

Silvilar 2010
September 14th - 17th
Freiburg, Germany

Estimation of forestry parameters in mountainous coppice stands using airborne laser scanning

Jean-Matthieu Monnet*[†], Eric Mermin[†], Jocelyn Chanussot[‡] & Frédéric Berger[†]

[†]UR EMGR, Cemagref, France

[‡]Gipsa-Lab, Grenoble Institute of Technology, France

An area-based method is implemented to predict forest stand parameters from airborne laser scanning data. Multiple regression models are calibrated with 31 field plots inventoried in a hillside dominated by coppice stands, located in the French Alps. Cross-validated prediction accuracies are respectively 13.2, 14.3, 19.4 and 25.5% for dominant height, mean diameter, basal area and stem density. Reducing calibration plot radius greatly influences prediction results. Although median values of forest parameters remain unchanged, field observations variability is higher for smaller plots. Depending on the forest parameter, prediction accuracy is significantly correlated with some distribution statistics (minimum and maximum values) of field observations computed with various radii. Indeed, greater variance or outliers may degrade the degree of fit of regression models.

1. Introduction

In alpine environments with high topographical constraints, evolutions in forestry practises and labour costs resulted in a progressive neglecting of mountainous stands. Near the footslopes, coppice stands which used to provided local inhabitants with fuelwood are quite frequent in the French Alps. They were among the first to be left over as a combination of the development of other energy sources and of the impossibility to grow high value timber products due to poor site quality. However, fossil fuels rarefaction and global warming alarms prompted public authorities to set ambitious objectives of increased woody biomass harvesting (Ginisty *et al.* 2007). Unfortunately, information about forest stands characteristics is now missing or outdated. Due to accessibility constraints, conventional field inventory methods can not provide the information required to forecast harvesting operations. Prospecting difficulties are all the more critical since forest stands display a high spatial heterogeneity linked with complex landform patterns encountered in mountainous areas.

High hopes have been set on airborne laser scanning (ALS) as this remote sensing technique was shown to be successful in stand parameters estimation and tree detection for coniferous forests (see review by Hyypä *et al.* 2008). Following works also demonstrated its accuracy in others contexts such as tempered deciduous forests (Popescu *et al.* 2002, Patenaude *et al.* 2004) and alpine environments (Heurich and Thoma 2008, Hollaus *et al.* 2009). To our knowledge, the case of mountainous coppice stands has not been investigated so far.

The aim of this paper is to evaluate the efficiency of ALS for forest parameters estimation in mountainous coppice stands. The area-based method proposed by Næsset (2002) is implemented. The effects of forest spatial heterogeneity are investigated by examining the influence of calibration plot size on field observations and on regression models accuracy.

*corresponding author. Email: jean-matthieu.monnet@cemagref.fr

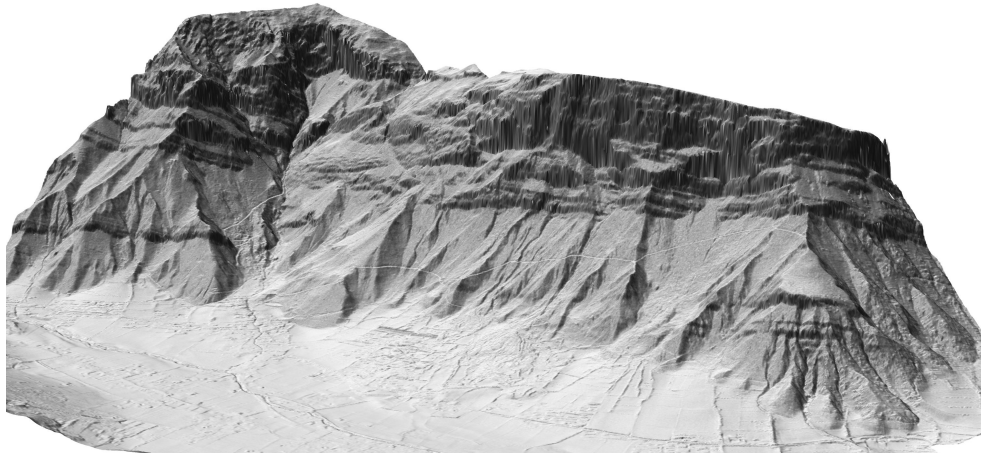


Figure 1. Shaded digital terrain model of the study area.

2. Material

2.1 Study area

The study area is a 4 km² hillside situated in the French Alps (town of Saint Paul de Varcès, 45°04'17"N, 05°38'25"E)(figure 1). The forest is mainly constituted of coppice stands and deciduous stands on poor quality sites, dominated by Italian maples (*Acer opalus*) and downy oaks (*Quercus pubescens*). Downslope, old chestnut (*Castanea sativa*) coppice stands are frequent. Common whitebeam is present in all the area, especially at the foot of cliffs. In thalwegs or in upper parts with better site quality, ash (*Fraxinus excelsior*) and beech (*Fagus sylvatica*) are common. Some areas have a dense understory of holly (*Ilex Aquifolium*), common hazel (*Corylus Avellana*) or box (*Buxus sempervirens*). Altitude ranges from 330 to 1270 m above sea level. High limestone cliffs overhang the area and rockfall events are frequent. No major silvicultural or harvesting operations have been performed in the area for more than fifty years.

2.2 Field data

From September to November 2009, $N = 31$ circular field plots were inventoried. Plots were distributed every 400 m along the 550, 750, 950 and 1150 m height contours, resulting in an irregular sampling scheme where horizontal distances between neighbouring plots ranged from 180 to 412 m with a mean value of 302 m. Plot centres were georeferenced using a Trimble GPS Pro XRS receiver. After differential correction with the Pathfinder® software, position precision (95% confidence radius) ranged from 0.6 to 1.5 m. All trees with diameter at breast height larger than 5 cm and located within 10 m horizontal distance from the plot centre had their diameter measured with a tape. Their positions to the plot centre were recorded using a Suunto KB-14 compass and Suunto PM-5 clinometer mounted on a tripod, and a Vertex III hypsometer. Maples (mainly *Acer opalus*), downy oak (*Quercus pubescens*) and common whitebeam (*Sorbus aria*) represented nearly 60 % of the stems. Ten tree heights were also measured on each plot with the hypsometer. Height sampling probability was proportional to stem basal area to ensure that dominant trees would be represented.

Table 1. Laser scanner acquisition parameters

Item	Value
Wavelength	1550 nm
Pulse repetition rate	200 kHz
Scan frequency	111.1 Hz
Scan angle	$\pm 30^\circ$
Flight height	550 m
Laser footprint	0.29 m
Theoretical point spacing	0.47 m

2.3 Laser data

Laser data was acquired on August 27th, 2009 over 8.6 km² with a fullwave RIEGL LMS-Q560 scanner. Acquisition parameters are summarised in table 1. Echoes were extracted from the binary acquisition files and georeferenced with the RIEGL software suite. The contractor also classified the resulting point cloud into ground and non-ground echoes using the TerraScan software, which implements an algorithm based on iterative surface reconstruction by triangulated irregular network (Axelsson 2000). Final echo density was 10 m⁻².

3. Methods

The area-based, two steps method proposed by Næsset (2002) is implemented to predict forest parameters from airborne laser scanning data. To evaluate the effect of calibration plot size on prediction accuracy, plot radius is reduced by excluding trees situated further than various distance thresholds before reiterating the method. Tested radii are $r \in \{5, 5.5, \dots, 10\}$.

3.1 Forest parameters

The following stand parameters are computed for each plot $j \in \{1, 2, \dots, 31\}$ and each radius r : basal area (G_r^j : surface occupied by the horizontal section of tree stems at 1.30 m height), stem density (N_r^j) and mean diameter at breast height (D_r^j). Dominant height (H_{10}^j : mean height of the 30 highest trees per hectare) is calculated for $r = 10$ m only, as the sampling scheme does not ensure that enough measured trees are included within each radius. Wilcoxon signed-rank tests are performed to compare forest parameters observations obtained with different plot sizes. Correlation between field observations statistics and plot radius is evaluated by computing Spearman ρ .

3.2 Extraction of laser metrics

For each plot j , laser points within r meters horizontal distance from the plot centre are extracted. Their relative heights are computed by subtracting the terrain height at their orthometric coordinates. Terrain surface is estimated by bilinear interpolation of points classified as ground points. Points with relative height lower than 2 m are excluded to avoid influence of dense shrubs understory. Three point groups are then constituted according to return positions: single echoes (only one echo for a given pulse), first echoes and last echoes. For each group two types of laser metrics are calculated. Height metrics correspond to the breakpoints of four height bins containing an equal number of points: minimum ($h_{g,0}$), first quartile ($h_{g,0.25}$), median ($h_{g,0.5}$), third quartile ($h_{g,0.75}$) and maximum ($h_{g,1}$) values, plus mean height ($h_{g,mean}$). Subscript $g \in \{s, f, l\}$ refers to the point groups: single, first or last echoes. Three density metrics are computed as the proportion of

echoes recorded below height thresholds corresponding respectively to 0.75, 0.5 and 0.25 of the maximum echo height recorded on the plot ($d_{g,0.75}$, $d_{g,0.5}$ and $d_{g,0.25}$ respectively). For each radius r , the predictors set $P_r = (P_r^i)_{i \in \{1, \dots, 27\}}$ consists of 31 observations of $3 \times 9 = 27$ laser metrics.

3.3 Multiple regression models

For each predictors set P_r and each corresponding dependent variables $y_r \in \{H_r, G_r, N_r, D_r\}$, a multiple regression model is fitted by ordinary least squares.

$$y_r = b_r + \sum_{i=1}^{27} a_r^i \times P_r^i \quad \text{with } ((a_r^i)_{i \in \{1, \dots, 27\}}, b_r) \text{ the model parameters} \quad (1)$$

Models including a maximum of three predictors are tested by exhaustive search among possible combinations. Models which do not fulfil the linear model assumptions or including a predictor with a partial p -value greater than 0.05 are discarded. The model with highest adjusted coefficient of determination ($adj-R^2$) is selected.

Prediction accuracy is evaluated in leave-one-out cross validation by computing the root mean square error ($RMSE$) and its coefficient of variation CV_{RMSE} .

$$RMSE = \sqrt{\frac{1}{N} \sum_{j=1}^N (y_j - \hat{y}_j)^2} \quad \text{with } \begin{cases} y_j & \text{the observed values} \\ \hat{y}_j & \text{the predicted values} \end{cases} \quad (2a)$$

$$CV_{RMSE} = \frac{RMSE}{\bar{y}} \quad \text{with } \bar{y} = \frac{1}{N} \sum_{j=1}^N y_j \quad (2b)$$

Differences between predicted values and field observations are evaluated by Wilcoxon signed-rank tests. Spearman ρ is computed to assess correlation between prediction accuracy and plot radius for each stand parameter.

4. Results

4.1 Field observations

Forest plots statistics for 5, 7.5 and 10 m radii are displayed in table 2. As site quality is rather poor in the area, average dominant height is only 17.8 m. Generally, stands with high values for basal area, dominant height and mean diameter are located on a few good quality sites in thalwegs with deep soil, such as the ash-dominated plot #14 with $H_{10} = 28.5$ m, $G_{10} = 59.7 \text{ m}^2 \cdot \text{ha}^{-1}$ and $D_{10} = 21.6$ cm. Small values are encountered on steep slopes at the bottom of cliffs with rockfall activity. For example, plot #21 is located on a scree and has $H_{10} = 13.7$ m, $G_{10} = 4.6 \text{ m}^2 \cdot \text{ha}^{-1}$ and $D_{10} = 8.3$ cm. In such areas stem density is highly variable.

Figure 2 plots field observations statistics of basal area, stem density and mean diameter as functions of plot radius. Although mean basal area seems to increase with plot radius, Wilcoxon signed-rank tests indicate that the medians of the observations differences ($y_{r1} - y_{r2}$ with $r1 \neq r2$ and $y_r \in \{H_r, G_r, N_r, D_r\}$) are not significantly different from zero at the $p < 0.05$ level.

Spearman correlation tests (table 3) show that minimum values of basal area and mean diameter are negatively correlated with plot radius, whereas stem den-

Table 2. Dominant height (H), basal area (G), stem density (N) and mean diameter (D) field observations statistics for calibration plot radius $r \in \{5, 7.5, 10\}$ ($N = 31$ plots).

Radius	H (m)		G (m ² .ha ⁻¹)				N (ha ⁻¹)			D (cm)	
	10	5	7.5	10	5	7.5	10	5	7.5	10	
Mean	17.8	32.3	34.6	34.8	1668	1714	1735	14.5	14.4	14.5	
Min	8.1	4.9	4.7	4.6	382	736	764	8.9	8.3	8.3	
Max	28.5	66.8	90.1	59.7	3310	3112	2833	30.0	23.3	22.7	
Sd	5.3	16.7	15.9	11.4	724	560	577	4.6	3.9	3.6	

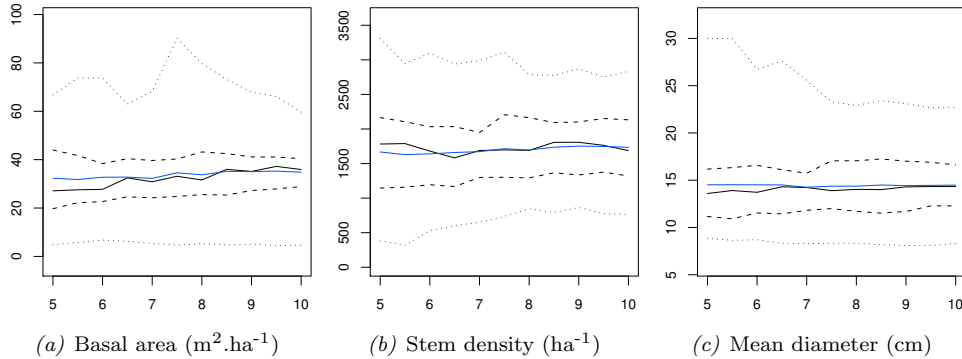


Figure 2. Influence of calibration plot radius (x-axis, meters) on field parameters observations. Dotted lines are the minimum and maximum values, dashed lines the first and third quartiles, black solid line is the median and blue solid line the mean value.

Table 3. Spearman correlation coefficient ρ between field observations statistics and plot radius for basal area (G), stem density (N) and mean diameter (D).

	Min	1 st quart.	Median	3 rd quart.	Max	Mean	Sd
G	-0.68*	0.98***	0.91***	-0.02	-0.31	0.85**	-0.89***
N	0.84**	0.90***	0.12	0.11	-0.76**	0.87***	-0.54
D	-0.87***	0.77**	0.75*	0.57	-0.93***	-0.52	-0.98***

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$ (two-sided test)

sity minimum is positively correlated. First quartile (all tested parameters) and mean values (basal area and stem density) are also positively linked to plot radius. Negative ρ values are found for maximum stem density and diameter, and standard deviation of the basal area and mean diameter. The median is correlated to plot radius for basal area and mean diameter, whereas the third quartile yields no significant correlation.

4.2 Multiple regression models

Selected multiple regression models for $r \in \{5, 7.5, 10\}$ are detailed in table 4. With 10 m radius calibration plots, prediction accuracies (cross validated coefficient of variation of the RMSE) are respectively 13.2, 14.3, 19.4 and 25.5 % for dominant height, mean diameter, basal area and stem density. For $r = 7.5$ m results slightly improve for mean diameter and stem density (respectively 13.6 and 23.0 %), but are worse for basal area (36.4 %). The lowest accuracies are achieved with 5 m radius: 32.9 % for stem density and 36.7 % for basal area. Wilcoxon signed-rank tests indicate that the median of the differences between predicted and observed values is not significantly different from zero at the $p < 0.05$ level for any forest parameter. Mean diameter models do not fulfil linear model assumptions for $r < 6$. It is noteworthy that laser metrics included in the models depend both on the forest parameter and on calibration plot radius.

Figure 3 displays prediction accuracy of multiple regression models as a function of calibration plot radius for basal area, stem density and mean diameter. Spearman

Table 4. Selected multiple regression models for radius $r \in \{5, 7.5, 10\}$.

Variable	Radius	Laser metrics in the model	$adj-R^2$ (%)	CV_{RMSE} (%)
Basal area (G)	5	$h_{s,0} + h_{s,1} + h_{f,mean}$	55.5	36.7
	7.5	$h_{f,0.75}$	46.0	36.4
	10	$h_{f,0.25} + h_{f,0.5} + d_{s,0.5}$	70.8	19.4
Stem density (N)	5	$h_{l,mean} + d_{f,0.75} + d_{l,0.25}$	51.0	32.9
	7.5	$h_{f,0.75} + h_{l,1} + d_{l,0.25}$	59.1	23.0
	10	$h_{l,0.5} + d_{f,0.25}$	58.0	25.5
Mean diameter (D)	5	linear model assumptions not satisfied		
	7.5	$h_{f,0} + h_{l,0.25} + d_{l,0.75}$	76.6	13.6
	10	$h_{s,0.75} + h_{f,0.75} + d_{l,0.25}$	71.4	14.3
Dominant height (H)	5, 7.5	no relevant due to sampling scheme		
	10	$h_{s,0.5} + h_{s,0.75} + h_{f,0}$	84.1	13.2

correlation coefficient ρ between the forest parameter and plot radius is significantly different from zero for basal area only ($\rho = -0.87$, $p < 0.001$, two-sided test). Stem density displays a similar decreasing trend but the coefficient of variation of the RMSE increases again when $r > 9$. For mean diameter the coefficient of variation of the RMSE decreases when plot radius increases from 6 to 7.5 m, and then remains stable around 14%.

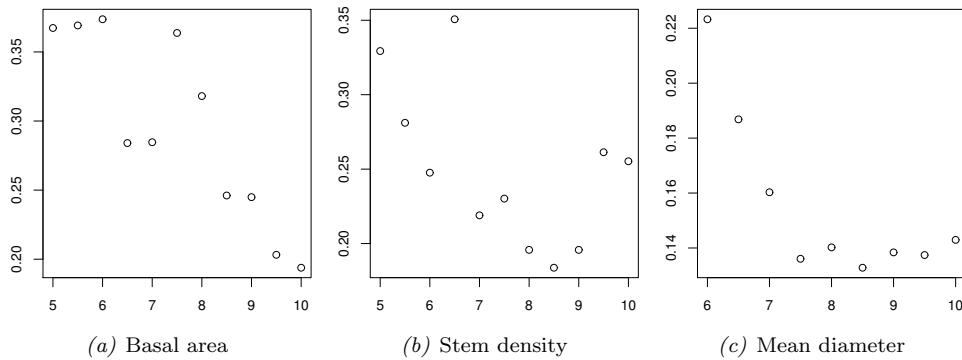


Figure 3. Influence of calibration plot radius (x-axis, meters) on the cross-validated coefficient of variation of the RMSE (y-axis) of multiple regression models.

5. Discussion

The ALS based method produces accurate estimates of forest stand parameters, showing that it is suitable for deciduous forests such as coppice stands. Results obtained with 10 m radius calibration plots are consistent with those obtained by Heurich and Thoma (2008) with 34 deciduous plots located in the Bavarian Forest National Park (Germany). In their study a multiple regression was also performed with forest parameters as dependent variables and laser metrics as predictors. Accuracy was similar for basal area (20.3%) and mean diameter weighted by basal area (13.2%). Stem density also yielded the greatest error with 29.8%, whereas a better result was achieved for dominant height with 8.1%.

Even though correlation is significant for basal area only, it is noteworthy that prediction accuracy varies greatly with plot calibration radius (figure 3). As the median differences between observed and predicted values are not significant for any of the plot radii, plot size may influence residuals variance or some higher order statistics.

Comparison of field parameters values obtained with different radii show that the median of differences are not significantly different from zero. However the

mean and standard deviation, as well as several distribution quantiles, are significantly correlated with plot radius for some of the forest parameters. Even though special attention is paid to ensure that field inventories are not biased, operational constraints and stand characteristics in such areas may explain such patterns, e.g.:

- need for a small stable platform to install and operate the tripod;
- minimum distance to big tree trunks to ensure visibility from the tripod and acceptable GPS signal;
- minimum distance to small trees and understory to unfold the 4 m GPS antenna;
- presence of compact groups of several stems in coppice stands.

Presence of outliers in field observations is all the more likely since small plots exhibit higher variability of forest parameters. Such data points may affect the coefficient of determination and prediction accuracy of multiple regression models. Besides, this effect is enhanced by GPS positioning errors. Differential GPS allows sub meter accuracy in open areas, but in mountainous forests its precision depends on canopy cover density and on the elevation mask. Indeed, propitious time intervals for acquisition are short and fragmented due to topographic conditions. Real precision is more likely to be around two to four meters. For small radii, there are high chances that the extracted point cloud is only partially located within the actual field plot, resulting in erroneously fitted linear models.

6. Conclusion

With an area-based method, forest stand parameters such as dominant height, basal area, stem density and mean diameter can be precisely estimated from airborne laser scanning data for mountainous, coppice stands. Comparison of regression models obtained with plot radii ranging from five to ten meters show that prediction accuracy depends on calibration plot size. This effect may be due to biases in field data collection or to high spatial variability of mountainous coppice stands.

However, another major factor that may greatly influence prediction models is GPS precision. Indeed, topographic constraints make GPS acquisition quite uncertain in alpine environments. Further investigation is required to quantify the influence of positioning accuracy on prediction results and optimise operational field protocols for calibration of airborne laser scanning models.

Acknowledgements

J.-M. Monnet held a doctoral fellowship from Cluster de Recherche Energies (région Rhône-Alpes, France).

References

- AXELSSON, P., 2000, DEM generation from laser scanner data using adaptive TIN models. In *Proceedings of the XIXth ISPRS Congress, IAPRS, XXXIII*, pp. 110–117.
- GINISTY, C., VALLET, P., CHABÉ FERRET, S., LEVESQUE, C. and CHAUVIN, C., 2007, Disponibilités en biomasse forestière pour des usages énergétiques et industriels en France. Note de synthèse. Technical report, Cemagref.

- HEURICH, M. and THOMA, F., 2008, Estimation of forestry stand parameters using laser scanning data in temperate, structurally rich natural European beech (*Fagus sylvatica*) and Norway spruce (*Picea abies*) forests. *Forestry*, **81**, pp. 645–661.
- HOLLAUS, M., DORIGO, W., WAGNER, W., SCHADAUER, K., HÖFLE, B. and MAIER, B., 2009, Operational wide-area stem volume estimation based on airborne laser scanning and national forest inventory data. *International Journal of Remote Sensing*, **30**, pp. 5159–5175.
- HYYPÄ, J., HYYPÄ, H., LECKIE, D., GOUGEON, F., YU, X. and MALTAMO, M., 2008, Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests. *International Journal of Remote Sensing*, **29**, pp. 1339–1366.
- NÆSSET, E., 2002, Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment*, **80**, pp. 88–99.
- PATENAUDE, G., HILL, R.A., MILNE, R., GAVEAU, D.L.A., BRIGGS, B.B.J. and DAWSON, T.P., 2004, Quantifying forest above ground carbon content using LiDAR remote sensing. *Remote Sensing of Environment*, **93**, pp. 368–380.
- POPESCU, S.C., WYNNE, R.H. and NELSON, R.F., 2002, Estimating plot-level tree heights with lidar: local filtering with a canopy-height based variable window size. *Computers and Electronics in Agriculture*, **37**, pp. 71–95.