Using airborne laser scanning to assess forest protection function against rockfall

Jean-Matthieu Monnet, Eric Mermin, Jocelyn Chanussot, Frédéric Berger

USING AIRBORNE LASER SCANNING TO ASSESS FOREST PROTECTION FUNCTION AGAINST ROCKFALL

Jean-Matthieu Monnet[1*], Eric Mermin[2], Jocelyn Chanussot[3], Frédéric Berger[4]

ABSTRACT

Forests situated on slopes in mountainous areas can provide protection against natural hazards such as avalanches and rockfall. Quantifying this mitigating effect requires accurate mapping of forest stands and estimation of their dendrometric characteristics. However, steep terrain and lack of accessibility hamper field surveys. Airborne Laser Scanning (ALS) is a remote sensing technique whose potential for the retrieval of forest parameters has been widely investigated in the past ten years. The objective of this study is to assess the potential of ALS for estimating stand parameters required as input data for rockfall simulation models or more generally for quantifying the rockfall protection function of forests. ALS data was acquired over an 8.6 km^2 area including coppice stands and deciduous stands on poor quality sites. Multiple regression models were established between laser-derived metrics and forest variables from 31 field plots. The coefficients of determination for stem density, mean diameter, dominant height and basal area ranged from 0.43 to 0.80. Comparison of cross validation results showed that laser-derived estimates are more accurate than values obtained by interpolation of field data.

Key Words: Protection Forest, Rockfall, Airborne Laser Scanning, Forest Attributes Mapping

INTRODUCTION

In the French Alps, a broad afforestation program was undertaken at the end of the nineteenth century by the French Forest Office to mitigate natural hazards such as erosion and torrential floods in degraded mountainous areas. Whereas the protective function of trees has been identified for a long time, research on the interactions between rockfall and tree stands only dates back to the 1980's (an overview of rockfall – forest interactions studies can be found in Dorren *et al.*, 2007). Since then, research has mainly focused on two levels of analysis: the potential for energy dissipation by single trees and the protection function of whole forest stands. Based on the knowledge acquired with field experiments and in combination with increasing processing power of computers, many rockfall simulation models have been developed since and are now used at local or regional scale to delineate risks zones. Although high precision digital terrain models based on remote sensing technologies such as airborne laser scanning (ALS) are now available, gathering detailed information about forest stand characteristics in mountainous areas remains a time consuming task.

---

1 Phd student, Cemagref, UR EMGR, 2 rue de la Papeterie-BP 76, F-38402 St-Martin-d'Hères, France
(*Corresponding Author; Tel.: +33 (0) 4 76 76 28 06; Fax.: +33 (0) 4 76 51 38 03;
Email: jean-matthieu.monnet@cemagref.fr)
2 Forestry technician , Cemagref, UR EMGR, 2 rue de la Papeterie-BP 76, F-38402 St-Martin-d'Hères, France
3 Professor, GIPSA-Lab, Grenoble Institute of Technology, BP 46, 38402 Saint Martin D'Heres, France
4 Scientific coordinator, Cemagref, UR EMGR, 2 rue de la Papeterie-BP 76, F-38402 St-Martin-d'Hères, France

The applications of airborne laser scanning for forestry have been identified since the late 1990's and numerous studies have shown its accuracy for the retrieval of tree and stand characteristics (an overview of the applications of small-footprint laser scanning in boreal forests can be found in Hyyppä *et al.*, 2008). Two main approaches can be distinguished. Image processing methods use a raster image of the vegetation height as input to delineate single trees. Statistical methods use the raw point cloud to establish relationships between laser metrics and stand attributes. Næsset (2002) proposed a two step statistical method that has been tested in various situations so far, including at operational level (Næsset, 2004). For this method lower point densities are required since individual tree identification is not performed. The studies mainly focused on coniferous stands with gentle topographic conditions, however Lim *et al.* (2003) and Heurich and Thoma (2008) have respectively shown its efficiency in hardwood forests in Canada, and mixed forests in an alpine environment.

As far as we know, this method has never been tested on deciduous forests constituted of coppice stands and stands on poor quality sites (poor site quality is assigned to stands whose dominant height is lower than 15 m at the age of 50 years). In an alpine region such as Rhône-Alpes in France, these types of stands are common. Coppice stands situated in the footslopes provided fuelwood whereas unattainable forests on poor quality sites were often left over. These compartments are now frequently abandoned but they still complete a major protection function. To obtain accurate risk zoning with present rockfall simulation models, it is necessary to map their stand attributes as precisely as possible. Unfortunately the information is now missing or outdated and intensive field campaigns are very expensive due to accessibility constraints. While precise digital elevation models are now frequently acquired with ALS, advantage should be taken of laser data availability to extract forest information.

The overall objective of this paper is to assess the potential of laser scanning for the estimation of forest attributes required in rockfall simulation models and consists of two sub objectives:
- to evaluate the accuracy of forest attributes estimation with statistical processing of laser cloud points in complex alpine deciduous forests;
- to compare laser-derived forest parameters to those obtained with usual interpolation methods (nearest neighbor and inverse distance weighted interpolation) from field data.

MATERIAL AND METHODS

Study area

The study area is located in the French Alps upstream the village of Saint Paul de Varces (1990 inhabitants). The valley is surrounded by steep slopes and cliffs (Fig. 1), and altitude ranges from 320 to 1300 m. In 2007 a 1500 ton block fell from the cliff, rolled 0.7 km through the forested slope and destroyed a micro power plant. It stopped 200 m before the first houses. The next year six blocks of 5 m diameter rolled down and stopped in agricultural fields and vineyards. These two extreme examples are only the visible extent of the high frequency of events in this area. Indeed, field observations reveal that most of the smaller rocks are stopped by the forest.

Field data

The forest is mainly constituted of Italian maple (*Acer opalus*), downy oaks (*Quercus pubescens*) and common white beam (*Sorbus aria*). In thalwegs with better site quality ash

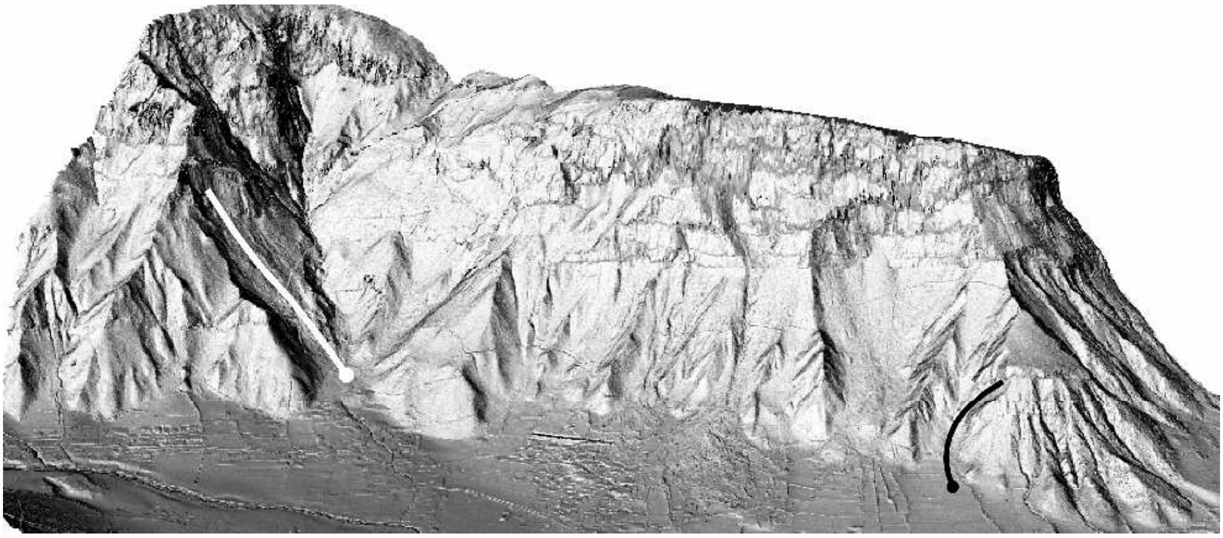

Fig. 1 3D overview of the shaded digital elevation model of the study area (St-Paul-de-Varces, France). The white and black lines represent respectively the 2007 and 2008 rockfall paths

(*Fraxinus excelsior*) and sycamores (*Acer pseudoplatanus*) are more common. Beech (*Fagus sylvativa*) is present in the upper parts. Old, coppiced chestnut trees (*Castanea sativa*) are found on the footslopes. 31 field plots were inventoried between September and October 2009. Plots were distributed every 400 m along the 550, 750, 950 and 1150 m height contours, resulting in an irregular sampling scheme were horizontal distances between neighboring plots ranged from 180 to 412 m with a mean value of 302 m. Plot centers were georeferenced using a Trimble® GPS Pro XRS receiver. After differential correction with the Pathfinder® software, the position precision (95% confidence radius) ranged from 0.8 to 1.2 m.

In each plot all trees with a diameter at breast height (*dbh*) larger than 5 cm and a horizontal distance of less than 10 m from the plot center were callipered. Additionally 10 tree heights were measured using a Vertex hypsometer. Each tree had a probability of being sampled proportional to its basal area in order to account for dominant trees. Stand attributes were then computed for each stand: mean diameter at breast height (*dbh*), its standard deviation (*stddbh*), basal area (*G*), stem density (*N*), and dominant height (h_{dom}). The dominant height was defined as the height of the 30 tallest trees per hectare and estimated as the height of the tallest sampled tree on each plot. All stand attributes are summarized in Table 1.

Laser data

The laser data was acquired with an airborne RIEGL LMS-Q560 fullwave scanner on August 27th, 2009, over 8.6 km^2. The acquisition parameters are summarized in Table 2. The fullwave files were pre-processed using the RIEGL software suite and the resulting point cloud was classified using TerraScan. The data was finally provided in the ASPRS (American Society for Photogrammetry & Remote Sensing) LAS format.

Table 1 Forestry stand parameters (number of plots *n*=31)

Variable	Mean	Minimum	Maximum
dbh (cm)	14.5	8.3	22.7
stddbh (cm)	6.8	3.0	15.0
G (m^2ha^{-1})	34.8	4.6	59.7
N (ha^{-1})	1735	764	2833
h_{dom} (m)	17.8	8.1	28.5

Table 2 System parameters of the acquisition flight

Variable	Value
Scanner model	RIEGL LMS-Q560
Scanner type	Full Wave Form
Wavelength	1550 nm
Pulse Repetition Rate	120 000 Hz
Scan frequency	77.3 Hz
Scan angle	± 30 degrees
Flight height	600 m
Laser footprint	32 cm
Point spacing in and across flight direction	60 cm

Point cloud pre-processing

For each plot, laser metrics were computed according to the following steps. All points located within a 10 m horizontal distance of the plot center were extracted. The local terrain elevation surface was estimated by linear interpolation of points classified as soil echoes. The relative height of each point was computed as the difference between the point height and the height of the interpolated surface at its planimetric coordinates. Points with a relative height of less than 2 m were then excluded from the set. Afterwards, three groups were constituted according to their return position:

- single echoes (only one echo for one emitted pulse);
- first of several echoes;
- last of several echoes.

For each group the breakpoints of 4 height bins containing each 25% of the points were calculated. These height metrics are abbreviated according to the h_{s1} scheme where the letter subscript indicates the point group (single *s*, first *f* or last *l*) and the digit subscript the number of the breakpoint (thus, subscripts 1, 3 and 5 correspond to the minimum, median and maximum height respectively). The mean height (h_{mean}) of the point groups was also computed.

Multiple regression

Multiple regression analysis was performed to establish relationships between stand parameters (dependent variables) and laser metrics (independent variables). The multiplicative model is expressed in Eq. (1):

$$Y = a_0 h_{s1}^{a_1} \dots h_{s5}^{a_5} h_{f1}^{a_6} \dots h_{f5}^{a_{10}} h_{l1}^{a_{11}} \dots h_{l5}^{a_{15}} h_{smean}^{a_{16}} h_{fmean}^{a_{17}} h_{lmean}^{a_{18}} \quad (1)$$

whose linear form is Eq. (2):

$$\ln Y = \ln a_0 + a_1 \ln h_{s1} + \dots + a_5 \ln h_{s5} + a_6 \ln h_{f1} + \dots + a_{10} \ln h_{f5} + a_{11} \ln h_{l1} + \dots$$
$$\dots + a_{15} \ln h_{l5} + a_{16} \ln h_{smean} + a_{17} \ln h_{fmean} + a_{18} \ln h_{lmean} \quad (2)$$

The linearization of the multiplicative model by logarithmic transformation was found suitable in other studies (Næsset, 2004). All statistical computing was performed with R version 2.10.0 (R Development Core Team, 2009). For each dependent variable (Y = *dbh*, *stddbh*, *G*, *N*, h_{dom}), the total number of predictors included in the model was reduced to a maximum of four by a two stage procedure:

- first the number of predictors was reduced using a stepwise model selection by AIC (function *stepAIC*, package *MASS*);
- then an exhaustive search among remaining combinations with less than four predictors was performed (function *regsubsets*, package *leaps*).

The linear model assumptions were checked using the *gvlma* package.

As no independent data was available to assess the prediction accuracy of the selected models, a cross validation was performed. Each plot was removed from the dataset at a time and the models fitted with the remaining data. The independent variables values for the leftover plot were then predicted and compared to the measured values. The procedure was repeated for each plot and the accuracy was evaluated using the root mean square error (RMSE, Eq. (3)) and the coefficient of variation of the RMSE (CV_{RMSE}, Eq. (4)):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Y_i - \hat{Y})^2}{n}} \tag{3}$$

$$CV_{RMSE} = \frac{RMSE}{\frac{1}{n}\sum_{i=1}^{n} Y_i} \tag{4}$$

where Y_i is the observed value, $\hat{Y}_i$ the estimated value and n the number of observations.

Comparison with interpolation methods

Nearest neighbor interpolation (NNI) and inverse distance weighted interpolation (IDW) are two simple interpolation methods to estimate forest attributes on a regularly spaced grid from irregularly sampled field plots. Prediction accuracy of these methods was also estimated using cross validation. The coefficient of variation of the RMSE for independent variables predicted with laser prediction (LP), NNI and IDW were then compared using Wilcoxon signed rank test.

To illustrate the prediction results of these methods, dominant height values were predicted over a 10 ha forested zone located in the southern part of the study area. NNI and IDW were used to predict dominant height on a regular 20 x 20 m spaced grid. For laser prediction, points located within each 20 x 20 m pixel where extracted and used to compute the laser metrics described above. The dominant height value for the pixel was then estimated with the selected multiple regression model.

RESULTS

Multiple regression models

Table 3 shows the selected models for each forest stand parameter. The best results are achieved for basal area (G) and dominant height (h_{dom}) with coefficients of determination of 0.8 and 0.74 respectively. The least accurate result concerns stem density (N) with 0.43.

Table 3 Statistical coherence from the multiple regression models (forest stand parameters as dependent variables and laser metrics as predictors, number of plots n=31)

Forest stand parameter	Model	Coefficient of determination (R^2)	CV_{RMSE} (%)
dbh (cm)	$1.97 \times h_{f1}^{-1.88} \times h_{f4}^{-0.49} \times h_{fmean}^{3.09}$	0.69	13.3
$stddbh$ (cm)	$0.53 \times h_{s0}^{0.28} \times h_{s2}^{-2.06} \times h_{l3}^{5.91} \times h_{lmean}^{-3.19}$	0.63	23.5
G (m^2ha^{-1})	$20.5 \times h_{f2}^{6.06} \times h_{l1}^{2.24} \times h_{fmean}^{-4.26} \times h_{lmean}^{-3.85}$	0.80	18.9
N (ha^{-1})	$49020 \times h_{f2}^{1.30} \times h_{l4}^{-2.32}$	0.43	29.2
h_{dom} (m)	$2.03 \times h_{s1}^{-0.86} \times h_{l3}^{1.65}$	0.74	17.6

Intermediate values were obtained for diameter related variables. Mean diameter (*dbh*) yielded a 0.68 coefficient of determination, and standard deviation of diameter (*stddbh*) 0.63. Regarding the coefficient of variation of the root mean square error, the stem density model was also the least accurate (30%). The best accuracy was achieved with mean diameter (13.3%) whereas dominant height, basal area and standard deviation of the diameter gave rougher values (17.6%, 18.9% and 23.5% respectively).

Cross validation and comparison with interpolation methods

The cross validation results are presented in Table 4. Laser prediction (LP) achieved the best accuracy with a coefficient of variation of the root mean square error ranging from 14.9% to 29.2%. Nearest neighbor interpolation (NNI) performed poorly: its best accuracy was obtained for mean diameter (CV_{RMSE} = 37%). Inverse distance weighted interpolation (IDW) yielded intermediate values (26.2% to 35.3%). For each method, the best individual result was obtained for the mean diameter and the second best for LP and IDW was achieved for dominant height.

As prediction accuracy is ordered for each forest parameter in the NNI>IDW>LP way, a one-tailed, paired Wilcoxon signed rank test gave a p-value of 0.031 for every pairwise model comparison. Regarding CV_{RMSE} for this set of attributes, LP hence gives significantly better estimates than IDW and NNI (0.05 significance level) and IDW performs significantly better than NNI. The results for the dominant height estimation over the 10 ha forested zone are displayed on Fig. 2.

DISCUSSION

Laser prediction accuracy

We obtained accurate estimates of forest stand parameters with the two step method, showing that it is suitable for complex deciduous forests such as coppice stands and stands on poor quality sites. Results are satisfactory for mean diameter and dominant height, with a coefficient of variation of the root mean square error less than 20% in the cross validation. Accuracy decreases slightly for basal area (22.6%), standard deviation of the diameter (28.2%) and for stem density (29.2%). These results are very similar to 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. Cross validation gave a coefficient of variation of 20.3% for basal area and 13.2% for mean diameter weighted by basal area. Stem density also yielded the greatest error with 29.8%, whereas a better result was achieved for dominant height with 8.1%. However, values of the coefficient of determination of the selected regression

Table 4 Comparison of nearest neighbor interpolation (NNI), inverse distance weighted interpolation (IDW) and laser prediction (LP) methods for the estimation of forest stand parameters from field data

Forest stand parameter	RMSE – CV_{RMSE} NNI	IDW	LP
dbh (cm)	5.4 - 37.0%	3.8 - 26.2%	2.2 - 14.9%
stddbh (cm)	3.5 - 46.1%	2.7 - 35.3%	2.1 - 28.2%
G (m^2ha^{-1})	17.4 - 49.9%	11.9 - 34.3%	7.9 - 22.6%
N (ha^{-1})	762 - 43.9%	594 - 34.2%	507 - 29.2%
h_{dom} (m)	7.7 - 43.4%	5.1 - 28.7%	3.1 - 17.6%

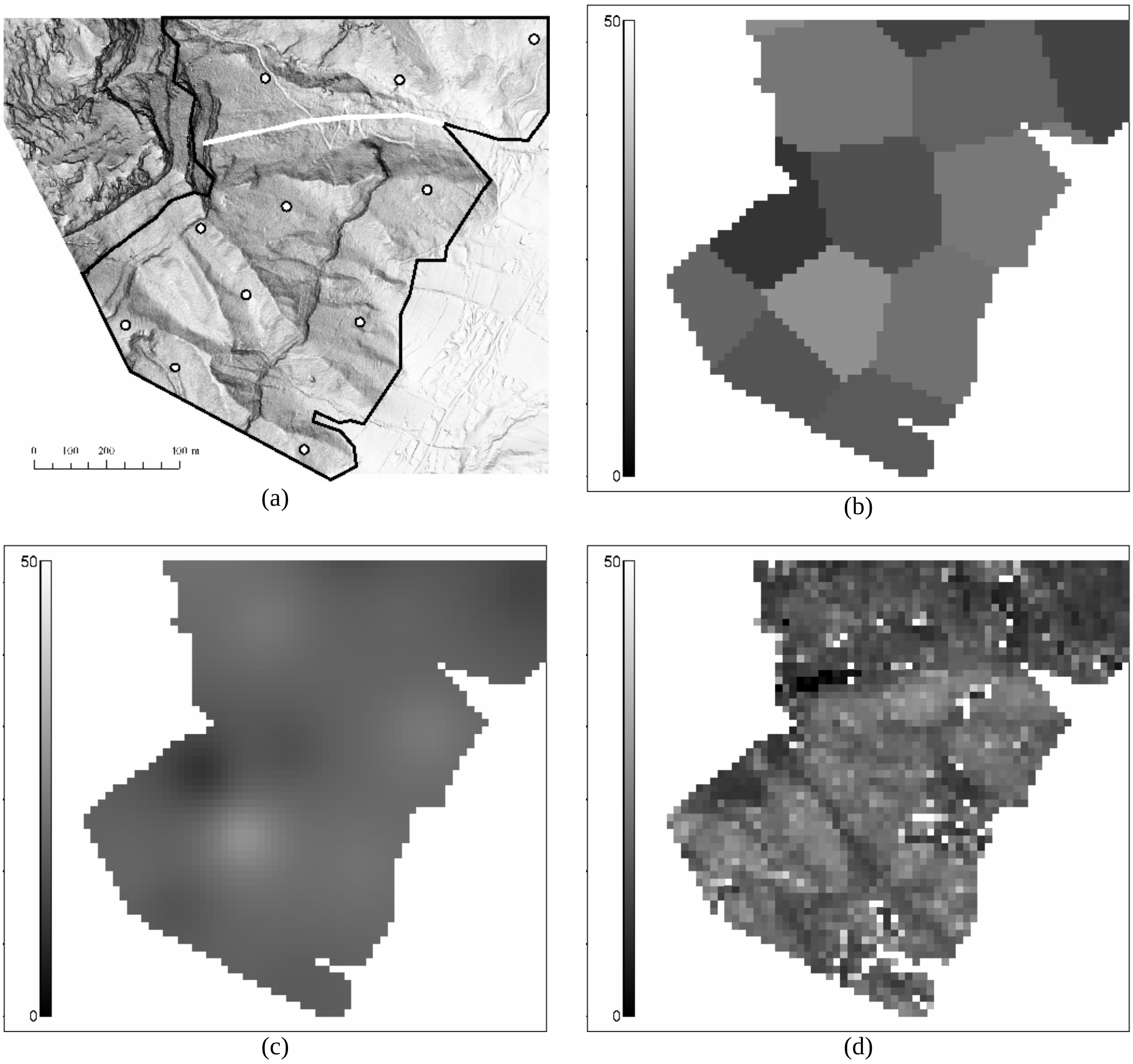

Fig. 2 Dominant height estimates over a 10 forested ha area. (a) Shaded digital elevation model with area localization (black envelope), path of the 2007 rockfall (white line) and positions of field plots (white dots). (b), (c) and (d) Prediction results for the NNI, IDW and LP methods respectively.

models were much higher in the Bavarian study. Indeed R^2 attained 0.90, 0.95, and 0.82 for stem density, dominant height and mean diameter respectively. In our study the regression analysis with basal area achieved a better result (0.80 compared to 0.66) but this may be due to the presence of an outlier. When removed, the coefficient of determination decreased to 0.59.

Various factors may explain this difference in model fitting. In this study the laser points were divided in three groups related to the return position of the echoes. Meanwhile, only 5 height quantiles and mean height were used as predictors, so that only 18 laser metrics were used in the multiple regression to avoid the risk of overfitting. In the Bavarian study a total of 108 field plots were considered and density-related metrics were also included in the model.

Despite the small size of our study area, forest stands display a great spatial heterogeneity. The alternation of ridges and thalwegs, the proximity of cliffs or rockfall stopping zones lead to changing soil and light conditions. For example, a young coppice stand in the footslope may have the same dominant height as an old stand located under a cliff, but will differ greatly with respect to other parameters. Moreover, due to the variability of forest patterns, two different forest structures may be included in a 10 m radius plot. A trade-off has to be found between the accuracy of the local stand parameters estimates and the resolution (scale of analysis) of the prediction model.

Besides the high differences in the coefficient of determination, the coefficients of variation of the RMSE obtained by cross validation are quite similar to those of the Bavarian study, indicating a similar predictive power. Regarding our objective of mapping forest attributes, this criterion is probably the most important.

Comparison of laser-derived forest parameters with estimates from interpolation methods

Cross validation results show that the introduction of laser models significantly improves the coefficient of variation of the root mean square error, compared to nearest neighbor or inverse distance weighted interpolations of field data such as basal area, mean diameter, standard deviation of the diameter, dominant height and stem density. Considering the small number of field plots, the results of interpolation methods for parameters whose high variability has been emphasized are not surprising. A good example is displayed in Fig. 2. Nearest neighbor interpolation is designed for factorial categories rather than continuous variables and does not yield satisfactory results with our data. It produces large polygons with uniform values (Fig. 2b) and would probably be more appropriate for well-delimited even-aged stands. Better results are achieved with inverse distance weighted interpolation: trends at global scale are somewhat better fitted (Fig. 2c). The decisive advantage of laser prediction is that the models are fitted at global scale but values are predicted using local information contained in the point cloud. The laser predicted dominant height (Fig. 2d) seems to exhibit spatial correlation with topographic features (Fig. 2a). Estimated heights are lower (darker pixels) on ridges, down cliffs and along rockfall paths and higher (lighter pixels) on regular hillsides, which is consistent with the visual observations on the field. A good example of the ability of laser prediction to map fine details is shown by the group of black pixels. It successfully identified an unforested area corresponding to the impact zone and path of the 2007 rockfall. However, patterns similar to salt and pepper noise (isolated bright or black pixels) can be observed in grey areas. Localized groups of high trees or windfalls could explain such variations, but they may also be linked to point cloud misclassification in steep terrain zones. Indeed, present algorithms often fail to correctly handle rugged terrain and especially cliffs (Kobler *et al.*, 2007). However these areas can be determined by slope analysis on the digital elevation model calculated with soil point cloud. More generally, others ways of further improving information extraction with refined laser data analysis should be investigated.

CONCLUSIONS

The results of this study show that forest parameters can be successfully estimated from airborne laser scanning data in complex deciduous forests such as coppice or poor site quality stands in an alpine environment. Levels of accuracy are similar to those obtained with relatively similar stands in Bavarian forests.

Airborne laser scanning data brings significant accuracy improvement to estimates achieved with field data alone. The scale of obtained details is also highly interesting for the assessment of forest protection function since present 3D rockfall simulation models work at resolutions less than 5 m and complex deciduous forest stands display a great spatial heterogeneity.

To reach an efficient compromise between field surveying intensity and mapping accuracy, further research should test the sensitivity of prediction to the number and size of field plots and investigate more complex methods to include laser metrics in forest attributes mapping in low accessibility areas.

ACKNOWLEDGMENTS

Jean-Matthieu Monnet held a doctoral fellowship from Région Rhônes-Alpes (France). We thank Sintegra Photo (Meylan, France) for the acquisition and pre-processing of airborne laser scanning data.

REFERENCES

Dorren, L., Berger, F., Jonsson, M., Krautblatter, M., Molk, M., Stoffel, M. and Wehrli, A. (2007). “State of the art in rockfall - forest interactions”, *Schweizerische Zeitschrift fur Forstwesen* 158(6), 128–141.

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(5), 645–661.

Hyyppä, J., Hyyppä, 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(5), 1339–1366.

Kobler, A., Pfeifer, N., Ogrinc, P., Todorovski, L., Ostir, K. and Dzeroski, S. (2007). “Repetitive interpolation: A robust algorithm for dtm generation from aerial laser scanner data in forested terrain”, *Remote Sensing of Environment* 108(1), 9–23.

Lim, K., Treitz, P., Baldwin, K., Morrison, I. and Green, J. (2003). “Lidar remote sensing of biophysical properties of tolerant northern hardwood forests”, *Canadian Journal of Remote Sensing* 29(5), 658–678.

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(1), 88–99.

Næsset, E. (2004). “Practical large-scale forest stand inventory using a small-footprint airborne scanning laser”, *Scandinavian Journal of Forest Research* 19(2), 164–179.

R Development Core Team (2009). R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. http://www.R-project.org