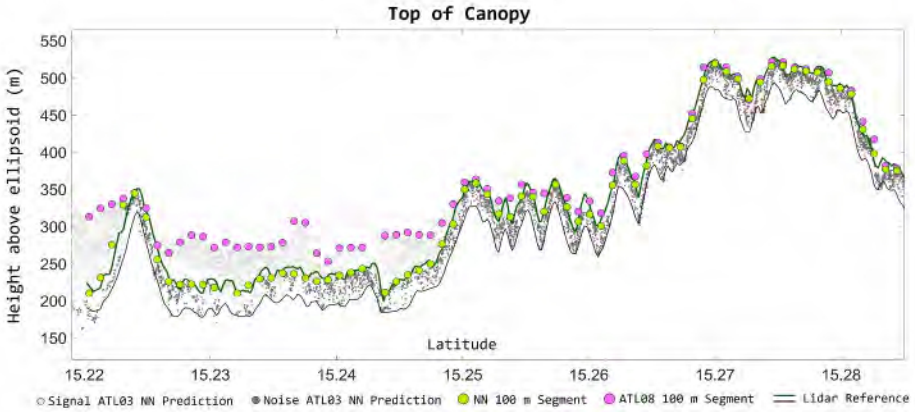


Fig. 9. Top of canopy estimates for test dataset 20181102 in Honduras, beam 1r. This region was not included in NN training.

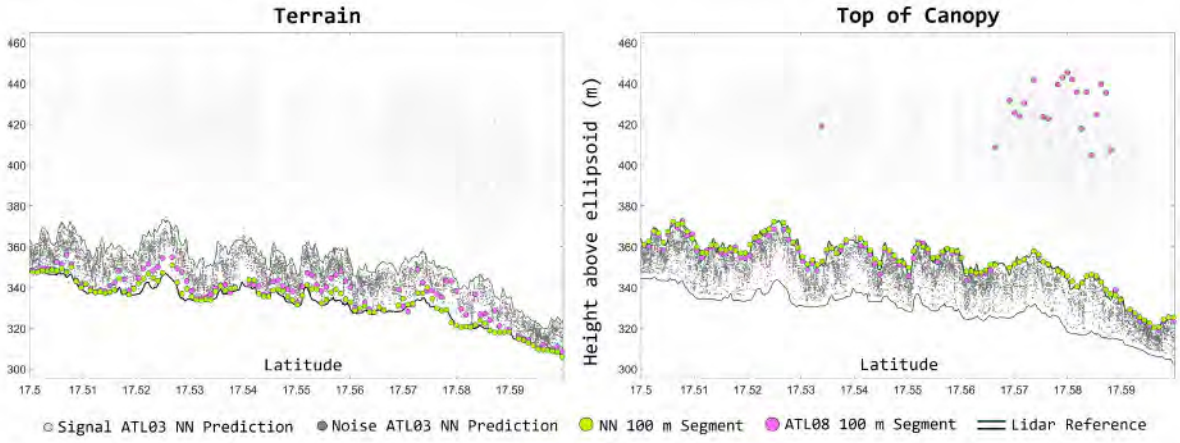


Fig. 10. Terrain (left) and top of canopy (right) estimates for Guatemala test dataset 20200125, beam 3r (training and testing area).

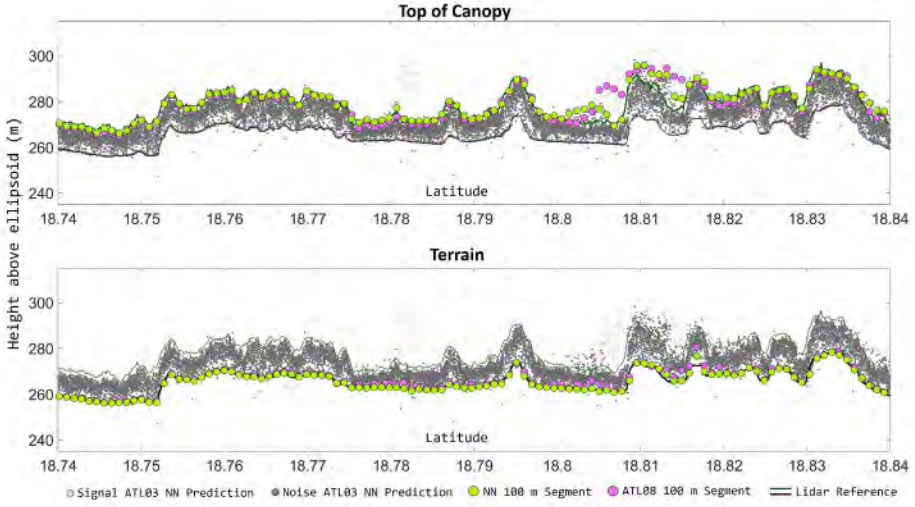


Fig. 11. Example of a test dataset in night conditions from a test only area in Mexico (20190127, beam 2 l).

lowered the accuracy of the test results if not manually reclassified. This was confirmed by training the network first with only automatically determined labels and then re-training using manually updated labels. After re-labeling, the training accuracy slightly improved, with 98% overall accuracy (OA) for the manually re-labeled training data, compared to about 96% OA when training on the automatically labeled clouds.

3.4. Robustness of the NN model

We first experimented with testing the network using profiles in geographic areas where it had not been trained to examine performance in a diverse range of tropical forests; as described above in this results section. Of the 18 test datasets, 13 are from test-only areas. The NN terrain estimates are lower in all test datasets while top of canopy estimates are lower in all but one (20200721, BZ, in Table 4). The higher top of canopy RMSE in this particular dataset is due primarily to localized differences between the HDL and ATL03 data caused by the 4 year time difference between their acquisition (2016 for HDL and 2020 for ATL03). The area around this profile has recent human activities and some variation in the vegetation cover between the acquisitions of the lidar and ICESat-2 data is clearly present. To verify this conclusion, a second experiment was performed with the 20200721, BZ dataset; two of the strong beams for this profile were included in the NN training dataset and the newly trained network was then used to examine the outlier test profile. This is to examine whether the network predictions were poor because the BZ dataset was not well represented by the original training data. The difference between the original and re-trained network results is minor, but when trained on additional site-specific files, the results are marginally better, as seen in Fig. 12. Using the original training data set, the predictions of the water surface are slightly noisier, but still acceptable, given that the network has not seen similar data and some "ringing" is clearly present. The ringing is the result of specular returns from standing water and is an instrument response to high signal rates, as described in Neuenschwander and Magruder (2019). In the paper, the authors mention the potential for ICESat-2 to detect inundated lands and flooded forests based on this ringing response. Future work could involve training a point-based neural network to specifically detect ringing in ATL03 data to isolate inundated areas.

3.5. Future directions

As described in section 2.6.1, the terrain and top of canopy detection algorithms we used are similar to those used for ATL08. Since the neural network separation of signal and noise closely follows the lidar ground reference even in day conditions, as seen in Fig. 13, we have used a moving minimum filtering for terrain and moving maximum for top of canopy together with successions of mean, median, and Savitzky–Golay filtering. We chose this approach for its simplicity to implement and use

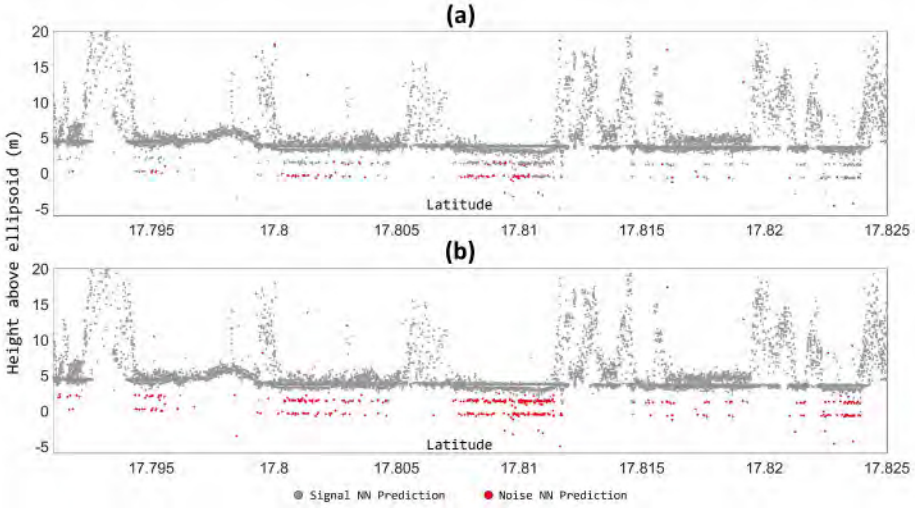


Fig. 12. An example of a result when: (a) the neural network has not been trained on similar type of data; (b) the training dataset has included data from the other beams for this ATL03 dataset. The network performs slightly better in scenario (b), as the "ringing" under the standing water surface is labeled noise.

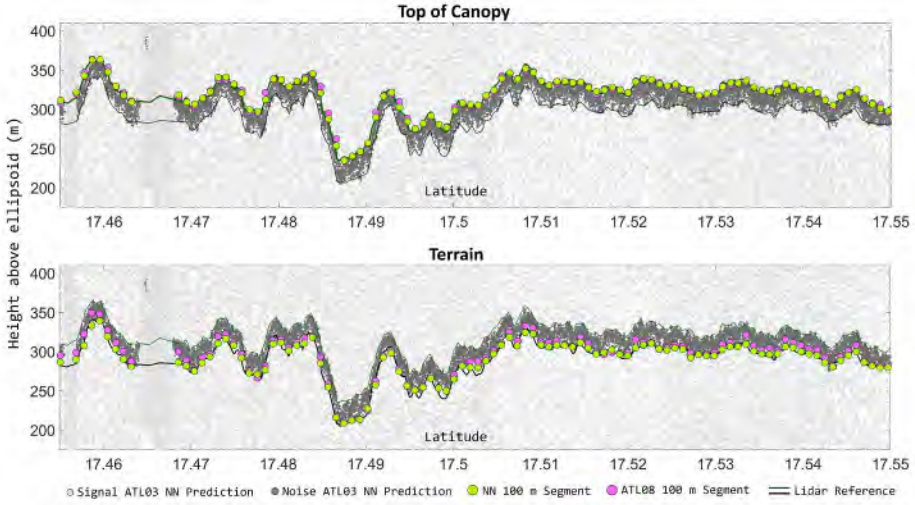


Fig. 13. Example of a test dataset in day noise conditions from Guatemala (20190727, beam 1 l).

as a comparison to ATL08 results, as our main focus was the noise filtering potential. In some cases, though, when the neural network predictions are outside the boundaries of the HDL reference, minimum filtering will cause the segmented predictions to be too low. Similarly, if a couple of photon events are predicted too high or too low compared to the main signal band, the minimum, and maximum filtering will cause the segment predictions to move to the outlying points. A step for future work could be a more elaborate terrain and top of canopy algorithm that would leverage the advantage of the cleaner noise-filtering provided by the ConvPoint neural network. Despite our best efforts to find datasets in both day and night conditions, most of the ICESat-2 ATL03 data available for the test sites was collected at night. It would be interesting to add more daytime acquisitions as the ICESat-2 mission continues and new data becomes available. Herein, only two classes were used for neural network training, however, it is feasible that more classes could be incorporated. This can be done by separating the signal into individual labels such as terrain, intra-canopy, and top of the canopy, for example. The ConvPoint implementation (or any other neural architecture) can be adjusted to contain as many classes as needed. The only downside would be that more labor would be involved to ensure that the training datasets are well-labeled. This architecture could also be an excellent tool for shallow bathymetry classification.

4. Conclusions

In this paper, a novel approach towards noise filtering of spaceborne photon counting data, namely, using a neural network for 3D point clouds coupled with airborne lidar data as a ground reference was proposed. Our results indicate that even with modest training from two areas, this approach achieves good results in six separate tropical forest regions with varied terrain, slope, and top of canopy characteristics. The neural network method could be used in combination with the two noise-filtering algorithms currently employed for classifying ATL03 photons. Its usefulness is reinforced by the continuous arrival of ATL03 data during the lifetime of the ICESat-2 satellite and the abundance of airborne lidar datasets for vegetated locations worldwide. Given the widespread availability of data and the potential of point-based deep learning algorithms, implementing this proposed noise filtering method may prove valuable not only for ICESat-2 but any future lidar satellite missions.

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

A majority of the ALS data was acquired by the NSF-supported facility the National Center for Airborne Laser Mapping (EAR #1830734). Partial support for the first and third authors was also provided by a grant from NASA (#80NSSC22K1098). The data training, validation, and testing were performed using the high-performance computing resources provided by the University of Houston Hewlett Packard Enterprise Data Science Institute.

References

Axelsson, P., 2000. Dem generation from laser scanner data using adaptive tin models. International Archives of Photogrammetry and Remote Sensing 33, 111–118.

Baccini, A., Goetz, S., Walker, W., Laporte, N., Sun, M., Sulla-Menashe, D., Hackler, J., Beck, P., Dubayah, R., Friedl, M., et al., 2012. Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. Nat. Clim. Change 2, 182–185.

Boulch, A., 2020. Convpoint: continuous convolutions for point cloud processing. Comput. Graph. 88, 24–34.

Bronstein, M.M., Bruna, J., LeCun, Y., Szlam, A., Vandergheynst, P., 2017. Geometric deep learning: going beyond euclidean data. IEEE Signal Process. Mag. 34, 18–42. https://doi.org/10.1109/MSP.2017.2693418.

Brown, R., Hartzell, P., Glennie, C., 2020. Evaluation of spl100 single photon lidar data. Rem. Sens. 12, 722.

Butler, H., Chambers, B., Hartzell, P., Glennie, C., 2021. Pdal: an open source library for the processing and analysis of point clouds. Comput. Geosci. 148, 104680.

Deng, L., Yu, D., et al., 2014. Deep learning: methods and applications. Foundations and trends® in signal processing 7, 197–387.

Diab, A., Kashef, R., Shaker, A., 2022. Deep learning for lidar point cloud classification in remote sensing. Sensors 22, 7868.

Dos-Santos, M., Keller, M., Morton, D., 2019. Lidar Surveys over Selected Forest Research Sites, Brazilian Amazon, 2008-2018. ornl daac, oak ridge, tennessee, usa.

Duncanson, L., Neuenschwander, A., Hancock, S., Thomas, N., Fatoyinbo, T., Simard, M., Silva, C.A., Armston, J., Luthcke, S.B., Hofton, M., et al., 2020. Biomass estimation from simulated gedi, icesat-2 and nisar across environmental gradients in sonoma county, California. Rem. Sens. Environ. 242, 111779.

Fernandez-Diaz, J.C., Velikova, M., Glennie, C.L., 2022. Validation of icesat-2 atl08 terrain and canopy height retrievals in tropical mesoamerican forests. IEEE J. Sel. Top. Appl. Earth Obs. Rem. Sens. 15, 2956–2970.

Gao, S., Li, Y., Zhu, J., Fu, H., Zhou, C., 2022. Retrieving forest canopy height from icesat-2 data by an improved dragann filtering method and canopy top photons classification. Geosci. Rem. Sens. Lett. IEEE 19, 1–5.

Glenn, N.F., Neuenschwander, A., Vierling, L.A., Spaete, L., Li, A., Shinneman, D.J., Pilliod, D.S., Arkle, R.S., McIlroy, S.K., 2016. Landsat 8 and icesat-2: performance and potential synergies for quantifying dryland ecosystem vegetation cover and biomass. Rem. Sens. Environ. 185, 233–242.

Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT press.

Guerra-Hernandez, J., Narine, L.L., Pascual, A., Gonzalez-Ferreiro, E., Botequim, B., Malambo, L., Neuenschwander, A., Popescu, S.C., Godinho, S., 2022. Aboveground biomass mapping by integrating icesat-2, sentinel-1, sentinel-2, alos2/palsar2, and topographic information in mediterranean forests. GIScience Remote Sens. 59, 1509–1533.

Guo, Y., Wang, H., Hu, Q., Liu, H., Liu, L., Bennamoun, M., 2020. Deep learning for 3d point clouds: a survey. IEEE Trans. Pattern Anal. Mach. Intell. https://doi.org/10.1109/TPAMI.2020.3005434, 1–1.

Gwenzi, D., Lefsky, M.A., Suchdeo, V.P., Harding, D.J., 2016. Prospects of the icesat-2 laser altimetry mission for savanna ecosystem structural studies based on airborne simulation data. ISPRS J. Photogrammetry Remote Sens. 118, 68–82.

Hackel, T., Savinov, N., Ladicky, L., Wegner, J.D., Schindler, K., Pollefeys, M., 2017. SEMANTIC3D.NET: a new large-scale point cloud classification benchmark. In: ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, pp. 91–98.

He, L., Pang, Y., Zhang, Z., Liang, X., Chen, B., 2023. Icesat-2 data classification and estimation of terrain height and canopy height. Int. J. Appl. Earth Obs. Geoinf. 118, 103233.

Herzfeld, U.C., McDonald, B.W., Wallin, B.F., Neumann, T.A., Markus, T., Brenner, A., Field, C., 2013. Algorithm for detection of ground and canopy cover in micropulse photon-counting lidar altimeter data in preparation for the icesat-2 mission. IEEE Trans. Geosci. Rem. Sens. 52, 2109–2125.

Huang, R., Xu, Y., Stilla, U., 2021. Granet: global relation-aware attentional network for semantic segmentation of als point clouds. ISPRS J. Photogrammetry Remote Sens. 177, 1–20.

Huang, X., Cheng, F., Wang, J., Duan, P., Wang, J., 2023. Forest canopy height extraction method based on icesat-2/atlas data. IEEE Trans. Geosci. Rem. Sens. 61, 1–14.

Hubau, W., Lewis, S.L., Phillips, O.L., Affum-Baffoe, K., Beeckman, H., Cuní-Sanchez, A., Daniels, A.K., Ewango, C.E., Fauset, S., Mukinzi, J.M., et al., 2020. Asynchronous carbon sink saturation in african and amazonian tropical forests. Nature 579, 80–87.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444.

Li, Y., Fu, H., Zhu, J., Wang, C., 2020. A filtering method for icesat-2 photon point cloud data based on relative neighboring relationship and local weighted distance statistics. Geosci. Rem. Sens. Lett. IEEE 18, 1891–1895.

Lin, Y., Knudby, A.J., 2023. Global automated extraction of bathymetric photons from icesat-2 data based on a pointnet model. Int. J. Appl. Earth Obs. Geoinf. 124, 103512.

Lin, Y., Vosselman, G., Cao, Y., Yang, M.Y., 2021. Local and global encoder network for semantic segmentation of airborne laser scanning point clouds. ISPRS J. Photogrammetry Remote Sens. 176, 151–168.

Liu, X., Su, Y., Hu, T., Yang, Q., Liu, B., Deng, Y., Tang, H., Tang, Z., Fang, J., Guo, Q., 2022. Neural network guided interpolation for mapping canopy height of China's forests by integrating gedi and icesat-2 data. Rem. Sens. Environ. 269, 112844.

Luo, Y., Qi, S., Liao, K., Zhang, S., Hu, B., Tian, Y., 2023. Mapping the forest height by fusion of icesat-2 and multi-source remote sensing imagery and topographic information: a case study in jiangxi province, China. Forests 14, 454.

Malambo, L., Popescu, S.C., 2021. Assessing the agreement of icesat-2 terrain and canopy height with airborne lidar over us ecozones. Rem. Sens. Environ. 266, 112711.

Maturana, D., Scherer, S., 2015. Voxnet: a 3d convolutional neural network for real-time object recognition. In: 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, pp. 922–928.

Meng, W., Li, J., Tang, Q., Xu, W., Dong, Z., 2022. Icesat-2 laser data denoising algorithm based on a back propagation neural network. Appl. Opt. 61, 8395–8404.

Musthafa, M., Singh, G., Kumar, P., 2023. Comparison of forest stand height interpolation of gedi and icesat-2 lidar measurements over tropical and sub-tropical forests in India. Environ. Monit. Assess. 195, 71.

Narine, L., Malambo, L., Popescu, S., 2022. Characterizing canopy cover with icesat-2: a case study of southern forests in Texas and Alabama, USA. Rem. Sens. Environ. 281, 113242.

Narine, L.L., Popescu, S., Neuenschwander, A., Zhou, T., Srinivasan, S., Harbeck, K., 2019. Estimating aboveground biomass and forest canopy cover with simulated icesat-2 data. Rem. Sens. Environ. 224, 1–11.

Narine, L.L., Popescu, S.C., Malambo, L., 2020. Using icesat-2 to estimate and map forest aboveground biomass: a first example. Rem. Sens. 12, 1824.

Neuenschwander, A., Guenther, E., White, J.C., Duncanson, L., Montesano, P., 2020. Validation of icesat-2 terrain and canopy heights in boreal forests. Rem. Sens. Environ. 251, 112110.

Neuenschwander, A., Pitts, K., 2019. The atl08 land and vegetation product for the icesat-2 mission. Rem. Sens. Environ. 221, 247–259.

Neuenschwander, A.L., Magruder, L.A., 2019. Canopy and terrain height retrievals with icesat-2: a first look. Rem. Sens. 11, 1721.

Neuenschwander, A.L., L.P, K., P.J, B., R, J., K, B., P, S.C., F.N, R., H, D., P, D., Sheridan, R., 2020. Ice, Cloud, and Land Elevation 1 Satellite 2 (ICESat-2) Algorithm Theoretical Basis Document (ATBD) for Land - Vegetation Along-Track Products (ATL08). NASA National Snow and Ice Data Center Distributed Active Archive Center. https://doi.org/10.5067/ATLAS/ATL08.003. URL: https://nsidc.org/data/atl08.

Neumann, T.A., B, A., H, D., R, J., S, J., H, K., G, A., L, J., B.L, S., R.e.a, T., 2020. Atlas/icesat-2 L2a Global Geolocated Photon Data, Version 3. NASA National Snow and Ice Data Center Distributed Active Archive Center, Boulder, Colorado USA. https://doi.org/10.5067/ATLAS/ATL03.003. URL: https://nsidc.org/data/atl03.

Neumann, T., Brenner, Anita, Hancock, David, J.R.J.S.K.H.A.G.J.L.S.L.T.R, 2021. Algorithm Theoretical Basis Document (Atbd) for Global Geolocated Photons Atl03. NASA National Snow and Ice Data Center Distributed Active Archive Center. URL: https://nsidc.org/data/atl03.

Peng, K., Xie, H., Xu, Q., Huang, P., Liu, Z., 2022. A physics-assisted convolutional neural network for bathymetric mapping using icesat-2 and sentinel-2 data. IEEE Trans. Geosci. Rem. Sens. 60, 1–13.

Popescu, S., Zhou, T., Nelson, R., Neuenschwander, A., Sheridan, R., Narine, L., Walsh, K., 2018. Photon counting lidar: an adaptive ground and canopy height retrieval algorithm for icesat-2 data. Rem. Sens. Environ. 208, 154–170.

Potapov, P., Hansen, M.C., Laestadius, L., Turubanova, S., Yaroshenko, A., Thies, C., Smith, W., Zhuravleva, I., Komarova, A., Minnemeyer, S., et al., 2017. The last frontiers of wilderness: tracking loss of intact forest landscapes from 2000 to 2013. Sci. Adv. 3, e1600821.

Qi, C.R., Su, H., Mo, K., Guibas, L.J., 2017a. Pointnet: deep learning on point sets for 3d classification and segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 652–660.

Qi, C.R., Su, H., Nießner, M., Dai, A., Yan, M., Guibas, L.J., 2016. Volumetric and multi-view cnns for object classification on 3d data. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 5648–5656.

Qi, C.R., Yi, L., Su, H., Guibas, L.J., 2017b. Pointnet : deep hierarchical feature learning on point sets in a metric space. Adv. Neural Inf. Process. Syst. 30.

Qin, N., Tan, W., Ma, L., Zhang, D., Li, J., 2021. Opengf: an ultra-large-scale ground filtering dataset built upon open als point clouds around the world. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 1082–1091.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 234–241.

Roynard, X., Deschaud, J.E., Goulette, F., 2018. Paris-lille-3d: a large and high-quality ground-truth urban point cloud dataset for automatic segmentation and classification. Int. J. Robot Res. 37, 545–557. https://doi.org/10.1177/0278364918767506.

Stoker, J.M., Abdullah, Q.A., Nayegandhi, A., Winehouse, J., 2016. Evaluation of single photon and geiger mode lidar for the 3d elevation program. Rem. Sens. 8 https://doi.org/10.3390/rs8090767. URL: https://www.mdpi.com/2072-4292/8/9/767.

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. Sci. Rep. 6, 28277.

Thomas, H., Qi, C.R., Deschaud, J.E., Marcotegui, B., Goulette, F., Guibas, L.J., 2019. Kpconv: flexible and deformable convolution for point clouds. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 6411–6420.

Turgeon-Pelchat, M., Foucher, S., Bouroubi, Y., 2021. Deep learning-based classification of large-scale airborne lidar point cloud. Can. J. Rem. Sens. 47, 381–395.

Urbazaev, M., Hess, L.L., Hancock, S., Sato, L.Y., Ometto, J.P., Thiel, C., Dubois, C., Heckel, K., Urban, M., Adam, M., et al., 2022. Assessment of terrain elevation estimates from icesat-2 and gedi spaceborne lidar missions across different land cover and forest types. Sci. Rem. Sens. 6, 100067.

Varney, N., Asari, V.K., Graehling, Q., 2020. Dales: a large-scale aerial lidar data set for semantic segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, pp. 186–187.

Wang, X., Pan, Z., Glennie, C., 2016. A novel noise filtering model for photon-counting laser altimeter data. Geosci. Rem. Sens. Lett. IEEE 13, 947–951.

Wu, Z., Song, S., Khosla, A., Yu, F., Zhang, L., Tang, X., Xiao, J., 2015. 3d shapenets: a deep representation for volumetric shapes. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1912–1920.

Xie, H., Sun, Y., Xu, Q., Li, B., Guo, Y., Liu, X., Huang, P., Tong, X., 2022. Converting along-track photons into a point-region quadtree to assist with icesat-2-based canopy cover and ground photon detection. Int. J. Appl. Earth Obs. Geoinf. 112, 102872.

Xie, Y., Tian, J., Zhu, X.X., 2020. Linking points with labels in 3d: a review of point cloud semantic segmentation. IEEE Geosci. Remote Sens. Magaz. 8, 38–59. https://doi.org/10.1109/MGRS.2019.2937630.

Ye, Z., Xu, Y., Huang, R., Tong, X., Li, X., Liu, X., Luan, K., Hoegner, L., Stilla, U., 2020. Lasdu: a large-scale aerial lidar dataset for semantic labeling in dense urban areas. ISPRS Int. J. Geo-Inf. 9, 450.

Yu, L., Yu, H., Yang, S., 2022. A deep neural network using double self-attention mechanism for als point cloud segmentation. IEEE Access 10, 29878–29889.

Zeng, T., Luo, F., Guo, T., Gong, X., Xue, J., Li, H., 2023. Recurrent residual dual attention network for airborne laser scanning point cloud semantic segmentation. IEEE Trans. Geosci. Rem. Sens. 61, 1–14. https://doi.org/10.1109/TGRS.2023.3285207, 5702614.

Zhang, R., Wu, Y., Jin, W., Meng, X., 2023. Deep-learning-based point cloud semantic segmentation: a survey. Electronics 12, 3642.

Zhu, X., Nie, S., Wang, C., Xi, X., Wang, J., Li, D., Zhou, H., 2020. A noise removal algorithm based on optics for photon-counting lidar data. Geosci. Rem. Sens. Lett. IEEE 18, 1471–1475.