Enhancing Forest Attribute Prediction by Considering Terrain and Scan Angles From Lidar Point Clouds: A Neural Network Approach

Karun R. Dayal, Sylvie Durrieu, Kamel Lahssini, Dino Ienco, and Jean-Matthieu Monnet

(Invited Paper)

***Abstract*—Sensitivity of lidar metrics to scan angle can affect the robustness of area-based approach (ABA) models, and modelling the interplay of scan geometry and terrain properties can be complex. The study hypothesizes that neural networks can manage the interplay of lidar acquisition parameters, terrain properties, and vegetation characteristics to improve ABA models. The study area is in Massif des Bauges Natural Regional Park, eastern France, comprising 291 field plots in a mountainous environment with broadleaf, coniferous, and mixed forest types. Field plots were scanned with a high overlap from multiple flight lines and the corresponding point clouds were considered independently to expand the standard ABA dataset (291 observations) to create a dataset containing 1095 independent observations. Computation of lidar, terrain, and scan angle metrics for each point cloud associated each observation in the expanded dataset with the scan information in addition to the lidar and terrain information. A multilayer perceptron (MLP) was used to model basal area and total volume to compare the predictions resulting from standard and expanded ABA datasets. With expanded datasets containing lidar, terrain, and scan information, the R^2 for the median predictions per plot were higher (R^2 of 0.83 and 0.85 for *BA* and V_{tot}) than predictions with standard datasets (R^2 of 0.66(*BA*) and 0.71(V_{tot})) containing only lidar metrics. It also outperformed an MLP model for a dataset with lidar and terrain information [R^2 of 0.77(*BA* and V_{tot})]. The MLP performed better than Random forest regression, which could not sufficiently exploit additional terrain and scan information.**

***Index Terms*—Area-based approach (ABA), artificial neural networks (ANN), forest attribute, lidar, Random forest (RF), topography.**

I. Introduction

THE ability of lidar technology to create dense 3-D representations of vegetation has been widely used to extract useful information to characterize forest properties [1]. With airborne lidar systems (ALS), it is possible to cover large areas to generate accurately measured 3-D point clouds. The lidar system involves a scanning mechanism that emits lidar pulses at a range of incident angles, i.e., vertical to inclined. An ideal lidar acquisition would involve a comprehensive scanning of the forest from multiple angles to obtain an accurate and sufficiently dense representation in the form of 3-D point clouds. Area-based approaches (ABA) utilize such point clouds wherein statistical descriptors (lidar metrics) of point clouds of representative forest plots are statistically linked with field measurements of desired forest attributes. In the interest of reliable predictions using lidar metrics, lidar metrics should be stable under different acquisition characteristics.

An essential requirement in lidar remote sensing of forests is detecting the ground surface beneath the canopy, enabling accurate measurement of vegetation heights. As a result, the lidar scan angle, or the (half) field of view, has been limited to 20° to ensure most lidar pulses reach the ground [2]. So far, most studies involving lidar remote sensing for forestry applications have followed this convention [3], [4], [5], [6], [7], [8], [9]. Recently, some studies [10], [11], [12], [13] have tried to assess the impact of scan angles greater than 20°, and many studies involving UAV-based lidar data routinely used much higher scan angles [14], [15].

It may be difficult for highly inclined lidar pulses to reach the ground surface owing to the increased occlusions. Nonetheless, it is also true that probing lidar canopies with inclined pulses may also lend newer insights or different perspectives [16]. A related study [13] observed that datasets comprising nadir point clouds did not always result in better ABA models, thereby emphasizing that forest canopies are not a homogenous medium, and the lidar-derived information (lidar metrics) depends on how the lidar pulses sample the canopy. Furthermore, two lidar acquisitions may not have identical properties, and the lidar metrics could be affected by the overall acquisition geometry as characterized by the acquisition properties (sensor properties, scan angle, scan azimuth, flying height), terrain properties, and vegetation structural characteristics.

In area-based approaches (ABA), there are generally numerous lidar metrics to choose from, and new metrics are constantly being developed to comprehensively summarize the vegetation

Manuscript received 6 January 2023; revised 13 March 2023; accepted 20 March 2023. Date of current version 12 April 2023. This work was carried out in the framework of the PROTEST Project, supported by the French Agency for Ecological Transition (ADEME) under Grant 1703C0069, GRAINE program, and the FRISBEE Project, supported by the TOSCA Continental Surface program of the Centre National d'Etudes Spatiales (order 4500070632). The work of K. R. Dayal was supported by the French Region of Occitanie. *(Corresponding author: Karun R. Dayal.)*

Karun R. Dayal, Sylvie Durrieu, Kamel Lahssini, and Dino Ienco are with the INRAE, UMR-TETIS, University of Montpellier, 34000 Montpellier, France (e-mail: karunrdayal@gmail.com; sylvie.durrieu@inrae.fr; kamel.lahssini@inrae.fr; dino.ienco@inrae.fr).

Jean-Matthieu Monnet is with the INRAE, LESSEM, University Grenoble Alpes, 38400 Grenoble, France (e-mail: jean-matthieu.monnet@inrae.fr).

Digital Object Identifier 10.1109/JSTARS.2023.3263595

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

structural information. Standard metrics used over the years include statistical descriptors such as the mean, standard deviation, variance, entropy, and percentiles of the height or intensity values, cover rate metrics, density metrics, and gap-fraction. A stepwise selection procedure is often employed to identify metrics useful in predicting forest attributes using multiple linear regression [17]. However, the final set of metrics may vary depending on the forest type or lidar acquisition parameters. Another approach is to use expert knowledge to define and select a short list of metrics that could explain most, if not all, of the variance of the dependent variables [18]. Still, assessing the influence of scan angle on selected metrics and, subsequently, on the forest attribute predictions may not always be practical.

Furthermore, the influence of lidar metrics may be site-specific, and it is advisable to assess the effects of scan angle before further analysis on a case-by-case basis [19], [20]. Traditionally, the modeling of forest attributes is done using parametric and non-parametric models. Due to their simplicity, parametric methods such as ordinary least squares (OLS) regression have been widely used by studies to model forest attributes [17], [21], [22]. However, parametric methods such as OLS use only a few metrics to avoid overfitting and the use of correlated variables. In non-parametric methods, there are no such limitations as these methods do not depend on any assumptions regarding the data and can accommodate nonlinear relationships between the dependent and independent variables [23]. Therefore, they are suited for modelling complex interactions between several lidar variables, acquisition geometry, and vegetation properties. Both k-nearest neighbors (KNN) and Random forest (RF) are among the most commonly used nonparametric approaches in ABA [24]. However, RF was found to have a higher level of transferability to new areas than KNN [21]. Artificial neural networks (ANN) are a popular nonparametric method to address inherent non-linearity in datasets [25], [26]. The feed-forward back-propagation multilayered perceptron (MLP) is often used with remote sensing data [25]. It consists of a network of several interconnected layers of neurons designed to mimic human brain capabilities, such as generalization and understanding complex patterns. Among the various nonparametric methods, the MLP has been demonstrated to have better generalization capabilities [27], [28].

However, MLP methods depend on the volume of data, which generally comprises large datasets with several thousand samples. In ABA approaches, which involve collecting labor-intensive field measurements in often-complex terrains, it is impossible to measure many field plots (samples) as field measurements make up a significant part of the costs. The number of field plots in ABA models typically ranges from a few tens to a few hundred. In addition, only a few field plots describe particular stand types. Generally, lidar acquisitions for forests are planned with multiple overlaps to thoroughly sample each forest area (or field plot) from multiple locations, and the point clouds acquired from each location may be considered independent observations. As vegetation can be quite heterogeneous when viewed from different perspectives, point clouds acquired with various angles also tend to reflect these differences. Therefore, the point clouds could be used independently to increase the number of observations. In other words, a point cloud obtained from a flight line is defined by the interplay of acquisition parameters, terrain properties, and vegetation characteristics, making it possible to consider it a unique and independent observation in ABA models. The combination of acquisition properties and (virtually limitless) lidar metrics for each observation pose challenges that are better handled with non-parametric methods.

A previous study over mountainous terrain [30] demonstrated that neural networks are well-suited to exploit terrain-related information to improve standard ABA predictions. In another study over the same area [13], single flight line datasets resulted in variable ABA predictions with parametric models comprising widely used metrics sensitive to scan angle. In [13], Dayal et al. attempted to normalize for the effects of inclined lidar scanning with the aim to homogenize the lidar metrics irrespective of the acquisition conditions. In the present study, we adopted an alternative perspective on the influence of acquisition geometry on lidar metrics and subsequent ABA predictions. The objective of this study was to evaluate the benefit of using individual point clouds obtained from different flight lines independently to 1) increase the number of observations, and 2) retain acquisition properties of each flight line to improve ABA predictions. We demonstrate 1) the benefits of expanding lidar datasets based on flight lines to build ABA models, and 2) the capacity of multilayer perceptron to model complex interactions between lidar signal and acquisition properties.

II. Materials and Methods

A. Study Area and Field Measurements

The study site is the Massif des Bauges Natural Regional Park in the French Alps. It is located between the two administrative departments of Savoie and Haute-Savoie and covers an area of approximately 850 km^2. The terrain is hilly (plot altitudes range from 420 to 1760 m). The most common tree species comprise silver fir (*Abies alba*), Norway spruce (*Picea abies*), and common beech (*Fagus sylvatica*). Field inventory involved measurements of 291 plots during the spring to autumn of 2018. The plot radius was 15 m. Plot center locations were measured using differential GNSS (DGNSS, Trimble, USA). Field inventory protocol involved measuring tree Diameter at Breast Height (DBH, measured 1.3 m above ground) of trees with DBH greater than 17.5 cm. Small trees (7.5 cm $\leq$ DBH $<$ 17.5 cm) were only counted within a plot radius of 10 m and classified as either coniferous or broadleaf.

Since DBH and height measurements were unavailable for all the trees with DBH greater than 7.5 cm, computation of basal area, stem, and total volumes at plot level required estimations for unmeasured trees. First, the number of small trees was extrapolated from the number of trees in 10 m radius plots to 15 m radius plots. Second, the nationwide tree inventory database (NFI) generated by IGN (Institut National de l'Information Géographique et Forestière) containing measurements of trees with DBHs in the 7.5 to 17.5 cm range was used to extrapolate DBH and height values for nonmeasured trees. All NFI plots in the ecoregion that includes the study site were selected to have forest plots with similar climatic and growing conditions

TABLE I
SUMMARY OF THE BASAL AREA (*BA*) AND TOTAL VOLUME (V_{TOT}) FOR THE 291 INVENTORY PLOTS

Basal area (m^2/ha)			Total volume (m^3/ha)		
Min	Mean	Max	Min	Mean	Max
0.36	30.2	89.7	2.52	312.1	1172

TABLE II
ACQUISITION PARAMETERS FOR THE TWO MISSIONS

	73	74
Date of acquisition	September 2016	September 2018
Sensor	Leica ALS70-HP	Riegl LMSQ780
Wavelength (nm)	1064	1064
Scan angle (deg)	46° (+23°/-23°)	60° (+30°/-30°)
Beam divergence (mrad)	0.15	≤0.25
Ground speed (m/s)	85	45
Point density (pts/m^2)	4	14
Flight height (AGL) (m)	1500*	1050

to those measured on the study site. For trees with DBH ranging from 7.5 to 17.5 cm, the median DBH value in the NFI database is 11.1 cm. This value was used to compute the basal area of the trees with DBHs lower than 17.5 cm. Using NFI measurements, allometric relationships were established for each species (or group of species when the number of trees was not high enough) to estimate the heights of all the trees when there were no available height measurements. Volumes were then computed using the allometric equations available in [29]. Dayal et al [13] and Lahssini et al. [30] followed the same protocol. A summary of field measurements is given in Table I.

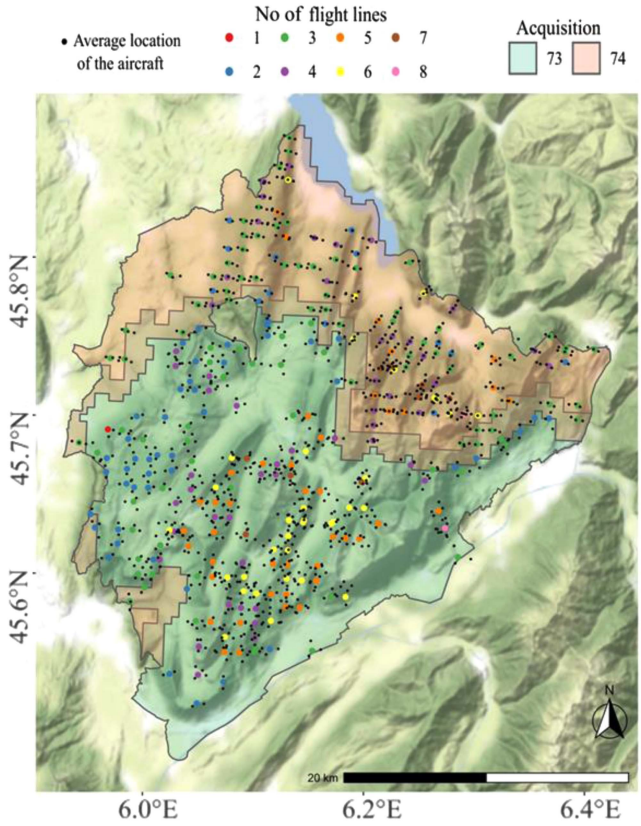


Fig. 1. Location of the study site, distribution of field plots, and coverage of lidar missions. Field plots are depicted in larger dots with the colors corresponding to the number of unique flight lines from which the plots were scanned completely. The black dots depict the approximate (average) location of the aircraft when it scanned a field plot (Map tiles by stamen design, under CC BY 3.0. Data by OpenStreetMap, under ODbL).

B. Lidar Data

Lidar data acquisition was carried out in two missions. The first mission (summer 2016) covering areas of department 73 resulted in a dataset of 4–5 points/m^2 (first returns) density on average and the second mission (summer 2018) covering areas in department 74 resulted in a dataset of approximately 14 points/m^2 (first returns) on average. Lidar acquisitions were carried out with multiple overlaps to scan each field plot from several locations with different azimuths and scan angles. Fig. 1 shows the locations of the field plots in the study area along with aircraft locations while scanning respective field plots. The acquisition parameters for the two missions are given in Table II.

C. Splitting of Point Clouds Based on Flight Lines

Point clouds corresponding to the field plots were clipped from the lidar data using the coordinates of the plot centers and plot radius (15 m). Due to flight line overlaps, the point cloud for a given plot is typically a composite of point clouds acquired with different scanning configurations. The point cloud was split for each plot based on the constituent flight lines. Each resulting constituent point cloud was characterized by the mean of the scan angles (*MSA*) with which it was scanned. We did not consider those point clouds acquired with *MSA* greater than 30° as they were most likely acquired when the aircraft made turns, and there were few such instances. The fundamental "unit" in our experiments is the point cloud for a plot acquired from a single flight line. We assessed pulse densities for each point cloud, and 90% of the constituent point clouds had a pulse density greater than one pulse per m^2. We computed the area covered by each constituent point cloud by fitting a two-dimensional hull to the points projected onto a horizontal plane. Then, an area threshold was used to drop any constituent point cloud that covered less than 90% of the total plot area (see Fig. 2).

The flight trajectory data was used to extract the locations of the aircraft while scanning respective field plots, and the average location of the aircraft was computed. The azimuth of the scan was calculated as the angle with respect to the geographic north between the average location of the aircraft and the respective field plot center. Each point cloud results from the unique scan geometry that is characterized by the mean scan angle, scanning azimuth, and scanning distance. Therefore, two kinds of datasets were considered. In the first kind of dataset, point clouds were not separated based on the flight lines. This dataset was called the *standard* dataset and contained as many point clouds as field plots in the study (291). Point clouds were separated based on the flight line information for the second dataset, called the *expanded* dataset. In the *expanded* dataset, there was one to

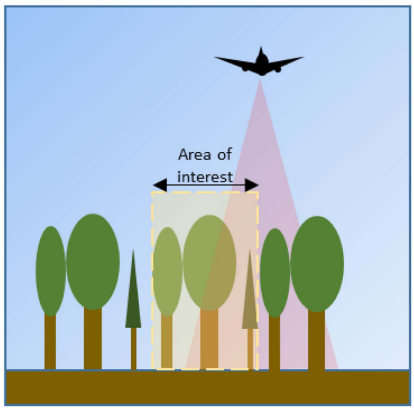


Fig. 2. Flight line that partially covers a plot.

eight point clouds per field plot resulting in 1095 point clouds and corresponding to a mean of 3.8 point clouds per plot (see Fig. 1).

D. Lidar Metrics

Lidar metrics were computed for each point cloud in both datasets after normalizing the point clouds in height, i.e., transforming the point elevation into height above the ground using the lidar-derived DTM. All points below a height threshold of five meters were considered lower vegetation and filtered out. Fifty-five metrics related to heights, intensities, and canopy were computed. The height-based metrics are the statistical distributions calculated for the Z values of the point cloud. The intensity metrics comprise statistical descriptors of the intensity values. Canopy metrics consist of gap fraction [31] and rumple index [32]. The gap fraction was computed as the ratio of the number of returns below the 5 m threshold to the total number of returns. Rumple index is the ratio of the 3-D surface area of the canopy to the surface area of the ground computed for the normalized point cloud. Gap fraction and rumple index were found to be very sensitive to the scan angle [33]. The summary of these metrics is given in Table III.

Point clouds can change according to the local topography and viewing configuration for a given forest plot. Depending on the slope, orientation, elevation, scan angle, and aircraft position, there could be several cases. In Fig. 3, the illustrations depict two possibilities of lidar scanning a plot on a slope with similar scan angles, 1) scanning along the slope and 2) scanning against the slope. An example of a point cloud scanned from different directions is shown in Fig. 3(c), wherein two point clouds with similar mean scan angles can have different properties due to the interaction between terrain properties and scanning parameters. Information on scan geometry and terrain properties were thus added as six additional variables, resulting in a total of 61 variables (see Table III). The three previously defined variables defined scan geometry, i.e., mean scan angle, scanning azimuth, and scanning distance (see Section II-C), differed for the same plot location according to the flight line. Terrain information was computed by generating digital terrain models (DTM) of a resolution of 1 m. The DTM of each plot was used to generate slope and aspect maps. The average slope, aspect, and elevation values were computed from the slope and aspect maps and DTMs, respectively, and were replicated for all the point clouds acquired from different flight lines over a given plot. As the overall point cloud geometry is defined by the acquisition parameters, terrain properties and vegetation properties, point clouds obtained from different flight lines were used independently to compute all the metrics mentioned earlier to retain the heterogeneities and homogeneities.

TABLE III
SUMMARY OF THE METRICS OBTAINED FROM LIDAR DATA, TERRAIN PROPERTIES, AND SCAN GEOMETRY

Name	Nos	Description	Type
pnth | n ∈{1,2,3,4,5}	5	% of nth echos	Echo-based
kurtosis_first	1	Kurtosis of first returns	Height-based
kurtosis_last	1	Kurtosis of intensity	
pzabovezmean	1	Percentage of points above mean	
skewness_first	1	Skewness of first returns	
skewness_last	1	Skewness of intensity	
zentropy	1	Entropy	
zmax	1	Maximum height value	
zmean	1	Mean height value	
zpcumn | n ∈{1,2...9}	9	Density	
zqn | n ∈{10, 15...95}	19	Percentile	
zsd	1	Standard deviation height value	
ikurt	1	Kurtosis of intensity	Intensity-based
imax	1	Maximum intensity	
imean	1	Mean intensity	
ipcumzqn | n ∈{10, 30...90}	5	Cumulative intensity proportion	
isd	1	Standard deviation of intensity	
iskew	1	Skewness of intensity	
itot	1	Total intensity	
gap_F	1	Gap fraction	Canopy
rumple_index	1	Rumple index	
asp	1	Aspect of plot	Acquisition geometry
az	1	Azimuth of acquisition	
dist	1	Distance of scanner	
ele	1	Elevation of plot	
meanang	1	Mean scan angle	
slope	1	Slope of plot	
Total	**61**		

For the *expanded* dataset, the values of the dependent variables, i.e., BA and V_{tot} were also replicated for each plot depending on the corresponding number of flight lines or point clouds. All the values were scaled between 0 and 1. All lidar metrics were computed using the lidR package in R [34].

E. Experiments and Cross-Validation Scheme

When the point clouds were considered per plot, i.e., the standard dataset, there were 291 samples. The same splits of data as those used for cross-validation of the models in [30] were used

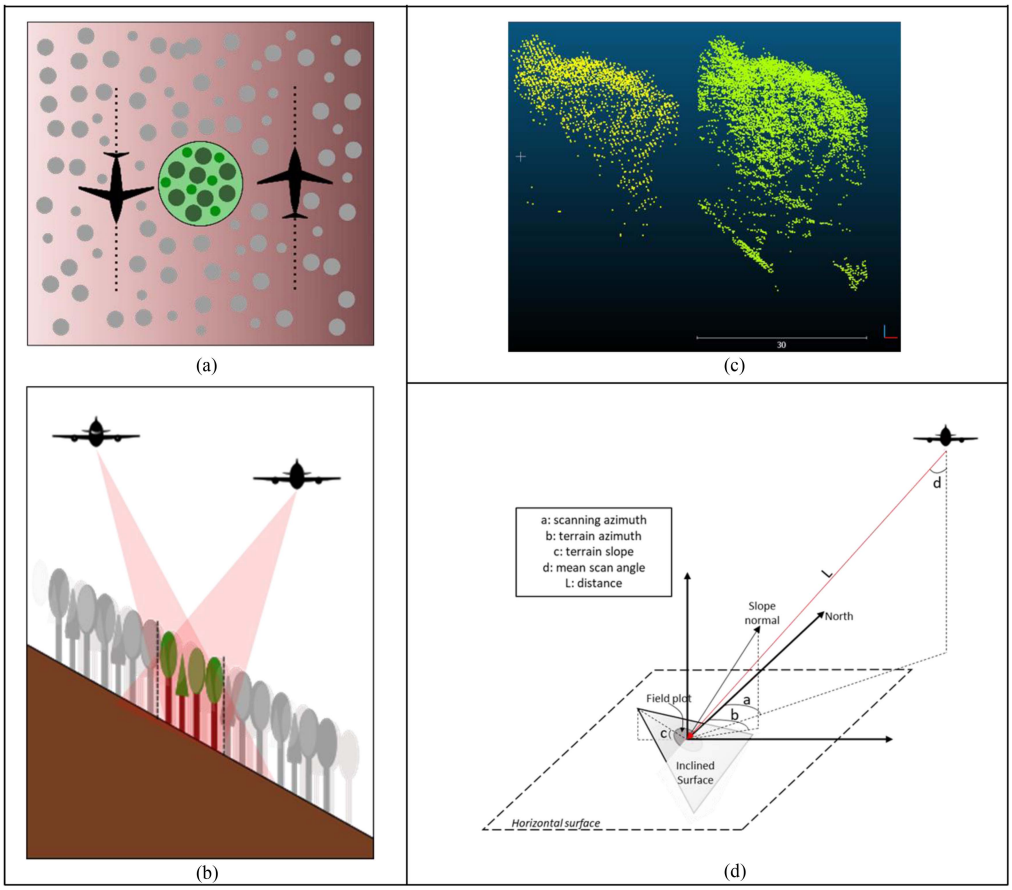


Fig. 3. Illustration of lidar scanning along the slope and against the slope. (a) Top view; (b) side view; (c) example point clouds with similar mean scan angles 26° (yellow) and 22° (green). (d) Example scan geometry with relevant parameters.

to compare results from our study directly. The standard dataset containing the field plot measurements and corresponding lidar metrics were split into training and test sets. The training set was further subdivided into training (191 field plots or samples) and validation set (50 field plots or samples), with roughly an 80: 20 ratio. The test set (50 field plots or samples) was completely blind to the training and validation sets. When the point clouds were considered per plot and flight line, there were 1095 samples in total in this expanded dataset, while the number of field plots was still the same (291). We used the field plot ids from the test set of the standard dataset to create the corresponding test (and training) set(s) of the expanded datasets to ensure that point clouds for the same sample field plots were not present in both the test and training sets. However, the training set was randomly divided into an 80:20 ratio. The cross-validation scheme is illustrated in Fig. 4. Thirty different splits containing training, validation, and test data were generated for all the datasets.

The benefit of the data expansion strategy was tested via the following three experiments.

1) The *standard* dataset *(std)* comprising only the lidar metrics was used to build a model and then compared with a model built with the *expanded* dataset *(exp)* comprising only the lidar metrics.

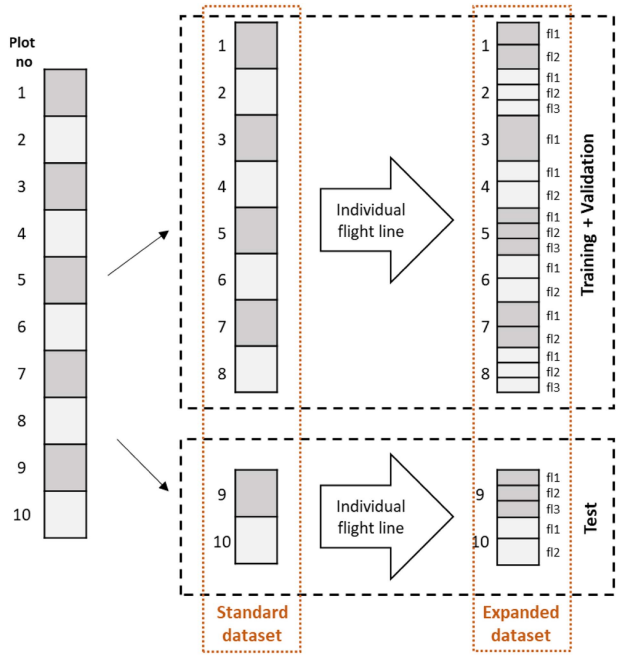


Fig. 4. Illustration of the cross-validation scheme for standard and expanded datasets. Each plot comprises one or more flight lines in the expanded dataset. The process was repeated 30 times (30 splits).

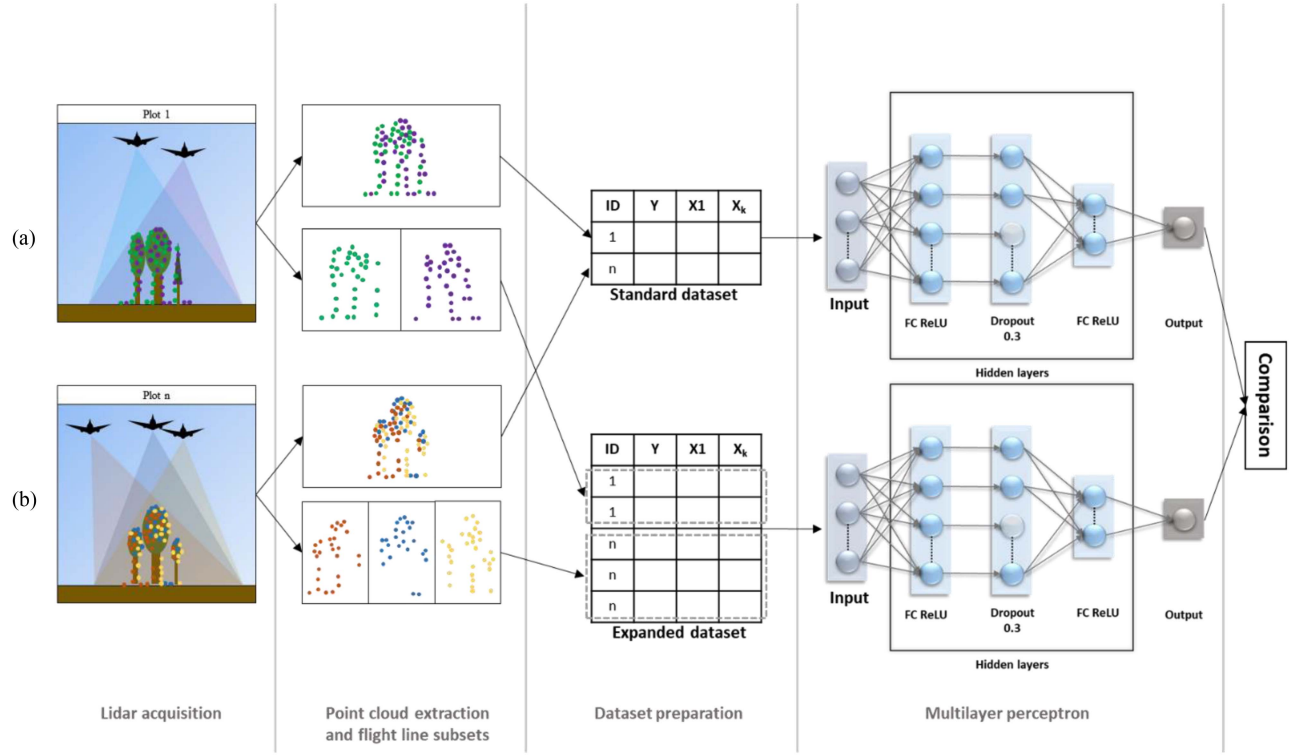


Fig. 5. Conceptual workflow employed in the study using a multilayer perceptron. (a) Standard ABA workflow. (b) Expanded dataset workflow. The MLP stage (i.e., modelling stage) comprises the tuning of the hyperparameters followed by model fitting for all the splits of the data. FC: fully connected, ReLU: rectified linear unit. The comparison stage comprises the computation of goodness-of-fit criteria.

2) The *standard* dataset was appended with terrain variables (slope, azimuth, and elevation) ($std_{terrain}$) to build a model and compared to a model built with the *expanded* dataset also appended with terrain variables ($exp_{terrain}$).
3) The *expanded* dataset was appended with both terrain and scan geometry variables ($exp_{terrain+\mathrm{scan}}$) and compared to both the *standard* dataset appended with terrain variables (scan geometry variables not applicable for this data set) and the *expanded* dataset appended with only terrain variables ($exp_{terrain}$).

Indeed, scan geometry features are not available along with standard lidar metrics as, in this case, the point cloud results from a merging of scanning configurations. The conceptual workflow used in the study is illustrated in Fig. 5.

F. Regression Models

We used the TensorFlow (2.6.0) library in Python (3.9.7) for the fully connected MLP [35]. The MLP network consisted of two hidden layers. Each neuron in a layer is fully connected (FC) to all the neurons in the following layer. The components of the MLP include the input layer, two hidden layers separated by a dropout, and an output layer. The KerasTuner [36] was used for the hyperparameter optimization to obtain three hyperparameters: number of neurons in the first and second hidden layers and the learning rate of the adaptive moment estimation (ADAM) optimizer. We used the RandomSearch tuner from Keras to find the optimum parameters. The candidate number of neurons in each layer were chosen from a set of possible values {32, 64, 128, 256, 512, 1024} and the learning rate was chosen from a set of possible values {1e−5, 1e−4, 1e−3}. The rectified linear unit (ReLU) function was used as the activation function. It defines how the input values it receives are output to the next neuron. A dropout rate of 0.3 was used to regularize the network to prevent overfitting. The ADAM optimizer was used to optimize the network with mean squared error regression loss. After tuning for the hyperparameters, the chosen model was the trained with each split (30 splits of training sets) with 200 epochs and batch size 1. The network is illustrated in Fig. 5.

RF models were also built for all the datasets to compare the performances with the MLP network. The number of trees built was 500, and mtry value was set to default, i.e., the number of independent variables divided by 3. The model was implemented using the randomForest package in R [37].

G. Model Accuracy Assessment

Once a model was trained with a training set of a split, it was applied to the corresponding test set to predict the forest attributes. The goodness-of-fit of the MLP and RF models was assessed using the determination coefficient (R^2), the root mean squared error (RMSE), the relative RMSE (rRMSE), and the mean percentage error (MPE). The formulae for these measures are as follows:

$$R^2 = 1 - \frac{\sum_1^n (y_i - \hat{y}_i)^2}{\sum_1^n (y_i - \bar{y})^2} \quad (1)$$

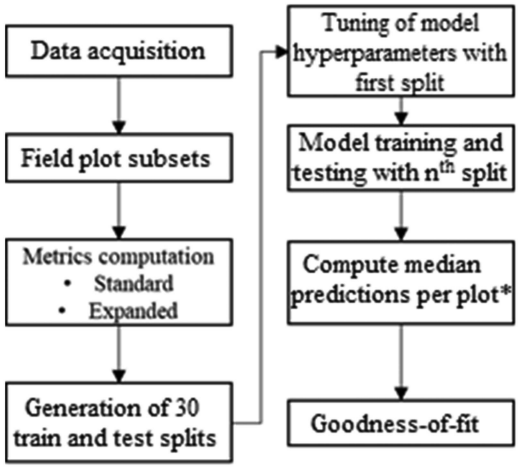


Fig. 6. Block representation of sequence of steps implemented with the MLP. The process is applied to each of the datasets described in Section II-E. There are 30 splits. ∗Applicable to expanded datasets only.

$$\text{RMSE} = \sqrt{\frac{\sum_1^n (\hat{y}_i - y_i)^2}{n}} \tag{2}$$

$$\text{rRMSE} = \frac{\text{RMSE}}{\bar{y}} * 100 \tag{3}$$

$$\text{MAE} = \frac{\sum_1^n |(y_i - \hat{y}_i)|}{n}. \tag{4}$$

Depending on the number of flight lines that scanned a given plot, there could be multiple predictions per plot for models built with the three kinds of expanded datasets (exp, $exp_{terrain}$, and $exp_{terrain+scan}$). The median value was considered for computing the goodness-of-fit criteria. The sequence of steps if given in Fig. 6.

H. Variable Importance

Notwithstanding the benefits of ANNs, one of the challenges is that they are considered "black boxes" with no clear indication of how they use the data provided to them to result in predictions. ANN methods are not ideal when interpreting the models is required. Lundberg and Lee [38] proposed the SHAP (Shapely Additive exPlanations) values to identify the feature importance in predictions based on the Shapely values [39].

SHAP values are calculated by measuring the impact of a given variable in various combinations with other dataset variables. It is based on the game theory wherein the marginal contribution of one variable (player) in the presence of other variables (players) is estimated. SHAP values were computed for the variables for each split of the data to understand how various lidar metrics contributed to our predictions. We also present the variable importance output (%incMSE) provided by the Random Forest models. The mean values across thirty splits were reported.

TABLE IV
SUMMARY OF THE TUNED HYPERPARAMETERS FOR DIFFERENT EXPERIMENTS (NEURONS IN THE FIRST HIDDEN LAYER, NEURONS IN THE SECOND HIDDEN LAYER, LEARNING RATE)

Dataset	Basal area	Total volume
Standard (metrics) or *std*	(256, 32, 1e-3)	(256, 64, 1e-3)
Standard (metrics + terrain) or $std_{terrain}$	(256, 64, 1e-3)	(256, 32, 1e-3)
Expanded (metrics) or *exp*	(1024, 32,1e-3)	(1024, 128, 1e-3)
Expanded (metrics + terrain) or $exp_{terrain}$	(1024, 128, 1e-3)	(1024, 128, 1e-3)
Expanded (metrics + terrain + scan geometry) or $exp_{terrain+scan}$	(1024, 64, 1e-3)	(1024, 64, 1e-3)

III. RESULTS

A. Hyperparameter Tuning

Tuning for each dataset resulted in different hyperparameters (see Table IV). For the standard datasets, there were 256 neurons, while for the expanded datasets, the tuning resulted in 1024 neurons in the first hidden layer. For the second hidden layer, the number of neurons varied between 32, 64, and 128. The learning rate was 1e−3.

B. Model Performances

The R^2, MAE, RMSE, and rRMSE are presented in Table V for both MLP and RF models. The observed and predicted values for *BA* and V_{tot} are shown in Figs. 7 and 8, comparing the results of the MLP ABA models built with 1) *std* and *exp* datasets, 2) $std_{terrain}$, and both $exp_{terrain}$ and $exp_{terrain+scan}$, respectively.

In Figs. 7 and 8, the regression lines reveal biases in the predictions with the MLP to different degrees for all the datasets. The plots with higher values of *BA* and V_{tot} have underestimated predictions, and those with lower values have slight overestimations, especially for *BA*. In all the cases, the MLP systematically outperformed RF. The RF R^2 was lower by 19% (*BA* and V_{tot}, std). All three error measurements, i.e., MPE, RMSE, and rRMSE%, were higher for RF than MLP. For example, the rRMSE values with $exp_{terrain+scan}$ datasets were higher for RF by approximately 60% for both *BA* and V_{tot}.

Regarding the datasets used, the trends were broadly similar for MLP and RF, but the rate of improvement with additional variables was higher for the MLP. The lowest model accuracies were observed for the model built with the std datasets (*BA*: $R^2 = 0.66$ and 0.53, rRMSE = 30.5% and 35.8%; V_{tot}: $R^2 = 0.71$ and 0.57, rRMSE = 34.4% and 35.8%, for MLP and RF, respectively). The *exp* datasets demonstrated relative improvements for the MLP (*BA*: $R^2 = 0.76$; rRMSE = 26%; V_{tot}: $R^2 = 0.78$, rRMSE = 30.5%) with 15% and 10% increase in R^2 and 15% and 11% percentage points reduction in the rRMSE for *BA* and V_{tot}, respectively. Incorporating terrain properties ($std_{terrain}$) resulted in better models with both MLP and RF. However, in contrast to the MLP, RF models only marginally benefited from the data expansion (exp datasets). *BA* predictions improved marginally (R^2:0.54 for exp versus 0.53 for std), while V_{tot}

TABLE V
COMPILATION OF THE GOODNESS-OF-FIT CRITERIA FOR ALL THE EXPERIMENTS

Forest attribute	Dataset	R^2		MPE		RMSE		rRMSE%	
		MLP	RF	MLP	RF	MLP	RF	MLP	RF
Basal area	*std*	0.66	0.53	6.7	7.9	8.9	10.4	30.5	35.8
	$std_{terrain}$	0.77	0.58	5.4	7.4	7.3	9.8	25	33.6
	exp	0.76	0.54	5.7	7.8	7.5	10.4	26	35.7
	$exp_{terrain}$	0.81	0.61	5	**7.2**	6.7	**9.5**	21.3	**32.6**
	$exp_{terrain+scan}$	**0.83**	**0.60**	**4.7**	**7.2**	**6.2**	**9.5**	**19.9**	**32.6**
Total volume	*std*	0.71	0.57	78.2	95.0	103	126.2	34.4	42.2
	$std_{terrain}$	0.77	0.62	71.7	89.2	96.2	117.5	32.2	39.3
	exp	0.78	0.55	68.2	96.8	91.3	129.7	30.5	43.2
	$exp_{terrain}$	0.83	0.64	57.6	86.5	77.9	115.1	26	38.3
	$exp_{terrain+scan}$	**0.85**	**0.64**	**54.6**	**86.2**	**71.8**	**114.6**	**24**	**38.1**

predictions deteriorated (R^2:0.55 for exp versus 0.57 for std). RF models with $exp_{terrain}$ were better than those with $std_{terrain}$ for both *BA* (R^2:0.58 for $std_{terrain}$ versus 0.61 for $exp_{terrain}$) and V_{tot} (R^2:0.62 for $std_{terrain}$ versus 0.64 for $exp_{terrain}$) with increases of 5% and 3% in R^2 values, respectively. Their errors were reduced in the range of 3%–5%.

The MLP models built with $std_{terrain}$ resulted in better goodness-of-fit values (*BA*: $R^2 = 0.77$, rRMSE = 25%; V_{tot}: $R^2 = 0.77$, rRMSE = 32.2%) than those built with std datasets (*BA*: $R^2 = 0.66$, rRMSE = 30.5%; V_{tot}: $R^2 = 0.71$, rRMSE = 42.2%). The goodness-of-fit of models built with $exp_{terrain}$ datasets were better for both *BA* and V_{tot} (*BA*: $R^2 = 0.81$, rRMSE = 25%; V_{tot}: $R^2 = 0.83$, rRMSE = 32.2%) with increases of 5% and 8%, respectively. The three error goodness-of-fit criteria (MAE, RMSE, and rRMSE) reduced in the 7% to 15% range and around 19% for *BA* and V_{tot}, respectively. Incorporating additional information about the scan geometry ($exp_{terrain+scan}$) resulted in slightly better MLP models with 3% higher R^2 values and lower errors in the 5%–8% range for both *BA* and V_{tot}. The RF models with $exp_{terrain+scan}$ datasets did not result in any improvements (less than a 1% reduction in errors). In addition, in Fig. 7, the saturation problem of underestimating large values is evident. It appears to be well-handled by the MLPs when combined with terrain and scan information (see Fig. 8).

C. Variable Importance

In MLP predictions, gap fraction and elevation were the two most important variables in predicting *BA* and V_{tot}. The distance metric was also ranked third in both instances, indicating its influence in the prediction with $exp_{terrain}$ datasets. The slope of the plot was the other acquisition geometry variable that was present among the top ten variables. The intensity metrics i_{tot} and i_{mean} were also present among the top ten variables in *BA* prediction while only i_{tot} was present among the top ten variables for V_{tot} prediction. Several height-based metrics such as the density metrics (zpcum3, zpcum5 and zpcum7 and zpcum6) were present among the top 15 variables. The second canopy metric, i.e., rumple index, was ranked only at the 25th and 22nd place for Basal area and Total volume predictions, respectively, and the mean scan angle was ranked at the 31st and 30th place.

In comparison, for the RF models of *BA* and V_{tot}, slope and aspect were ranked within the top five variables. Gap fraction and plot elevation were also highly ranked. Compared to the MLP variable importance, the distance variable was less critical, and the rumple index (the second canopy variable) and the gap fraction were ranked better. The height-based density metrics were relatively lower ranked, with only z_{pcum8} and z_{pcum9} in the top ten variables. Some metrics found important in the MLP predictions, such as i_{tot}, i_{mean}, and p2th were of lower importance in RF predictions. The traditional metrics, such as z_{max}, z_{mean}, or z_{q95} were found to be more important metrics.

IV. DISCUSSION

A. Model Tuning

We rebuilt the models with std and $std_{terrain}$ [30] and observed that the goodness-of-fit criteria were higher for models with both std and $std_{terrain}$ (*BA*: R^2 of 0.66 and 0.71; V_{tot}: 0.71 and 0.77) compared to [30] (*BA*: R^2 of 0.61 and 0.69; V_{tot}: 0.67 and 0.74). This could be due to variations in hyperparameter tuning resulting in better models and underlining the importance of the tuning the various hyperparameters to obtain better model performances.

The random initializations of the algorithms used in the models may yield varying hyperparameters. In our experiments, we observed the learning rate was often tuned to 1e−3. In contrast, the learning rate was fixed at 1e−4 based on expert knowledge and was not included in the parameter optimization step in [30], which could explain the differences in model performances in the present study for std and $std_{terrain}$ datasets.

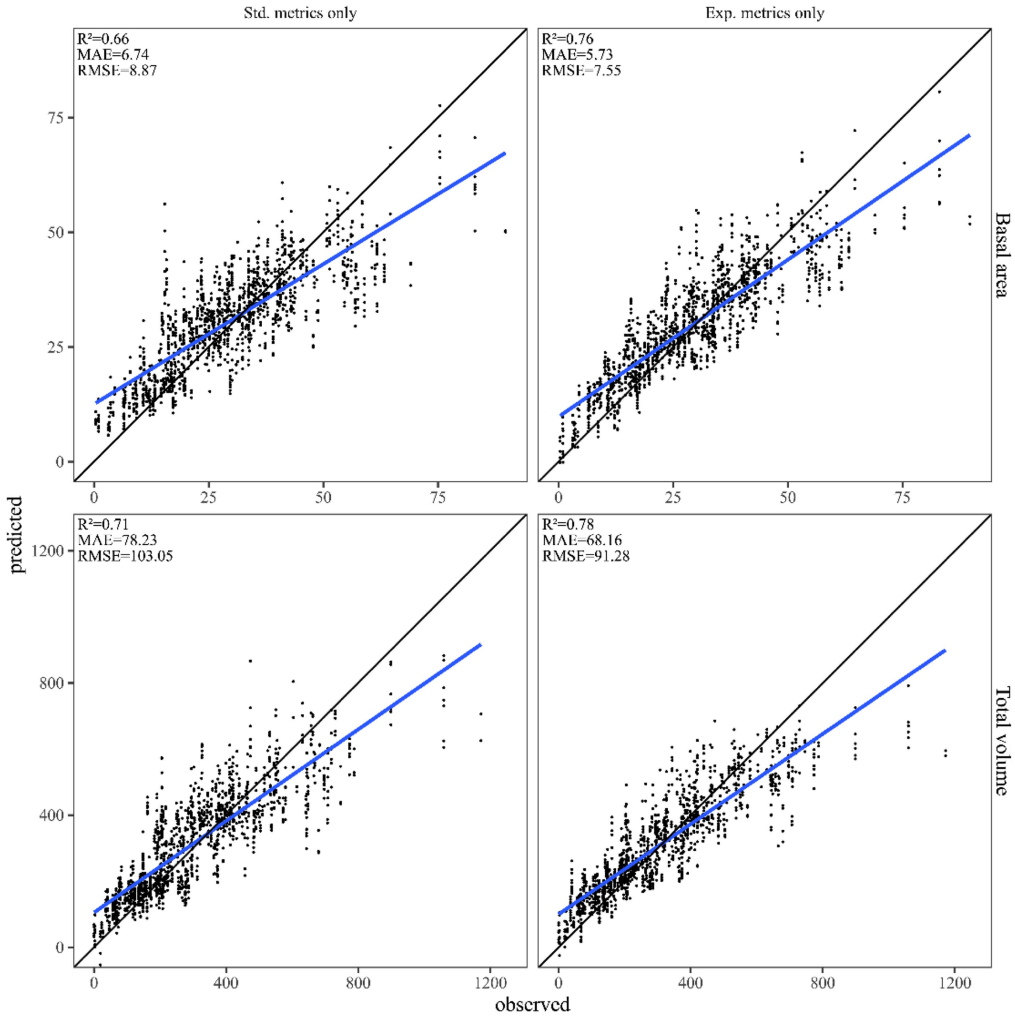


Fig. 7. Multilayer perceptron scatterplots of predicted and observed values for models built with std and exp datasets with only lidar metrics. Std datasets comprise metrics computed from standard lidar point clouds per plot obtained from one or many flight lines. Exp datasets include metrics computed by considering lidar point clouds obtained from each flight line as independent observations. Multiple point clouds per plot in expanded datasets are not to be confused with multiple *Y* values for the same *X* values in this figure, as this indicates the repeated occurrence of a plot in the test sets in different data splits.

Additionally, NNs that were wider in the first hidden layer compared to the second hidden layer resulted in consistent predictions. It was indeed possible that in some tuning results, the hyperparameters of the best-tuned model were different from what were eventually employed in this study. Therefore, selecting a suitably tuned model involved some degree of empirical assessment of the architecture of the NN via a trial and error procedure.

B. Multiangle Perspectives of the Same Vegetation

A prevalent problem regarding saturation was also observed in this study with *std* datasets for both basal area (*BA*) and total volume (V_{tot}). The saturation issue was handled well by a deep-learning-based fusion strategy using lidar and optical (Sentinel-2) datasets [30]. In this study, however, the saturation effects appear less apparent for models built with expanded datasets consisting of terrain properties ($exp_{terrain}$ and $exp_{terrain+scan}$ datasets). The changes in lidar point cloud properties due to different acquisitions may convey information on species composition at the plot level, as done by the Sentinel-2 time series [30].

The models with expanded datasets consistently outperformed those with corresponding standard datasets. Between $exp_{terrain}$ and $exp_{terrain+scan}$, the R^2 values improved for MLP models from 0.81 to 0.83 and 0.83 to 0.85 for *BA* and V_{tot}, respectively. It appears that the improvement was not of a large magnitude. However, it is worth noting that all error measurements decreased (rRMSE reduced by 1.4 and 2 percentage points for *BA* and V_{tot}) and that the saturation effect was better handled with acquisition geometry variables. On the other hand, RF could not provide comparable results (see Fig. 11). The goodness of fit criteria for RF did not change between the two data sets, which could be attributable to the fact that there may be some redundancy in the information offered by the point clouds that were considered independent observations. Fig. 3(a) and (b) illustrate the differences in point clouds due to slope, even if the scan parameters are nearly similar due to the steep slope.

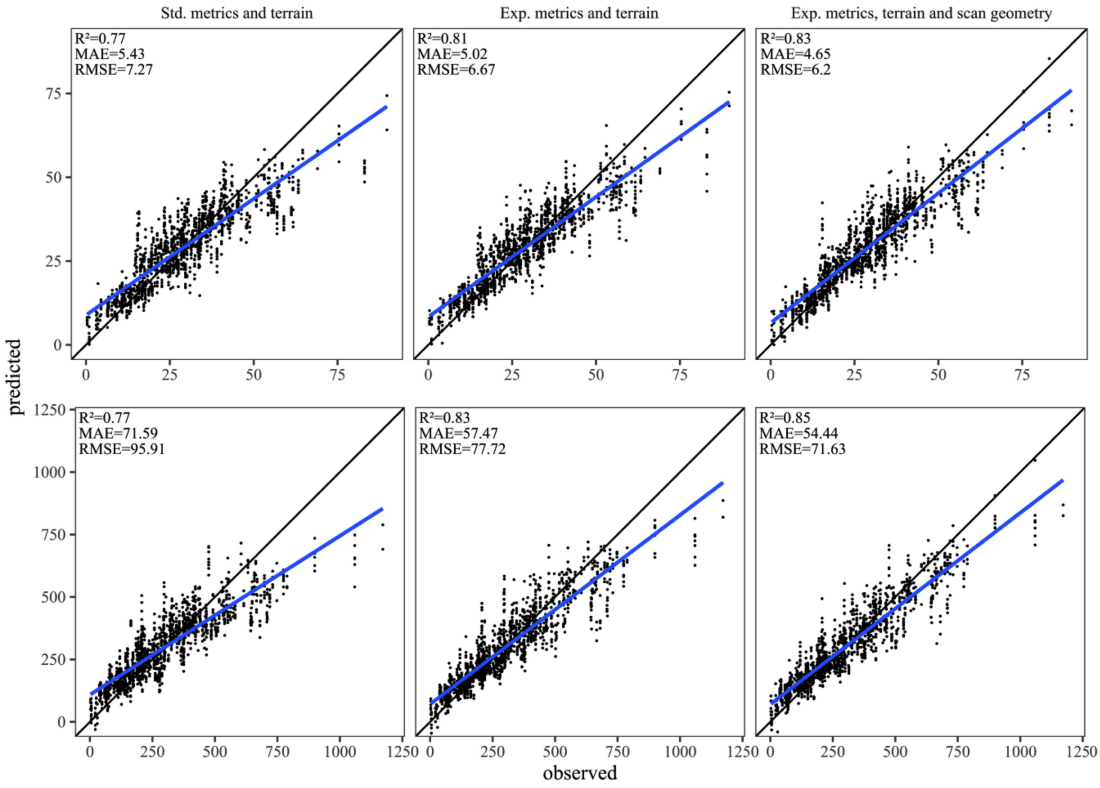


Fig. 8. Multilayer perceptron scatterplots of predicted and observed values for models built with std and exp datasets with lidar metrics and terrain properties (left and middle panels) and with expanded datasets with lidar metrics, terrain and scan properties (right panels) for *BA* (top panels) and V_{tot} (bottom panels). Std datasets comprise metrics computed from standard lidar point clouds per plot obtained from one or many flight lines. Exp datasets comprise metrics computed by considering lidar point clouds obtained from each flight line as independent observations. Multiple point clouds per plot in expanded datasets are not to be confused with multiple *Y* values for the same *X* values in this figure as this indicates the repeated occurrence of a plot in the tests sets in different splits of data.

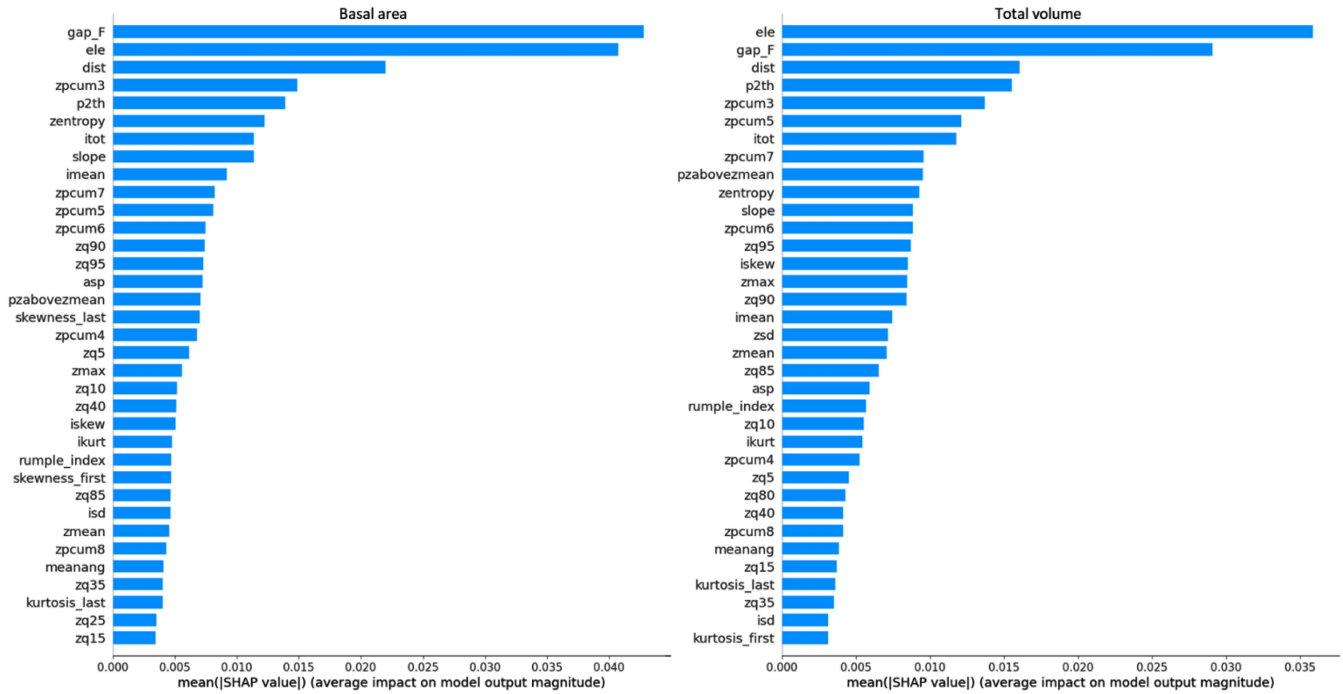


Fig. 9. SHAP values summary plot of all the metrics used in this study ($exp_{terrain+scan}$) for *BA* (left) and V_{tot} (right) MLP predictions. The values are ranked based on their importance in predictions.

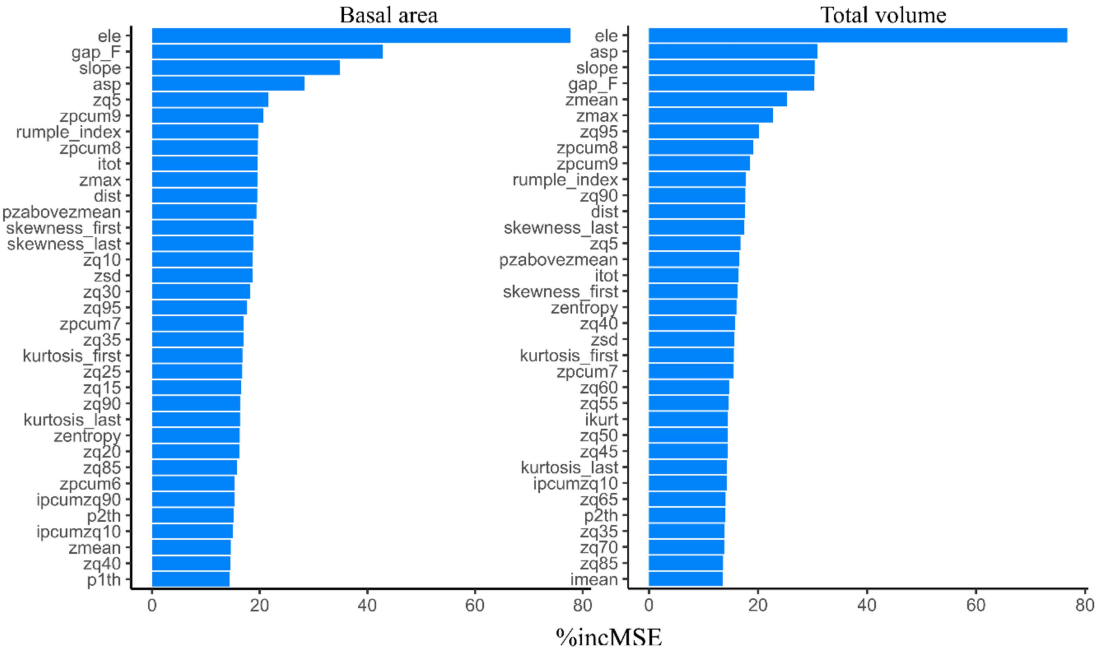


Fig. 10. %incMSE values summary plot of all the metrics used in this study ($exp_{terrain+scan}$) for *BA* (left) and V_{tot} (right) RF predictions.

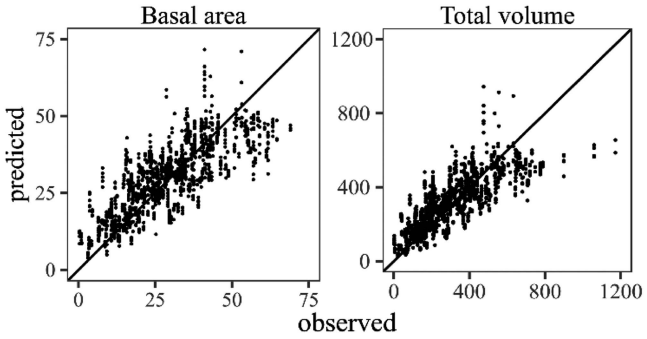


Fig. 11. Scatter plots of *BA* and V_{tot} predictions for expanded dataset with terrain and scan information using RF regression. Refer to Table V for prediction scores.

Irrespective of the variations due to differences in scan angle, the slope directly affects the point cloud and the resulting lidar metrics [see Fig. 3(c)]. The MLP could learn this complex and nonlinear relationship in this instance since, in the *expanded* datasets, we retained both heterogeneity (metrics sensitive to acquisition properties) and homogeneity (metrics not sensitive to acquisition properties) in lidar metrics.

As the addition of scan geometry demonstrated improvements with the MLP, a qualitative assessment of the scatterplots in Fig. 8 reveals that a model built with the $exp_{terrain+scan}$ dataset could deal with the issue of saturation commonly observed with large values. The scatterplots are comparable to those obtained by [30] after implementing a fusion of lidar and optical information, which creates interesting possibilities for future studies. Furthermore, while the MLP outperformed the RF models, some of the relative improvements across the datasets, though marginal, were apparent even in the RF models.

C. Interpretating SHAP Values

Regarding the SHAP values (see Fig. 9), the improvement in the ABA predictions corresponds to introducing new variables across the datasets. ANNs benefitted from including terrain variables (slope, aspect and elevation), yielding improved results [30]. The SHAP values for the elevation metric, which were among the highest, confirm that the terrain-related properties are crucial, especially in highly varying terrain. This is supported by the fact that slope and aspect also figured in the top 30 variables. Similarly, the distance of the scanner (considered an acquisition-related metric in this study) was among the top three metrics. Based on its importance, it is apparent that it was mainly responsible for resulting in improvements between $exp_{terrain}$ and $exp_{terrain+scan}$ datasets. The mean scan angle was ranked lower at around 30 (out of 61), whereas the azimuth of the acquisition was of lower importance. It is, however, crucial to note the improvements in ABA predictions on account of the data expansion. The rRMSE values for the $exp_{terrain}$ datasets were lower by 9% for *BA* and 20% for V_{tot} compared to the $std_{terrain}$ dataset. Metrics, such as the gap fraction, provided unique and diverse perspectives of the same point cloud to the NN. Although in this study we did not implement range normalization of the intensity, variables such as i_{tot} and i_{mean} were ranked high based on their SHAP values (see Fig. 9). In ABA models, virtually infinite metrics are being (or, have been) developed to capture vegetation properties comprehensively. SHAP values could be useful to inform this process.

On comparing with the ranking of the variables used in RF models ($exp_{terrain+scan}$) (see Fig. 10), it is apparent that metrics, such as elevation, gap-fraction, and terrain properties figure among the top-ranked variables (the distance, i.e., dist, of the scanner was ranked lower). However, this did not translate into

better predictions, thereby demonstrating the capability of neural networks to utilize the additional geometry information provided via the expansion of ABA datasets or additional metrics or both.

D. Potential of Different Modeling Strategies

Modeling strategies certainly influence the results. A few studies have explored different deep-learning methods to predict forest attributes from lidar data. Martins-Neto et al. [40] used an MLP architecture with the principal components of a set of metrics similar to our study. They observed an rRMSE of 22.5% for the predictions of *BA* in heterogeneous tropical forests, but with gentle or no relief. In our study, despite the mountainous relief, which is known to add issues in ABA modeling, the best-performing model was the $exp_{terrain+scan}$ dataset with an rRMSE of 19.9%. However, Martins-Neto et al. [40] did not consider metrics such as the gap fraction with proven explanatory power for forest structure characterization. Additionally, the gap fraction and the rumple index (used in this study) are metrics sensitive to lidar scan angle [33], [41]. The data expansion strategy may have benefited from additional information from these two metrics, among other sensitive metrics. In contrast, RF models were unable to utilize the additional information as is apparent in the saturation of large values with $exp_{terrain+scan}$ datasets (see Fig. 11).

Liu et al. [27] observed lower rRMSE values of 14.5% in volume predictions in predominantly Eucalyptus, and Chinese-fir-dominated stands. In comparison, an rRMSE of 24% was observed in this study. Even if they used a more advanced modeling framework that combines a fully connected neural network and an optimized radial basis neural network, the result difference is also likely to be linked to relatively simpler forest stands under study. In addition, studies using other modelling methods, such as OLS or RF, reported rRMSE values of basal area (*BA*) and volume (V_{tot}) predictions were in the range of 23% to 29% and 22% to 34%, respectively [19]. In this study, we observed an rRMSE for *BA* and V_{tot} of 19.9% and 24% for a model built with $exp_{terrain+scan}$ datasets.

The choice of lidar metrics is also a part of the modelling strategy. We used the intensity information provided by the data provider without implementing a normalization step to enhance it, as demonstrated in different studies [40], [42], [43], [44] showed that intensity information is more effective than height-based metrics in discriminating tree species. The forests in PNR Bauges comprise forest plots of broadleaved, coniferous, and mixed types of forests with different tree species. Calibrated intensity information could further help in improving the accuracies of the models. In addition, the intensity information is also known to be affected by the scan angle [40], thus potentially providing multiple perspectives of the vegetation.

Dayal et al. [13] demonstrated the benefits of using voxel-based metrics. Voxel-based metrics contributed by reducing rRMSE values for OLS regression ABA models based on forest type (riparian, broadleaf, coniferous, and mixed). For OLS ABA models, using voxel-based metrics improved the predictions by reducing the rRMSE by 10%, 4%, and 14% for riparian, broadleaf, and mixed types, respectively. The rRMSE values for different forest types and with only four lidar metrics (including voxel-based metrics) ranged between 30% and 40%, comparable to those observed with MLP models built with std, $std_{terrain}$, and exp datasets in this study. Therefore, including normalized intensity information along with stand-type characteristics and voxel-based metrics could be beneficial for building more accurate models using deep-learning approaches and a possible future area of exploration.

E. Impact of Data Sample Characteristics

The distribution of *BA* and V_{tot} across different test sets reveals that the dataset was not balanced, as there are very few plot measurements with very high values. Many of the cross-validation splits used in the study suffered from a lack of balanced training and testing samples. As a result, for all models, the impact of a few plots with high amounts of wood resource, i.e., *BA* greater than 70 m^2/ha and V_{tot} greater than 800 m^3/ha, can be observed. Such stand types are scarce, and the predictions for those plots might be outside the range of the training set values for some models and explain the saturation effect. MLP models are known for their capacity to generalize [27], [28], which could also explain why they performed better than RF. We believe that field plot measurement representing diverse forest stands will further help build robust models. In this study, the sampling strategy used to collect field measurements followed a systematic sampling scheme to establish sites for periodic monitoring. A stratified sampling scheme would be more suitable for building models [45], [46]. Moreover, some differences may arise because small trees (DBH < 7.5 cm) were not measured as per the field measurements protocol. The contribution of such trees in estimating the signal could be significant.

In this study, we opted to describe our datasets based on the flight lines as "expanded datasets." The common practice combines all these observations in lidar ABA models to generate the 'standard' datasets. Essentially, we assumed each observation in the "expanded" dataset to be an independent physical observation (or lidar scanning) of a given field plot. As the vegetation in the field plots is rarely identical when viewed from multiple directions, it may be argued that the resulting differences from different scans are comparable to the data augmentation procedure that is commonly used to increase the number of samples when dealing with images [47], [48], [49], [50]. To avoid confusion with commonly practiced data augmentation strategies, we referred to our modified datasets as 'expanded' datasets. Nonetheless, we would like to emphasize the similarities.

V. Conclusion

This study demonstrated that considering point clouds from different flight lines as independent observations in nonparametric models can improve ABA predictions for forest attributes. By considering the point clouds as independent observations, we retained the heterogeneity in the lidar metrics due to variations in the acquisition geometry in the form of an expanded dataset with

a significantly higher number of observations than a standard ABA dataset. An MLP could harness the expanded information to predict forest attributes in a complex forest environment with higher accuracies than a RF model, which is commonly used in ABA approaches. The present MLP model also demonstrated the potential to result in predictions comparable to methods involving optical and lidar data fusion. Optical data may be incorporated to improve further the results observed in this study.

ACKNOWLEDGMENT

The authors would like to thank H. Evers, C. Feppon, and É. Mermin for field measurements. The authors would also like to thank the reviewers for their valuable criticisms and suggestions.

REFERENCES

[1] R. O. Dubayah and J. B. Drake, "Lidar remote sensing for forestry," *J. Forestry*, vol. 98, no. 6, pp. 44–46, Jun. 2000, doi: 10.1093/jof/98.6.44.

[2] R. Nelson, "How did we get here? An early history of forestry lidar 1," *Can. J. Remote Sens.*, vol. 39, pp. S6–S17, Jun. 2013, doi: 10.5589/m13-011.

[3] P. A. Fekety, M. J. Falkowski, A. T. Hudak, T. B. Jain, and J. S. Evans, "Transferability of Lidar-derived basal area and stem density models within a northern Idaho ecoregion," *Can. J. Remote Sens.*, vol. 44, no. 2, pp. 131–143, Mar. 2018, doi: 10.1080/07038992.2018.1461557.

[4] T. Gobakken and E. Næsset, "Assessing effects of laser point density, ground sampling intensity, and field sample plot size on biophysical stand properties derived from airborne laser scanner data," *Can. J. Forest Res.*, vol. 38, no. 5, pp. 1095–1109, May 2008, doi: 10.1139/X07-219.

[5] T. Gobakken and E. Næsset, "Weibull and percentile models for lidar-based estimation of basal area distribution," *Scand. J. Forest Res.*, vol. 20, no. 6, pp. 490–502, Dec. 2005, doi: 10.1080/02827580500373186.

[6] R. Leiterer, R. Furrer, M. E. Schaepman, and F. Morsdorf, "Forest canopy-structure characterization: A data-driven approach," *Forest Ecol. Manage.*, vol. 358, pp. 48–61, 2015, doi: 10.1016/j.foreco.2015.09.003.

[7] B. Mitchell, H. Fisk, J. Clark, and E. Rounds, "Lidar acquisition specifications for forestry applications," *Nat. Leader Strategy*, vol. 9, pp. 1–27, 2018.

[8] E. Næsset, "Estimating timber volume of forest stands using airborne laser scanner data," *Remote Sens. Environ.*, vol. 61, no. 2, pp. 246–253, Aug. 1997, doi: 10.1016/S0034-4257(97)00041-2.

[9] H. O. Ørka, O. M. Bollandsås, E. H. Hansen, E. Næsset, and T. Gobakken, "Effects of terrain slope and aspect on the error of ALS-based predictions of forest attributes," *Forestry*, vol. 91, no. 2, pp. 225–237, 2018, doi: 10.1093/forestry/cpx058.

[10] O. R. van Lier, J. E. Luther, J. C. White, R. A. Fournier, and J. - F. Côté, "Effect of scan angle on ALS metrics and area-based predictions of forest attributes for balsam fir dominated stands," *Forestry. Int. J. Forest Res.*, vol. 95. no. 1, pp. 49–72, Jun. 2021, doi: 10.1093/forestry/cpab029.

[11] J. Lopatin, K. Dolos, H. J. Hernández, M. Galleguillos, and F. E. Fassnacht, "Comparing generalized linear models and random forest to model vascular plant species richness using LiDAR data in a natural forest in central Chile," *Remote Sens. Environ.*, vol. 173, pp. 200–210, Feb. 2016, doi: 10.1016/j.rse.2015.11.029.

[12] M. Fedrigo, G. J. Newnham, N. C. Coops, D. S. Culvenor, D. K. Bolton, and C. R. Nitschke, "Predicting temperate forest stand types using only structural profiles from discrete return airborne lidar," *ISPRS J. Photogrammetry Remote Sens.*, vol. 136, pp. 106–119, Feb. 2018, doi: 10.1016/j.isprsjprs.2017.11.018.

[13] K. R. Dayal et al., "An investigation into lidar scan angle impacts on stand attribute predictions in different forest environments," *ISPRS J. Photogrammetry Remote Sens.*, vol. 193, no. 8, pp. 314–338, 2022, doi: 10.1016/j.isprsjprs.2022.08.013.

[14] K. Ma et al., "Performance and sensitivity of individual tree segmentation methods for UAV-LiDAR in multiple forest types," *Remote Sens.*, vol. 14, no. 2, 2022, Art. no. 298, doi: 10.3390/rs14020298.

[15] A. P. D. Corte et al., "Measuring individual tree diameter and height using gatoreye high-density UAV-lidar in an integrated crop-livestock-forest system," *Remote Sens.*, vol. 12, no. 5, 2020, Art. no. 392, doi: 10.3390/rs12050863.

[16] A. G. Kamoske, K. M. Dahlin, S. C. Stark, and S. P. Serbin, "Leaf area density from airborne lidar: Comparing sensors and resolutions in a temperate broadleaf forest ecosystem," *Forest Ecol. Manage.*, vol. 433, no. 8, pp. 364–375, Feb. 2019, doi: 10.1016/j.foreco.2018.11.017.

[17] E. Næsset, "Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data," *Remote Sens. Environ.*, vol. 80, no. 1, pp. 88–99, Apr. 2002, doi: 10.1016/S0034-4257(01)00290-5.

[18] M. Bouvier, S. Durrieu, R. A. Fournier, and J. P. Renaud, "Generalizing predictive models of forest inventory attributes using an area-based approach with airborne LiDAR data," *Remote Sens. Environ.*, vol. 156, pp. 322–334, Jan. 2015, doi: 10.1016/j.rse.2014.10.004.

[19] N. C. Coops et al., "Modelling lidar-derived estimates of forest attributes over space and time: A review of approaches and future trends," *Remote Sens. Environ.*, vol. 260, no. 11, 2021, Art. no. 112477, doi: 10.1016/j.rse.2021.112477.

[20] J. R. Roussel, M. Béland, J. Caspersen, and A. Achim, "A mathematical framework to describe the effect of beam incidence angle on metrics derived from airborne lidar: The case of forest canopies approaching turbid medium behaviour," *Remote Sens. Environ.*, vol. 209, pp. 824–834, May 2018, doi: 10.1016/j.rse.2017.12.006.

[21] P. Tompalski, J. C. White, N. C. Coops, and M. A. Wulder, "Demonstrating the transferability of forest inventory attribute models derived using airborne laser scanning data," *Remote Sens. Environ.*, vol. 227, pp. 110–124, Jun. 2019, doi: 10.1016/j.rse.2019.04.006.

[22] J. C. White et al., "A model development and application guide for generating an enhanced forest inventory using airborne laser scanning data and an area-based approach," 2017. [Online]. Available: https://publications.gc.ca/collections/collection_2018/rncan-nrcan/Fo148-1-18-eng.pdf

[23] D. N. Cosenza et al., "Effects of numbers of observations and predictors for various model types on the performance of forest inventory with airborne laser scanning," *Can. J. Forest Res.*, vol. 52, no. 3, pp. 385–395, Mar. 2022, doi: 10.1139/cjfr-2021-0192.

[24] F. E. Fassnacht et al., "Importance of sample size, data type and prediction method for remote sensing-based estimations of aboveground forest biomass," *Remote Sens. Environ.*, vol. 154, no. 1, pp. 102–114, 2014, doi: 10.1016/j.rse.2014.07.028.

[25] P. M. Atkinson and A. R. L. Tatnall, "Introduction neural networks in remote sensing," *Int. J. Remote Sens.*, vol. 18, no. 4, pp. 699–709, 1997, doi: 10.1080/014311697218700.

[26] S. Gopal and C. Woodcock, "Remote sensing of forest change using artificial neural networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 34, no. 2, pp. 398–404, Mar. 1996, doi: 10.1109/36.485117.

[27] H. Liu et al., "Deep learning in forest structural parameter estimation using airborne LiDAR data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 1603–1618, 2021.

[28] R. Özçelik, M. J. Diamantopoulou, F. Crecente-Campo, and U. Eler, "Estimating crimean juniper tree height using nonlinear regression and artificial neural network models," *Forest Ecol. Manage.*, vol. 306, pp. 52–60, 2013, doi: 10.1016/j.foreco.2013.06.009.

[29] C. Deleuze, T. Senga Kiessé, J. - P. Renaud, F. Morneau, M. Rivoire, and P. Santenoise, "Rapport final EMERGE sur les modèles de volumes," Programme BIOE, Tech, Rep. ANR-08- BIOE-003, 2008.

[30] K. Lahssini, F. Teste, K. R. Dayal, S. Durrieu, D. Ienco, and J. M. Monnet, "Combining LiDAR metrics and sentinel-2 imagery to estimate basal area and wood volume in complex forest environment via neural networks," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 4337–4348, 2022.

[31] C. Hopkinson, L. Chasmer, C. Young-Pow, and P. Treitz, "Assessing forest metrics with a ground-based scanning lidar," *Can. J. Forest Res.*, vol. 34, no. 3, pp. 573–583, Mar. 2004, doi: 10.1139/x03-225.

[32] J. S. Jenness, "Calculating landscape surface area from digital elevation models," *Wildlife Soc. Bull.*, vol. 32, no. 3, pp. 829–839, 2004, doi: 10.2193/0091-7648(2004)032[0829:clsafd]2.0.co;2.

[33] K. R. Dayal et al., "Scan angle impact on lidar-derived metrics used in ABA models for prediction of forest stand characteristics: A grid based analysis," *ISPRS - Int. Arch. Photogrammetry Remote Sens. Spatial Inf. Sci.*, vol. XLIII-B3-2, no. B3, pp. 975–982, Aug. 2020, doi: 10.5194/isprs-archives-XLIII-B3-2020-975-2020.

[34] J. R. Roussel et al., "lidR: An r package for analysis of airborne laser scanning (ALS) data," *Remote Sens. Environ.*, vol. 251, no. 2, 2020, Art. no. 112061, doi: 10.1016/j.rse.2020.112061.

[35] M. Abadi et al., "{TensorFlow}: Large-scale machine learning on heterogeneous systems," 2015, [Online]. Available: https://www.tensorflow.org/

[36] T. O'Malley et al., "KerasTuner," 2019. [Online]. Available: https://github.com/keras-team/keras-tuner
[37] A. Liaw and M. Wiener, "Classification and regression by randomForest," *Latest News*, vol. 2, no. 3, pp. 18–22, 2002, [Online]. Available: https://cran.r-project.org/doc/Rnews/
[38] S. M. Lundberg and S. I. Lee, "A unified approach to interpreting model predictions," *Adv. Neural Inf. Process. Syst.*, vol. 2017, no. 2, pp. 4766–4775, May 2017, doi: 10.48550/arXi.v.1705.07874.
[39] L. S. Shapley, *A Value for N-Person Games*. Santa Monica, CA, USA: RAND Corporation, 1952.
[40] R. P. Martins-Neto et al., "Identification of significative LiDAR metrics and comparison of machine learning approaches for estimating stand and diversity variables in heterogeneous brazilian atlantic forest," *Remote Sens.*, vol. 13, 2021, Art. no. 2444.
[41] A. Montaghi, "Effect of scanning angle on vegetation metrics derived from a nationwide airborne laser scanning acquisition," *Can. J. Remote Sens.*, vol. 39, pp. 37–41, 2013, doi: 10.5589/m13-052.
[42] D. Gatziolis, "Dynamic range-based intensity normalization for airborne, discrete return lidar data of forest canopies," *Photogrammetric Eng. Remote Sens.*, vol. 77, no. 3, pp. 251–259, 2011, doi: 10.14358/PERS.77.3.251.
[43] C. Hopkinson, "The influence of flying altitude, beam divergence, and pulse repetition frequency on laser pulse return intensity and canopy frequency distribution," *Can. J. Remote Sens.*, vol. 33, no. 4, pp. 312–324, 2007, doi: 10.5589/m07-029.
[44] Y. Shi, T. Wang, A. K. Skidmore, and M. Heurich, "Important LiDAR metrics for discriminating forest tree species in central Europe," *ISPRS J. Photogrammetry Remote Sens.*, vol. 137, pp. 163–174, Mar. 2018, doi: 10.1016/j.isprsjprs.2018.02.002.
[45] T. Gobakken, L. Korhonen, and E. Næsset, "Laser-assisted selection of field plots for an area-based forest inventory," *Silva Fennica*, vol. 47, no. 5, 2013, Art. no. 943, doi: 10.14214/sf.943.
[46] M. Maltamo, O. M. Bollandsås, E. Næsset, T. Gobakken, and P. Packalén, "Different plot selection strategies for field training data in ALS-assisted forest inventory," *Forestry*, vol. 84, no. 1, pp. 23–31, 2011, doi: 10.1093/forestry/cpq039.
[47] A. Mikolajczyk and M. Grochowski, "Data augmentation for improving deep learning in image classification problem," in *Proc. Int. Interdiscipl. PhD Workshop*, 2018, pp. 117–122, doi: 10.1109/IIPHDW.2018.8388338.
[48] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *J. Big Data*, vol. 6, no. 1, 2019, Art. no. 60, doi: 10.1186/s40537-019-0197-0.
[49] L. Taylor and G. Nitschke, "Improving deep learning with generic data augmentation," in *Proc. IEEE Symp. Ser. Comput. Intell.*, 2018, pp. 1542–1547, doi: 10.1109/SSCI.2018.8628742.
[50] J. Wang, L. Perez, and J. Wang, "The effectiveness of data augmentation in image classification using deep learning," *Convolutional Neural Netw. Vis. Recognit*, vol. 2016, pp. 5079–5082, 2017, doi: 10.48550/ARXIV.1712.04621.



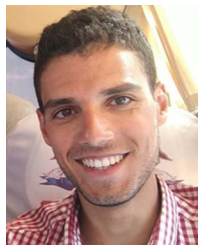
Karun R. Dayal received the master's degree in remote sensing and geographic information systems from the Indian Institute of Remote Sensing, Dehradun, India, in 2017, and the Ph.D. degree in lidar remote sensing of forests from AgroParisTech, Paris, France, in 2022.

He worked as a GIS analyst in Bengaluru, India. He is currently working as a Research Engineer with the UMR-TETIS, INRAE, Montpellier, France. His research interests include forest structure characterization using remote sensing data.



Sylvie Durrieu received the master's degree in forest and plant biology from the University of Nancy, Lorraine, France, in 1989, the engineer degree (French "baccalauréat" degree + 6 years), and the Ph.D. degree in remote sensing from the French Institute of Forestry, Agricultural and Environmental Engineering (ENGREF-AgroParisTech), Paris, France, in 1989 and 1994, respectively..

She received an accreditation to supervise research in Earth and Water Sciences from the University of Montpellier, Montpellier, France, (ED N° 584), in 2018. From 1994 to 1999 she worked at the French National Forest Inventory (IFN) and carried out R&D in the field of remote sensing applied to forest inventory and mapping, focusing on optical imagery technology. In 1999 she joined IRSTEA, the National Research Institute of Science and Technology for the Environment and Agriculture, which was merged in 2020 with INRA (the National Institute for Agricultural Research) to create INRAE. She is a Research Scientist in remote sensing and environmental sciences with INRAE (the France's National Research Institute for Agriculture, Food and Environment), Montpellier, France, and is part of the Territories, Environment, Remote Sensing & Spatial Information (TETIS) Research Unit in Montpellier. Her research interests include Lidar technology for forest ecosystem monitoring and she has developed approaches for characterizing and monitoring forest resource and biodiversity from 3-D Lidar data.



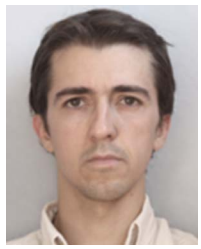
Kamel Lahssini received the master's degree in aerospace engineering from SUPAERO, Toulouse, France, in 2015, with a specialization in signal and image processing. He is currently working toward the Ph.D. degree in the use of multisensor satellite data to estimate tree height and biomass in tropical forest ecosystems with UMR-TETIS, INRAE, Montpellier, France.

He worked in Total as a Remote Sensing Engineer and in Safran as a UAV System Engineer. He joined INRAE in 2020 and earned a Specialized Master in Geomatics and Remote Sensing from AgroParisTech, Montpellier, France.



Dino Ienco received the M.Sc. and Ph.D. degrees in computer science both from the University of Torino, Torino, Italy, in 2006 and 2010, respectively.

He joined the TETIS Laboratory, IRSTEA, Montpellier, France, in 2011, as a Junior Researcher. His main research interests include machine learning, data science, graph databases, social media analysis, information retrieval and spatio-temporal data analysis with a particular emphasis on remote sensing data and Earth Observation data fusion. Dr. Ienco served in the program committee of many international conferences on data mining, machine learning, and database including IEEE ICDM, ECML PKDD, ACML, IJCAI as well as served as a Reviewer for many international journals in the general field of data science and remote sensing



Jean-Matthieu Monnet received the engineering degree from Ecole Polytechnique, Palaiseau, France, in 2004, and the Ph.D. degree in signal processing from Université de Grenoble, Grenoble, France, in 2011.

He has worked with AFOCEL as a Research Engineer in wood supply optimization (2005–2008). Since 2012, he has been with the LESSEM research unit of INRAE, Grenoble. His research interests include forest remote sensing with the specific objective of mapping forest ecosystem services in mountain environments.