iForest
Biogeosciences and Forestry

Research Article
doi: 10.3832/ifor2980-012
vol. 12, pp. 323-329

Prediction of stem diameter and biomass at individual tree crown level with advanced machine learning techniques

Salim Malek [1-3],
Franco Miglietta [1],
Terje Gobakken [2],
Erik Næsset [2],
Damiano Gianelle [3],
Michele Dalponte [3]

Knowledge about the aboveground biomass (AGB) and the diameters at breast height (DBH) distribution can lead to a precise estimation of carbon density and forest structure which can be very important for ecology studies especially for those concerning climate change. In this study, we propose to predict DBH and AGB of individual trees using tree height (H) and crown diameter (CD), and other metrics extracted from airborne laser scanning (ALS) data as input. In the proposed approach, regression methods, such us support vector machine for regression (SVR) and random forests (RF), were used to find a transformation or a transfer function that links the input parameters (H, CD, and other ALS metrics) with the output (DBH and AGB). The developed approach was tested on two datasets collected in southern Norway comprising 3970 and 9467 recorded trees, respectively. The results demonstrate that the developed approach provides better results compared to a state-of-the-art work (based on a linear model with the standard least-squares method) with RMSE equal to 81.4 kg and 92.0 kg, respectively (compared to 94.2 kg and 110.0 kg) for the prediction of AGB, and 5.16 cm and 4.93 cm, respectively (compared to 5.49 cm and 5.30 cm) for DBH.

Keywords: Aboveground Biomass, Diameter at Breast Height, Airborne Laser Scanning (ALS), Remote Sensing (RS), Support Vector Machine for Regression (SVR), Random Forests (RF)

Introduction

Forests are considered a major component of the global carbon cycle. A precise characterization of forest ecosystems in terms of carbon stock density and forest structure is an important key in international efforts to mitigate climate change. Carbon density can be estimated directly from the aboveground biomass (AGB) of trees, while the knowledge about the distribution of diameter at breast height (DBH) can be useful in understanding the forest structure (Slik et al. 2010). Having precise information about the distribution of those two parameters can help to understand the structure and the dynamics of forests. In the past, assessing those characteristics was primarily done with field-based inventory data and sometimes combined with conventional remote sensing (RS) data such as aerial photography and optical satellite images (Dalponte & Coomes 2016, Dalponte et al. 2018). ALS sensors, also referred to as airborne LiDAR (Light Detection And Ranging), are nowadays the most accurate remote sensing technology for monitoring forest carbon (Lefsky et al. 2002, Asner et al. 2012), as they can produce highly detailed 3D point clouds pinpointing locations on branches and the forest floor (Dalponte & Coomes 2016) and they measure surface elevation within a precision of a few centimeters, which offers the potential for studying forests at tree level.

In forest inventories in general, DBH and height (H) are measured and registered in the field in order to predict the AGB using allometric models. AGB is then converted to carbon density for each field-reference tree (Chave et al. 2014, Mensah et al. 2016, Zhang et al. 2016, Peng et al. 2017). With ALS, it became possible to measure the heights of trees in large forests in a short time, which makes it more practical compared to filed-measured methods. However, DBH cannot be measured directly with ALS sensors. Therefore, many studies have been carried out to try to predict DBH from airborne remote sensing data (ARS). The work of Gobakken & Naesset (2004) is considered the first where the DBH (and also basal area) distribution was predicted by using ALS data at the plot level. Their approach was based on a Weibull density function (Weibull 1951) and regression analysis was used to estimate the corresponding parameters. Recent studies have extracted ARS variables from each individual tree crown (ITC) detected in ARS data (Hauglin et al. 2013, 2014, Jucker et al. 2017, 2018, Mareya et al. 2018, Dalponte et al. 2018). Among this last group of studies, Jucker et al. (2017) proposed new allometric models to predict DBH and AGB based on H and crown diameter (CD) extracted from ARS data. Dalponte et al. (2018) in a recent study successfully linked field-reference DBH and AGB with H and CD extracted from ALS data using linear models.

The objective of the current study is to analyze the use of machine learning meth-

(1) Institute of Biometeorology, CNR, 50145 Firenze (Italy); (2) Faculty of Environmental Sciences and Natural Resource Management, Norwegian University of Life Sciences, P.O. Box 5003, NO-1432 Ås (Norway); (3) Dept. of Sustainable Agro-ecosystems and Bioresources, Research and Innovation Centre, Fondazione E. Mach, v. E. Mach 1, 38010 San Michele all'Adige, TN (Italy)

@ Michele Dalponte (michele.dalponte@fmach.it)

Received: Oct 22, 2018 - Accepted: Apr 06, 2019

Citation: Malek S, Miglietta F, Gobakken T, Næsset E, Gianelle D, Dalponte M (2019). Prediction of stem diameter and biomass at individual tree crown level with advanced machine learning techniques. iForest 12: 323-329. – doi: 10.3832/ifor2980-012 [online 2019-06-14]

Communicated by: Carlotta Ferrara

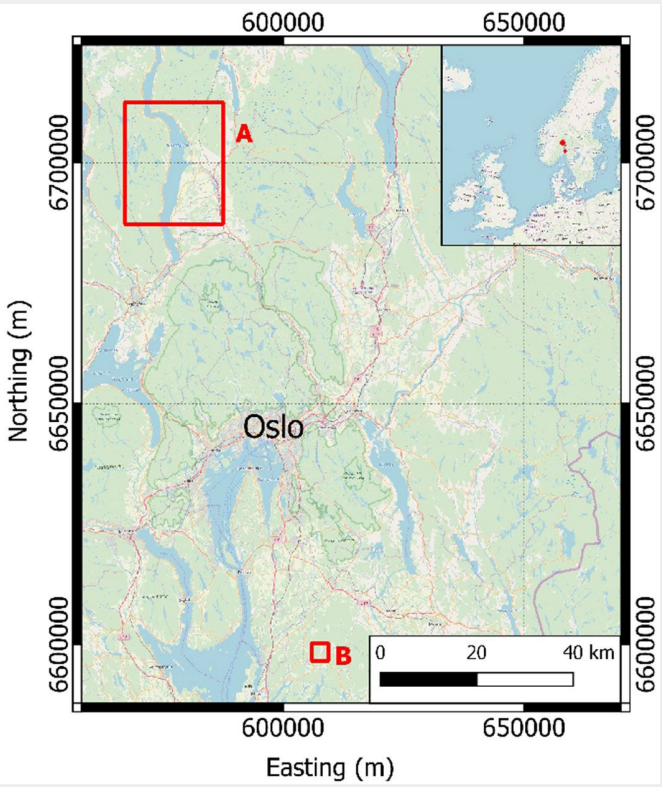


Fig. 1 - Location of the two study areas. (A) Hadeland; (B) Våler.

ods, such as Support Vector Machines for Regression (SVR) and Random Forest (RF) to predict DBH and AGB at the ITC level using metrics extracted from ALS data. In the first part of the experiment, only H and CD are used as input in order to compare with the work of Dalponte et al. (2018). After that, additional ALS metrics are used as input in order to see their impact on the quality of the prediction.

Materials and methods

Datasets description

In this study, two datasets located in boreal forests of southeastern Norway were used: Hadeland and Våler (Fig. 1). The main tree species in the two areas are Norway spruce (*Picea abies* [L.] H. Karst), Scots pine (*Pinus sylvestris* L.), and deciduous tree species, such as birch (*Betula* spp. L.) and aspen (*Populus tremula* L.). A summary of the field data of the two datasets is presented in Tab. 1 where only the trees chosen for the experiments are showed and grouped according their species.

Hadeland dataset

The field data acquired in the Hadeland district (Fig. 1) were collected on 13 circular sample plots of size 500 m^2 and 21 circular sample plots of size 1000 m^2 over a total area of about 1300 km^2. Within each sample plot, tree species, DBH, and tree coordinates were recorded for all trees with DBH > 3 cm. A total of 3970 trees were recorded. AGB of each tree was calculated using the allometric models of Marklund (1988).

ALS data were acquired on 21st and 22nd of August 2015 using a Leica ALS70 laser scanner operated at a pulse repetition frequency of 270 kHz. The flying altitude was of 1100 m above ground level. Up to four echoes per pulse were recorded and the resulting density of single and first echoes was 5 m^{-2}.

Våler dataset

The data were acquired in the Våler municipality in the southern part of Norway (Fig. 1). The field data were collected on 152 circular sample plots of size 400 m^2. Within each sample plot, tree species, DBH, and tree coordinates were recorded for all trees with DBH > 5 cm. A total of 9467 trees were recorded. AGB of each tree was calculated using the allometric models of Marklund (1988).

The ALS data were acquired on 9th September 2011 using a Leica ALS70 system operating with a pulse repetition frequency of 180 kHz. The flying altitude was of 1500 m above ground level. Up to four echoes per pulse were recorded and the resulting density of single and first echoes was 2.4 m^{-2}.

Methods

In Fig. 2 the architecture of the prediction system used is provided, and in the following paragraphs each step is detailed.

ITC delineation

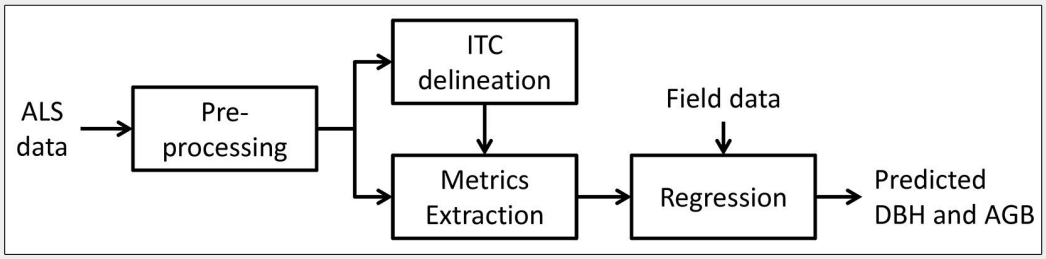
ITCs were delineated using an approach based on the ALS data and the delineation algorithm of the R package "itcSegment". The algorithm starts first by finding the local maxima within a rasterized Canopy Height Model (CHM) and designates them as tree tops, and then uses a decision tree method to grow individual crowns around the local maxima. The different steps for this adopted approach are as follows (Dalponte & Coomes 2016):

1. apply a 3 × 3 low-pass filter to the rasterized CHM in order to smooth the surface and reduce the number of local maxima;
2. localization of local maxima by using a circular moving window of variable size. The user provides a minimum and maximum size of the moving window; the window size is adapted according to the central pixel of the window: the size of the window is linearly related to the CHM height. A pixel of the CHM is considered as local maximum if its value is greater than all other values in the window, and if it is greater than some minimum height above ground. The window size is adapted according to the height of the central pixel of the window;
3. labeling each local maximum as an "initial region" around which a tree crown can grow;
4. extraction of the heights of the four neighboring pixels from the CHM and

Tab. 1 - Summary statistics of the field data for all datasets. For the tree height, DBH and AGB the data range and the mean (in brackets) are provided. For the species the number of trees and the percentage (in brackets) are provided.

Variable	Species	Hadeland	Våler
Tree height (m)	Spruce	5.8 - 25.4 (15.8)	3.5 - 33.3 (18.6)
	Pine	4.7 - 23.1 (16.0)	4.4 - 26.0 (15.1)
	Broadleaves	5.1 - 22.9 (13.8)	5.8 - 26.3 (15.3)
DBH (cm)	Spruce	5.1 - 44.1 (19.3)	4.3 - 50.3 (21.1)
	Pine	4.7 - 51.1 (25.6)	4.0 - 47.9 (20.2)
	Broadleaves	4.0 - 49.5 (14.8)	4.7 - 38.9 (16.2)
Tree AGB (kg)	Spruce	6.1 - 681.4 (155.1)	4.2 - 1232.9 (216.3)
	Pine	3.1 - 691.0 (214.3)	2.4 - 728.4 (146.3)
	Broadleaves	2.4 - 738.2 (103.2)	4.1 - 680.4 (125.5)
Species	Spruce	737 (59.7%)	1326 (50.2%)
	Pine	315 (25.5%)	956 (36.2%)
	Broadleaves	182 (14.8%)	361 (13.6%)

Fig. 2 - Architecture of the prediction system used.



adding them (the pixels) to the region if their vertical distance from the local maximum is less than a predefined percentage of the local maximum height, and less than a predefined maximum difference;

5. reiteration of the previous step for all the neighboring cells included in the region until no further pixels are added to the region;
6. extraction of single and first echoes from the ALS data from each identified region (having first removed low elevation echoes, *i.e.*, below 2 m);
7. application of a 2D convex hull to these echoes. The resulting polygons become the final ITCs. For each ITC CD and H are provided. The CD is computed as $2 \cdot \sqrt{(ITC_{area}/\pi)}$, while the height is computed as the 99th percentile of the elevation of the single and first return ALS echoes inside each ITC.

The delineated ITCs were automatically matched to the trees in the field data sets. If only one field-measured tree was included inside an ITC, then that tree was associated with that ITC. In the case that more than one field-measured tree was included in a segmented ITC, the field-measured tree with the height most similar to the ITC height was chosen.

ALS metrics extraction

From each delineated ITC, metrics were extracted in order to build the regression models. In particular, two sets of metrics were considered. The first set, called H+CD, contained two geometric metrics of the extracted ITCs, the height and crown diameter. The second set contains metrics extracted from the ALS points falling inside each ITC. This set of metrics comprised 50 statistics and they are summarized in Tab. 2. They were extracted from both elevation (Z) and intensity (I) of ALS points. We used the function "lasmetrics" of the library "lidR" of the software R to extract those ALS metrics.

Support vector machine for regression

Let us consider a matrix of training observations $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N]'$, where N is the number of observations and each vector $\mathbf{x}_i$ is represented in the d-dimensional measurement space. In our case, N corresponds to the number of ITCs used for the training and the measurement space is their corresponding H, CD, and the ALS metrics. Let us also consider the output vector $\mathbf{y} = [y_1, y_2, \ldots, y_N]'$ associated with $\mathbf{X}$ and corresponds to the measured DBH and the AGB. The aim of our proposed method is to estimate the relationship between the input vectors $\mathbf{x}_i$ and their target values y_i.

Support Vector machine for Regression (SVR – Vapnik 1998, Smola & Schölkopf 2004) performs linear regression in a feature space using an epsilon-insensitive loss (ε-SVM). This technique is based on the idea of deducing an estimate $g'(\mathbf{x}_i)$ of the true but unknown relationship $y_i = g(\mathbf{x}_i)$ ($i = 1, \ldots, N$) between the vector of observations $\mathbf{x}_i$ and the target value y_i such that: (i) $g'(\mathbf{x}_i)$ has, at most, ε deviation from the desired targets y_i; and (ii) it is as smooth as possible. This is performed by mapping the data from the original feature space of dimension d to a higher d'-dimensional transformed feature space (kernel space), *i.e.*, $\Phi(\mathbf{x}_i) \in \Re^{d'}$ ($d' > d$), to increase the flatness of the function and, by consequence, to approximate it in a linear way as follows (eqn. 1):

$$g'(\mathbf{x}_i) = \omega^* \Phi(\mathbf{x}_i) + b^* \quad (1)$$

Therefore, SVR is formulated as minimization of the following cost function (eqn. 2):

$$\psi(\omega, \xi) = \frac{1}{2}\|\omega\|^2 + c\sum_{i=1}^{N}(\xi_i + \xi_i^*) \quad (2)$$

subject to (eqn. 3):

$$\begin{cases} y_i - [\omega \cdot \Phi(\mathbf{x}_i) + b] \le \varepsilon + \xi_i \\ [\omega \cdot \Phi(\mathbf{x}_i) + b] - y_i \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0 \end{cases} \quad (3)$$

where ξ_i and ξ_i^* are the slack variables that measure the deviation of the training sample $\mathbf{x}_i$ outside the ε-insensitive zone. c is a parameter of regularization that allows tuning the tradeoff between the flatness of the function $g'(\mathbf{x})$ and the tolerance of deviations larger than ε.

The aforementioned optimization problem can be transformed through a Lagrange function into a dual optimization problem expressed in the original dimensional feature space in order to lead to the following dual prediction model (eqn. 4):

$$g'(\mathbf{x}) = \sum_{i \in U}(\alpha_i - \alpha_i^*) K(\mathbf{x}_i, \mathbf{x}) + b^* \quad (4)$$

where K is a kernel function, U is a subset of indices ($i = 1, \ldots, N$) corresponding to the nonzero Lagrange multipliers α_i's or α_i^*'s. The training observations that are associated to nonzero weights are called *SVs*. The kernel $K(\cdot,\cdot)$ should be chosen such that it satisfies the condition imposed by the Mercer's theorem, such as the Gaussian kernel functions (Vapnik 1998, Smola & Schölkopf 2004). In this study, the SVM implemented

Tab. 2 - Metrics extracted from the ALS points.

Metric	Description
Zmax	Maximum Z
Zmean	Mean Z
Zsd	Standard deviation of Z distribution
Zskew	Skewness of Z distribution
Zkurt	Kurtosis of Z distribution
Zentropy	Entropy of Z distribution
ZqP	Ph percentile of height distribution, with P from 5 to 95 at steps of 5
ZpcumP	Cumulative percentage of points in the Pth layer, with P from 5 to 95 at steps of 5
Itot	Sum of intensities for each return
Imax	Maximum intensity
Imean	Mean intensity
Isd	Standard deviation of intensity
Iskew	Skewness of intensity distribution
Ikurt	Kurtosis of intensity distribution
IpcumzqP	Percentage of intensity returned below the Pth percentile of Z, with P from 5 to 95
pRth	Percentage of Rth return, with R from 1 to 4

Tab. 3 - Accuracy statistics for DBH and AGB predictions using H+CD as input.

Dataset	Method	DBH		AGB	
		RMSE (cm)	*PIR* (%)	*RMSE* (kg)	*PIR* (%)
Hadeland	Dalponte et al. (2018)	5.49	-	94.19	-
	RF	5.42	1.28	84.35	10.45
	SVR	5.16	6.01	81.43	13.55
Våler	Dalponte et al. (2018)	5.30	-	109.99	-
	RF	5.19	2.08	95.46	13.21
	SVR	4.93	6.98	92.04	16.32

in the "kernlab" library of the software R was used.

Random Forest (RF)

The Random Forest (RF) method, which was proposed by Breiman (2001), is considered as one of the most effective machine learning method for predictive analytics. It consists of many decision trees trained on different parts (selected randomly) of the same training set at each node with the goal of reducing the variance. Each node is split using the best among a subset of its corresponding predictors. The strategy of randomness used in RF has been demonstrated to be robust against over-fitting problems (Breiman 2001, Liaw & Wiener 2002).

Given training set $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N]'$, with response $\mathbf{y} = [y_1, y_2, \ldots, y_N]'$ where N is the number of observations which is used in building a forest. Inside the forest a set of K trees $T_i = 1, \ldots, K$ is constructed. The output of each tree predicts the outputs for the actual value $\{y'_1 = T_1(\mathbf{x}), \ldots, y'_m = T_m(\mathbf{x})\}$, where $m = 1, \ldots, K$. The final result of the RF is the average of all tree predictions and is calculated as follows (Hannan et al. 2017, Liu et al. 2015 – eqn 5):

$$y_{RF}(\mathbf{x}) = \frac{1}{K}\sum_{k=1}^{K} \hat{y}_k(\mathbf{x}) \quad (5)$$

The evaluation of the Random Forest regression is done through the minimization of the mean square error (MSE) in order to select the optimum trees in the forest. In this study, the Random Forest classifier implemented in the "randomForest" library of the software R was used.

Parameter setting

In order to evaluate our methods, each dataset (*i.e.*, Hadeland, and Våler) was divided in two subsets (same as in Dalponte et al. 2018). The first set was used for calibration and the second for the validation phase. The split in the two sets was carried out in order to have similar characteristics in both sets in terms of spatial distribution, and DBH and AGB variation. For the Hadeland dataset, 607 observations were used for calibration and 627 observations were used for validation, while for the Våler dataset, 1398 and 1245, respectively.

Regarding the SVR, the Radial Basis Function (RBF) was used as kernel functions. To compute the best parameter values, we use a cross-validation technique with a number of folds equal to 3. During the cross validation, the parameter of regularization of the SVR c and the width of its kernel function γ were varied in the range $[1, 10^4]$ and $[10^{-3}, 5]$, respectively. The ε value of the insensitive tube was fixed to 10^{-3}.

For the RF method, we fix the number of trees to grow to 100, while the number of variables which will be randomly sampled as candidates at each split is fixed to 1 when using only CD and H as features and to 25 when using all the features (CD+ H+ALS data).

Performance evaluation

In order to evaluate the developed method of prediction and perform a direct comparison with results of the state-of-the-art methods, we adopted the Root Mean Square Error (RMSE) which measures the differences between values predicted by our model and the ground-reference values (eqn. 6):

$$RMSE = \sqrt{\frac{1}{N_t}\sum_{i=1}^{N_t} (y_{t_i} - y'_{t_i})^2} \quad (6)$$

where N_t is the total number of test observations, yt_i is the ground-reference target value and $y't_i$ is the predicted value of the developed regression method (ARS-predicted). Both yt_i and $y't_i$ correspond to the i^{th} test observation $\mathbf{x}t_i$.

We also adopted the percentage improvement ratio measure (*PIR*) in order to evaluate the level of improvement of our method compared to those of the state-of-the-art (SOA). It is formulated as follows (eqn. 7):

$$PIR = \frac{RMSE_{SOA} - RMSE_{our}}{RMSE_{SOA}} \cdot 100 \quad (7)$$

Results and discussion

In the first part of the analysis, only H and CD were used as input in order to compare the obtained results with those reported in Dalponte et al. (2018). Tab. 3 reports the results obtained with the set of metrics H+CD in terms of *RMSE* and *PIR*.

From Tab. 3, it can be seen that the proposed method using SVR and RF provide better results in terms of *RMSE* compared to the state-of-the-art method (Dalponte et al. 2018) for both DBH and AGB predictions. SVR shows substantial improvements while the RF results are located between those of SVR and the state-of-the-art method (Dalponte et al. 2018).

In greater detail, considering the AGB prediction, *RMSE* for the Hadeland dataset was 81.43 kg with SVR and 84.35 kg with RF, while it was 92.04 kg with SVR and 95.46 kg with RF in the Våler dataset. The improvement is significant compared to the results obtained by the state-of-the-art method (94.19 kg and 109.99 kg for Hadeland and Våler, respectively – Dalponte et al. 2018) with a maximum *PIR* equal to 13.55% for Våler dataset and 16.32% for Hadeland dataset. Regarding the DBH prediction, the *RMSE* by using SVR was 5.16 cm for Hadeland dataset and 4.93 cm for Våler dataset. Those results show also significant improvements compared to the state-of-the-art method (Dalponte et al. 2018) where *RMSE* was 5.49 cm for Hadeland and 5.30 cm for Våler. However, RF did not show considerable improvement and provided results close to those of the reference method (5.42 cm for Hadeland and 5.19 cm for Våler). In term of *PIR*, the maximum improvement was equal to 6.98% and 16.32% for Hadeland and Våler datasets, respectively.

To see visually the quality of the results, we show in Fig. 3 the field-reference DBH *vs.* ARS predicted DBH, and in Fig. 4 the field predicted AGB *vs.* ARS predicted AGB. The ARS predicted values are close to the regression line only until a certain value, while afterwards the bias is increasing. For example, if we consider the worst case corresponding to the prediction of DBH for the Hadeland dataset, the prediction values were close to the regression line when the DBH was inferior to 35 cm, while above this value, the predictions were giving values around 30 cm. This can be explained by the fact that among the 607 observations presented in the training, there were only 23 observations which have value higher than 35 cm which represents only 3.8% of the total training data. Moreover, the relationship between the tree DBH (and AGB) and its height is not linear as, after a certain age, trees stop to growth in height and they grow mainly in DBH. Thus, models that are based on ALS data have problems in modelling the DBH and AGB of old trees. Additionally, the traditional ground-based inventories require a big effort in terms of time and they may not be the best choice to provide a balanced dataset regarding the tree height, stem diameter and other features. It can be more efficient to select remotely the best samples (trees) to be later annotated in the field by experts. In terms of *RMSE*, our method showed good results since most of the predicted values fall close to the regression line except a few observations, but those few ones occupy a large range (between 35 and 50 cm which represents 30% of the total occupied range).

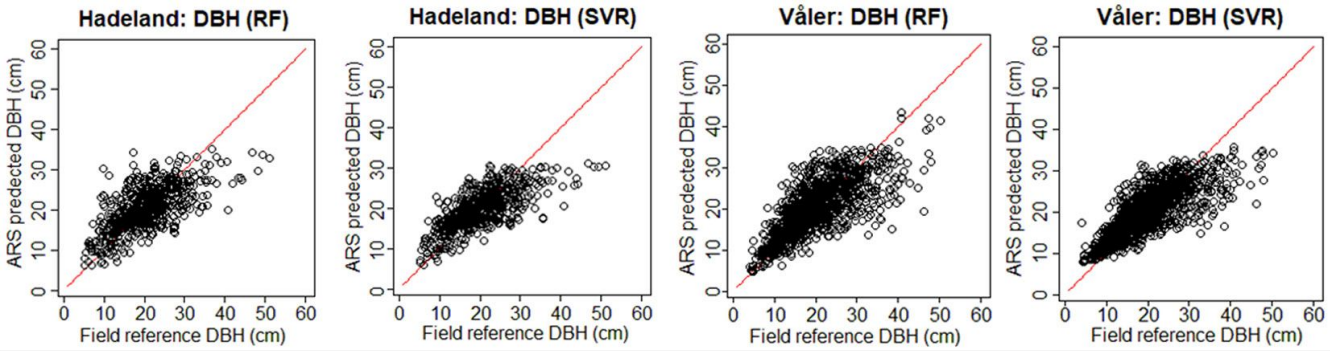


Fig. 3 - Field-reference DBH *vs.* ARS-predicted DBH for the two datasets using H+CD as input. In red the 1:1 line.

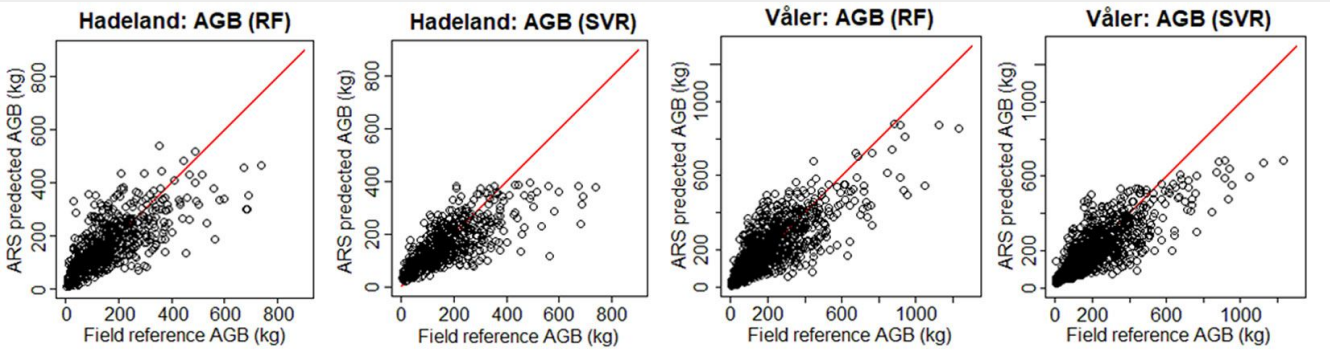


Fig. 4 - Field-reference AGB *vs.* ARS-predicted AGB for the two datasets using H+CD as input. In red the 1:1 line.

In order to improve the results, a set of ALS metrics were used together with the previous metrics (H and CD). The obtained results (Tab. 4) show that the ALS metrics help to improve the results, especially when using the RF method. For example, regarding the prediction of AGB, the *RMSE* reached a value less than 80 for the Hadeland dataset (78.5 with SVR and 76.0 by RF)

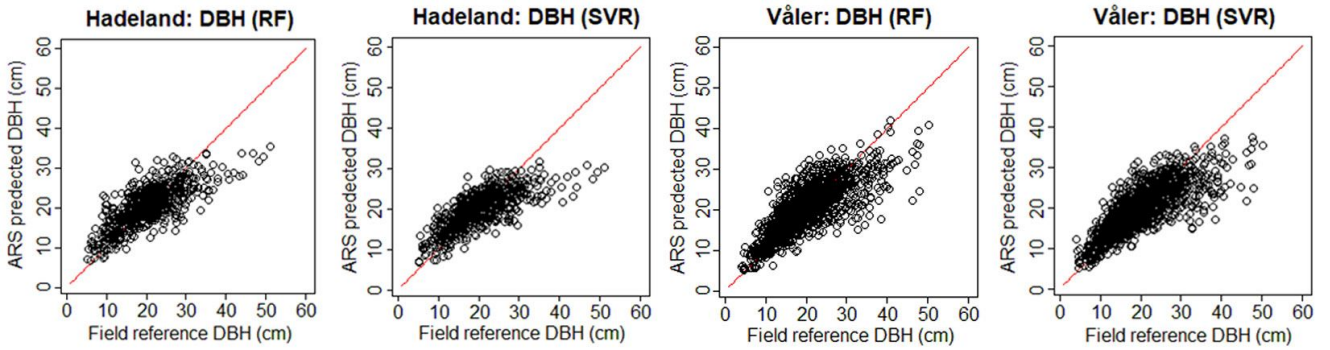


Fig. 5 - Field-reference DBH *vs.* ARS-predicted DBH using H+CD+ALS data. In red the 1:1 line.

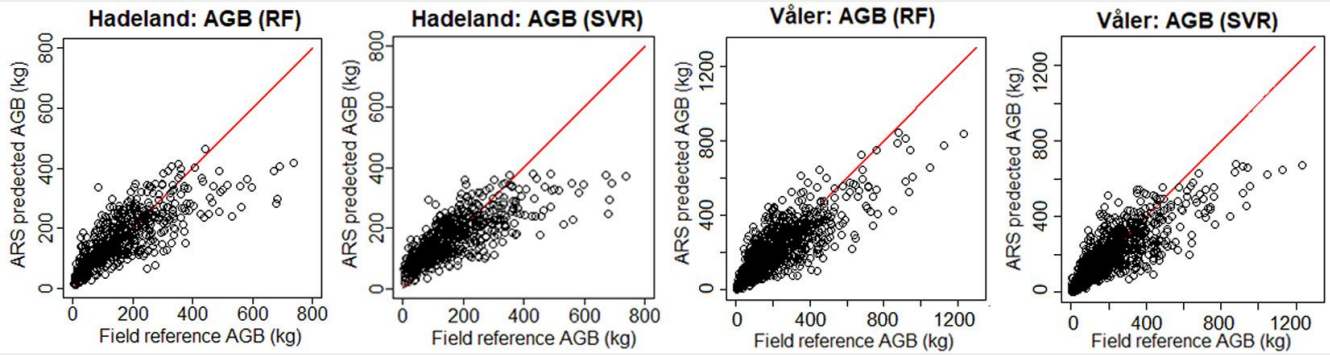


Fig. 6 - Field-reference AGB *vs.* ARS-predicted AGB using H+CD+ALS data. In red the 1:1 line.

Tab. 4 - Accuracy statistics for DBH and AGB predictions using H+CD+ALS metrics.

Dataset	Method	DBH		AGB	
		RMSE (cm)	*PIR* (%)	*RMSE* (kg)	*PIR* (%)
Hadeland	RF	4.79	12.75	75.97	19.34
	SVR	4.93	10.20	78.54	16.62
Våler	RF	4.88	7.92	88.93	19.15
	SVR	4.87	8.11	91.15	17.13

which represents an improvement of 16.62% by using SVR and 19.34% by using RF. For the Våler dataset RMSE was lower than 90 (88.9 with RF) with a *PIR* equals to 19.15. The same remarks can be given for the DBH prediction part where the improvement was significant and the *PIR* was 12.75% and 8.11% for Hadeland and Våler datasets, respectively.

In Fig. 5 and Fig. 6, we present the field-reference DBH *vs.* ARS-predicted DBH and the field-predicted AGB *vs.* ARS-predicted AGB, respectively, by using ALS metrics along with H and CD. From the different graphs, a slight improvement can be noticed compared to the previous ones (using only H and CD metrics in Fig. 3 and Fig. 4). However, the problem of inaccurate prediction of high values of AGB and DBH appeared also in this part of the experiments. We think that such problem can be investigated in a separate work in order to explore better the possible challenges that it may present.

Conclusion

In this work, we proposed an approach to predict DBH and AGB of trees from remote sensing data by using SVR and RF regression methods. The developed approach was tested on two datasets. On the first part of the experiments, the metrics H and CD were used in order to predict DBH and AGB. The obtained results were promising and the improvements were noticeable, especially in terms of *RMSE*. In order to improve the results, we proposed to introduce ALS metrics and use them together with H and CD. The obtained results were encouraging, especially with RF where the improvement was large. However, our method was not able to predict the isolated samples with highest values of AGB and DBH. This was probably due to the fact that the relationship among the height of a tree (and thus of the majority of the ALS metrics) and the DBH and AGB was saturating, and also to the fact that there was a limited number of training observations for the larger trees.

Finally, in order to improve the quality of the results and to get better predictions for old trees with large values of DBH and AGB, we think it can be more advantageous to use techniques that preprocess the data in order to yield a balance for their distribution over all the scale of the different metrics.

Acknowledgements

This work was supported by the HyperBio project (project 244599) financed by the BIONR program of the Research Council of Norway and TerraTec AS, Norway.

References

Asner GP, Mascaro J, Muller-Landau HC, Vieilledent G, Vaudry R, Rasamoelina M, Hall JS, Van Breugel M (2012). A universal airborne LiDAR approach for tropical forest carbon mapping. Oecologia 168 (4): 1147-1160. - doi: 10.1007/s00442-011-2165-z

Breiman L (2001). Random forests. Machine Learning 45 (1): 5-32. - doi: 10.1023/A:1010933404324

Chave J, Réjou-Méchain M, Búrquez A, Chidumayo E, Colgan MS, Delitti WBC, Duque A, Eid T, Fearnside PM, Goodman RC, Henry M, Martínez-Yrízar A, Mugasha WA, Muller-Landau HC, Mencuccini M, Nelson BW, Ngomanda A, Nogueira EM, Ortiz-Malavassi E, Pélissier R, Ploton P, Ryan CM, Saldarriaga JG, Vieilledent G (2014). Improved allometric models to estimate the aboveground biomass of tropical trees. Global Change Biology 20: 3177-3190. - doi: 10.1111/gcb.12629

Dalponte M, Coomes DA (2016). Tree-centric mapping of forest carbon density from airborne laser scanning and hyperspectral data. Methods in Ecology and Evolution 7 (10): 1236-1245. - doi: 10.1111/2041-210X.12575

Dalponte M, Frizzera L, Orka HO, Gobakken T, Naesset E, Gianelle D (2018). Predicting stem diameters and aboveground biomass of individual trees using remote sensing data. Ecological Indicators 85: 367-376. - doi: 10.1016/j.ecolind.2017.10.066

Gobakken T, Naesset E (2004). Estimation of diameter and basal area distributions in coniferous forest by means of airborne laser scanner data. Scandinavian Journal of Forest Research 19: 529-542. - doi: 10.1080/02827580410019454

Hannan MA, Ali JA, Mohamed A, Uddin MN (2017). A random forest regression based space vector PWM inverter controller for the induction motor drive. IEEE Transactions on Industrial Electronics 64 (4): 2689-2699. - doi: 10.1109/TIE.2016.2631121

Hauglin M, Dibdiakova J, Gobakken T, Naesset E (2013). Estimating single-tree branch biomass of Norway spruce by airborne laser scanning. ISPRS Journal of Photogrammetry and Remote Sensing 79: 147-156. - doi: 10.1016/j.isprsjprs.2013.02.013

Hauglin M, Gobakken T, Astrup R, Ene L, Naesset E (2014). Estimating single-tree crown biomass of Norway spruce by airborne laser scanning: a comparison of methods with and without the use of terrestrial laser scanning to obtain the ground reference data. Forests 5: 384-403. - doi: 10.3390/f5030384

Jucker T, Caspersen J, Chave J, Antin C, Barbier N, Bongers F, Dalponte M, Van Ewijk KY, Forrester DI, Haeni M, Higgins SI, Holdaway RJ, Iida Y, Lorimer C, Marshall PL, Momo S, Moncrieff GR, Ploton P, Poorter L, Rahman KA, Schlund M, Sonké B, Sterck FJ, Trugman AT, Usoltsev VA, Vanderwel MC, Waldner P, Wedeux BMM, Wirth C, Wöll H, Woods M, Xiang W, Zimmermann NE, Coomes DA (2017). Allometric equations for integrating remote sensing imagery into forest monitoring programmes. Global Change Biology 23: 177-190. - doi: 10.1111/gcb.13388

Jucker T, Asner GP, Dalponte M, Brodrick PG, Philipson CD, Vaughn NR, Teh YA, Brelsford C, Burslem DFRP, Deere NJ, Ewers RM, Kvasnica J, Lewis SL, Malhi Y, Milne S, Nilus R, Pfeifer M, Phillips OL, Qie L, Renneboog N, Reynolds G, Riutta T, Struebig MJ, Svátek M, Turner EC, Coomes DA (2018). Estimating aboveground carbon density and its uncertainty in Borneo's structurally complex tropical forests using airborne laser scanning. Biogeosciences 15: 3811-3830. - doi: 10.5194/bg-15-3811-2018

Lefsky MA, Cohen WB, Harding DJ, Parker GG, Acker SA, Gower ST (2002). Lidar remote sensing of above-ground biomass in three biomes. Global Ecology and Biogeography 11: 393-399. - doi: 10.1046/j.1466-822x.2002.00303.x

Liaw A, Wiener M (2002). Classification and regression by random forest. R News 2 (3): 18-22. [online] URL: http://www.researchgate.net/publication/228451484

Liu M, Liu X, Liu D, Ding C, Jiang J (2015). Multivariable integration method for estimating sea surface salinity in coastal waters from *in situ* data and remotely sensed data using random forest algorithm. Computers and Geosciences 75: 44-56. - doi: 10.1016/j.cageo.2014.10.016

Mareya HT, Tagwireyi P, Ndaimani H, Gara TW, Gwenzi D (2018). Estimating tree crown area and aboveground biomass in Miombo woodlands from high-resolution RGB-only imagery. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 11 (3): 868-875. - doi: 10.1109/JSTARS.2018.2799386

Marklund LG (1988). Biomass functions for pine, spruce and birch in Sweden. Report 45, Department of Forest Survey, Swedish University for Agricultural Sciences, Uppsala, Sweden, pp. 73.

Mensah S, Veldtman R, du Toit B, Kakaï RG, Seifert T (2016). Aboveground biomass and carbon in a South African mistbelt forest and the relationships with tree species diversity and forest structures. Forests 7: 1-17. - doi: 10.3390/f7040079

Peng S, He N, Yu G, Wang Q (2017). Aboveground biomass estimation at different scales for subtropical forests in China. Botanical Studies 58: 45. - doi: 10.1186/s40529-017-0199-1

Slik JWF, Aiba SI, Brearley FQ, Cannon CH, Forshed O, Kitayama K, Nagamasu H, Nilus R, Payne J, Paoli G, Poulsen AD, Raes N, Sheil D, Sidiyasa K, Suzuki E, Van Valkenburg JLCH (2010). Environmental correlates of tree biomass, basal area, wood specific gravity and stem density gradients in Borneo's tropical forests. Global Ecology and Biogeography 19:

50-60. - doi: 10.1111/j.1466-8238.2009.00489.x
Smola AJ, Schölkopf B (2004). A tutorial on support vector regression. Statistics and Computing 14 (3): 199-222. - doi: 10.1023/B:STCO.0000035301.49549.88
Vapnik VN (1998). Statistical learning theory, vol. 1. Wiley, New York, USA, pp. 1-768.
Weibull W (1951). A statistical distribution function of wide applicability. Journal of Applied Mechanics 18: 293-297. [online] URL: http://web.cecs.pdx.edu/~cgshirl/Documents/Weibull-ASME-Paper-1951.pdf
Zhang Y, Chen HYH, Taylor AR (2016). Aboveground biomass of understorey vegetation has a negligible or negative association with overstorey tree species diversity in natural forests. Global Ecology and Biogeography 25: 141-150. - doi: 10.1111/geb.12392