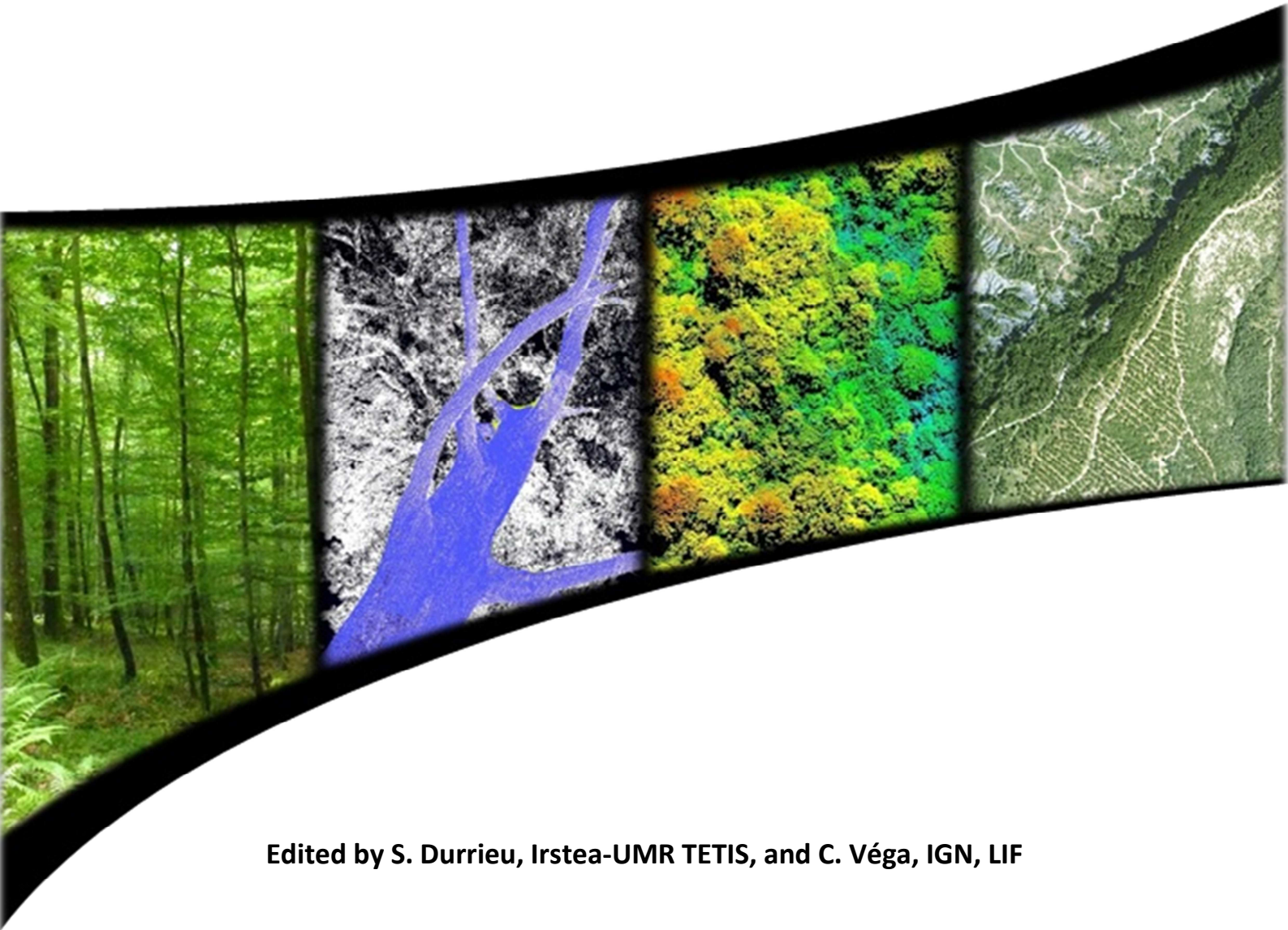
Proceedings of SilviLaser 2015

14th conference on Lidar Applications for Assessing and Managing Forest Ecosystems

September 28-30, 2015 - La Grande Motte, France



Edited by S. Durrieu, Irstea-UMR TETIS, and C. Véga, IGN, LIF

SilviLaser 2015

SilviLaser 2015 was the 14th conference in a series focusing on Lidar Applications for assessing and Managing Forest Ecosystems. The Conference aimed to bring together research scientists and practitioners from around the world to share their experience in the development and application of Lidar to improve our understanding of forest ecosystem functioning and facilitate their sustainable management through improved forest assessment and inventory. It also aimed to strengthen and develop new linkages between researchers, data providers and product end users.

SilviLaser 2015 was organized by Irstea-UMR TETIS, IGN-LIF and Société Française de Photogrammétrie et de Télédétection (SFPT) and took place in the frame of the ISPRS Geospatial Week 2015, a bundle of ISPRS and non ISPRS events on spatial information.

The scientific organization was supported by IUFRO and by Agropolis Fondation (under the reference ID 1500-012 through the « Investissements d'avenir » programme - Labex Agro: ANR-10-LABX-0001-01).

The French national forestry board (ONF - Office National des Forêts) supported the organization of a field trip dedicated to forest management issues in France, in relation with the Conference.

Scientific Committee

Co-chairs : Sylvie **Durrieu**, Irstea, France
Cédric **Véga**, IGN, France

Members :

Hans-Erik **Andersen**, USDA Forest Service Pacific Northwest Research Station, USA
John **Armston**, Remote Sensing Centre, DSITIA, Australia
Marc **Bouvier**, Irstea, France
Johannes **Breidenbach**, Norwegian Forest and Landscape Institute, Norway
Warren **Cohen**, USDA Forest Service Pacific Northwest Research Station, USA
Bruce **Cook**, NASA's Goddard Space Flight Center, USA
Nicholas **Coops**, University of British Columbia, Canada
Jean-François **Côté**, Canadian Wood Fibre Centre, Natural Resources Canada, Canada
Darius **Culvenor**, Environmental Sensing Systems, Australia
Mark **Danson**, University of Salford, UK
Richard **Fournier**, Université de Sherbrooke, Canada
Terje **Gobakken**, Norwegian University of Life Sciences, Norway
Eloi **Grau**, Irstea/CNES, France
David **Harding**, NASA Goddard Space Flight Center, USA
Ross **Hill**, Bournemouth University, UK
Yasumasa **Hirata**, Forestry and Forest Products Research Institute, Japan
Markus **Hollaus**, Vienna University of Technology, Austria
Chris **Hopkinson**, University of Lethbridge, Canada
Akira **Kato**, Chiba University, Japan
Barbara **Koch**, University of Freiburg, Germany
Ronald **McRoberts**, USDA Forest Service
Zengyuan **Li**, Chinese Academy of Forestry, China
Matti **Maltamo**, University of Eastern Finland, Finland
Jean-Matthieu **Monnet**, Irstea, France
Felix **Morsdorf**, University of Zürich, Switzerland
Erik **Naesset**, Norwegian University of Life Sciences, Norway
Håkan **Olsson**, Swedish University of Agricultural Sciences, Sweden
Norbert **Pfeifer**, Vienna University of Technology, Austria
Sorin **Popescu**, Texas A&M University, USA
Jean-Pierre **Renaud**, ONF, France
David **Riaño**, University of California Davis, USA
Jan **Rombouts**, ForestrySA, Australia
Jackie **Rosette**, Swansea University, UK
Benoît **St-Onge**, Université du Québec à Montréal, Canada
Paul **Treitz**, Queen's University, Canada
Grégoire **Vincent**, IRD, France
Mike **Wulder**, Canadian Forest Service, Canada
Randolph **Wynne**, Virginia Polytechnic Institute and State University, USA

Contents

POSTER SESSION

KEYNOTE PRESENTATION

Tree architecture and tropical forest structure

Francis Hallé, Professor Emeritus

University of Montpellier (France)

Prof. Francis Hallé is a famous botanist who has dedicated his life to researching tropical forests. He developed the fundamentals of tree architecture and forest structure dynamics, together with Prof. Roelof Oldeman. He is a pioneer in exploring life in the canopy of tropical forests and cofounder of the famous canopy raft (Radeau des Cimes).

Presentation content:

Prof. Francis Hallé shared his thorough knowledge on tropical forests with insight on the importance of structure in ecosystem functioning. His talk was an opportunity to provide the SilviLaser science community with invaluable information on tropical forests and food for thought with regard to the use of lidar to study tropical forest ecosystems.

KEYNOTE PRESENTATION

Lidar and photogrammetric point clouds for describing forest canopies: competing or synergistic technologies?

Benoît St-Onge, Professor (Ph.D.)

Université du Québec à Montréal (UQAM), (Canada)

Benoît St-Onge is a Professor at the Geography Department of UQAM since 1995. He obtained his Ph.D. in Geography at University of Montréal in 1994. He specializes in the use of 3D remote sensing and geomatics methods for the analysis of forest environments. Funded by the Natural Sciences and Engineering Council of Canada (NSERC) and the Fonds Québécois de la recherche sur la nature et les technologies (FQRNT), he develops image and laser altimetry data processing techniques for forest mapping. He also designs new digital photogrammetry approaches for the 3D mapping of forest canopies. He currently works on species identification using standard or multispectral lidar data, and photogrammetric point clouds.

Presentation content:

Image matching has reached a level of sophistication that enables us to create dense photogrammetric point clouds of the surface of forest canopies in which even individual trees can be resolved. These not only contain 3D data but also color intensities. Moreover, airborne laser scanners are evolving towards multispectral systems tightly integrated to precise metric cameras. Will these new systems altogether replace conventional aerial photography? Because image matching over dense forests does not provide information on ground elevation, lidar will always be needed to first establish a DTM. This enables future updates, as well as long-term retrospective studies, with air photography alone. This being given, how do the image matching and lidar data compare in terms of accuracy and density of points? Of particular importance in complex forests are the occlusion patterns. Image matching requires that a forest surface element be visible from at least two points of views, whereas only one is needed for lidar. How is this affected by acquisition parameters? Lidar intensities depend on the reflectance of the intercepted material, and on the very variable cross-section area of interception. In the case of images, the sun-target-sensor geometry determines intensities, also in a very variable way. In which case do we find the best signal to noise ratio? Can, and should we, combine these data? Based on numerous examples and on theoretical considerations, tentative answers to these questions we provided, hoping to raise awareness regarding this complex and evolving issue, and stir up discussions and further research.

KEYNOTE PRESENTATION

The evolution of NASA airborne and spaceflight lidars for measurement of forest structure and the Earth' s surface: the past, present and future.

Dr. Bruce Cook on behalf of Dr. David J. Harding

NASA Goddard SpaceFlight Center (USA)

Dr. David Harding conducts research in the topographic expression of land surface processes and the physical properties of vegetation, snow, ice and water, in particular by utilizing advanced airborne and space-based laser altimeter systems. He is principal investigator for the airborne Slope Imaging Multi-polarization Photon-Counting Lidar (SIMPL), developed through the NASA Instrument Incubator Program. He was a member of the Ice, Cloud and land Elevation Satellite (ICESat) Science Team and is now a member of the ICESat-2 Science Definition Team. He is also Study Scientist for the Lidar Surface Topography (LIST) mission that the NRC recommended be conducted by NASA in their Earth Science Decadal Survey report.

Presentation content:

NASA has significantly advanced methods to measure forest structure and the elevation of the Earth&apos s surface through its leadership role in the development of lidar technologies and in their scientific use. Advances are driven by the development and airborne demonstration of instrumentation with increasing capabilities that serve as technology and science pathfinders for satellite missions. Evolving capabilities include large footprint lidars, high power and micropulse laser transmitters, full-waveform and photon counting receivers, simultaneous measurement of laser and solar reflectance and systems with multiple beams, wavelengths and polarizations. The incorporation of new capabilities in satellite missions and their benefits, with a focus on forest remote sensing, were discussed in the context of the ICESat mission, the upcoming ICESat-2 and GEDI missions and those to follow.

KEYNOTE PRESENTATION

The national LIDAR mapping program in the Philippines: early applications and latest developments focussing on forest applications.

Enrico Paringit, Professor

University of the Philippines Diliman (Philippines)

Dr. Enrico Paringit is Professor at the UP Diliman Department of Geodetic Engineering (DGE). He received the Geospatial World Excellence in Policy Implementation Award for 2014 in Geneva, Switzerland, for the Disaster Risk Assessment, Exposure and Mitigation-Light Ranging and Detection technology (DREAM-LiDAR) 3D mapping.

Presentation content:

As a consequence of the global change, typhoons and floods are expected to increase in number and intensity, exposing island-states like the Philippines to major human and material losses with relative negative impact such as the setback of national development. The Disaster Risk and Exposure Assessment for Mitigation (DREAM) Program was initiated in 2011 with the aim to better prepare the country and its people for such natural disasters. By being proactive in providing Filipino communities with science-based information in an era of a rapidly changing environment, through the use of state-of-the-art technologies backed by skilled, committed, and driven Filipino workforce, this project envisions to make the DREAM of a more resilient Philippines a reality (https://dream.upd.edu.ph/). Beside the Phil-LiDAR 1 Program, an expansion of the DREAM Program that aims at producing 3D flood and hazard maps for the 2/3 of the Philippine river systems from a national LiDAR coverage, the Phil-LiDAR 2 Program aims to produce detailed resource maps using LiDAR for various applications, including forest monitoring and protection. Pr. Enrico Paringit presented the DREAM and the Phil-LIDAR 2 programs, focusing on forest applications.

Estimating and mapping forest structural diversity using airborne laser scanning data

Matteo Mura[1], Ronald E. McRoberts[2], Gherardo Chirici[3], Marco Marchetti[1]

[1]*Dipartimento di Bioscienze e Territorio, University of Molise, Italy*
[2]*Northern Research Station, U.S. Forest Service, Saint Paul, Minnesota, USA*
[3]*Department of Agricultural, Food and Forestry Systems, University of Florence, Florence, Italy*

Highlights: The paper presents the results of estimating indices of structural diversity for a study study area in Molise, Italy, and the construction of the relative maps depicting the spatial pattern of the structural diversity indices. Our results demonstrate the utility of using ALS data for improving areal estimates of indices.

Key words: *ALS, forest structural diversity, GREG estimator, model-assisted estimator*

Introduction

Structure is a well-known important factor in assessing forest diversity [1], and the availability of maps depicting spatially-explicit patterns of structural diversity would be of great use when planning conservation strategies. Further, mapping forest structure would facilitate habitat and diversity assessments for large, remote and steep areas that cannot be safely reached by field crews. Recently, Airborne Laser Scanning (ALS) have been shown to be reliable and valuable sources of information for estimating and assessing forest structural parameters ([2-3-4-5]) thus contributing great potential support for forest biodiversity studies. For management and planning purposes, the accuracy and precision of biodiversity estimates for the area of interest are essential, thus information regarding the accuracy or precision of parameters estimated from the map unit predictions is required [6]. Motivated by these considerations, the aim of this paper is twofold: (i) to construct inferences in the form of confidence intervals for estimates of common indices of forest structural diversity using LiDAR data obtained via airborne laser scanning (ALS) as auxiliary information, and (ii) to construct maps depicting the spatial pattern of these structural diversity indices using gridded ALS data metrics. The study area is in Molise, Italy, and the analyses include comparison of estimates obtained using design-based, simple random sampling (SRS) and model-assisted estimators.

Materials and methods

Study area

The study area is located in the southwestern part of Molise Region in central Italy and includes 36,360 ha.

Field data

A two-phase tessellation stratified sampling (TSS) was carried out in the study area. The area was covered by 437 hexagons, each with an area of 1 km2. During the first phase, a point was randomly selected in each hexagon and classified as "forest" or "non-forest" Of the 437 points, 197 were classified as "forest". In the second phase, 62 points were randomly selected from the 197 "forest" points (sampling rate ≈ 30%) and surveyed in the field by a circular plot of 13-m radius where the diameters at breast-height (1.30 m) (DBH) of all trees with DBH of at least 9.5 cm were measured. Height (H) was measured for a sub-sample of plot trees and estimated for the remaining trees using a model of the H-DBH relationship constructed using data for the measured trees.

Structural diversity indices

Using the data collected in the field, two common indices of forest structural diversity were calculated for each plot. Horizontal structural diversity was assessed using the standard deviation of DBH (σ_{DBH}), while plot-level vertical structural diversity was assessed using the standard deviation of H (σ_H).

Airborne Laser Scanning (ALS) data

ALS data were acquired under leaf-on canopy conditions in June 2010. The LiDAR instrument was an Opthech Gemini LiDAR, a two-return range detection system set with a maximum scan angle of 15° and a pulse frequency of 70 KHz, resulting in an average density of 1.5 pulses/m2.

Common procedures for pre-processing ALS data include removal of outliers, ground/non-ground classification, and computation of normalized height. Firstly, air points that are clearly higher than the median

elevation of surrounding points and isolated points with few neighbors resulting from sensor errors or backscatter by flying objects were removed. Subsequently a ground surface was generated by classifying ground points on the basis of the adaptive triangular irregular network (TIN) model algorithm [7]. Lastly, the relative height above ground of each echo was calculated and used to extract 22 height and density canopy metrics for each sample plot measured in the field. The metrics calculated were Canopy cover (cov) as the proportion of first returns above 1.30 m on all first returns. Canopy density metrics were the proportions of all returns above 1.30 m (dns), the proportion (d_{00}) and the count (c_{00}) of returns between 1.30 m and 10 m. Canopy height metrics were the percentiles of the height canopy distribution ($p_{10}, p_{20}, \cdots, p_{90}, p_{99}$) and height summary statistics such as minimum (H_{min}), maximum (H_{max}), average (H_{avg}), standard deviation (H_{std}), coefficient of variability (H_{cv}), skewness (H_{ske}) and kurtosis (H_{kur}). Further, the canopy relief ratio (*CRR*) was also calculated [8]. Such metrics were used as independent variables for constructing prediction models. The same metrics were calculated for the 23-m × 23-m pixels that tessellated the study area and were approximately of the same area as the plots. Maps were constructed by calculating model predictions of the structural diversity indices for these pixels.

Model development and spatial predictions

Simple linear models using the 22 ALS-derived metrics as candidate independent variables were developed to predict the two indices. Regression diagnostics were used to check for potential outliers and multicollinearity problems. The presence of outliers was assessed by the leverage influence of an observation on the overall model behavior [9]. Problems of multicollinearity were minimized by using an F-test to select the model with the fewest number of predictors that still retained statistical significance, even at the expense of a decrease in R^2. Once the most statistically significant models with the fewest predictor variables were selected, they were used with the gridded ALS metrics to predict the structural diversity indices for all pixels in the study area.

Inference

The mean value of each index was attributed to the entire forested area estimated using the model-assisted, generalized regression (GREG) estimators [10] calculated as

$$\hat{\mu}_{GREG} = \frac{1}{N}\sum_{i=1}^{N}\hat{y}_i - 1/n[\sum_{i=1}^{n}(\hat{y}_i - y_i)] \quad (1)$$

where N is the number of population units (the 23-m × 23-m forest cells in the study area), $\hat{y}_i$ is the model prediction for the *i-th* population unit, *n* is the sample size, $\hat{y}_i$ is the model prediction for the *i-th* sample plot and y_i is the observed value for the *i-th* plot.

Results

The sum of squared errors of the developed models were compared using the F-Test to keep the most statistically significant model with the fewest number of predictor variables. From this test, the ultimate selected models had four predictors for the σ_{DBH} model, and the three predictors for the σ_H model (Table 1), yielding a R^2 of 0.63 for σ_{DBH} and of 0.52 for σ_H. The models were then applied upon the gridded ALS metrics selected as independent variables in order to map the diversity indices in the study area (Figure 1). The spatial predicted values ranged between 0-29 for σ_{DBH} and between 0-10 for σ_H. For the entire study area, the SRS estimate of mean σ_{DBH} was 6.56 with standard error (SE) of 0.58, and the GREG estimate was 6.36 with SE=0.06. The SRS estimate of mean σ_H was 2.90 with SE=0.17, and the GREG estimate was 2.93 with SE=0.02.

Table 1: Parameter estimates for the final regression model for σ_{DBH} (left) and σ_H (right).

R^2=0.63; RMSE=2.52		R^2=0.52; RMSE=0.91	
Predictors	Parameter estimated	Predictors	Parameter estimated
$Intercept$	7.06**	$Intercept$	3.46***
H_{min}	-7.93***	H_{min}	-2.48***
H_{cv}	6.13**	H_{cv}	2.68***
H_{ske}	2.39*	p_{20}	0.42***
p_{20}	1.76***		

[a] Level of significance: *<0.05; **<0.01; ***<0.001

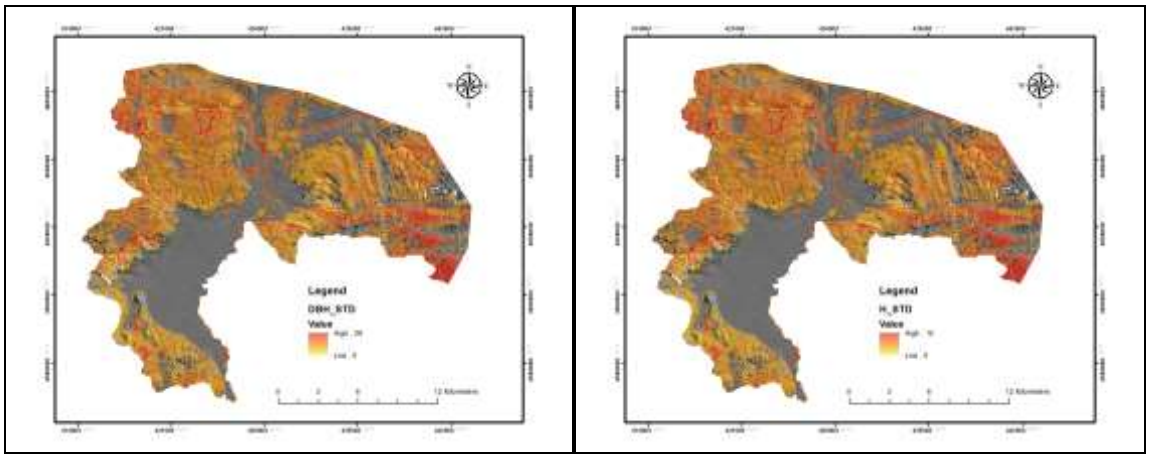


Figure 1: Spatial predictions of σ_{DBH} (left) and σ_H for the forested portion of the study area.

Conclusions

Several conclusions can be drawn from the study. Firstly, ALS data are a valid source of information for estimating and spatially predicting forest structural diversity. Secondly, simple linear models were sufficient to characterize the relationship between the structural diversity indices and the ALS metrics used as predictor variables. Thirdly, the GREG estimator worked well for estimating mean structural diversity.

References

[1] McElhinny, C., Gibbons, P., Brack, C., & Bauhus, J. (2005). Forest and woodland stand structural complexity: Its definition and measurement. Forest Ecology and Management, 218, 1-24.

[2] Lefsky, M.A., Cohen, W.B., Parker, G.G., & Harding, D.J. (2002). Lidar Remote Sensing for Ecosystems Studies. BioScience 52(1), 19-30.

[3] Lim, K., Treitz, P., Wulder, M., St-Onge, B., & Flood, M. (2003). LiDAR remote sensing of forest structure. Progress in Physical Geography, 27(1), 88-106.

[4] Zimble, D.A., Evans, D.L., Carlson, G.C., Parker, R.C., Grado, S.C., & Gerard, P. D. (2003). Characterizing vertical forest structure using small-footprint airborne LiDAR. Remote Sensing of Environment, 87(2-3), 171-182.

[5] Wulder, M.A., Bater, C.W. Coops, N.C. Hilker, T. & White, J.C. (2008). The role of LiDAR in sustainable forest management. The Forestry Chronicle, 84(6), 807-826.

[6] McRoberts R.E., (2011). Satellite image-based maps: Scientific inference or pretty pictures? Remote Sensing of Environment, 115(2), 715-724.

[7] Axelsson, P. (2000). DEM generation from laser scanner data using adaptive TIN models. *International Archives of Photogrammetry and Remote Sensing*, *33*(B4), 110-117.

[8] Parker, G.G., & Russ, M.E. (2004). The canopy surface and stand development: Assessing forest canopy structure and complexity with near-surface altimetry. *Forest Ecology and Management*, *189*, 307-315.

[9] Rowlings, J.O., Pantula, S.G., & Dickey, D.A. (1998). *Applied regression analysis: a research tool, 2nd ed*. Springer-Verlag New York, Inc.

[10] Särndal, C.E., Swensson, B., Wretman, J. (1992). Model assisted survey sampling. Springer-Verlag, Inc. New York.

Comparing generalized linear models and random forest to model vascular plant species richness using LiDAR data in a natural forest in central Chile

Javier Lopatin[1,2], Klara Dolos[1], Jaime Hernández[2], Mauricio Galleguillos[3], Fabian E. Fassnacht[1]

[1]*Institute of Geography and Geoecology, Karlsruhe Institute of Technology (KIT), Kaiserstraße 12, 76131 Karlsruhe, Germany. E-mails: javier.lopatin@kit.edu, klara.dolos@kit.edu, fabian.fassnacht@kit.edu.*
[2]*Laboratory of Geomatics and Landscape Ecology, Faculty of Forest and Nature Conservation, University of Chile, 11315 Santa Rosa, Santiago, Chile. E-mail: jhernand@uchile.cl.*
[3]*Department of Environmental Sciences, School of Agronomic Sciences, University of Chile, 11315 Santiago, Chile. E-mail: mgalleguillos@renare.uchile.cl.*

Highlights: Vascular plant richness estimation in a Mediterranean forest. Generalized linear models (GLM) outperformed random forest (RF) when fed with LiDAR-based predictors representing micro- and macro-topography as well as vegetation structure. Outcomes show that RF deals less efficiently with biodiversity count data due to asymmetry of the error distribution and heteroscedasticity.

Key words: *species richness, LiDAR data, GLM, random forest, bootstrapping validation.*

Introduction

Today, biodiversity is considered to be an essential element of the Earth system from which all humans benefit directly or indirectly [1]. Theoretical and empirical studies have suggested that local biodiversity is strongly influenced, in a positively correlated fashion, by environmental heterogeneity [2]. More diverse environments can host more ecological niches, which in turn can be colonized and inhabited by a greater number of species [3-4]. As all imaging remote sensing sensors are able to deliver information on environmental heterogeneity, they have a high potential to support the identification and monitoring of species rich sites, as requested above.

The application of remote sensing data to map richness of plants has been examined since approximately the turn of the millennium and the potential of the technique has been confirmed in several studies [5]. So far conducted studies on richness of plants show a strong tendency towards the application of optical remote sensing sensors, while the number of studies that investigated the potential of active optical sensors, such as Light Detection and Ranging (LiDAR), is still sparse. LiDAR data have proven to be one of the most powerful data acquisition systems to obtain topographical and vegetation-structural information. Both of these information were found to be an estimate of environmental heterogeneity in earlier studies [6-7].

In the present study, the suitability of airborne discrete-return LiDAR data for the mapping of vascular plant species richness within a Sub-Mediterranean second growth native forest ecosystem was examined. As species richness data constitute count data and therefore is prone to show asymmetric error distribution and heteroscedasticity, the model approach has to be selected carefully. In this realm, we compared a widely used machine learning method (random forest) with a generalized linear model (GLM) which we optimized for the response variables by assuming a negative binomial data distribution. Predictive models for vascular plant richness of four different layers were performed: total, tree, shrub and herb richness.

Study area and ground data

The study area, Monte Oscuro, is located in the Andes foothills of central Chile in the Maule region (35°07‟00” S, 70°55‟30” W). This mountainous area is associated with the Sub-Mediterranean Temperate bioclimatic zone, with a mean annual precipitation of 1000 mm which is mainly concentrated between April and October. The site has an extent of 1295 ha, a mean altitude of 1075 m above sea level and mostly southern aspect. Monte Oscuro is covered by secondary deciduous forest dominated by *Nothofagus obliqua* (Mirb.) Oerst., mixed with sclerophyll species.

We used the vascular plant richness as response variable due to the clear dominance of this taxonomic group in Mediterranean forests and also because it is essential for the trophic network and ecosystem functioning [8]. A botanic survey was performed between January of 2013 and January of 2014. A 200 x 200 m regular grid was applied to locate the square nested plots [9]. In each sampling unit, species within three height layers were registered and determined: trees (height of 2 m or more), shrubs (less than 2 m of height) and herbs (non-woody plants). Each plot was composed by 6 nested subplots. The largest one, in which only trees were registered, occupied 225 m² (15 m x 15 m) and it had its vertices aligned to the four cardinal points (N, E, S, and W). Shrubs were identified in two subplots of 25 m² (5 m x 5 m). Herbs were registered in three sub-subplots of 1 m².

Methods

A discrete return LiDAR dataset (average point cloud density of 4.6 points per m² and a footprint of 29 cm) was used to obtain five topographical (macro- and micro-topography) predictors and seven predictors related to canopy structure. The predictors were derived from a 1 m² digital terrain model (DTM) and digital canopy model (DCM) respectively. These predictors were used to estimate total, tree, shrub and herb plant vascular richness using two different approaches: 1) random forest (RF) algorithm as representation of one of the most use machine learning (ML) methods in remote sensing (RS), and 2) generalized linear model (GLM) as an example for traditional statistical models. As many ML methods are described as being non-parametric, it is frequently assumed that there are no requirements concerning the data distribution of the response variable. However, this is not true for regression trees and random forest methods which fit standard linear (Gaussian) regressions for tree nodes or are based on measures for node impurity such as the sum of squared deviations to the mean [10], as such approaches do not account for asymmetry and heteroscedasticity [11]. For this reason, GLMs may be preferable over standard machine learning methods, ordinary least squares and some PLS-based models for modeling count data, due to the option to choose an appropriate model family for the particular residual distribution (in this case we optimized the GLM by assuming a negative binomial data distribution with log-link function).

The applied GLM models were combined with a feature selection method as GLM cannot handle multi-colinearity among the predictors. Since the relation between explaining variables and the response appeared to be linear, hierarchical partitioning selection was applied. The resulting variable importances were used to select the final models. This variable importance calculated as percentage of total explained variance was also compared with the rank of variable importance given by RF. Additionally, models for all possible combinations of response variables were calculated and their AIC and explained deviances were compared. For all response variables the most relevant explaining variables regarding AIC and explained deviances were identical.

For validation purposes, a bootstrap with 500 boots was performed. That is, in each bootstrap iteration, we drew n-times from the n available samples with replacement. In this procedure a certain number of samples (on average, 36.8% of the total number of samples) are not drawn. These samples were subsequently used as holdout sample for an independent validation [12].

The performances of the models were compared based on differences in the coefficients of determination (R^2 – calculated as squared Pearson's correlation coefficient) and the normalized root mean square error (nRMSE) between predicted and observed richness values. To enable sound comparisons between the four response variables, the normalized RMSE (nRMSE, calculated as [RMSE/[max(N° spp) − min(N° spp)]]*100) was used. High values of R^2 and low values of nRMSE indicate high model quality. Bias of prediction was measured as one minus the slope of a regression without intercept of the predicted versus observed values.

Results and discussion

Variable importance was determined for both modeling approaches. They agreed that the three most important predictor variables to predict species richness of all forest layers were: mean altitude above sea level (DTM), standard deviation of slope (sd slope) and mean canopy height. In all layers, mean canopy height was chosen as the best predictor by the GLM approach of hierarchical partitioning.

The best parameterization of the two models were used to build the final predictive models. As already mentioned, GLMs cannot deal with co-linearity among the predictor variables. For this reason, only the first three variables were used in the final predictive model. In the case of RF models, all variables were used as RF typically works better with a larger number of predictors.

Results for model performances were summarized in terms of R^2 and nRMSE of all 500 bootstrap values. NB models show systematically better performances than RF (higher R^2 and lower nRMSE). In all but one layer, GLMs showed higher coefficient of determination (R^2) lower errors and a lower bias. The only exception is the herb layer which shows minimally higher errors for the NB compared to RF. LiDAR data has a precision of about 15-20 cm, which is a significant value for the height of this kind of plants, and therefore LiDAR data is expected to be less related to this layer. The best model fit was found when predicting total richness (R^2 of 0.62 and nRMSE of 14.76%), while the worst fit was observed for herb richness (R^2 of 0.47 and nRMSE of 23.35%). The GLM of herb richness prediction displayed high accuracy for low values and high dispersion for high values. Furthermore, both models show a systematic tendency to overestimate small values and underestimate high values. Moreover, this effect is higher in RF models.

The spatial distribution of species richness within the study area shows distinct patterns. For lower altitudes and areas close to the riversides we generally observed higher species richness. This is presumably related to favourable hydrological and climatic conditions. That is, warmer temperatures combined with a sufficient water supply. These areas were found to be highly heterogeneous in terms of micro-topography (sd slope). From an ecological perspective, the observed patterns agree well with the known vegetation composition of the area. Sclerophyll species (more common in warmer environments) are more abundant in the lower altitudes, while the Valdivian evergreen species are usually located near the riversides due to their water needs. Finally, *Nothofagus* are the dominant, more common and well distributed species in the area [13]. In low elevations typical species from each forest types add to an overall increased species number.

Conclusions

LiDAR derived variables on stand properties and forest structure have proven to be able to predict vascular plant species richness in forest ecosystems. It appeared that the in remote sensing studies frequently applied RF algorithm does show drawbacks when modeling count data. Here, we found already available methods such as GLMs to be better in terms of precision and bias, due to their capability to model asymmetric error distributions as typical for count data and as also observed for our dataset.

The selected applied predictor variables found during the variable selection incorporate the three information types available in LiDAR data. Altitude above sea level describes the macro-topography and is a proxy for climatic conditions, soil properties and, in the study region also for management. Standard-deviation of slope describes the variation in micro-topography when taking into account the 1 m pixel size applied in this study. Mean canopy height can be considered a proxy for forest type and light availability at the forest floor influencing species richness in the understory.

Finally, lower altitudes and areas close to the riversides exhibited higher species richness. These areas appeared to be highly heterogeneous in terms of micro-topography resulting in high micro-habitat heterogeneity. Additionally in this elevation belt sclerophyll, *Nothofagus* and Valdivian evergreen forest converge together in this transitional zone where the richness had shown to be high.

References

[1] Duffy, E. (2009). Why biodiversity is important to the functioning of real-world ecosystems. Frontiers in Ecology and the Environment, 7, 437-444.

[2] Díaz, S.; Fargione, J.; Chapin, F. & Tilman, D. (2006). Biodiversity loss threatens human well-being. PLoS Biology, 4, 1300–1305.

[3] Dufour, A.; Gadallah, F.; Wagner, H.; Guisan, A. & Butler, A. (2006). Plant species richness and environmental heterogeneity in a mountain landscape: Effects of variability and spatial configuration. Ecography, 29, 573–584.

[4] Bergen, K.; Goetz, S.; Dubayah, R.; Henebry, G.; Hunsaker, C.; Imhoff, M.; Nelson, R.; Parker, G. & Radeloff, V. (2009). Remote sensing of vegetation 3-D structure for biodiversity and habitat: Review and implications for LiDAR and radar spaceborne missions. Journal of Geophysical Research, 114, 1-13.

[5] Rocchini, D.; Balkenhol, N.; Carter, G. A.; Foody, G. M.; Gillespie, T. W.; He, K. S.; Kark, S.; Levin, N.; Lucas, K.; Luoto, M.; Nagendra, H.; Oldeland, J.; Ricotta, C.; Southworth, J. & Neteler, M. (2010). Remotely sensed spectral heterogeneity as a proxy of species diversity: Recent advances and open challenges. Ecological Informatics, 5, 318-329.

[6] Ceballos, A.; Hernández, J.; Corvalán, P. & Galleguillos, M. (2015). Comparison of Airborne LiDAR and Satellite Hyperspectral Remote Sensing to Estimate Vascular Plant Richness in Deciduous Mediterranean Forests of Central Chile. Remote Sensing, 7, 2692-2714.

[7] Lopatin, J.; Galleguillos, M.; Fassnacht, F. E.; Ceballos, A. & Hernández, J. (2015). Using a Multistructural Object-Based LiDAR Approach to Estimate Vascular Plant Richness in Mediterranean Forests With Complex Structure. IEEE Geoscience and Remote Sensing Letters, 12, 1008-1012.

[8] Palmer, M.; Earls, P.; Hoagland, B.; White, P. & Wohlgemuth, T. (2002). Quantitative tools for perfecting species lists. Environmetrics, 13, 121–137.

[9] Stohlgren, T.Stohlgren, T. (Ed.), 2007. Measuring Plant Diversity: Lessons from the Field. Oxford University Press Inc.

[10] Loh, W.-Y. (2014). Fifty Years of Classification and Regression Trees. International Statistical Review, 82, 329–348.

[11] Chaudhuri, P.; Lo, W.-D.; Loh, W.-Y. & Yang, C.C. (1995). Generalized Regression Trees. Statistica Sinica, 5, 641–666.

[12] Fassnacht, F., Neumann, C., Förster, M., Buddenbaum, H., Ghosh, A., A., C., P.K., J. & Koch, B.(2014). Comparison of feature reduction algorithms for classifying tree species with hyperspectral data on three central European test sites. IEEE Journal of Selected Topics in Applied Earth Observation and remote Sensing, 7, 2547-2561

[13] Gajardo, R., 1994. La vegetación natural de Chile. Editorial Universiaria.

Individual tree properties from ALS data as input to habitat analysis in boreal forest

Eva Lindberg [1,2], Jean-Michel Roberge [3], Therese Johansson [3],
Markus Hollaus [2], Johan Holmgren [1] and Joakim Hjältén [3]

[1] *Department of Forest Resource Management, Swedish University of Agricultural Sciences (SLU), 901 83 Umeå, Sweden; E-mail: eva.lindberg@slu.se, johan.holmgren@slu.se*
[2] *Vienna University of Technology, Department of Geodesy and Geoinformation, Research Group Photogrammetry, Gußhausstraße 27 –29, 1040 Vienna, Austria; E-mail: markus.hollaus@geo.tuwien.ac.at*
[3] *Department of Wildlife, Fish, and Environmental Studies, Swedish University of Agricultural Sciences (SLU), 901 83 Umeå, Sweden; E-mail: jean-michel.roberge@slu.se; therese.johansson@slu.se; joakim.hjalten@slu.se*

Highlights: This study shows examples of detailed analysis of forest canopy from ALS data with potential use as input to habitat analysis in forests. This includes delineation of individual tree crowns and analysis of the distribution of tree heights and the tree species composition.

Key words: *canopy structure, forest age, biodiversity, species richness, LiDAR.*

Introduction

Planning and management of forests for biodiversity conservation requires knowledge about local stand conditions such as forest structure and tree species composition in relation to habitat demands for different species, as well as the amount and distribution of diversity hotspots in the surrounding landscape [1] Airborne laser scanning (ALS) have proven useful for predicting species richness and composition in a range of taxonomic groups [e.g., 2] Until now, habitat studies including ALS data have mostly used area-based analysis where features are extracted as statistics from the ALS data in raster cells with size corresponding to field plots, typically in the order of 100-400 m^2. However, more detailed information about objects and structures can be derived from the ALS data. With dense ALS data (>5 returns/m^2), individual tree crowns (ITC) can be delineated from the data.

The objective of this study is to show examples of detailed analysis of the forest canopy from ALS data and use the derived information for habitat analysis. The canopy analysis includes identification of individual tree crowns and estimation of the height distribution and information for tree species classification. This information can be used, for example, to predict patterns of species richness in the forest landscape.

Materials

The study area is a 30 × 40 km forest landscape in the middle boreal zone of northern Sweden (64°05' - 64°10'N, 19°05' - 19°30' E). The age of the forest stands was 8-130 years and the dominating tree species were Scots pine (*Pinus sylvestris*) and Norway spruce (*Picea abies*) with a minor component of deciduous trees. Two ALS datasets were included: dense ALS data (5 returns m^{-2}) and sparse ALS data (0.5-1 m^{-2}). We sampled species richness and abundance of flying and epigaeic (i.e., ground-living) beetles in 33 study stands and birds in 47 study stands, all covered with dense ALS data.

We have used the same study area in an earlier study where we derived area-based metrics from dense ALS data and forest estimates from satellite images [3]. This provided insights in the approximate suitable scale and relevant variables for habitat analysis. The most important variable for species richness and abundance of both birds and insects was the maximum height of the forest, which was correlated with the total stem volume. For birds, the second most important variable was the amount of low vegetation (0.5-3 m) with positive sign, while for beetles the second most important variables was the proportion of Scots pine with negative sign.

For the analysis in this pilot study, we compared two middle-aged pine-dominated forest stands with labels S-M03 and S-M06. According to field observations, these two forest stands had similar basal area. The area-based variables derived from remotely sensed data (e.g., maximum height and fraction of returns between 0.5 and 3 m above the ground) also had similar values for the two stands. However, the species richness of flying and epigaeic beetles was higher in S-M03 (Table 1).

Analysis

Delineation of individual tree crowns (dense ALS data)

The delineation was done by segmentation of a correlation surface (CS) model followed by ellipsoidal tree model clustering of the ALS data in 3D (Figure 1). More details are given in [4]. The aim of the segmentation was to establish one segment for each tree in the topmost canopy layer. The aim of the clustering was to establish one cluster for each tree in the topmost canopy layer as well as one cluster for each tree below. The algorithm was based on k-means clustering using ellipsoidal tree crown models. Two categories of clusters were defined: Fixed clusters corresponding to trees already identified by segmentation of the CS and additional flexible clusters corresponding to trees below the topmost canopy layer. For each cluster, the mean height above the ground (i.e., the centroid) as well as the maximum and minimum height above the ground of the ellipsoid was calculated. The maximum height was assumed to be the tree height and the minimum height was assumed to be the crown base height.

Summary variables (i.e., individual tree-based variables) were calculated from the clusters in circles with a radius of 50 m centred on the field plot in each study stand. The variables were the mean and standard deviation of the centroid and maximum height above the ground of the returns assigned to each cluster as well as the relative live crown height, which was calculated as the height of the lowest point of each ellipsoid divided by the highest point of each ellipsoid (Table 1).

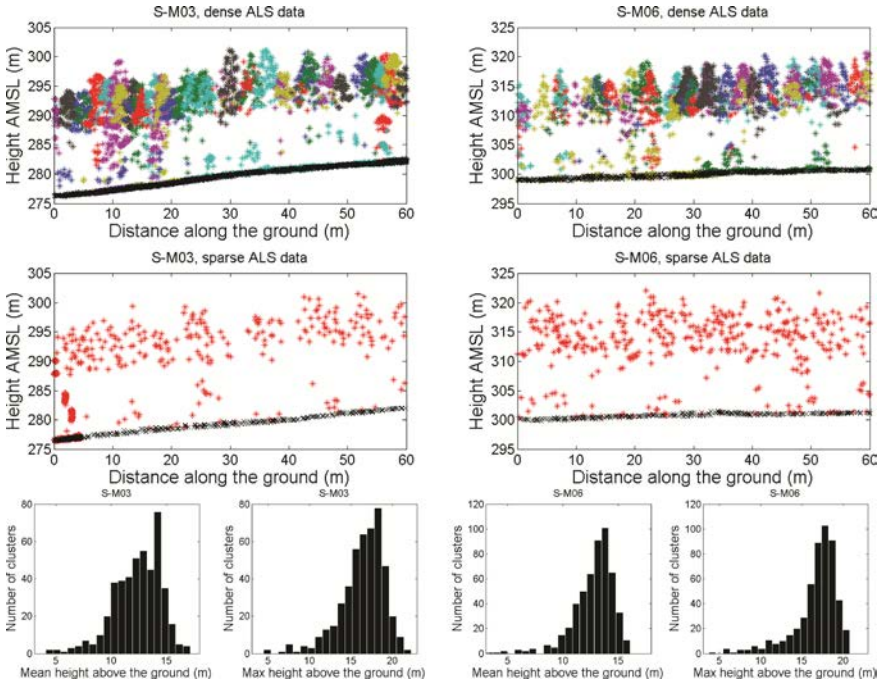


Figure 1: Transects with width 10 m from dense ALS data in S-M03 (top left) and S-M06 (top right) and sparse ALS data in S-M03 (middle left) and S-M06 (middle right) and histograms of centroid and maximum height of clusters in S-M03 (bottom left) and S-M06 (bottom right). The dense ALS data have been delineated into individual tree crowns. The ALS returns assigned to one cluster are shown in the same colour.

Area-based analysis (sparse and dense ALS data)

Area-based metrics were derived from the sparse and dense ALS data in raster cells including all returns (i.e., first and last). The maximum height above the ground was derived in 2.5×2.5 m raster cells and the number of returns was derived in 12.5×12.5 m raster cells. Summary variables were calculated as the mean of the metrics in circles with a radius of 50 m centred on the field plot in each study stand (Table 1).

Table 1: Field observations and variables derived from ALS data for the two forest stands.

Field observations	S-M03	S-M06
Basal area (m^2/ha)	25.7	28.3
Proportion of basal area pine/spruce/deciduous	83.2%/15.6%/1.2%	90.5%/3.5%/6.0%
Bird species richness	5	5
Flying beetle species richness	151	96

Epigaeic beetle species richness	75		56	
Area-based variables	Sparse	Dense	Sparse	Dense
Maximum height (m)	13.5	15.4	14.9	15.7
Low vegetation < 3 m	5.0%	3.7%	6.2%	5.4%
Middle vegetation 3-10 m	9.8%	10.8%	7.1%	5.3%
High vegetation > 3 m	46.7%	55.6%	52.6%	57.0%
ITC-based variables				
Mean of centroid height (m)	-	12.40	-	12.80
Standard deviation of centroid height (m)	-	2.14	-	1.87
Mean of maximum height (m)	-	16.43	-	16.73
Standard deviation of maximum height (m)	-	2.70	-	2.72
Mean relative live crown height	-	0.52	-	0.55

Conclusions

The result of the individual tree-delineation was a representation of each tree crown in three dimensions. Based on this, various properties of the trees can be derived such as height, live crown height, crown diameter, and crown shape. This can in turn be used for classification of the species of the individual trees using allometric relationships [5]. Area-based analysis of ALS data alone does generally not allow for tree species classification since the height distribution of a forest stand with vertically elongated tree crowns (e.g., spruces) might be similar to that of a forest stand with dense low vegetation.

The area-based variables in the two study stands were similar despite that the species richness of flying and epigaeic beetles was higher in S-M03. However, S-M06 was more pine-dominated, while S-M03had a higher proportion of spruce, and tree species diversity is known to affect overall species diversity. Visual inspection of the ITC revealed more trees with vertically elongated tree crowns and more low vegetation in S-M03. This resulted in a different distribution of the height above the ground of the returns assigned to each cluster and a different mean relative live crown height (Table 1). This means that the ITC provides additional information to separate these two stands, which is not feasible only from the area-based variables.

In our previous area-based study, the most important variable for both birds and insects was the maximum height of the forest. This was consistent with other studies showing a positive correlation between forest age and the species richness of these species. In the current study, ITC delineation made it possible to derive the distribution of tree heights, which can be used as a proxy for the age distribution of the trees. The height distribution is also a measure of the layering, which is important for birds in boreal forests. Additionally, tree species composition is important for habitat analysis. To study the importance of information derived from the ITC clusters, we will derive ITC-based variables for all study stands and relate them to the field observations of birds and beetles.

Acknowledgments

This study was financed by The Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (Formas) and the Mistra research programme Future Forests. The field inventory and the acquisition of ALS data were financed by the Swedish Energy Agency, VINNOVA, and the research programme Future Forests.

References

[1] Paillet, Y., Berges, L., Hjältén, J., Odor, P., Avon, C., Bernhardt-Roemermann, M., Bijlsma, R.-J., De Bruyn, L., Fuhr, M., Grandin, U., *et al.* (2010). Biodiversity differences between managed and unmanaged forests: Meta-analysis of species richness in europe. Conservation Biology, 24(1), 101-112.

[2] Zellweger, F., Braunisch, V., Baltensweiler, A., & Bollmann, K. (2013). Remotely sensed forest structural complexity predicts multi species occurrence at the landscape scale. Forest Ecology and Management, 307, 303-312.

[3] Lindberg, E., Roberge, J.-M., Johansson, T., & Hjältén, J. Can airborne laser scanning or satellite images, or a combination of the two, be used to predict the abundance and species richness of birds and beetles at a patch scale? In International Workshop on Remote Sensing and GIS for Monitoring of Habitat Quality, Vienna, Austria, September 24–25, 2014; Pfeifer, N.; Zlinszky, A., Eds.

[4] Lindberg, E., Eysn, L., Hollaus, M., Holmgren, J., & Pfeifer, N. (2014). Delineation of tree crowns and tree species classification from full-waveform airborne laser scanning data using 3-d ellipsoidal clustering. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7(7), 3174-3181.

[5] Widlowski, J.-L., Verstraete, M., Pinty, B., & Gobron, N. *Allometric relationships of selected european tree species*; EUR 20855 EN; Office for Official Publications of the European Communities: Ispra, Italy, 2003; p 61.

Using LiDAR data to model forest structure for canopy-associated wildlife species in old-growth forest

Joan C. Hagar[1], Bianca N.I. Eskelson[2], Patricia Haggerty[1], S. Kim Nelson[3], David G. Vesely[4]

[1] *U.S. Geological Survey - Forest & Range Ecosystem Science Center; joan_hagar@usgs.gov*
[2] *University of British Columbia*
[3] *Oregon Cooperative Wildlife Research Unit, Oregon State University*
[4] *Oregon Wildlife Institute*

Highlights: We used airborne LiDAR data to model habitat for three threatened wildlife species associated with structurally complex, old-growth forests. LiDAR metrics describing 3-dimensional structure provided a new way to quantify canopy habitat with variables that are ecologically relevant to species highly influenced by canopy structure.

Key words: *biodiversity, forest canopy, habitat modelling, LiDAR, old-growth forest*

Introduction

Accurate and reliable wildlife habitat models are important for predicting species distributions in order to identify priority areas for conservation and restoration, as well as to guide survey efforts. Development of such models for species that inhabit the canopy of mature forests can be especially challenging because of the difficulties of discerning and measuring the relevant features at an appropriate spatial scale in complexly structured, 3-dimensional space. This problem is exemplified in the productive forests in the Pacific Northwestern U.S., which support both valuable timber resources and wildlife species of conservation concern that are strongly associated with the structurally complex canopy of mature and old-growth forests. Recovery and management plans for old-growth associated species such as the marbled murrelet (*Brachyramphus marmoratus*), northern spotted owl (*Strix occidentalis caurina*), and red tree vole (*Arborimus longicaudus*), identify the need for refined measures of habitat structure, to more accurately estimate the availability of suitable habitat. Until recently, acquiring data that characterize habitat for species that are responsive to forest structure has not been feasible because traditional methods of accurately quantifying habitat have been limited by the inaccessibility of the canopy of trees that can be more than 70 m tall. However, LiDAR (Light Detection And Ranging) is an emerging remote-sensing tool that can provide fine-scale data describing vertical complexity of vegetation, offering a promising application for improving characterization of forest structure at a sub-meter scale of resolution to inform habitat models for landscape scale applications.

The goals of our study were to test the utility of LiDAR derived variables to accurately identify suitable habitat for species associated with the canopy of old-growth forest, and to provide a predictive tool for developing better occupancy maps for threatened and endangered species to guide habitat management. Our specific objectives were to 1) identify LiDAR-derived variables that were most strongly associated with occupancy at multiple spatial scales; 2) assess the usefulness of LiDAR in refining understanding of habitat relationships for each species by comparing models incorporating LiDAR-derived variables to models using traditionally-derived variables; and 3) develop predictive maps of probability of occurrence for each of 3 species of conservation concern within a management district (1,700 km^2) using LiDAR-derived data.

Materials and Methods

The study was conducted in western Oregon, USA, on forest lands managed by the Bureau of Land Management (BLM) under the Northwest Forest Plan [1], which established extensive late-successional forest reserves across the study area to protect habitat for native forest-associated species. Wildlife occurrence data used to develop habitat models were derived from a BLM database of survey records. Survey sites were not randomly selected because protocols require that all potential habitat be surveyed prior to tree harvesting or other forest management projects that are likely to affect sensitive species. Thus, all the sites used in our analyses were considered potential habitat that also were suitable for timber harvest.

Discrete return LiDAR data were collected from an airborne platform. Average pulse density was 8.1/m^2 and up to 4 returns/pulse were recorded. Raw LiDAR point files (LAS format) were processed to produce a set of metrics describing forest cover, forest height, and topographic characteristics for the entire project area. LiDAR metrics were produced using FUSION software [2].

We evaluated the performance of LiDAR-derived variables against those derived from field measurements and other remote-sensing methods that have been used to develop the current best available habitat models. Current habitat models for spotted owls, marbled murrelets, and red tree voles include variables from modeled relationships between data from field plots and spatial predictor variables derived from Landsat satellite imagery, climate variables and topographic variables (derived from the 2006 Gradient Nearest Neighbor imputation model created for the Northwest Forest Plan effectiveness monitoring [3] and referred to as "GNN variables"). We compared the explanatory power of field-measured and GNN variables to selected LiDAR-derived variables describing gradients believed to be important to old-growth forest species, including those related to tree size and structural complexity (e.g., Table 1). We fit logistic regression models developed with each group of variables (LiDAR-derived and non-LiDAR variables) and assessed model performance with AUC, leave-one-out cross-validation, overall accuracy, sensitivity, specificity, and Cohen's k.

Table 1. Comparison of models developed with 1) lidar-derived variables, and 2) other variables derived from field plot and other remote sensing methods (GNN/Landtrendr model) to predict occurrence of an old-growth associated species, the marbled murrelet (*Brachyramphus marmoratus*), in western Oregon, USA.

Model	Variable Description	AUC	Cohen's k
1. LiDAR model		*0.74*	*0.24*
ALLCVABVMN_max	cover in upper portion of canopy		
EL_p99_max	maximum height of tallest trees		
EL_p10_max	maximum height of bottom of canopy		
FRSTCVABVMD_std	variation in cover in upper canopy		
EL_kurt_min	distribution of vegetation across canopy height intervals		
Distance-to-coast (km)	Distance to the Pacific Ocean from stand center		
2. GNN/Landtrendr model		*0.64*	*0.12*
SLOPE_PCT	Percent slope		
PCTMATURE_50	Percent area classified as closed canopy, large to giant conifer		
CANCOV_CON	Percent canopy cover of all conifers		

Results and Discussion

Our work showed that incorporating LiDAR-derived variables into models can more accurately represent habitat for wildlife species, especially those highly influenced by canopy structure such as the three species we studied. We found that models with LiDAR-derived variables improved prediction accuracy of habitat occupancy for canopy-associated wildlife species over models using variables derived from traditional remote sensing. Our final model for stand-level occupancy by marbled murrelet included 5 LiDAR-derived variables describing canopy structure (Table 1). With an area under the curve value (AUC) of 0.74, this model had acceptable discrimination and fair agreement (Cohen's k = 0.24), especially considering that all sites in our sample were regarded by managers as potential habitat. The LiDAR model provided better discrimination between occupied and unoccupied sites (Figure 1) than did a model using variables derived from other remotely sensed data that were previously reported as important predictors of murrelet occupancy. We also found two LiDAR metrics that accurately discriminated murrelet nest sites from random sites (average AUC=0.91). The LiDAR-derived variables selected in our models described dimensions of canopy structure that reflect age-related stand characteristics and are consistent with the nesting ecology of the species we examined.

We believe that LiDAR variables had better discriminatory power than the previously used 2-dimensional, categorical variables because of the ability to quantify 3-dimensional gradients in canopy complexity with continuous metrics. The ability of LiDAR to provide continuous, rather than categorical, metrics describing vegetation structure allows for more realistic representation of the ecological gradients that influence species' distributions. For example, kurtosis of height distribution, a LiDAR-derived variable describing vertical stratification of canopy vegetation, was important in our model of murrelet habitat. The LiDAR-derived kurtosis metric provides a directly measured index of canopy layering as an alternative to a variable (number of canopy layers) that had previously been defined categorically and estimated subjectively. Evaluated in total, the variables identified by our models as important correlates of occupancy describe ecologically meaningful, 3-dimensional features of canopy structure that have not been easily or accurately quantified by other mensuration methods.

None of the variables that entered our models allowed perfect discrimination of occupied and unoccupied stands. However, when evaluating model performance, it is important to consider that all stands in our sample were considered potential habitat. In spite of the difficulty in discriminating among occupancy classes related to a limited range of variation in stand characteristics, our models nonetheless were able to differentiate habitat

based on fine-scale structural attributes. These results indicate that the detailed descriptions of habitat features at relevant spatial scales available through LiDAR can provide superior discriminatory power among habitats that may appear similar from a human perspective. By providing new ways to quantify three-dimensional forest structure, LiDAR expands the lexicon that can be used to describe and understand the physical features that influence habitat selection and use, thereby increasing our capacity to understand habitat relationships.

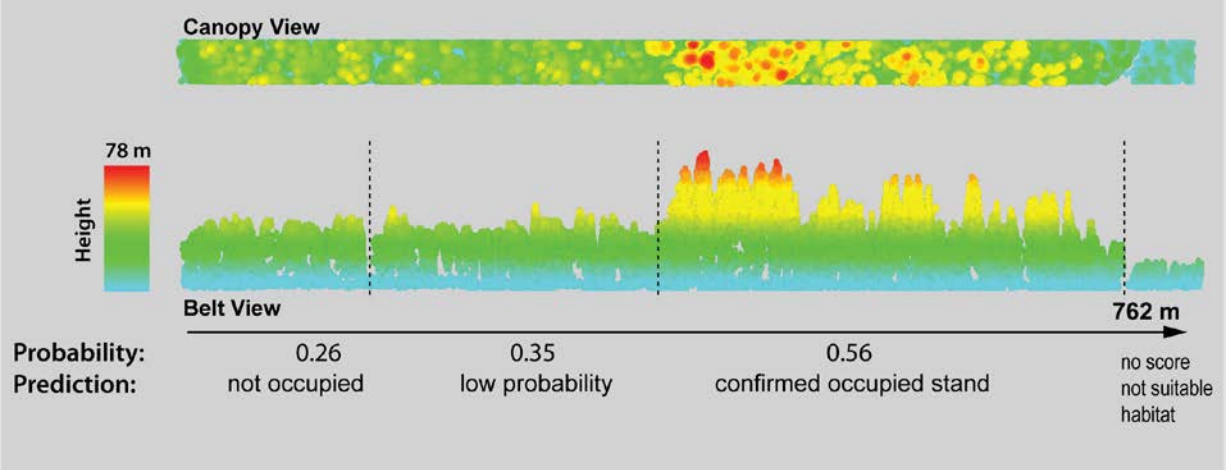


Figure 1: Comparison of structural profiles derived from airborne LiDAR data for example forest stands representing 3 occupancy classes and unsuitable habitat for the marbled murrelet (*Brachyramphus marmoratus*) in western Oregon, USA.

Conclusion

LiDAR can substantially improve the accuracy and precision of characterizations of habitat relationships for wildlife species known to be responsive to forest structure by providing fine-scale vegetation data that are inaccessible by both ground plot and other remote sensing methods. Our work demonstrates the advantages of using LiDAR over other remote sensing data for estimating availability of habitat for canopy-associated species in structurally complex forests. Our models can be used to identify sites that have high probability of occupancy by murrelets (Figure 1), northern spotted owls, or red tree voles, based on the fine-scale canopy characteristics to which these species are known to respond.

References

[1] U.S. Department of Agriculture [USDA], and U.S. Department of Interior [USDI]. 1994. Record of decision for amendments to Forest Service and Bureau of Land Management planning documents within the range of the northern spotted owl: standards and guidelines for management of habitat for late-successional and old-growth forest related species within the range of the northern spotted owl. U.S. Department of Agriculture [USDA], and U.S. Department of Interior [USDI], Washington, D.C., USA.

[2] McGaughey, R. J. 2009. FUSION/LDV: software for LIDAR data analysis and visualization; Version 2.8. Available from http://www.fs.fed.us/eng/ rsac/fusion/. Accessed 15 July 2015.

[3] Moeur, M., J. L. Ohmann, R. E. Kennedy, W. B. Cohen, M. J. Gregory, Z. Yang, H. M. Roberts, T. A. Spies, and M. Fiorella. 2011. Northwest Forest Plan–the first 15 years (1994–2008): status and trends of late-successional and old-growth forests. U.S. Department of Agriculture, Forest Service General Technical Report PNW-GTR- 853, Pacific Northwest Research Station, Portland, Oregon, USA.

Remote sensing maps of post-fire forest structure facilitate the modeling and mapping of Lewis's woodpecker nesting habitat

Jody C. Vogeler[1], Zhiqiang Yang[1], Warren B. Cohen[1, 2]

[1] *Department of Forest Ecosystems and Society, 321 Richardson Hall, Oregon State University, Corvallis, OR, USA, 97331*
[2] *USDA Forest Service, PNW Research Station, 3200 NW Jefferson Way, Corvallis, OR, USA, 97331*

Highlights: Our study found promise in the utility of lidar and Landsat time series modeled snag and shrub products in conjunction with additional lidar structure and topographic metrics for modeling and mapping nesting habitat for a species of conservation concern, the Lewis's woodpecker, across a post-fire landscape in central Oregon.

Key words: *Lewis's woodpecker, post-fire, nesting habitat, lidar, predictive mapping.*

Introduction

As remote sensing mapping products representing important habitat components increase in availability, detail, and accuracy, new opportunities arise to map habitat for forest species of conservation concern, providing spatially explicit management and conservation resources. One such species yet to be represented in validated fine-scale habitat maps is the Lewis's woodpecker (*Melanerpes lewis*), which is a species of conservation interest in multiple U.S. states and areas of Canada with habitat loss and degradation proposed as contributing factors in the species decline [1]. Previous Lewis's woodpecker breeding season habitat studies, including the published Habitat Suitability Model (HSM) [2], have identified the importance of a woody shrub component along with moderate canopy cover and a density of large snags [3], although these relationships have not been calibrated with nesting locations and mapped across a post-fire landscape. In the context of source-sink meta-populations, the high reproductive success of post-fire forests have led them to be considered as population sources for the Lewis's woodpecker [4], thus representing critical habitat for this species of concern.

Studies have begun to explore the use of airborne lidar occasionally fused with satellite remote sensing data to map important yet difficult to characterize habitat components [5]. Among these are the fine-scale distributions of snags of varying size thresholds and woody shrubs [5]. The incorporation of such products into habitat mapping efforts are in the early stages and have yet to be validated with on the ground Lewis's woodpecker nesting data or explored in a post-fire landscape. The purpose of this study was to incorporate primary (directly derived) and secondary (calibrated with ground data) remote sensing products in the fine-scale modeling and predictive mapping of Lewis's woodpecker nesting habitat in a post-fire landscape.

Material and Methods

We focused our study on a 9-year post-fire complex in the East Cascades of Oregon. We conducted surveys and monitored nests for Lewis's woodpeckers in the breeding season of 2012. Following fledge/fail, we collected nest tree and cavity data and recorded snag locations with a Trimble GeoExplorer GPS unit, also sampling non-use plots for comparison within modeling efforts.

We included primary and secondary remote sensing metrics in our evaluation of Lewis's woodpecker nesting habitat selection. Primary metrics were directly derived from lidar data acquired in 2009 and 2010 by Watershed Sciences using a Leica ALS50 Phase II and a ALS60 sensor with a pulse rate of > 105 kHz and an average pulse density of 8.6/m2 processed at a 10m grid scale in FUSION. We used previously created maps of the probability of occurrence of snags ≥ 50 cm dbh, the size threshold containing all but 3 of our Lewis's nests, and a map of mean shrub cover, also at the 10 m grid scale, created using the fusion of lidar structure and topography data with time series Landsat products [6]. Field data stratified by plant association group and fire severity were used to calibrate the models and map the relationships across the lidar-fire overlap area.

A 50 m radius buffer was created around the nests and non-use snags to evaluate relationships between Lewis's woodpecker nest patch selection and a variety of previously noted habitat features of importance at this scale [3]. We summed the number of pixels within the nest or non-use patch that were predicted to contain a snag with dbh ≥ 50 cm and divided by the total number of pixels to create a large snag density metric. After initial exploratory analyses including testing for high collinearity and evaluating variable interactions, we included the following metrics in the statistical analyses: mean canopy cover (COVER%); standard deviation of canopy cover (COVER_SD); standard deviation of the lidar-derived 10m mean height raster to represent the variability in heights across the patch (HEIGHT_SD); degree slope (SLOPE); mean shrub cover (SHRUB%);

and density of snags ≥50 cm dbh (SNAGS). We evaluated 20 candidate models using generalized linear models. For all competing models following AIC model selection, we plotted Receiver-Operating-Curves (ROC) and calculated Area-Under-the-Curve (AUC) values, model and parameter Akaike weights. Ultimately the competing model with high overall predictive capability while minimizing prediction errors was selected as the top model to then apply to the landscape to map suitable Lewis's woodpecker nest patches (Figure 1).

Results

We found models of primary and secondary remote sensing metrics as good predictors of Lewis's woodpecker nest patch selection. Several of the competing models (model with the lowest AIC value and those with delta AIC ≤2) had comparable AUC values, so we chose the model that minimized the false positive and false negative prediction errors as the top model with a percent correctly classified rate of 83% and containing the following habitat metrics: COVER%, COVER_SD, SLOPE, SNAGS, SHRUB%. In single variable candidate models, COVER%, COVER_SD, and SLOPE had significant negative relationships and SNAGS had a significant positive influence on the probability of Lewis's selecting a nest patch. SNAGS also exhibited the highest parameter Akaike weight.

Of the 26 active Lewis's woodpecker nests that we located, all succeeded. We often observed Lewis's modifying existing cavities although it was difficult to determine exactly how many of the nest cavities were fully created or were modified existing cavities as many of the nests were located after initiation. All nests were in broken top snags. Most often nests occurred within the top quarter of the snag (84%), with an average nest height of 17.0 m (SE = 1.2 m). The average nest snag dbh and height was 71.0 cm (SE = 3.1 cm) and 21.6 m (SE = 1.4 m), respectively.

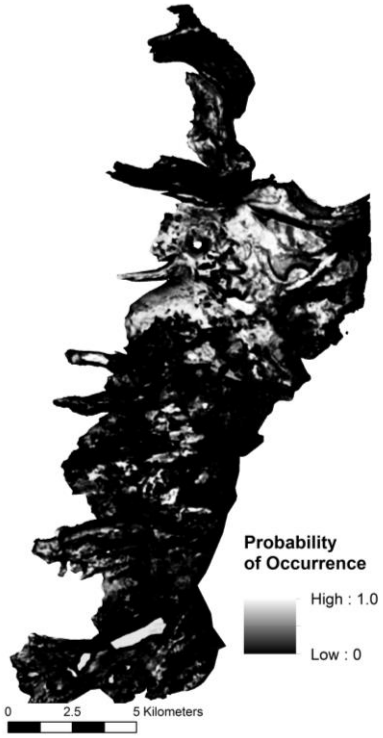


Figure 1. Predicted Lewis's woodpecker nest patch suitability within the central Oregon B&B Fire Complex.

Discussion

Through the use of primary and secondary remote sensing products, we were able to successfully model and map Lewis's woodpecker nesting habitat. To our knowledge, this represents the first time fine-scale secondary snag products have been used in the modeling and mapping of cavity-nester habitat calibrated using known nesting locations, and we found promise in this application. Of note, we have created predictive maps for nest patches for a species of concern in an identified source habitat type, a potentially valuable tool for local conservation efforts and adding to the regional knowledge of species distributions and habitat utilization.

The density of snags ≥50 cm dbh had a strong positive relationship with the probability of nest occurrence. While Lewis's woodpeckers have been found to inhabit and even prefer partially salvaged post-fire stands (Saab et al. 2007), an important management consideration is the retention of large enough snags to meet habitat needs. In our study area we observed Lewis's nests in both salvaged and unlogged units, although the majority of nests were in unlogged stands or along the edge of salvaged and retention patches. As time since fire progresses, snag fall rates increase, with fall rates lower for larger snags [7]. In our 9-yr old burn, snags have begun to fall opening up stands than may have been unfavourable for Lewis's woodpeckers in the early years post-fire, yet

retaining the preferred and longer lasting larger snags. Some salvage logging guidelines utilize a lower snag size threshold than those selected as nest trees in our study as criteria for snag retention. The removal of large snags could have negative impacts for this species of concern which has shown a preference for snags ≥50 cm dbh within several studies [3] including this one. While we have demonstrated one important application of these secondary remote sensing products, the ability to map the distribution of large snags in post-fire landscapes using remote sensing techniques may prove valuable for multiple habitat mapping, forest management, and monitoring applications.

Conclusions

While our models identified the density of snags as the strongest predictor of nesting patches with canopy cover, variability, slope, and shrub cover also included in the top model, these relationships may vary along a post-fire chronosequence and in other unburned Lewis's Woodpecker breeding season habitats. Our study demonstrated the utility of secondary remote sensing metrics, such as fine-scale snag and shrub distributions, in predicting nesting habitat for this species. Future studies should test and validate these relationships with Lewis's Woodpecker nesting locations in other forest types and explore the utility for mapping habitat relationships for additional post-fire wildlife species.

Acknowledgements

Our research was funded by the National Aeronautics and Space Administration, Carbon Cycle Program (NASA Grant 10-CARBON10-45). We thank our field technicians for assisting in data collection.

References

[1] Vierling, K. T., Saab, V. A., and Tobalske, B. W. (2013). Lewis's Woodpecker (Melanerpes lewis), The Birds of North America Online. <http://bna.birds.cornell.edu/bna/species/284>. Accessed 28 October 2014.

[2] Sousa, P. J. (1983). Habitat suitability index models: Lewis' woodpecker. U. S. Fish and Wildlife Service Publication FWS/OBS-82/10.32, Fort Collins, CO, US.

[3] Zhu, X., Srivastava, D. S., Smith, J. M. N., and Martin, K. (2012). Habitat selection and reproductive success of Lewis's Woodpecker (*Melanerpes lewis*) at its northern limit. PloS One, 7(9): e44346.

[4] Saab, V. A., and Vierling, K. T. (2001). Reproductive success of Lewis's woodpecker in burned pine and cottonwood riparian forests. The Condor, 103: 491-501.

[5] Martinuzzi, S., Vierling, L. A., Gould, W. A., Falkowski, M. J., Evans, J. S., Hudak, A. T., and Vierling, K. T. (2009). Mapping snags and understory shrubs for a LiDAR-based assessment of wildlife habitat suitability. Remote Sensing of Environment 113: 2533-2546.

[6] Vogeler, J. C., Yang, Z., and Cohen, W. B. *In Review*. Mapping post-fire habitat characteristics through the fusion of remote sensing tools. Remote Sensing of Environment.

[7] Raphael, M., Morrison, M., and Yoder-Williams, M. (1987). Breeding bird populations during twenty-five years of post-fire succession in the Sierra Nevada. Condor, 89: 614–626.

Calibration of a full-waveform, dual-wavelength terrestrial laser scanner

Zhan Li[1], Alan H. Strahler[1], David L. B. Jupp[2], Crystal B. Schaaf[1,3], Glenn Howe[4], Kuravi Hewawasam[4], Supriya Chakrabarti[4], Timothy A. Cook[4], Ian Paynter[3], Edward J. Saenz[3]

[1] *Department of Earth and Environment, Boston University, Boston MA, USA*
[2] *CSIRO Marine and Atmospheric Research, Canberra, ACT, Australia*
[3] *School for the Environment, University of Massachusetts Boston, Boston MA, USA*
[4] *Department of Physics and Applied Physics, University of Massachusetts Lowell, Lowell, MA, USA*

Highlights: A model for calibration of a full-waveform, dual-wavelength terrestrial lidar (DWEL) couples a logistic telescope efficiency function with an inverse power function to provide apparent reflectance from intensity with range. Separate models for each laser are fitted together and constrained to provide dual-wavelength spectral fidelity with a normalized difference spectral index.

Key words: *Terrestrial lidar calibration; TLS; DWEL; full-waveform; normalized difference*

Introduction

A new and important application of terrestrial laser scanning is the quantification of vegetation structure, principally measures of the physical dimensions of trees, amount and location of leaves, and gaps between and within tree canopies. While point clouds of laser returns recording the presence and location of vegetation objects can provide some elements of structure, complete inference of vegetation structure requires using the intensity of the scattered return, which in turn requires calibration. The calibration of terrestrial lidar scanners provides unique challenges, including (1) a very large variation in intensity with range that can induce saturation of the detector system by bright targets in the near field and reduced intensities that merge with the noise field in the far range; and (2) strong telescopic effects, with defocusing producing weak signals at near range. This paper summarizes how these and other challenges have been addressed for a dual-wavelength, full-waveform terrestrial scanner, the Dual-Wavelength Echida® Lidar (DWEL). The simultaneous calibration of returns from DWEL's two lasers, pulsing at wavelengths of 1064 and 1548 nm, also demonstrates how calibration can ensure both radiometric and spectral fidelity in a unified process, thus providing a pathway for calibration of other dual and multiple wavelength terrestrial lidars now in various stages of development and application [1-2].

Apparent reflectance

In its most useful form, calibration provides a signal expressed as a property of the scattering event itself, such as the *apparent reflectance* of the scattering surface. This measure, defined as the reflectance of a perfectly diffuse target orthogonal to the lidar beam and completely filling the field of view that returns the same energy from the same range as an actual target, is directly useful in retrieving vegetation properties such as leaf area index, foliage density with height, and separation of returns of leaves from those of trunks, branches and ground. For terrestrial lidars, it also provides additional value in more consistent measurements of mean tree diameter, canopy height, stem count density, and indirectly, above-ground biomass [3-4].

To further illustrate the properties of apparent reflectance, if we assume a known diffuse reflectance for the target surface (e.g., leaves) and a random orientation of target surfaces (*i.e.*, mean value of $\cos(\theta)$ of leaves at different orientations equals 0.5), the apparent reflectance is then directly related to the illuminated fraction F and the gap fraction $(1 - F)$. Alternatively, we may assume that the target fully intercepts the beam ($F = 1$); the apparent reflectance is then directly proportional to the diffuse reflectance of the target [5].

The Dual-Wavelength Echidna Lidar (DWEL)

The primary scientific objective of DWEL instrument is to separate leaves and woody materials in forests readily in three dimensional space using their different spectral reflectances. Based on the design of the Echidna® Validation Instrument (EVI) [3-4], DWEL adds a second coaxial pulsed laser and acquires full-waveform scans at both near-infrared (NIR, 1064 nm) and shortwave infrared (SWIR, 1548 nm) wavelengths with simultaneous laser pulses. At the SWIR wavelength, laser power returned from leaves is much lower than from woody materials, such as trunks and branches, due to absorption by liquid water in leaves. In contrast, returned power from leaves and woody materials is similar at the NIR wavelength. A more complete description of the DWEL is found in [6-7].

Basic processing

Before point cloud generation and radiometric calibration, the raw waveforms from DWEL are preprocessed to remove background noise; convert digitizer time to apparent range by aligning each waveform to the peak of the outgoing pulse using the signals from a scattering wire or internal Spectralon panel [7]; and correct laser power drift, typically due to instrument temperature change, by scaling recorded intensities according to mean intensities observed from the internal Spectralon panel for each mirror rotation. Further processing removes strong, near-range pulses that saturate the instrument's digitizers and replaces them with unsaturated equivalents. Point clouds of scattering events are then constructed from the preprocessed waveforms. A waveform deconstruction process identifies individual hits found in the waveform and associates each with its peak intensity, creating a cloud of points representing hits and intensities at coordinates determined from range, zenith, and azimuth angles recorded by the instrument.

Calibration model

The purpose of our calibration is to convert the intensity associated with each pulse peak to apparent reflectance, a measure that removes the effects of telescope efficiency and fall-off with range. The calibration model is:

$$\rho_{app}(R_j) = I_j \cdot \frac{R_j^b}{C_0 \cdot K(R_j)} \quad , \quad K(r) = \frac{1}{\left(1 + C_1 \cdot e^{-C_2 \cdot r}\right)^{C_3}}$$

where $\rho_{app}(R_j), I_j,$ and R_j are the apparent reflectance, intensity, and range of the *j*th peak; C_0 is a constant scaling intensity to apparent reflectance; $C_1, C_2,$ and C_3 are constants of the telescope efficiency function $K(r)$ with range *r;* and *b* is the exponent of range. The efficiency function has a logistic shape rising from 0 to 1; parameters C_1 and C_3 together determine the range at which the function approaches its asymptote of one, while C_2 controls the rate at which telescope efficiency rises from 0 to 1 in the near range. For DWEL, the functions rise rapidly from 0 m range to 0.5 at about 3-4 m (NIR, SWIR) and reach 1 asymptotically at about 10-15 m.

To provide data to fit the model, we scanned three panels of different reflectance values from a nearly perpendicular direction at 33 range locations from 0.5 m to 70 m. The three panels included a white Spectralon panel and two foam boards painted with flat interior wall paint in light and dark gray tones. We used the manufacturer's specification for the Spectralon panel and determined the reflectances of the gray panels by the average ratio of each gray panel to the Spectralon panel at all ranges. All panels were sufficiently large to intercept the whole laser beam at 70 m.

To fit the model, we estimated the calibration parameters of the two wavelengths together in a joint calibration model using an objective error function that minimized relative errors in ρ_{app} from both wavelengths and a spectral constraint that minimized the differences in a normalized difference index from both wavelengths ($NDI = [(\rho_{1064} - \rho_{1548}) / (\rho_{1064} + \rho_{1548})]$). The constraint optimized the calibration for vegetation study and also improved the stability of parameter fitting. To assess the goodness of fit of the model, we randomly divided the intensities (about 24,000 samples for each range) into a training set (80 percent) and a validation set (20 percent). In the training set, return intensities were normalized by the corresponding panel reflectance to provide equivalent target reflectances of 1.0 and then averaged together for each range to reduce noise in the data. Table 1 documents the fit of the model to both training and validation datasets; R^2 values are high and RMSE values are quite low. More definitive studies are needed to confirm these results. It should be noted, however, that no instrument is perfectly stable and recalibration will most likely produce slightly different values.

Table 1: Assessment of calibration fitting and validation.

Wavelength	1064 nm		1548 nm	
	Training	Validation	Training	Validation
Adjusted R^2, measured vs. modeled intensity	0.954	0.948	0.983	0.964
RMSE, apparent reflectance	0.108	0.081	0.092	0.064

Example of apparent reflectance processing

Figure 1 provides two images of a scan of a deciduous hardwood forest stand in the northeastern United States, acquired in September, 2014. The site, located within the Harvard Forest, Petersham, Massachusetts, is a 1-ha plot dominated by red maple (*Acer rubrum*), red oak (*Quercus rubra*), and white birch (*Betula papyrifera*) with an understory of these species accompanied by American beech (*Fagus grandifolia*), American chestnut (*Castanea dentata*) and others. Several large white pines (*Pinus strobus*) also occur in the stand. The images are color composites, with 1548 nm values in red, 1064 nm values in green, and blue as dark constant. Values displayed are the sums of intensity or apparent reflectance returns, averaged for all waveforms in the bins of an equal-angle projection. In the upper image, intensity clearly decreases with range and colors are inconsistent. In the lower image, trunks are visible at far range and colors are more uniform.

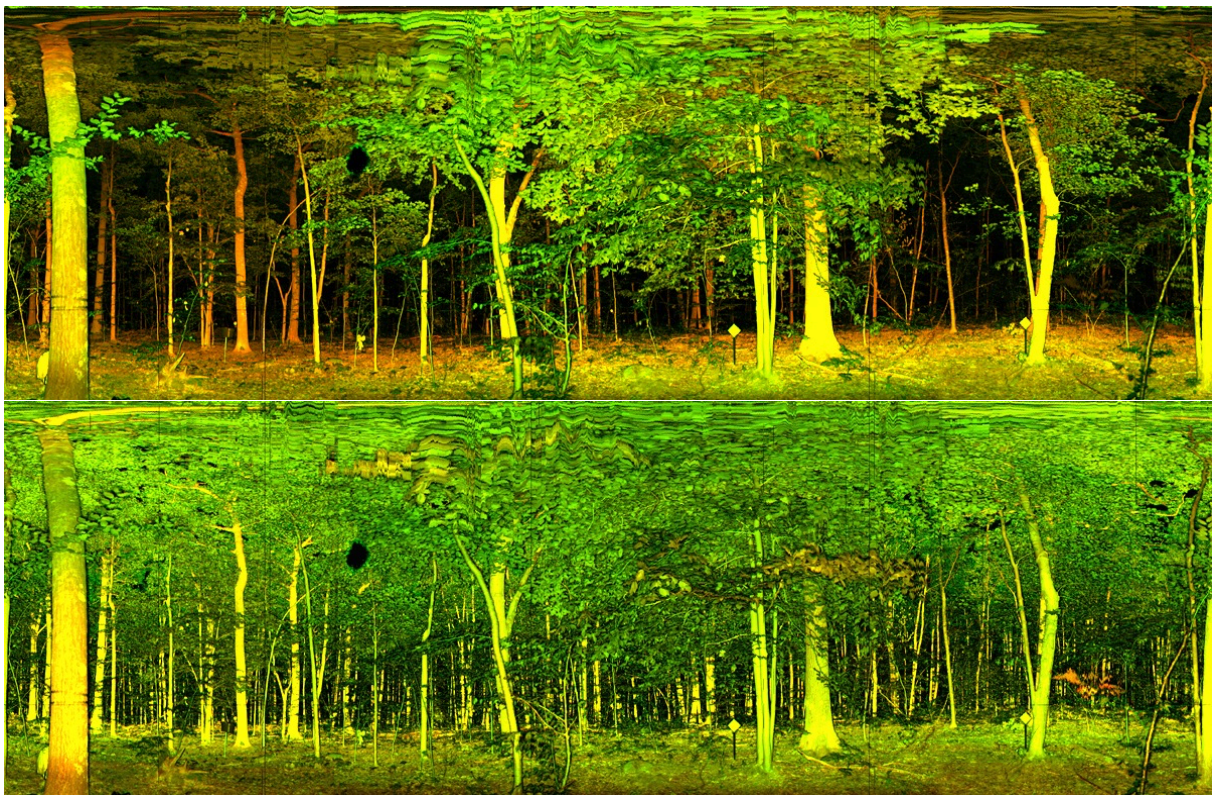


Figure 1: An example of intensity values (upper image) of a dual-wavelength scan calibrated to apparent reflectance (lower image). Equal-angle panoramic images: azimuth angle, abscissa; zenith angle, ordinate.

Conclusion

In summary, our work addresses the challenges posed in calibrating multiple-wavelength, full-waveform terrestrial laser scanners by formulating a flexible calibration model, acquiring appropriate calibration data, fitting the model with a constraint providing spectral consistency, and testing the results. The next step is to use calibrated data to retrieve forest structural parameters with the new dual-wavelength data of the DWEL. This is the subject of additional papers now in preparation. This work was supported by the US National Science Foundation, under grant number MRI-0923389.

References

[1] Danson, F. M., R. Gaulton, R. P. Armitage, M. Disney, O. Gunawan, P. Lewis, G. Pearson, & F. A. Ramirez. (2014). Developing a dual-wavelength full-waveform terrestrial laser scanner to characterize forest canopy structure. Agricultural and Forest Meteorology, 198-199, 7–14.

[2] Hakala, T., J. Suomalainen, S. Kaasalainen, & Y. Chen. (2012). Full waveform hyperspectral lidar for terrestrial laser scanning. Optics Express, 20(7), 7119-7127.

[3] Yao, T., X. Yang, F. Zhao, Z. Wang, Q. Zhang, D. L. B. Jupp, D. Culvenor, G. J. Newnham, W. Ni-Meister, C. Schaaf, C. Woodcock, J. Wang, X. Li, & A. H. Strahler. (2011). Measuring forest structure and biomass in New England forest stands using Echidna ground-based lidar. Remote Sensing of Environment, 115(11), 2965-2974.

[4] Yang, X., Strahler, A. H., C. B. Schaaf, D. L. B. Jupp, T. Yao, F. Zhao,Z. Wang, D. Culvenor, G. J. Newnham, J. L. Lovell, R. O. Dubayah, C. E. Woodcock, & W. Ni-Meister. (2013). Three-dimensional forest reconstruction and structural parameter retrievals using a terrestrial full-waveform lidar instrument (Echidna®). Remote Sensing of Environment, 135, 36-51.

[5] Jupp, D. L. B., D. S. Culvenor, J. L. Lovell, G. J. Newnham, A. H. Strahler, & C. E. Woodcock. (2009). Estimating forest LAI profiles and structural parameters using a ground-based laser called 'Echidna®.' Tree Physiology, 29(2), 171–181.

[6] Douglas, E. S., A. Strahler, J. Martel, T. Cook, C. Mendillo, R. Marshall, S. Chakrabarti, C. Schaaf, C. Woodcock, Z. Li, X. Yang, D. Culvenor, D. L. B. Jupp, G. J. Newnham, & J. Lovell. (2102) DWEL: A dual-wavelength Echidna lidar for ground-based forest scanning. Proceedings IEEE International Geoscience and Remote Sensing Symposium (IGARSS) 2012, 4998-5001.

[7] Douglas, E. S., J. Martel, Z. Li, G. Howe, K. Hewawasam, R. A. Marshall, C. L. Schaaf, T. A. Cook, G. J. Newham, A. Strahler, & S. Chakrabarti. (2015). Finding leaves in the forest: The Dual-Wavelength Echidna Lidar. IEEE Geoscience and Remote Sensing Letters, 12(4), 776-780.

Improving estimates of leaf area index with a dual-wavelength full-waveform TLS: The Salford Advanced Laser Canopy Analyser (SALCA)

Author names: Lucy A. Schofield[1], F. Mark Danson[1], Neil S. Entwistle[1], Eric Casella[2], Mat I. Disney[3], Rachel Gaulton[4]

[1]*Ecosystems and Environment Research Centre, School of Environment and Life Sciences University of Salford, Salford, M5 4WT, UK*

[2]*Centre for Sustainable Forestry and Climate Change, Forest Research Agency, Alice Holt Lodge, Wrecclesham, Surrey GU10 4LH, UK*

[3]*Dept. of Geography, University College London, Gower Street, London WC1E 6BT, UK*

[4]*School of Civil Engineering & Geosciences, Newcastle University, Newcastle upon Tyne, NE1 7RU, UK*

Highlights: Calibrated laser reflectance in two wavelengths was used to extract returns resulting from leaf hits and corrected for amount of material in each footprint and clumping of material. True LAI was calculated using an inversion of gap fractions, validated for individual trees, and analysed for multi-temporal forest plots.

Key words: *TLS, LAI, phenology, dual-wavelength, calibration, SALCA*

Introduction

In the past decade, Light Detection And Ranging (LiDAR) systems have proven their ability to collect three-dimensional (3D) datasets with high detail and accuracy. Applications for forest structure measurements and associated biophysical parameters have been demonstrated from both airbourne (ALS) and ground-based (TLS) laser scanning systems. However, challenges still remain that limit the full characterisation of tree structures from these sensors. One of these is the inability to directly distinguish returns resulting from woody and leaf material. As a result, important ecological parameters such as leaf area index (LAI) and foliage profiles are often calculated from the entire forest point cloud leading instead to the generation of plant area index (PAI) and whole plant profiles. The accurate separation of tree leaf and wood returns would have far-reaching benefits including increasing accuracy in forest structural measurements, providing a validation tool for other indirect approaches, and in reducing uncertainty in hydrological, ecological, and climate modelling. Carbon allocation among tree organs is a topic of significant ecological interest due to the differing functioning roles of wood as slow decomposable carbon pools with low metabolic activity, and leaves as fast decomposable dynamic carbon pools which control moisture, gas exchange, and collect radiation. The relationship between processes (respiration and production) and Gross Primary Production (GPP) differs between the tree components and therefore their potential effect on carbon sequestration in times of rising carbon dioxide. It has been found that total biomass alone is not a good predictor of carbon flux in forests [1]. Therefore rather than assuming whole plant carbon use, it has been recommended that estimates are made by component. Forests are also sensitive indicators of climate change *via* vegetation phenology. Improved measurements of the seasonal accumulation and loss of leaf material would provide a significant resource with which to examine how species and ecosystems have responded to temperature variations and how they may respond to future climate change [2].

Although the physiological mechanisms involved in leaf and woody material growth are relatively well understood, directly measuring how much of each pool is present at one time, and its spatial and temporal characteristics, poses a significant challenge in structurally complex heterogeneous forest environments. Recent development of innovative multi-spectral sensors could provide a solution. Multi-spectral TLS instruments measure their surroundings using different wavelengths of radiation exploiting the spectral properties of forest target features. The aim of this paper is to offer an improved approach to estimating leaf area index with dual-wavelength full-waveform TLS.

The Salford Advanced Laser Canopy Analyser

The Salford Advanced Laser Canopy Analyser (SALCA) [3] is a dual-wavelength full-waveform TLS. The sensor was developed by the University of Salford and Halo Photonics Ltd as an experimental research

instrument to measure forest canopies using a hemispherical scanning pattern. Two lasers are fired for every one 'shot' at wavelengths 1545nm and 1063nm and the full backscattered signal from both is recorded. Information is extracted from the return signal by decomposing each full waveform into a series of individual echoes using custom built algorithms.

An improved method of LAI estimation from TLS

The ability to generate apparent reflectance from a laser scanning system is essential for inferring information on target properties. In forest environments, this relates to data of increased ecological value and is the first step to distinguishing leaves from woody material, as well as parameters relating to the health of vegetation. For the SALCA instrument, a novel approach has been applied to perform a radiometric calibration using artificial neural networks to convert recorded intensity into reflectance with a RMSE of 6% for both wavelengths. Based on the reflective properties of leaf and woody components in the wavelengths inherent to SALCA, and demonstrated in [3] and [4], taking a reflectance ratio of the wavelengths at each point should allow separation of these materials. Returns were matched between wavelengths and the SALCA Normalised Ratio Index (SNRI) calculated according to (equation 1):

$$SNRI = \frac{(\rho_{1063} - \rho_{1545})}{(\rho_{1063} - \rho_{1545})} \quad (1)$$

where ρ indicates reflectance. A threshold was then applied to separate the point cloud into two components representing leaves and wood. This allowed LAI to be calculated directly from the photosynthetically active material alone. To take into account the amount of material in each shot a weighting was applied based on the leaf reflectance of a 'full hit'. Leaf area index, the most important variable for modelling a range of ecological processes, was then derived using an inversion of the gap fraction model (equation 2):

$$LAI(z) = -2 \int_0^{\pi/2} \frac{\ln\left(P_{gap}(\theta, z)\right) cos\theta}{\Omega(\theta)} sin\theta d\theta \quad (2)$$

where LAI is the leaf area index; $P_{gap}(\theta, z)$ is the gap fraction in the direction of the zenith angle θ at height z, and; $\Omega(\theta)$ is the canopy element clumping index. The clumping index measures the spatial aggregation of foliage elements, and was calculated by applying nearest neighbour point pattern analysis to each leaf point cloud in a 3D domain. Incorporating the clumping index corrects for the non-randomness in the 3D distribution of the foliage that often leads to underestimation of LAI.

From a single tree to multi-temporal forest plots

This approach was first tested on SALCA scans of three individual oak trees. Accuracy was assessed using destructive sampling of foliage carried out at the Forest Research UK site at Alice Holt, Farnham, in July 2014 as part of a collaboration initiative with University College London (UCL), University of Newcastle, and University of Salford.

Secondly, the spatial and temporal characteristics of LAI was examined for a range of common UK forest plots. Data was acquired over 32 field visits to Delamere Forest, Cheshire, UK, over a full annual phenological cycle between March 2014 and April 2015, and included broadleaf deciduous, evergreen conifer, and deciduous conifer species. Coincident digital hemispherical photographs were also acquired. This paper reports the first results of this experiment demonstrating accurate, fully validated, estimates of true LAI at different forest stands.

References

[1] Litton, C.M., Raich, J.W. & Ryan, M.G. (2007). Carbon allocation in forest ecosystems. Global Change Biology, 13, 2089-2109.

[2] IPCC, (2007). *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M, Averyt, K. B., Tignor, M. and Miller, H. L. (eds.) Cambridge University Press. Cambridge.

[3] Danson, F. M., Gaulton, R., Armitage, R. P., Disney, M., Gunawan, O., Lewis, P., Pearson, G., & Ramirez, A. F. (2014). Developing a dual-wavelength full-waveform terrestrial laser scanner to characterize forest canopy structure. Agricultural and Forest Meteorology, 198-199, 7-14.

[4] Gaulton, R., Danson, F. M., Ramirez, F. A., & Gunawan, O. (2013). The potential of dual-wavelength laser scanning for estimating vegetation moisture content. Remote Sensing of Environment, 132, 32-39.

Optimizing vegetated coastal environment observations with terrestrial remote sensing

Ian Paynter, Crystal Schaaf[1], Edward Saenz[1], Francesco Peri[1], Jen Bowen[1], Bob Chen[1], Bruce Cook[2], Lola Fatoyinbo[2], Andres Vega[3], Alan Strahler[4]

1. School for the Environment, University of Massachusetts Boston
2. NASA Goddard Space Flight Center
3.Sirena Biological Station, Costa Rica
4. Earth and Environment, Boston University

Highlights: Fine-scale, simultaneously acquired classification from terrestrial lidar and imagery may characterize the tidal and phenological components of LandSat 7 and 8 imagery. Terrestrial lidar provides a way to quantify geomorphological changes in saltmarshes in response to nitrogen fertilization. Adaptation of forest analysis techniques allows biomass models of saltmarshes and mangroves.

Key Words: Lidar, coastal, mangroves, saltmarshes, LandSat

Introduction

Coastal environments play substantial roles in geochemical cycles, and are essential for carbon sequestration, coastal armoring, and storm energy mitigation. They serve as the front line of global climate change, subject to and sensitive to potential changes in freshwater, oceanic and atmospheric chemical conditions as well changes in the frequency and intensity of storms. Therefore it is extremely important to determine, monitor and eventually model conditions and function in coastal environments globally. Remote sensing has provided the ability to observe general ecosystem conditions and change at meaningful spatial and temporal resolutions for large spatial extents. However, coastal ecosystems present considerable challenges in the use of remote sensing resources due to the high temporal dynamism resulting from the strong, regular influence of tides and vegetation phenology. This dynamism can be exacerbated by influences such inundation by extremely strong, episodic flooding or storm surge, and snow cover at some latitudes. Characterizing these sources of variation and quantifying their effect on remote sensing observations clarifies the investigation of underlying, long-term ecological changes.

Content

Through a synthesis of terrestrial and airborne remote sensing resources, validated with traditional field methods, we have developed classification, reconstruction and quantification methods for vegetated coastal ecosystems. In saltmarshes, we are establishing fine-scale classification techniques by combining terrestrial lidar observations with RGB imagery. A highly-portable, rapid-scanning, highly-resilient terrestrial lidar, the 905nm, 4mrad resolution, 40m range Compact Biomass Lidar (CBL), built around a SICK LMS151 lidar unit, was developed by the University of Massachusetts Boston in collaboration with Rochester Institute of Technology to meet the challenges of deploying terrestrial lidar in coastal environments. The CBL scans from a tripod in mangrove environments, or a mobile tram system to maximize overhead views of saltmarsh vegetation and creeks while minimizing disturbance. While mangrove classification follows traditional forestry methods, saltmarsh classification techniques utilize observable structure, reflective properties and imagery to separate water, soil and vegetation species.

Throughout the 2015 growing season, fine-scale classifications are being acquired for seven adjacent areas co-located with LandSat 7 and 8 pixels, by scanning with terrestrial lidar at a unique-coverage sampling distance of 5m within each pixel, and combining with overhead optical imagery captured from a portable rig suspending a GoPro Hero3+ camera at a height of 2.5m. The acquisition times for these data corresponds to the times of the satellite overpasses at the Plum Island Long Term Ecological Research site in Massachusetts, USA. Relating the biological and structural components of the pixels to the intensities of the various LandSat spectral bands allows characterization of the various components of the signal, especially the relative contribution of water, vegetation, and exposed soil,

depending on the tidal state. We will present validation by manual classification and sensitivity studies in terms of frame-shift of the pixels and component level identification, as well as preliminary comparisons to LandSat data in terms of relating prominent ecosystem components such as water and exposed soil to LandSat per-channel reflectance values.

These classification techniques are also being transferred to NASA Goddard's Lidar, Hyperspectral and Thermal (G-LiHT) airborne instrument package, which acquired data over Plum Island LTER during 2014. The potential to augment the lidar structural classification of water, soil and vegetation species with the hyperspectral and thermal instruments is under investigation. Future saltmarsh overpasses by G-LiHT will be accompanied by simultaneous terrestrial data acquisition, facilitating validation.

In conjunction with the ongoing TIDE (Trophic Interactions in Detrital Ecosystems) project to study the effects of nitrogen fertilization on saltmarsh ecology we employ the CBL to examine creek geomorphology. Nitrogen fertilization has been associated with loss of conserved creek bank structural form, via changes in the root morphology of creek bank vegetation [1]. Preliminary studies in 2014 suggested that analyzing lateral sections of creek banks extracted from lidar scans can provide a quantification of structure loss. Throughout 2015 we are acquiring a temporal series of scans of the complete experimental and control creeks, and investigating geomorphological differences with a spatial function-fitting method. Scans are acquired with a purpose-built tram system allowing the lidar to be deployed above the center of each creek. Scans are then co-located by manual alignment and segmented perpendicular to the predominant creek direction. Segments from each slice of the control creeks are fitted to quadratic functions and the fit of experimental creek segments to these functions is tested to describe geomorphological differences.

Species-level classification of saltmarsh vegetation also facilitates biomass estimation through a combination of individual unit reconstruction to estimate density, and volumetric analysis of polygons formed by fitting facetized meshes to lidar data. Preliminary studies at Thompson Island, Nantucket Island, Plum Island and Waquoit Bay (Massachusetts, USA) throughout 2013 and 2014 suggest these techniques are particularly effective in saltmarshes where the influence of salinity gradients results in highly conserved areas of vegetation species. Validation at a fine scale will be facilitated by traditional field measurements of vegetation at Plum Island during 2015. We present the biomass estimation methods, as well as extrapolative modeling techniques under development.

In addition to the saltmarsh deployments, the CBL was deployed alongside the collection of traditional field measurements as part of a validation campaign for NASA's EcoSAR (Synthetic Aperture Radar) instrument in mangroves in Drake's Bay, Costa Rica in 2014. This also served as an initial study of the capabilities of terrestrial lidar to provide reconstructions of mangrove vegetation for quantified analysis. Stems, branches and leaves were captured effectively by high-density scanning schemes, and reconstructed for biomass quantification utilizing an adaption of the Quantitative Structure Modeling (QSM) [2] which is typically used for terrestrial forests. We are investigating two alternative approaches for quantification of the biomass of the morphologically distinct exposed root structures of the mangrove species. We show initial results from both structural inversion, which allows the utilization of QSM to model the roots of each tree, and a volumetric voxelization analysis applied only to the root layer.

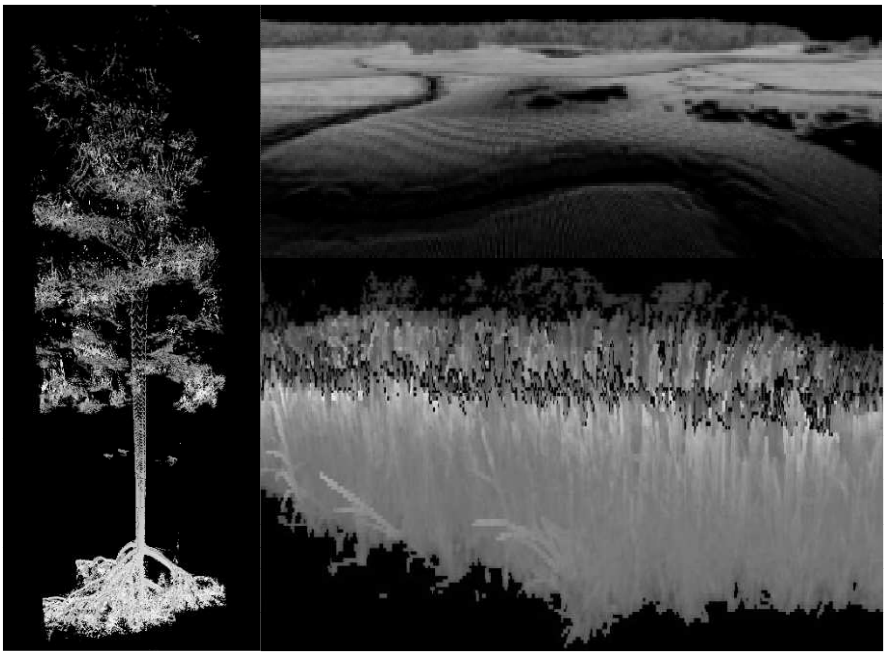


Figure 1: (Clockwise from top-left) Point cloud extracted from multiple CBL scans for a mangrove tree in Drake's Bay, Costa Rica; from G-LiHT lidar instrument for Plum Island saltmarsh, MA, USA; and reprojected from CBL scan for Thompson Island saltmarsh, MA, USA.

References

[1] Deegan, L. A., Johnson, D. S., Warren, R. S., Peterson, B. J., Fleeger, J. W., Fagherazzi, S., & Wollheim, W. M. (2012). Coastal eutrophication as a driver of salt marsh loss. Nature, 490(7420), 388-392.

[2] Raumonen, P., Kaasalainen, M., Åkerblom, M., Kaasalainen, S., Kaartinen, H., Vastaranta, Holopainen, M., Disney, M., & Lewis, P. (2013). Fast automatic precision tree models from terrestrial laser scanner data. Remote Sensing, 5(2), 491-520.

A method addressing signal occlusion by scene objects to quantify the 3D distribution of forest components from terrestrial lidar

Richard A. Fournier [1], Jean-François Côté[2], Florentin Bourge[1], Sylvie Durrieu[3], Alexandre Piboule[4], and Martin Béland[5]

[1] *CARTEL, Départ. de géomatique appliquée, Université de Sherbrooke, Sherbrooke (QC) CANADA*
[2] *Natural Resources Canada, Canadian Wood Fibre Centre, Quebec City (QC), G1V 4C7, CANADA*
[3] *Irstea, UMR TETIS, Centre de Montpellier, F-34196, Montpellier, FRANCE*
[4] *Office National des Forêts (ONF), Départ. de recherche et développement, Nancy, FRANCE*
[5] *Centre de Recherche en Géomatique, Départ. de géomatique, Université Laval, Quebec City (QC) CANADA*

Highlights: Estimating exact 3D distribution of canopy components using terrestrial lidar in forest is limited by signal occlusion. We propose a method to address this limitation: it uses voxels, beam returns and beam propagation through the scene. The proposed method was validated using simulated forest scenes and a lidar simulator.

Key words: *terrestrial laser scanner, signal occlusion, forest components, 3D spatial distribution, point cloud processing.*

Introduction

Terrestrial lidar scanner (TLS) can be used in forest stands to investigate fine structural attributes at the stand- and tree-level. Visual inspection of the point cloud from a single scan taken by a TLS shows many areas with no points due to signal occlusion. The impact of signal occlusion is more important as the distance from the scanner increases. It also increases with stand structural complexity. One way to reduce the occluded areas in the point cloud of a forest plot is to produce a co-registered (or aligned) point cloud from several scans at different locations inside and around the plot. This way to proceed reduces but does not eliminate occluded areas. However merging multiple scans introduces oversampling: i.e. the possibility that one component may be probed from several directions. Oversampling is also likely to occur in a single scan for targets close from the scanner for which laser beams may overlap. Overall, signal occlusion and oversampling is present in all point clouds of forest plots taken by TLS and it alters the estimation of an exact 3D spatial distribution of the forest components if this estimation is based on point density. This situation is so prevalent that one can assume that the number of lidar returns is biased compared with true component distribution: it is proportionally higher for objects closer to the sensor and conversely lower with increasing distance from the sensor. This results in a misrepresentation of the number of returns compared with the real distribution of the forest components. For instance the vertical profile drawn from TLS data artificially inflates the number of returns from lower canopy and does not provide a proper account of the components from the upper part of the canopy for which the laser beam is strongly occluded. A similar effect exists with airborne lidar data [1]. At this point no correction method exists for this bias of representation of component distribution.

Enhanced forest inventory are expected from new methods of processing TLS data designed to estimate forest attributes [2]. For example, assessing stem shape (DBH and stem taper) is not affected by point oversampling from multiple scans, if the scans are well aligned, but it may be affected by signal occlusion. Several methods proposed replicating detailed tree structure from raw TLS data ([3]). These methods are promising but they require extensive computing to provide a 3D spatially-explicit structural representation of the forest. A simpler method to process the TLS point cloud is required to help estimating the exact 3D distribution of forest components. Voxel representation may be a way to reduce the complexity of point clouds while retaining a sufficient level of details for further analyses. The representation of the point cloud into voxels can allow estimation of many structural attributes such as LAI [4] or metrics describing asymmetry [5] or linked to the quality of wood fiber attributes [6]. The main objective of this study is to develop and validate a method to process the TLS point cloud taken in forest stands, using voxel representation, to allow estimating reliable 3D forest component distribution, consequently removing the impact of signal occlusion and oversampling.

Materials and Methods

The study objective was met by following four methodological steps: (1) simulating detailed forest structure for a range of documented conditions, (2) simulating the point cloud resulting from multiple TLS scans acquisition for the simulated forest scenes, (3) developing a two-step method to deal with signal occlusion and oversampling, and (4) estimating error of the results from the two-step method. These efforts were therefore concentrated towards the development of a method to estimate the component distributions in 3D using the Plant Area Density (PAD in m^2/m^3) in a voxel-based environment.

Simulation of forest canopy: It is virtually impossible to have direct in situ measurements of forest component distribution in 3D to serve as a reference. Consequently we decided to simulate realistic forest scenes with the fine-scale architectural model L-Architect [3]. We reproduced 12 realistic forest scenes over a range of documented conditions typical of boreal forests of Eastern Canada largely dominated by black spruce (Picea mariana) and balsam fir (Abies balsamea) [3].

Simulation of the TLS point cloud: All the simulated forest scenes were *"*scanned*"* by a TLS-simulator based on a ray-tracing algorithm and in similar conditions typically performed during field acquisitions. The TLS-simulator replicated a hemispherical geometry and technical characteristics similar to the Faro© and the Zoller+Fröhlich© sensors in high resolution mode, which recorded only the first return. All plots were scanned from 5 positions, including one at the plot centre and the 4 others placed at an equal distance near the plot edge. The resulting five scans of each plot were aligned, clipped on the plot dimension and assembled.

A two-step method to address signal occlusion and oversampling: The datasets were processed and analysed for voxel sizes of 5, 10, 15 and 20 cm edges, respectively, to assess if voxel resolution had an impact on the quality of the results. The method was composed of two distinct steps. The first step provided an estimate of the PAD within each voxel according to the number of lidar beams crossing the voxel and the number of returns. This step combined two procedures: procedure 1 used the L-Vox algorithm that estimates the theoretical number of lidar beams crossing each voxel and calculates a density index using the number of returns ([7], [3]), and procedure 2 processed the number of incident lidar beams and their path in the voxel following the method suggested by [4] to estimate the PAD for every voxel. These procedures both required defining a minimum number of lidar beams entering the voxel. After several preliminary tests, we set the threshold at ten. This allowed processing voxels with moderate beam occlusion, but not voxels severely occluded. The second step of the method involved assigning missing value for voxels severely occluded with at least one return but less than 10 lidar beams crossing them. All these voxels were assigned with the average PAD value of the corresponding voxels located within a 1-m slice in height (z).

Error estimation: The simulated scenes with L-Architect provided an exact account considered as the reference values of PAD and LAD (Leaf Area Density). It was therefore possible to use that information to compare the reference vertical profiles from the ones of the simulated scenes produced from the TLS-simulated datasets. Differences between vertical profiles quantified the impact of beam occlusion and object oversampling. Then, the reference PAD values of all scenes could be compared with the results from each step of the two-steps method. Lastly, the two-step method was adapted to process only returns associated with foliage. This allowed comparing how the two-step method could also be used to produce realistic spatial LAI distribution of a forest scene. Overall the error estimation quantified the error of the two-step method to process TLS point clouds for its ability to provide exact distribution of forest components, expressed in PAD and LAI.

Results and Discussion

Preliminary results suggest that the two-step method corrects the raw TLS data for the impact of beam occlusion and object oversampling by at least 95%. When vegetation density was assessed directly from point frequencies of raw TLS point clouds, most profiles showed important differences with the reference and the normalized profiles of the simulated scenes. For example, the peak of point density for the TLS-simulated data occurs mostly at 4 to 5 m below the one from the simulated scenes. This concurs with the general hypothesis that TLS returns tend to overestimate the quantity of component area on the lower part of the canopy and understorey (closer to the sensor) and underestimate this quantity in the upper part of the canopy [1] (Fig 1). Applying the first step of the method already improved greatly the normalized vertical profiles compared from the raw TLS data by repositioning the vertical maximum close to the one provided by the reference. Similarly, applying the first step of the method also improved greatly the general shape of the vertical profile compared to the raw TLS data with the reference values (Fig 1). Both the position of the maximum and the parametrized shape of the vertical profile are being improved to a 90% match or better for four varied scenes we tested so far. Furthermore, applying the second step of the method improved even more the results with a match reaching and beyond 95% for the tested scenes (Fig 1). These encouraging results will be further tested on twelve scenes and documented to assess the limitations of the methods. However, the preliminary results over four diversified scenes indicated that the two-step method is a good way to preprocess raw TLS data of forest scenes to recover the exact distribution of components through the PAD and LAD. We plan to implement these procedures in Computree (http://computree.onf.fr), an open source web platform to process TLS data. The two-step method will therefore be potentially useful to deal with signal occlusion and oversampling, but it will need to be adapted to specific sensors and for other forest scene configurations yet not tested.

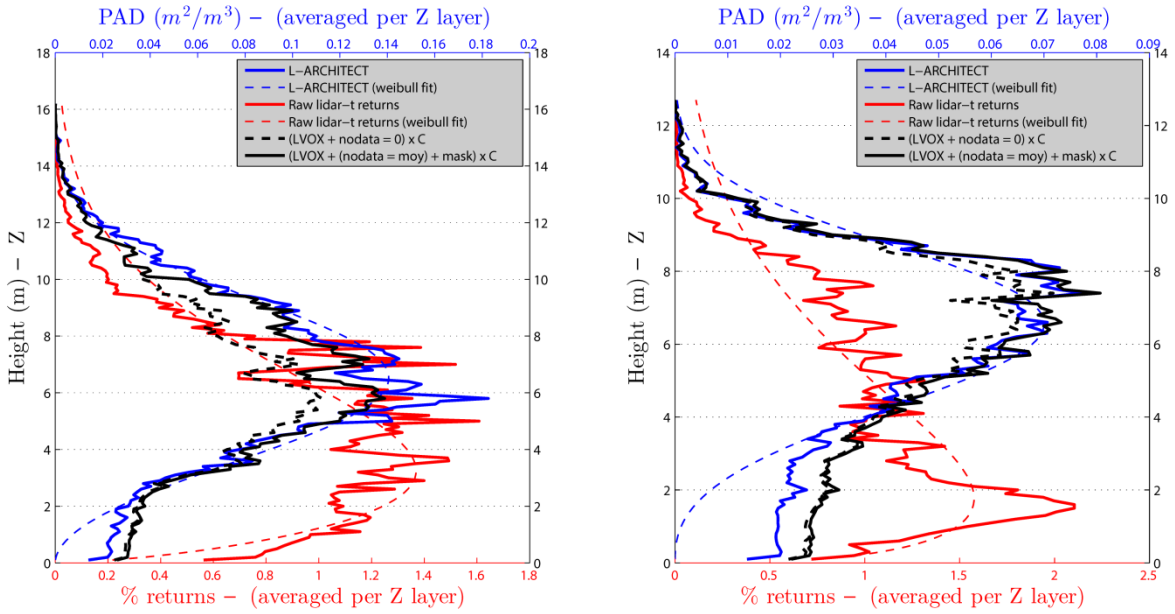


Figure 1. Four vertical profiles for two forest scenes simulated with 10 cm voxels: (blue line) the reference values in PAD from the simulated scenes with L-Architect, (red line) the raw TLS-simulated data normalized (density values) of the forest scenes, (dotted black line) the normalized values (density values) resulting from the first step of the method and (solid black line) the normalized values (% returns) from the two steps of the method.

References

[1] Chasmer, L., Hopkinson, C., Treitz, P. (2006) Investigating laser pulse penetration through a conifer canopy by integrating airborne and terrestrial lidar. Cdn J. Rem. Sens., 32: 116-125.

[2] Dassot, M., Constant, T., and Fournier, M. (2011) The use of terrestrial LiDAR technology in forest science: application fields, benefits and challenges. Ann. For. Sc., 68(5): 959-974.

[3] Côté, J.-F., R.A. Fournier and R. Égli (2011) An architectural model of trees to estimate forest structural attributes using terrestrial LiDAR. Envir. Model. Soft., 26: 761-777.

[4] Béland, M., Widlowski, J.-L., Fournier, R.A., Côté, J.-F., Verstraete, M.M. (2011) Estimating leaf area distribution in savanna trees from terrestrial LiDAR measurements. Agric. For. Meteorol., 151: 1252-1266.

[5] Seidel, D., Leuschner, C., Müller, A., Krause, B. (2011) Crown plasticity in mixed forests: Quantifying asymmetry as a measure of competition using terrestrial laser scanning. For. Ecol. Manag., 261: 2123-2132.

[6] Blanchette, D., Fournier, R.A., Luther, J.E. and Côté, J.-F. (2015) Predicting wood fibre attributes using local scale metrics from terrestrial LiDAR data: a case study over the conifer stands of Newfoundland. For. Ecol. Manag.. 347: 116-129.

[7] Durrieu, S., Allouis, T., Fournier, R.A., Véga, C., Albrech, L. (2008) Spatial quantification of vegetation density from terrestrial laser scanner data for characterization of 3D forest structure at plot level. In SilviLaser 2008, Edinburgh, UK, 17-19 September.

Quantitative structure tree models from terrestrial laser scanner data

Pasi Raumonen

Department of Mathematics, Tampere University of Technology, Finland, pasi.raumonen@tut.fi

Highlights: Accurate quantitative structure models (QSM) of trees can be reconstructed from terrestrial laser scanner measurements (TLS). QMSs are fast to reconstruct and are very useful for interpreting the TLS measurements and therefore have lots of applications, such as non-destructive above-ground biomass estimation and change quantification and tree structure metrics.

Key words: *TLS, QSM, tree model, tree extraction, biomass estimation, buttress roots*

Introduction

Terrestrial laser scanners (TLS) can provide detailed non-destructive measurements of the tree structure. Computational methods are required to transform the (xyz) point cloud into useful and compact information. One approach is to reconstruct a quantitative structure model (QSM) of the tree from the data [1-2]. A QSM is a hierarchical collection of geometric primitives or building blocks, which give the local geometric information of the tree structure such as diameters, lengths, and directions (see Figure 1). The simplest and the most often used and robust choice of the building block is the circular cylinder, but other shapes such as cones and elliptic cylinders can be also used [3]. QSMs also contain the topological branching structure information, i.e., what is the branching order of a given branch and which branch is its parent branch. Thus from a QMS lots of useful tree information can be computed such as total, stem, branch, 1st-order branch, etc. volumes and lengths, including dbh and tree height. Also volumes, lengths, taper and branching angles for individual branches can be computed. Notice that the building block approach results in discontinuous surface models. However, using continuous surface models would make the model reconstruction harder without contributing anything substantial to the quantitative structure information.

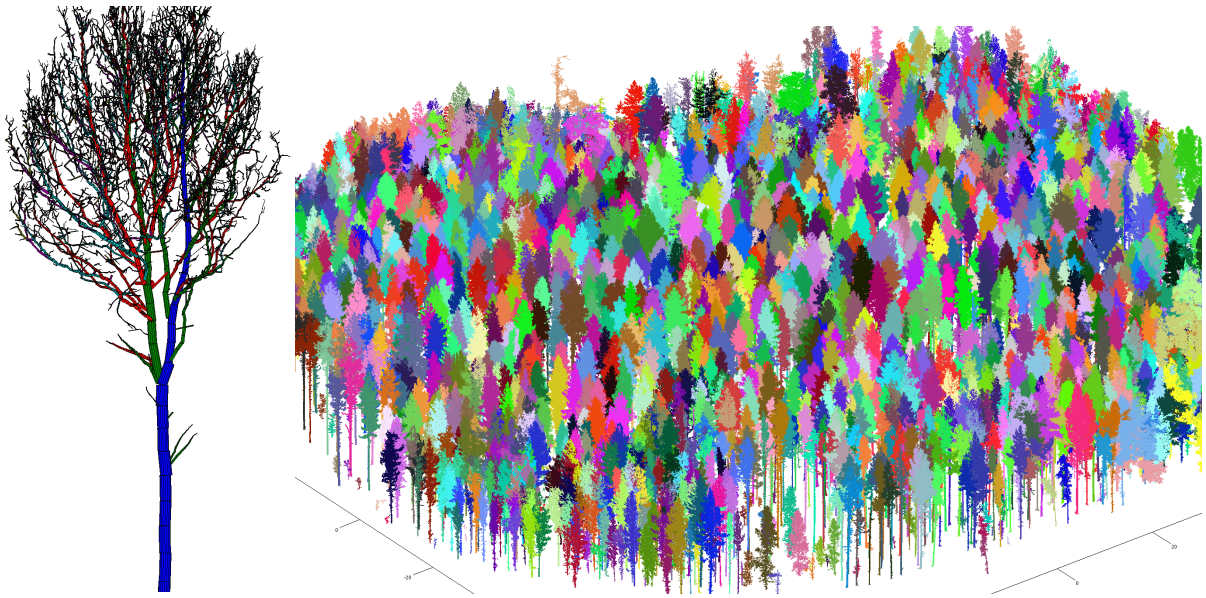


Figure 1: A QSM of an oak tree located in Southern England (left) and QSMs of a Scots pine forest located in Punkaharju, Finland.

Reconstruction of QSMs from TLS data

Details of the QSM reconstruction from TLS data are presented in [1-3]. In a nutshell, first the point cloud is segmented into its individual branches and then the QSM is reconstructed by fitting geometric primitives to the segments. For the segmentation process the point cloud is partitioned into small sets that correspond to small surface patches in the tree surface. These patches are natural surface elements, which allow fast and accurate segmentation based on surface growing and local connectivity of the neighbouring patches. The point cloud may also contain measurements from ground and understory, and these are automatically removed. The whole reconstruction process is very general in the sense that no species-specific assumptions are used.

The duration of the reconstruction process of a QSM from a point cloud with millions of points from a single tree is usually few minutes with a laptop computer. User needs to specify few input parameters defining the patch diameters and cylinder lengths. As a rule of thumb, the patch size, which can vary along the tree according to automatically approximated local branch size, needs to be smaller than or comparable to the details to be modelled. Therefore if the sizes of patches are small enough, the results should be quite robust and not change much if the parameter values are changed a little. On the other hand, very small patch sizes, particularly with 1 or 2 scan positions and low scan resolution, can result in large underestimation of the volume.

Applications

Above-ground biomass

Given basic wood density, it is possible to convert the volumes from QSMs to biomass. Thus with TLS and QSMs it is possible to measure the above ground tree biomass accurately, non-destructively, cheaply and fast as our experience with Australian Eucalyptus plots shows [2]. With the eucalyptus the biomass was also estimated much more accurately than with species-specific allometry [2].

Automatic tree extraction and massive scale tree modelling

The segmentation process used in the QSM reconstruction can be also used for automatic tree extraction from point clouds [4]. We have shown that in a day you can extract and reconstruct QSMs for thousands of trees in multi-hectare plots (see Figure 1) with a laptop computer. The number of scanner positions does not need to be high as a few tens of scans should be enough to cover multi-hectare areas, assuming the forest is not very dense.

Tree structure metrics and distributions

The geometrical and topological structure information of QSMs can be used to compute e.g. the branch size distribution. This distribution tells how much volume there is in tree parts within given diameter class. This kind of metrics and distributions can give new insight for tree growth and structure. For example, we found that the branch size distribution has similar shape for trees of different height in similar growing conditions [5]. We have also used metrics from different branch orders for estimation of plant scaling laws for metabolism [6].

Change detection and quantification

Non-destructiveness of the TLS measurement allows repeated measurements of the whole tree during multi-year periods or a growth season and therefore follow-up of growth and change of trees. The change in the biomass, branch length, number of branches, etc. is then estimated from QSMs reconstructed from the repeated measurements [7]. Not only growth, but also dying and storm and other damage can be detected.

Below-ground biomass and structure

Stump-root systems can be uncovered from ground and then measured with TLS. From the resulting point clouds we can reconstruct hybrid QSMs, where the complex stump part is modelled with a triangulation and the roots with cylinders [8]. The hybrid QSM of a stump-root system contains the same geometrical and topological information as the cylindrical tree QSMs.

Buttress roots

Big tropical trees can have large buttress roots with complicate shapes, which are not possible to model accurately with simple shapes such as cylinders. Triangulations fitted to point clouds can estimate the volume and surface area of these roots accurately. The reconstruction of triangulations is robust as it uses a priori information about the data and shape, allowing realistic interpolation over large gaps in the measurement cover. A hybrid QSM, combining the triangulated buttress roots and cylinder model of the upper parts of the stem and branches, can be made to model the whole tree accurately.

Functional structural tree models

Functional structural plant models (FSPM) aim to replicate the geometrical and topological structure of the plant and the growth dynamics using some biological, physical, etc. rules. The plant structure and particularly also its random variability, i.e., distributions, need to be measured for developing realistic FSPMs. The measurements should be also non-destructive to allow follow-up of the development of an individual plant. With TLS and QSMs we can, for the first time for FSPMs of large trees, do this kind of measurements in large-scale, fast, and cheaply [9].

Tree response to wind forcing

The detailed structure information of QSMs can be used to simulate the response of the tree parts to given wind forcing. The simulations can predict wind damage on tree plantations and natural forests. Conversely, one can simulate the wind resistance caused by the trees and used this information for example for optimisation of wind turbine locations in forests.

Forest fires and fuel structure modelling

Structure information of the QSMs provides ways to estimate and model fuel distribution in the forest for forest fires [10]. The fuel distribution information helps the management of forests to effectively prevent or minimise fires in forests close to residential areas.

QSMs as supports for eco-physiological data

QSMs can be combined with or used as realistic structural support for eco-physiological data such as leafs, chlorophyll and water content. Multi- or hyper-spectral TLS measurement may be needed to measure these eco-physiological data.

Determination of tree phenotype

A tree species can have a large number of structural phenotypes. Determination of the phenotypes may require accurate and comprehensive measurements of large number of trees to recognize possibly quite subtle differences. TLS and QSMs could offer a useful tool for structural phenotype determination.

Acknowledgement

Dr Eric Casella from UK Forest Research (oak tree) and Dr Raisa Mäkipää from Natural Resources Institute Finland (pine forest) provided the TLS data used for models shown in Figure 1.

References

[1] Raumonen, P., Kaasalainen, M., Åkerblom, M., Kaasalainen, S., Kaartinen, H., Vastaranta, M., Holopainen, M., Disney, M., & Lewis, P. (2013). Fast Automatic Precision Tree Models from Terrestrial Laser Scanner Data. Remote Sensing, 5, 491-520.

[2] Calders, K., Newnham, G., Burt, A., Murphy, S., Raumonen, P., Herold, M., Culvenor, D., Avitabile, V., Disney, M., Armston, J., & Kaasalainen, M. (2015). Non-destructive estimates of above-ground biomass using terrestrial laser scanning. Methods in Ecology and Evolution, 6, 198-208.

[3] Åkerblom, M., Raumonen, P., Kaasalainen, M., & Casella, E. (2015). Analysis of geometric primitives in quantitative structure models of tree stems. Remote Sensing, *7*, 4581-4603.

[4] Raumonen, P., Casella, E., Calders, K., Murphy, S., Åkerblom, M., & Kaasalainen, M. (2015). Massive-scale Tree Modelling from TLS Data. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume II-3/W4, 189-196.

[5] Krooks, A., Kaasalainen, S., Kankare, V., Joensuu, M., Raumonen, P., & Kaasalainen, M. (2014). Predicting tree structure from tree height using terrestrial laser scanning and quantitative structure models. Silva Fennica, 48, Article Id 1125.

[6] Lau Sarmiento, A. I. (2014). Using T-LiDAR as an alternative measurement technique for plant-scaling modeling in tropical forest. MSc thesis, Wageningen University and Research Centre.

[7] Kaasalainen, S., Krooks, A., Liski, J., Raumonen, P., Kaartinen, H., Kaasalainen, M., Puttonen, E., Anttila, K., & Mäkipää, R. (2014). Change Detection of Tree Biomass with Terrestrial Laser Scanning and Quantitative Structure Modelling. Remote Sensing, 6, 3906-3922.

[8] Smith, A., Astrup, R., Raumonen, P., Liski, J., Krooks, A., Kaasalainen, S., Åkerblom, M., Kaasalainen, M. (2014). Root system characterization and volume estimation by terrestrial laser scanning. Forests, 5, 3274-3294.

[9] Kaasalainen, M., Potapov, I., Raumonen, P., Åkerblom, M., Sievänen, R., Kaasalainen, S. (2013). Bayes trees and forests: combining precise empirical and theoretical tree models. Proc. of FSPM 2013, Saariselkä, Finland, 9-14 Jun. 2013, pp. 61-63.

[10] Hom, J., Pang, Y., Zhao, D., Xu, G., Krooks, A., Raumonen, P., Kaasalainen, M., Dandois, J., Clark, K., Skowronski, N., Patterson, M., Gallagher M. (2013). Characterization of fuel structure and estimations of biomass using 3D terrestrial laser scanning. Proc. of Fourth Fire Behavior and Fuels, Raleigh, North Carolina, USA, February 18 -22, 2013.

Three-phase model-based estimation of growing stock volume utilizing Landsat, LiDAR and field data in large-scale surveys

Svetlana Saarela[1], Anton Grafström[2], Göran Ståhl[2]

[1]*Department of Forest Sciences, University of Helsinki, FINLNAD*
[2]*Department of Forest Resource Management, Swedish University of Agricultural Sciences, SWEDEN*

Highlights: The study presents a new model-based estimation approach for forest resources assessments combining field sample data and full cover Landsat data and sampled LiDAR data as auxiliary information. The results are of importance for developing forest inventory methodologies under REDD+ requirements.

Keywords: *LiDAR, Landsat, large-scale forest inventory, model-based, three-phase sampling.*

Introduction

Over the previous decades the interest in large-scale forest inventories utilizing several sources of data has increased considerably. In case design-based samples are selected, the model-assisted estimation framework (Särndal *et al.* 1992) allows combining several sources of auxiliary information through, e.g., multi-stage or multi-phase surveys. In this case model relationships between auxiliary and target variables are used to improve the precision of the estimators.

Model-based inference is an alternative mode of inference that can be applied in case auxiliary data are available. Described briefly, model-based inference relies more heavily on the correctness of the model(s) applied in the estimators. While the dependence on the model is a drawback, this mode of inference also has advantages over design-based approaches; e.g., in some cases smaller sample sizes might be applied for reaching a certain level of accuracy. For a detailed presentation and discussion about design-based and model-based inference in forest surveys, see Gregoire (1998).

While several sources of auxiliary information straightforwardly can be applied in the case of model-assisted estimation, following established sampling theory, (e.g., Saarela *et al.* 2015a) this issue is less explored for model-based inference for the case when the different auxiliary variables are not available wall-to-wall. For example, recent studies by Ståhl *et al.* (2011) and Corona *et al.* (2014) suggest how probability samples of auxiliary data can be incorporated in connection with model-based inference. This approach was termed "hybrid inference" by Corona *et al.* (2014) to clarify that auxiliary data were collected according to design-based sampling.

In this study we present how the model-based estimation technique can be applied in a three-phase sampling framework utilizing Landsat full cover auxiliary data, LiDAR sample data and field data. New estimators for the population mean and its variance were developed, and analysed through Monte Carlo simulation. The precision of the estimator was compared to the precision of an estimator utilizing LiDAR full cover or Landsat full cover, only.

Material and methods

An artificial population was applied in this study; it was created using a multivariate probability distribution copula technique (Ene *et al.* 2012). The artificial population resembles the Kuortane study area, western Finland, and consists of Landsat spectral values of green (B20), red (B30) and shortwave infra-red (B50) bands, LiDAR metrics maximum height (h_{max}), the height of the 80^{th} percentile of height distribution (h_{80}), canopy relief ratio (CRR) and the percentage of first returns above 2 m (p_{veg}) as a crown cover estimate, and growing stock volume values. Details of the population can be found in Saarela *et al.* (2015b). Figure 1 provides an overview of the study area.

Our objective was to estimate the population mean μ of the grid cell volumes in the study area. Our first phase "sample" is a wall-to-wall coverage of Landsat-based auxiliary information. The second phase sample was obtained from *M* grid cells. Each second phase sample has two sets of auxiliary data: the Landsat- and the LiDAR-based. The third phase sample comprises *m* grid cells, subsampled from the second phase sample. For these grid cells, growing stock volume was assessed. Different estimators were compared in the study:

Case A: three-phase model-based estimation, where Landsat data were collected in the first sample phase LiDAR data in the second phase, and field data about growing stock volume in the third phase.

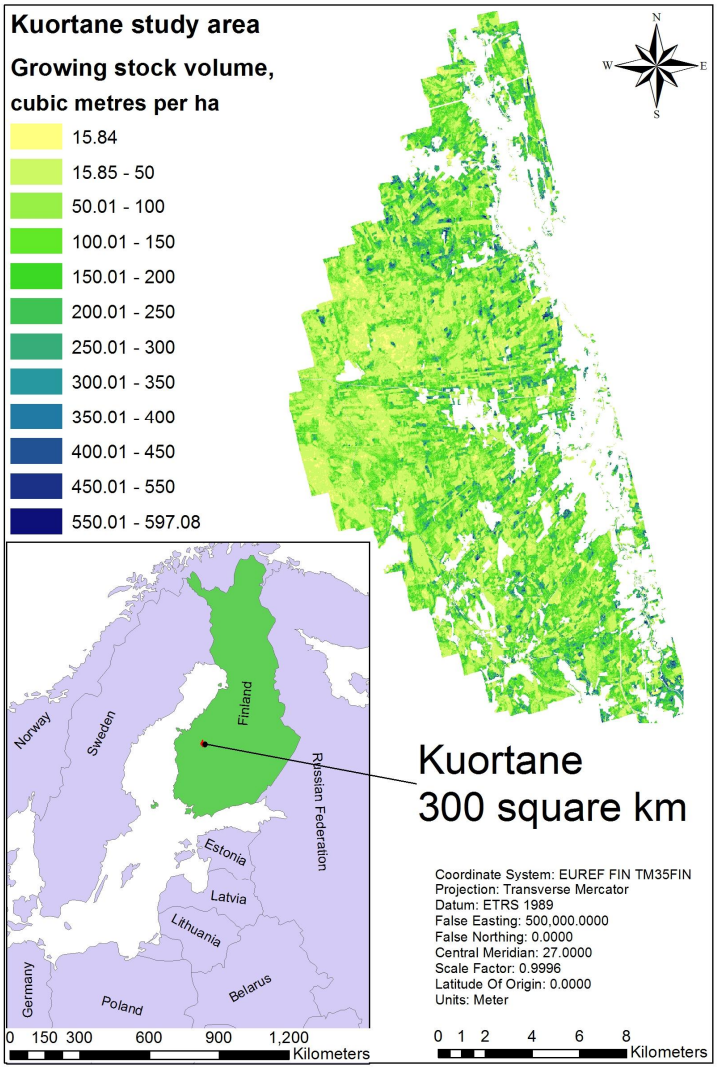


Figure 1: The Kuortane study area.

Case B: two-phase model-based estimation with Landsat data and field data.

Case C: this case has the same sampling design as Case B, but with LiDAR auxiliary information available instead of Landsat.

We used ordinary least square regression for developing the needed models, assuming that our response variables were independent and normally distributed with homogeneous variance.

For each case, Monte Carlo sampling simulation with 1000 repetitions was performed. We used simple random sampling without replacement for all cases and phases of sampling. Based on the outcomes of the simulations, the variance of an estimator was approximated through empirical variance as

$$V(\hat{\mu})_{emp} = \frac{1}{1000-1}\sum_{i=1}^{1000}(\hat{\mu}_i - \bar{\hat{\mu}})^2$$

where $\hat{\mu}_i$ is the estimated population mean after i^{th} iteration, and $\bar{\hat{\mu}}$ is the average of estimated population means over all iterations. This variance was compared with the average of the estimated variances.

Results

In Table 1 the averages of estimated population means $\bar{\hat{\mu}}$ and its variance $\bar{\hat{V}}(\hat{\mu})$, and empirical variances $V(\hat{\mu})_{emp}$ for each case are presented.

The population size was *N*=818016 grid cells, the second phase sample was *M*=1000 grid cells, and the third phase sample was *m*=100 grid cells.

Table 1: The averages of estimated population mean and its variance, and empirical variance.

Case	$\bar{\hat{\mu}}$, m^3ha^{-1}	$\bar{\hat{V}}(\hat{\mu})$, $(m^3ha^{-1})^2$	$V(\hat{\mu})_{emp}$,$(m^3ha^{-1})^2$
A	101.217	14.173	13.575
B	100.925	64.478	66.106
C	101.173	8.598	8.792

Discussion

Non-surprisingly, Case C with LiDAR auxiliary information available wall-to-wall led to the highest precision. LiDAR metrics are highly correlated with forest attributes such as growing stock volume and biomass (Næsset, 2002). Case B, utilizing full cover of Landsat data, resulted in the lowest precision of the population mean estimation. Case A shows how different sources of auxiliary information can be utilized in the model-based estimation framework. The main advantage of the approach is that it doesn't require a probability sample of field data; hence purposive sampling can be applied, which leads to decreased costs.

The method can be used as a base for developing technologies for large-scale surveys utilizing several sources of auxiliary information where probability sampling cannot be established due to the forest areas remote location. The method is one step forward to decreasing costs on surveys maintaining or even improving given precision.

References

[1] Corona, P., Fattorini, L., Franceschi, S., Scrinzi, G., & Torresan, C. (2014). Estimation of standing wood volume in forest compartments by exploiting airborne laser scanning information: model-based, design-based, and hybrid perspectives. Canadian Journal of Forest Research, 44(11), 1303-1311.

[2] Ene, L. T., Næsset, E., Gobakken, T., Gregoire, T. G., Ståhl, G., & Nelson, R. (2012). Assessing the accuracy of regional LiDAR-based biomass estimation using a simulation approach. Remote Sensing of Environment, 123, 579-592.

[3] Gregoire, T. G. (1998). Design-based and model-based inference in survey sampling: appreciating the difference. Canadian Journal of Forest Research, 28(10), 1429-1447.

[4] 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.

[5] Saarela, S., Grafström, A., Ståhl, G., Kangas, A., Holopainen, M., Tuominen, S., Nordqvist, K., & Hyyppä, J. (2015a). Model-assisted estimation of growing stock volume using different combinations of LiDAR and Landsat data as auxiliary information. Remote Sensing of Environment, 158, 431-440.

[6] Saarela, S., Schnell, S., Grafström, A., Tuominen, S., Nordkvist, K., Hyyppä, J., Kangas, A., & Ståhl, G. (2015b). Effects of sample size and model form on the accuracy of model-based estimators of growing stock volume in Kuortane, Finland. Canadian Journal of Forest Research. (In press).

[7] Ståhl, G., Holm, S., Gregoire, T. G., Gobakken, T., Næsset, E., & Nelson, R. (2011). Model-based inference for biomass estimation in a LiDAR sample survey in Hedmark County, Norway This article is one of a selection of papers from Extending Forest Inventory and Monitoring over Space and Time. Canadian Journal of Forest Research, 41(1), 96-107.

[8] Särndal, C. E., Swensson, B., & Wretman, J.(1992). Model Assisted Survey Sampling (1th ed.).

Model-based Inference of Aboveground Biomass over a California Conifer Forest

Qi Chen

Department of Geography, University of Hawaii at Manoa, Honolulu, HI, 96822

Highlights: This study addresses the impacts of errors of tree-level biomass estimates and plot locations on model-based biomass estimation. It also explores the integration of ALOS PRISM DSM and lidar DTM for biomass estimation.

Key words: *model-based inference, lidar, ALOS PRISM, tree AGB errors, plot location errors*

Introduction

Model-based inference has become increasingly important in recent years for estimating forest attributes, especially over remote areas where accessibility is limited and field data collection based on rigorous sampling design is difficult. However, several prominent issues remain to be thoroughly investigated for broader application of model-based methods for biomass estimation. First, most of the current methods assume that the tree aboveground biomass (AGB) estimates derived from field measurements (e.g. DBH and tree height) and allometric models are free of errors. In reality, both field measurements and allometric methods have errors that can contribute to the errors of the estimation of tree AGB. Incorporating errors of individual tree AGB estimates into model-based estimation is a recent practice (Chen et al., 2015; McRoberts and Westfall, 2015) and thus more research is needed along this line. Second, there is a pressing need to combine the national forest inventory (NFI) plots with remotely sensed data for estimate forest AGB over large areas. However, most NFI plots in many countries were located with recreational GPS. We need to assess how the errors in plot locations affect the AGB model constructed from the field data and the corresponding AGB estimation over the whole study area. Third, the recent availability of high spatial resolution satellite Digital Surface Model (DSM) constructed from stereo imagery (e.g., ALOS PRISM), when combined with a high quality Digital Terrain Model (DTM), provides a promising opportunity to map canopy height over large areas. The use of the derived canopy height from the integration of such products for accurate biomass estimation is yet to be assessed.

Objectives

Using airborne lidar data and ALOS PRSIM DSM over a conifer forest in California, this study attempts to address the following issues. First, impacts of the errors of individual tree AGB estimates on model-based estimation will be investigated. Second, this study will compare the use of (1) airborne lidar height metrics and (2) the canopy height calculated as the difference between ALOS PRISM DSM and lidar DTM for AGB estimation and assess their uncertainty. Third, the effects of using recreational instead of survey-grade GPS in model-based estimation are assessed. Lastly, model-based estimations will be compared with model-assisted and probability-based estimation of AGB.

Materials and Methods

Study Area

Our study area is located in the Sagehen Experimental Forest in California, which covers approximately 3,925 ha and is on the eastern slope of the Sierra Nevada approximately 32 km north of Lake Tahoe (Fig. 1). Conifer species present include white fir (*Abies concolor*), red fir (*Abies magnifica*), mountain hemlock (*Tsuga mertensiana*), lodgepole pine (*Pinus contorta*), Jeffrey pine (*Pinus jeffreyi*), sugar pine (*Pinus lambertiana*), and western white pine (*Pinus monticola*). Non-forested areas include fens, wet and dry montane meadows and shrub fields. Elevation ranges from 1862 m to 2670 m with slopes averaging 18% but can reach 70% in parts of the watershed.

Field data

A total of 523 circular plots of 0.05 ha was installed with a random starting location over the study area in 2004-2006. These field plots were located with handheld Garmin eTrex recreational GPS with horizontal accuracy of 3 to 11 meters. In 2006, a Trimble® GeoXH™ handheld GPS with Zephyr Geodetic antenna was used to re-measure the center of 81 individual plots. The average horizontal accuracy of the new GPS measurements is 0.1 m with the majority <0.2 m and, at the worse case, 1.5 m. At each plot, all trees greater than

5 cm in diameter at breast height (DBH, breast height = 1.37m) were measured with a nested sampling design. Canopy trees (≥19.5 cm DBH) were tagged and measured in the whole plot; Understory trees (≥5 cm DBH to <19.5 cm DBH) were measured in a randomly selected third of the plot. Tree measurements include species, DBH, tree height, canopy base height, vigor, and crown position.

Remotely sensed data

Lidar data were collected from September 14 to 17, 2005 for the study area using an Optech ALTM 2050 system on an airplane flying at an altitude of ~800 m and average velocity of 260 km per hour. The ALTM 2050 acquired up to three returns per pulse at a pulse frequency of 50 kHz, scan frequency of 38 Hz, and a maximum scan angle of 15°, creating a swath width of ~580 m. The point density is about 2-4 returns per square meters. Optech, Inc. rates the RMSE precision of individual point locations surveyed by the ALTM 2050 as ±15 cm vertical and ±50 cm horizontal. ALOS PRISM DSM of 5 m resolution was acquired.

Methods

Airborne lidar data were processed using the Tiffs (Toolbox for Lidar Data Filtering and Forest Studies) software (Chen, 2007) to filter the raw lidar points and separate them into ground and non-ground returns, to interpolate the ground returns into a DTM, and to extract canopy height metrics from the lidar points within individual forest plots. ALOS PRISM DSM will be compared with a lidar-derived DSM. Correspondingly, the derived canopy height based on ALOS PRISM DSM and lidar DTM is compared with the canopy height derived from lidar alone. Errors in tree measurements and allometric methods are simulated. Sensitivity analysis was done to explore the impacts of tree AGB estimate errors on biomass models based on remotely sensed data following Chen et al. (2015). AGB estimates and uncertainties were compared based on different predictors (heights from lidar alone vs. heights from the integration of ALOS and lidar) and different plot locations (recreational vs. survey grade GPS).

References

[1] Chen, Q. (2007). Airborne lidar data processing and information extraction. Photogrammetric Engineering and Remote Sensing, 73(2), 109-112.

[2] Chen, Q., Vaglio Laurin, G., Valentini, R. (2015). Uncertainty of remotely sensed aboveground biomass over an African tropical forest: Propagating errors from trees to plots to pixels. Remote Sensing of Environment, 160, 134-143.

[3] McRoberts, R. E., & Westfall, J. A. (2015). Propagating uncertainty through individual tree volume model predictions to large-area volume estimates. Annals of Forest Science, 1-9.

Using airborne laser scanning data with small pulse densities to enhance national forest inventory large area estimation

Ronald E. McRoberts[1], Qi Chen[2], Grant M. Domke[1], Daniel J. Kaisershot[1]

[1]Northern Research Station, U.S. Forest Service, Saint Paul, Minnesota, USA
[2]Department of Geography, University of Hawai'i at Mānoa, Honolulu, Hawai'i, USA

Precedents for combining inventory plot data with metrics for airborne laser scanning (ALS) data with small pulse densities for enhancing forest inventory estimates are beginning to be reported in the literature. When combining data from multiple sources, accurate co-registration to a common coordinate system is essential. The accuracy of co-registration depends to a large degree on the accuracy of plot locations acquired using global positioning system (GPS) receivers. [1] reported that the effects of plot location errors of 5 m and greater contributed to greater deviations between plot observations and ALS-based model predictions of growing stock volume. Thus, common practice for ALS studies is to use GPS receivers with sub-meter accuracies. However, for large-area national forest inventory (NFI) applications, providing a \$5000-\$10,000 survey grade GPS receiver to each of a large number of field crews would be a financial challenge. The objective of the study was to compare estimates of mean above ground biomass per unit area (AGB, Mg/ha) by combining ALS data with small pulse densities and data for FIA plots acquired using both recreation grade GPS receivers (RECR) with accuracies on the order of 5-20 m and survey grade GPS receivers (SURV) with sub-meter accuracies.

The study area was Itasca County in north central Minnesota, USA, which consists of 7583 km^2 including 5540 km^2 of forest land. ALS data with mean density of approximately 0.9 pulses/m^2 were acquired in April 2012, and standard distributional, height, and density metrics were calculated. AGB estimates were obtained for 168-m^2, circular, NFI plots measured in 2012; of these plots, 80 were on forest land and 53 were on non-forest land. The study area was tessellated into square, 169-m^2 grid cells that served as population units.

The model-assisted, generalized regression (GREG) estimators were used to enhance precision ([2], Section 6.5),

$$\hat{\mu}_{GREG} = \frac{1}{N}\sum_{i\in U}\hat{y}_i - \frac{1}{n}\sum_{i\in S}\varepsilon_i$$

and

$$\hat{Var}(\hat{\mu}_{GREG}) = \frac{1}{n(n-1)}\sum_{i\in S}(\varepsilon_i - \bar{\varepsilon})^2,$$

where U is the population, S is the sample, i indexes population units or plots, y_i is plot AGB, $\hat{y}_i$ is the corresponding ALS-assisted model AGB prediction, $\varepsilon_i = \hat{y}_i - y_i$, and $\bar{\varepsilon} = \frac{1}{n}\sum_{i\in S}\varepsilon$

Consideration was given to both linear models of the form,

$$y_i = \beta_1 \cdot x_{i1} + \cdots + \beta_p \cdot x_{ip} + \varepsilon_i,$$

and nonlinear models of the form,

$$y_i = \frac{\alpha}{1+\exp\left(\beta_0+\beta_1\cdot x_{i1}+\cdots+\beta_p\cdot x_{ip}\right)}+\varepsilon_i\,,$$

where x_{ij} is the j^{th} ALS metric, and α and the βs are parameters to be estimated. Linear models are more familiar but can produce negative and unrealistically large predictions. The nonlinear model has the potential to circumvent these prediction problems because it includes a lower asymptote at V=0 and an upper asymptote at V= αwhose estimate is based on the sample data. The linear models were fit using stepwise selection techniques, and the nonlinear models were fit using the same ALS metrics selected for the corresponding linear models.

Some plots had been harvested between the 2009 plot measurement date and the 2012 ALS data acquisition date. For the SURV analyses, four plots were removed as outliers because of disturbance between the dates. In addition, for the RECR analyses, one additional plot was removed due to the effects of obvious substantial plot location error.

Deletion of outliers substantially improved the quality of fit for both the linear and nonlinear models for both the RECR and SURV datasets. For both the linear and nonlinear models, qualities of fit of the SURV models to the data were similar with respect to all three measures: root mean square error (RMSE), R^2, and mean residual; qualities of fit for the RECR models were also similar. However, for the RECR linear models, approximately 20% of predictions were negative; for estimation purposes, these predictions were changed to 0.

Despite the similar qualities of fit, estimates mean AGB per unit area obtained using the RECR dataset were substantially greater than estimates obtained using the SURV dataset (Table 1). This result is attributed to the effects of RECR plot location errors. The RECR plot coordinates tended to associate the few large AGB plots with locations for which the ALS data had lower pulse return heights than would usually be associated with large AGB plots (Figure 1); similar results were observed for smaller AGB plots, but not to the same degree. The RECR models then tended to predict erroneously large AGB for lower pulse return heights throughout the population which, in turn, produced larger overall estimates of mean AGB per unit area. The conclusion is that RECR plot coordinates should not be used with the model-assisted regression estimators for estimating AGB.

References.

[1] Gobakken, T., & Næsset, E. (2009). Assessing effects of positioning errors and sample plot size on biophysical stand properties derived from airborne laser scanner data. Canadian Journal of Forest Research 39: 1036-1052.

[2] Särndal, C.-E., Swensson, B., & Wretman, J. (1992). Model assisted survey sampling. Springer-Verlag, Inc. New York. 694 p.

Table 1. Estimation statistics.

Dataset	No. outliers deleted	Quality of model fit to data			GREG Estimates	
		RMSE	R^2	Mean residual	Mean	SE
Linear model						
SURV	0	41.56	0.63	2.95	35.37	3.74
	4	29.59	0.81	0.58	34.84	2.68
RECR	0	43.17	0.61	0.47	57.70	3.84
	5	29.55	0.78	0.61	53.59	2.67
Nonlinear model						
SURV	0	40.98	0.64	1.74	31.88	3.69
	4	31.81	0.78	3.90	32.13	2.86
RECR	0	44.31	0.59	2.40	54.49	3.94
	5	30.42	0.77	2.94	48.37	2.74

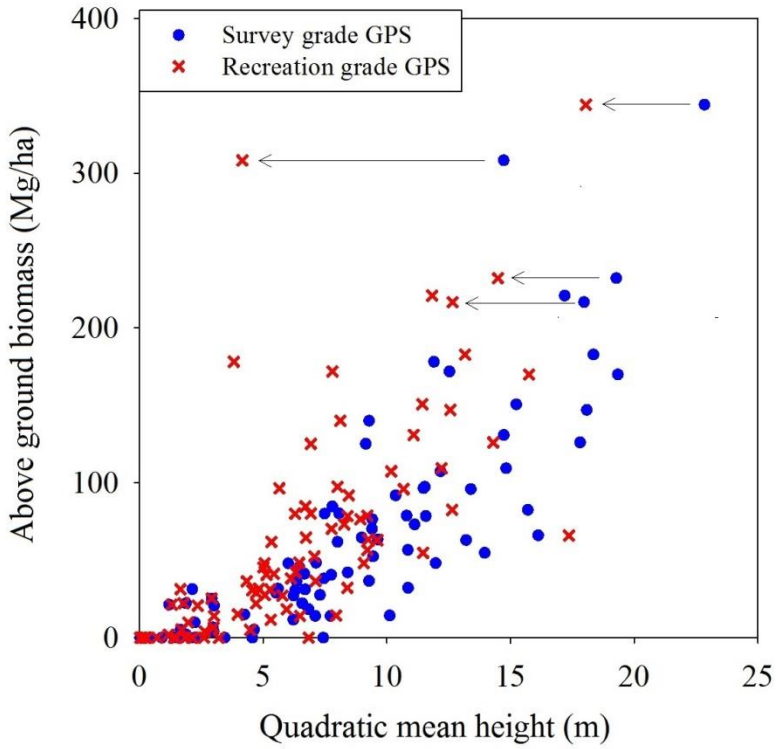


Figure 1. Plot AGB observations versus quadratic mean height for both SURV and RECR locations.

The 2014 Tanana Inventory Pilot: A USFS-NASA partnership to leverage advanced remote sensing technologies for forest inventory

Hans-Erik Andersen[1], Chad Babcock[2], Robert Pattison[1], Bruce D. Cook[3], Douglas C. Morton[3], and Andrew O. Finley[4]

[1]*USDA Forest Service Pacific Northwest Research Station, Seattle, WA & Anchorage, AK, USA. handersen@fs.fed.us, rrpattison@fs.fed.us*
[2]*School of Environmental and Forest Sciences, University of Washington, Seattle, WA, USA. babcoc76@uw.edu*
[3] *Code 618, Biospheric Sciences Branch, NASA/Goddard Space Flight Center, Greenbelt, MD, USA. bruce.cook@nasa.gov, douglas.morton@nasa.gov*
[4]*Department of Forestry, Michigan State University, East Lansing, MI, USA. finleya@msu.edu*

Highlights: The 2014 Tanana Inventory Pilot project established 100 FIA plots and sampled remote sensing information for the Tanana valley inventory unit in interior Alaska. Here, we examine the efficacy of three potential forest inventory estimation procedures in terms of estimation accuracy, uncertainty and bias.

Keywords: *design- and model-based inference, forest inventory, model-assisted, multilevel models*

Introduction

Interior Alaska (approx. 50 million forested hectacres in size) is the last remaining forested area in the United States (US) where the Forest Inventory and Analysis (FIA) program is not currently implemented. A joint NASA-FIA inventory pilot project was carried out in 2014 to increase familiarity with interior Alaska logistics and evaluate the utility of state-of-the-art high-resolution remote sensing information (lidar, hyperspectral and thermal airborne imaging) to support future FIA inventory procedures for interior Alaska. Here, standard (design-unbiased, plot-based) FIA estimation approaches are compared with model-assisted (i.e. approximately design-unbiased) and model-based (spec. Bayesian hierarchical) approaches which utilize relationships between field measurements and auxiliary structural and spectral metrics.

Data description

FIA plots were established at a 1:4 intensity (or 1 plot per 9,715 hectares) on a regular (i.e. systematic) hexagonal grid across the Tanana Valley State Forest and Tetlin National Wildlife Refuge; both of which fall within the Tanana valley of interior Alaska. Modified field measurement protocols were developed and implemented in order to better characterize boreal forest conditions, including 1) additional micro-plots to account for higher densities of small diameter trees, 2) soil samples to assess organic soil carbon storage, and 3) lichen/moss measurements. The relatively sparse FIA field plot sample collection was augmented with samples of airborne remotely sensed data acquired with Goddard's Lidar Hyperspectral and Thermal (G-LiHT) imager to increase the precision of inventory parameter estimates. G-LiHT is a portable, airborne imaging system, developed at NASA-Goddard Space Flight Center, that simultaneously maps the composition, structure, and function of terrestrial ecosystems using lidar, imaging spectroscopy, and thermal imaging. G-LiHT provides high-resolution (<1 m) data that is well suited for studying forest ecosystem dynamics, including assessments health and productivity that operate at spatial scales ranging from individual trees to landscapes. G-LiHT data supports local-scale mapping and regional-scale sampling of plant biomass, photosynthesis, and disturbance. The data is accurately georeferenced and can be matched precisely with field plot data that are georeferenced using survey-grade GPS. G-LiHT data was acquired in July-August, 2014 along single swaths (250 meters wide) spaced 9.3 km apart over the entire Tanana inventory unit (135,000 km^2). Preliminary analyses indicate a strong relationship between lidar derived variables and forest inventory metrics and also suggest that the addition of a second micro-plot at subplot locations further improves model fit a predictive ability for interior Alaska (Figure 1).

Analysis methods

We examine three methodological approaches to estimate forest inventory variables of interest; focusing initially on aboveground biomass (AGB) estimation. The three estimation procedures include 1) the standard, fully design-based approach currently used by the FIA for forest inventory estimation in the contiguous US[1]; 2) A model-assisted technique where sample collections of remote sensing data can be incorporated into the estimation procedure to potentially decrease uncertainty while still being approximately design-unbiased[2-3]; and 3) a Bayesian multilevel modeling approach where the sampling design can be explicitly accommodated within the modeling framework via appropriately structured random effects to potentially obtain approximate design-unbiased estimates of AGB via summarization of the posterior predictive distribution of AGB integrated over the forested domain[4-5]. We plan to assess the accuracy and bias of the three approaches experimentally via simulation and applicability of the three proposed methods for interior Alaska by implementing them for the Tanana Valley State Forest and Tetlin National Wildlife Refuge using the field and remote sensing data collected during the 2014 Tanana Inventory Pilot project.

Concluding Remarks

Results from this research will inform forest inventory estimation procedures for interior Alaska. It is prohibitively expensive to estimate forest inventory variables using field-only methods in the vast and remote interior Alaskan landscape. Currently proposed approaches for augmenting field campaigns with remote sensing data while remaining design-unbiased are promising but still lack the required flexibility in data collection procedures (e.g., require systematic field sampling designs for efficient remote sensing data collection, need complete collections of remote sensing data for double-sampling) necessary for realistic implementation in a large scale forest monitoring protocol required for interior Alaska. By comparing design- and model-based approaches we will gain understanding about how estimation bias influences forest inventory estimates for interior Alaska and potentially discern if it is possible to obtain approximate design-unbiased inventory estimates while leveraging the flexibility of model-based approaches.

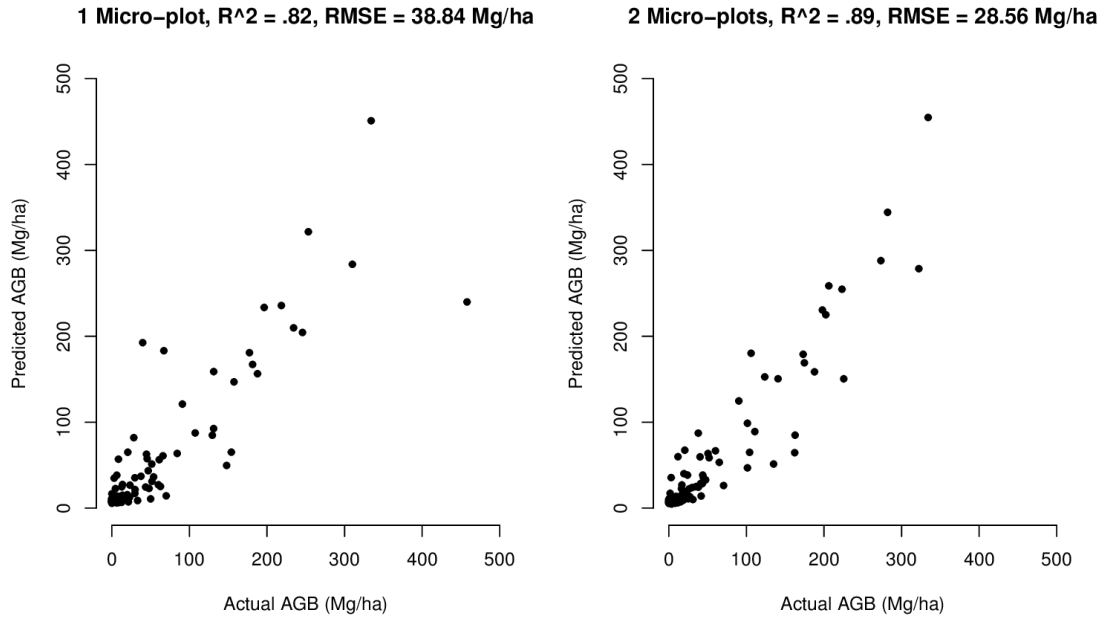


Figure 1. Scatterplots of Predicted vs. Actual Aboveground Biomass (AGB) for plots within Tetlin NWR, interior Alaska.

References

[1]Bechtold, W.A., & Patterson, P.L. (2005). The enhanced forest inventory and analysis program: national sampling design and estimation procedures. US Department of Agriculture Forest Service, Southern Research Station Asheville, North Carolina.

[2]Sarndal, C., Swensson, B., & Wretman, J. (2003). Model Assisted Survey Sampling. Springer.

[3] Gregorie, T.G., Stahl, G., Naesset, E., Gobakken, T., Nelson, R., & Holm, S. (2011). Model-assisted estimation of biomass in a LiDAR sample survey in Hedmark County, Norway. Canadian Journal of Forest Research, 41(1), 83-95.

[4] Banerjee, S., Carlin, C.P., & Gelfand, A.E. (2014). Hierarchical Modeling and Analysis for Spatial Data. (2nd ed.). Chapman and Hall/CRC.

[5] Chen, C., Wakefield, J., & Lumely, T. (2014). The use of sampling weights in Bayesian hierarchical models for small area estimation. Spatial and Spatio-temporal Epidemiology, 11, 33-43.

Comparison of unit- and area-level small area estimation methods – The influence of spatial autocorrelation

Johannes Breidenbach, Rasmus Astrup

Norwegian Forest and Landscape Institute

Highlights: A comparison of different SAE methods

Key words: *NFI, SAE, variance estimation*

Summary

Estimating the properties of small domains such as forest stands is a traditional field of use of airborne laser scanning (ALS) in forest inventories and is currently still the biggest market for ALS-based products. The statistical methods used in this field of application are subsumed under the term small area estimation (SAE).

SAE methods can generally be classified as unit-level or area-level approaches. In the unit-level approach, observed (field) data need to be exactly co-registered with the ALS data in order to derive a regression model. Exact coordinates of the field data (e.g., sample plots) are needed in this case. The regression model is used to map the response variable of interest, such as biomass or timber volume. The mean of the estimated map pixels within a stand is known as the synthetic estimate which is the currently most commonly used SAE technique.

In the area-level approach, estimates of the response variable need to be available for some domains in order to derive a regression model on the domain level. In forest inventories, these data can be obtained by clustering some sample plots within some stands. Exact coordinates of the sample plots are not required. Estimates are made directly on the stand-level using the regression model. Model-based variance estimators are defined for both the unit- and area-level approach.

While the unit-level approach appears to be more intuitive for a forester with experience in ALS-based inventories on the first sight, area-level SAE methods can be expedient especially in areas where exact coordinates of the field sample plots are difficult to obtain. Furthermore, often the field survey does not allow a direct link to the remote sensing data as for example if line transect sampling or variable radius plots (Bitterlich sampling) are applied. In these cases area-level methods may be the correct choice. Advantages and disadvantages of the two approaches are discussed based on example data where the effect of spatial autocorrelation is discussed.

A rule-based method using L-moments for direct detection of shade-intolerant regeneration in Boreal forests from airborne laser scanning

Rubén Valbuena, Lauri Mehtätalo, Matti Maltamo, Petteri Packalen

[1] *University of Eastern Finland, School of Forest Sciences & School of Computing, PL 111. Joensuu, Finland*
rubenval@uef.fi, lauri.mehtatalo@uef.fi, matti.maltamo@uef.fi, petteri.packalen@uef.fi

Highlights: Stratification by two simple rules, L-coefficient of variation $Lcv > 0.5$ and L-skewness $Lskew > 0$, discriminates forest areas under (potential) shade-intolerant regeneration (not fully stocked). These rules can be applied directly, using neither auxiliary prior information nor field data support, providing an unambiguous procedure for low-density national datasets, with multiple applications.

Key words: *L-coefficient of variation, Gini coefficient, forest structure; tree size inequality.*

Objectives

The objective of this study was to research whether airborne laser scanning (ALS) metrics may be applied directly for stratifying the forest area into zones of similar structural complexity, e.g. even-sized or multi-layered stands. Therefore, the objective was to test a similar method as the classification outlined in Valbuena et al. [1] but, instead of carrying out predictive modelling, using L-moments from the ALS returns directly ($L1$, $L2$ and $L3$, are first, second and third L-moments) [2]. We selected mathematical rules based on L-coefficient of variation (Lcv) as a descriptor of relative dispersion of ALS heights and the L-skewness ratio ($Lskew$) as a descriptor of asymmetry. Thus, we aimed at developing a method for stratification of ALS scans using simple rules constructed in the $L1$-$L2$-$L3$ space, and testing whether such stratification would be significant for describing forest structure.

Introduction

In forest ecosystems, the Lorenz curve is a representation of the relative dominance that each individual have in a tree assemblage [1]. Consequently, indicators directly related to the Lorenz curve have been selected for obtaining a consistent and objective procedure for forest structure characterization. Most relevant indicators have been found to be: the Gini coefficient (GC) and basal area larger than mean ($BALM$) [3]. These indicators have been used for many applications in forest management and inventory, especially due to their capacity to describe the shape of tree diameter distributions. GC has been used as a descriptor of the degree of evenness among the tree diameters, for stratifying ALS-assisted estimations of tree diameter distributions [4]. $BALM$ have also been used in ALS remote sensing for evaluating relative understory development in multi-layered forest structures [1]. Overall, GC can be considered as a measure of relative dispersion of tree diameters, whereas $BALM$ can be used to describe the asymmetry of the diameter distribution. ALS have been used for predicting these forest indicators, GC [5] and $BALM$ [3], using a wide range of methods based on statistical modelling (i.e. supported by training with field data). In this contribution we show the mathematical equivalence between the GC and the Lcv, reflecting on the convenience of using L-moments for describing distributions of ALS heights. We tested the possibility of characterizing forest structure using descriptions of the distributions of ALS heights, instead of tree diameter distributions (Table 1). The forest area was classified using mathematical properties of these ALS metrics directly, while field data is used only for validation of the method. A success in this approach would release the need for field support information for model training in ALS applications for forest assessment.

Materials and Methods

This research was carried out in the context of a project pursuing the automated detection of multi-layered silvicultural development classes in Finland (see [6] for details on the materials used). For developing these rules, we studied the relations between the L-moments with forest indicators and the target development classes determined in the field. It is noteworthy to emphasize that ALS employed was low-density datasets which are being acquired at a national scale [7], as similar surveying programmes area also being carried out in many other countries around the world, which brings potential for direct and consistent replication.

Table 1: Correspondence among descriptors for distributions of tree diameters and ALS heights.

Descriptor	Describing	Forest response	ALS metrics
Location	Tree size	Quadratic mean diameter (QMD)	1st L-moment
Dispersion	Tree size inequality	Gini coefficient (GC)	L-coefficient of variation
Asymmetry	Tree size dominance	Basal area larger than mean ($BALM$)	L-skewness

The L-coefficient of variation

The L-moment ratios have the advantage of being bounded by finite intervals [2], providing relative descriptions for the distribution of ALS return heights. The L-coefficient of variation is obtained as the ratio of the second to the first L-moments:

$$Lcv = \frac{L2}{L1} = \frac{E(X_{2:2}) - E(X_{1:2})}{2E(X)} \quad (1)$$

For positive random variables, the theoretical values for the second L-moment ratio are bounded by the [0, 1] range. This brings the advantage that concentration measures are comparable among distributions differing in their location ($L1$). In this communication, we include a formal demonstration for $GC = Lcv$. This equivalence reveals that the properties observed for GC in relation to describing empirical distributions must stand for the Lcv as well. For instance, the middle value of $Lcv = 0.5$ corresponds asymptotically to a theoretical uniform distribution [8]. Thus, this value can be used to stratify the ALS dataset into two groups, according to whether the variance of the ALS heights is higher or lower than the variance of a corresponding uniform distribution over its empirical range.

The L-skewness

The third L-moment ratio is obtained by division between the third and the second L-moments. It is called the L-skewness ($Lskew$), as it has been found to be a robust descriptor for the asymmetry of the distribution:

$$Lskew = \frac{L3}{L2} = \frac{E(X_{3:3}) - 2E(X_{2:3}) - E(X_{1:3})}{E(X_{3:3}) - E(X_{1:3})} \quad (2)$$

The numerator ($L3$) expresses the expected difference between the maximum-median and minimum-median differences in a sample of size 3, which provides a L-measure for the asymmetry of the distribution of ALS heights. Hence, $Lskew = 0$ corresponds to a symmetric distribution, $Lskew > 0$ describes positive asymmetry (left-skewed distribution) and $Lskew < 0$ describes negative asymmetry (right-skewed distribution). Additionally, $Lskew$ has the advantage of presenting theoretical bounds within the [-1, 1] interval. Consequently, $Lskew$ is a descriptor of asymmetry relative to dispersion, and thus independent of the spread of the ALS heights distribution.

Results and Discussion

Each L-moment related with a different property of the forest (Table 1). The $L1$ showed a clear linear relation to the QMD in such a way that would, by itself, rank even-sized development classes in order of increasing QMD. However, this relation no longer stood for forest areas containing parent trees (seed stands), while other multi-layered forest classes overlapped in this relation with the even-sized ones. Therefore, the linear relation between the forest variable and the ALS metric was blurred by the presence of this type of forest areas, which we identified as being under conditions for shade-intolerant regeneration. Therefore, discriminating this type of stands automatically would improve the estimation of other tree-size dependent forest attributes.

The relation between the Lcv of ALS heights and the GC of tree size inequality [1] was studied as well. Results showed the threshold at $Lcv > 0.5$ alone may suffice to discriminate seed-tree stands by itself. This rule, however, failed to identify forest areas with shade-tolerant regeneration in the understory under a closed dominant canopy. In view of this result, we propose that $Lcv > 0.5$ may correspond to the tree size inequality of a forest in presence of shade-intolerant regeneration, whereas for shade-tolerant regeneration it does not. In other words, a correspondence between GC and Lcv may only happen when the elevated value of GC is due to the presence of a gap in the canopy, which allows a large proportion of the laser footprint to get through and disperse its corresponding returns along the vertical profile of the canopy. High values of Lcv of ALS heights may therefore be a sign for shade-intolerant regeneration, whether factual (in presence of seedlings) or potential (just a gap). The question remains, to be clarified by future research, on whether the inability of Lcv to detect undergrowth was due to the low density of the dataset employed, and therefore higher densities would be necessary for this task. Using the $Lcv = 0.5$ rule for discriminating even-sized vs. multi-layered classes, obtained an overall accuracy of 87.0% and a coefficient of agreement $\kappa = 0.53$. A total of 97.0% of the even-sized plots were correctly classified by this rule, while just 48.9% of the plots belonging to multi-layered plots were correctly identified as such, due to the inability for this rule to identify shade-tolerant regeneration. Consequently, when we considered the possibility for discriminating only shade-intolerant regeneration against the remaining forest areas the overall accuracy increased to a 92.1% with a $\kappa = 0.62$, which we considered a significant achievement for an rule-based method not requiring field support for training.

The next step was to observe the capacity of $Lskew$ to incorporate additional information for forest structural class discrimination. The threshold at $Lskew = 0$ further assisted in discriminating multi-layered from even-sized development classes. Positive skewness ($Lskew > 0$) was observed when there were high proportions of ALS heights with relatively lower heights. It was therefore related to the presence of uneven-sized stands which had a larger number of returns backscattering from the lower strata. We also observed positive asymmetry in seedling and sapling stand, as well as in areas having gaps in the dominant canopy. However, areas with shade-tolerant regeneration under closed canopies were also concealed within the even-sized group at $Lskew < 0$. We therefore suggest that, again, values of $Lskew > 0$ were also related to shade-intolerant regeneration, and therefore the rule could as well be used for masking out these areas automatically.

Combining both rules lowered the overall accuracy in classifying even-sized and multi-layered areas of the forest down to 79.1% and κ = 0.42. Even though the plot-level accuracy lowered, the method was more successful in terms of detecting multi-layered development classes. The proportion of correctly-classified multi-layered plots was increased to 63.8%, while the amount of correctly classified among the even-sized development classes was just slightly lowered to 83.1%. A very relevant property of the combined-rule approach was its capacity for virtually delineating stands directly. Although the method was carried out at 16x16 meters cell level, the resulting maps seemly identified entire stands quite sharply (Figure 1). This partially releases the need for segmentation procedures that may involve more subjective steps, trial-and-error, or manual delineation. We therefore regarded this combined approach to be the most useful in identifying forest areas with potential for shade-tolerant regeneration. These zones may be masked out with the intention to lower the signal-to-noise ratio in the prediction of tree size-dependent forest variables like QMD in the remaining forest areas.

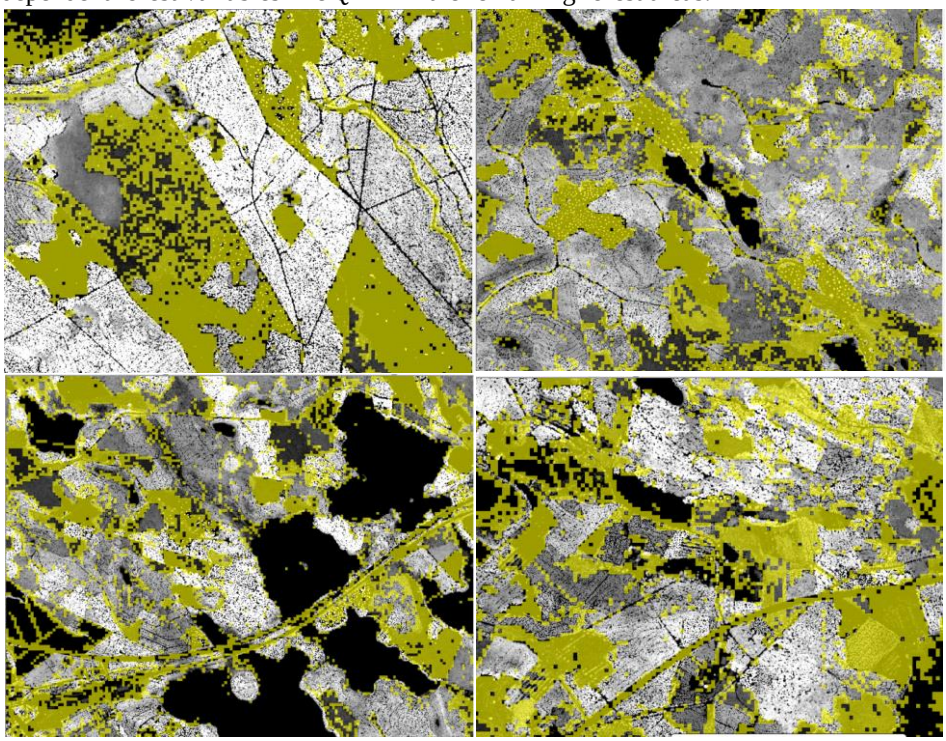


Figure 1: Examples of maps obtained by direct classification rules (yellow denotes $Lcv > 0.5$ OR $Lskew > 0$).

Acknowledgements

This research was funded by Suomen Metsäkeskus (SMK Finnish Forest centre). Special thanks to Juho Heikkilä and Jussi Lappalainen (SMK), Heli Laaksonen (NLS), and Aki Suvanto (Blom Kartta Oy) for their support at different stages of this study.

References

[1] Valbuena, R., Packalen, P., Mehtätalo, L., García-Abril, A., & Maltamo, M. (2013) Characterizing forest structural types and shelterwood dynamics from Lorenz-based indicators predicted by airborne laser scanning. Canadian Journal of Forest Research, 43, 1063–1074.

[2] Hosking, J.R.M. (1990) L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics. Journal of the Royal Statistical Society. Series B (Methodological), 52, 105–124

[3] Valbuena R., Vauhkonen J., Packalen P., Pitkänen J. & Maltamo M. (2014) Comparison of Airborne Laser Scanning Methods for Estimating Forest Structure Indicators Based on Lorenz Curves. ISPRS Journal of Photogrammetry & Remote Sensing, 95, 23–33.

[4] Bollandsås, O.M., & Næsset, E. (2007). Estimating percentile-based diameter distributions in uneven-sized Norway spruce stands using airborne laser scanner data. Scandinavian Journal of Forest Research, 22, 33–47.

[5] Ozdemir, I., & Donoghue, D.N.M. (2013). Modelling tree size diversity from airborne laser scanning using canopy height models with image texture measures. Forest Ecology and Management, 295, 28–37.

[6] Valbuena, R., Maltamo, M., Heikkilä, J., & Packalen, P (2015) Detection of multi-layered forest development classes using airborne laser scanning. Silvilaser 2015 Conference, La Grande Motte, France.

[7] NLS - National Land Survey of Finland 2013. Laser scanning data (available online at maanmittauslaitos.fi). Visited in Sep. 2013.

[8] Valbuena, R., Packalen, P., Martín-Fernández, S., & Maltamo, M. (2012) Diversity and equitability ordering profiles applied to study forest structure. Forest Ecology and Management, 276, 185–195.

Producing an understory cover map over a large region using heterogeneous lidar datasets

Ranjith Gopalakrishnan, Dr. Valerie Thomas and Dr. Randolph Wynne

Dept. of Forest Resources and Environment Conservation, Virginia Tech, USA.

Highlights: We assess the efficacy in using an heterogeneous lidar dataset (over 90 separate projects) to estimate forest understory cover over a large region. We developed a random forest based model that yielded RMSEs of ~18% live shrub cover, and helped identify significant understory lidar metrics.

Key words: Forestry, lidar, understory cover, wall-to-wall mapping, national forest inventories.

Introduction

Accurate and spatially explicit maps of the understory are important to forest resource managers and scientists for the following reasons: 1) Estimating ladder fuel (by which surface fires reach the tree canopy) load is crucial for robust forest wildfire modeling; 2) Good spatial maps of understory shrub distribution are extremely useful for forest wildlife habitat and biodiversity management; 3) Understory is an important component while accounting for total forest stand biomass. The ability to penetrate the canopy and generate returns from the understory is an important added advantage of lidar remote sensing over traditional optical methods. But currently, there is no good strategy to get understory estimates over very large areas (such as the entire southeastern United States) using a combination of airborne LIDAR and ground based national forest inventory data. Combining disparate lidar datasets with national forest inventories has shown promising initial results for canopy height [1]. This work tries to extend a similar approach to the understory.

Methods

Study area and lidar data used

The study area considered here is relatively large and includes forested regions with lidar coverage for the 13 southern states of the US. The forests in this region are quite heterogeneous, and may be broadleaf (e.g., oak-hickory and maple-beech-birch forests), or conifer forests (eg: planted pines) or even mixed forests (e.g., oak-pine, oak-gumcypress). Due to the size of the continental US, there has not been any large-scale, national- level lidar data acquisition projects. Meanwhile, there has been multiple regional acquisitions with data in the public domain, but with varying sensor configurations and flight parameters. We obtained many such regional lidar datasets by approaching three coordinating government agencies: (1) The U.S. Geological Survey (USGS); (2) National Oceanic and Atmospheric Administration (NOAA); (3) Natural Resources Conservation Service (Alabama).

Field data used

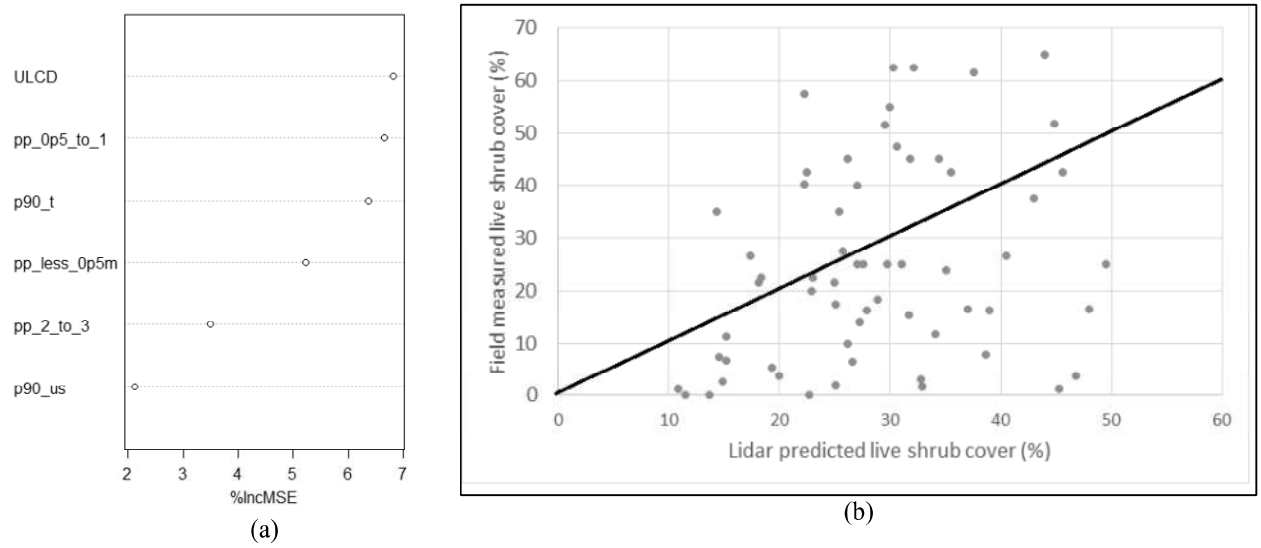
The Forest Inventory and Analysis (FIA) program of the United States Forest Service (USFS) has been collecting extensive field data on a nationally-standardized sampling plot, since the 1990s. The FIA uses a multi-phase inventory, and during the third phase, trained crews measure numerous forest health indicators on 1/16 subset of the original FIA forest plots [2]. Such plots are known as phase 3 (P3) plots. The FIA P3 plot (henceforth FIA plot) layout is shown in figure 1(b). Understory measurements are problematic and the repeatability of such measurements are usually called into question. In [2], the authors describe a study where they inspected data from 139 FIA P3 plots that were re-measured for quality control reasons. They reported that across the suite of measurement variables that comprise the fuels inventory of the FIA program, 15 of the 27 variables did not attain the desired repeatability levels. We choose live shrub cover percent as our field-measured dependent variable as it had reasonable repeatability, as per that study. Live shrub cover is measured at each forested microplot, and is defined as the percent of the microplot area that is covered with live shrubs (shrub being defined as a woody plant).

Lidar metrics and model used

We identified 205 FIA P3 plots that had good corresponding lidar coverage, and cut out 120m (square) plots around these FIA plots (see [1] for plot shape and size rationale). Thus, we had 205 plots which had both LIDAR and understory FIA field data. We then calculated several non-collinear lidar metrics (as independent variables) from these lidar plots (see table 1). Note that the return labeling (first, second, last) were ignored and only the height above ground were taken into consideration in the calculation of these metrics. We then used the Random Forest (RF) algorithm, with a random set of 70% of the plots for training and the rest for validation. A reduced model using only the important variables were used for the validation phase. We then generated an understory cover map for a relatively small area (9.2 x 9.2 square kilometers), to test out the predictive model developed. Specifically, an area straddling Darlington and Marlboro counties of the state of South Carolina, USA, was chosen.

Table 1: Lidar-derived metrics of understory presence (top 5) and canopy height (last one).

Sl. num.	Lidar metric	Metric description
1.	pp_less_0p5m	Percentage of points (returns) less than 0.5 meters. These are mostly ground returns.
2.	pp_0p5_to_1	Percentage of points in the height bin of 0.5 to 1.0 meters.
3.	pp_2_to_3	Percentage of points in the height bin of 2.0 to 3.0 meters.
4.	p90_us	The 90th percentile of the height of understory returns. Here, returns in the height bin of 0.0 to 3.0 meters are considered understory returns.
5.	ULCD	Percent points in the 0.5 to 2.0 meter bin, relative to all points in the 0.0 to 2.0 bin. For more, see [3].
6.	p90_t	The 90th percentile of the height of canopy returns. Here, returns in the height bin of 3.0 meters and above are considered canopy returns.



(a) (b)

Figure 2 (a): Variable importance graph from the initial random forest model. The increase in RMSE observed when the variable is "dropped" is given; (b) Scatter-plot showing the agreement between the predicted and field-measured live shrub cover (final model). The solid black line is the 1:1 fit line. In general, Lidar under-predicts the live shrub cover.

Results

A random forest model was first made including all six predictor variables and using 143 randomly chosen plots. The variable importance graph obtained can be seen in figure 2(a). Using this, we refined the model using only the following variables: ULCD, pp_0p5_to_1 and p90_t. This final model gave an RMSE of 17.4% when it was tested with the remaining 62 plots (averaged over 20 runs). A scatterplot showing the predicted and the observed live shrub height using this final model for these test plots can be seen in figure 2(b). The results are promising: the authors in [3] reported comparable understory measurement accuracies ranging from ± 7 to 22%. Height bin related metrics like

pp_0p5_to_1 and ULCD have been found to be significant by other studies, too. The p90_t metric seems to point to the influence of the overstory on the understory. The understory cover map (and related information) generated using the random forest model for the region of interest (in the eastern US) can be seen in figure 3. It can be seen that there is good interrelationship between the estimated understory cover and the vertical returns profile. We are in the process of analyzing the residuals from the scatterplot and attributing high residuals to probable causes. Some of them can be: 1) The error in measurement of shrub cover, as detailed idn [2]; 2) The plot location is recorded with GPS devices of variable accuracies, so registration with lidar data may not be always accurate; 3) The microplots are quite small, and may not be representative of the ground condition on the entire plot.

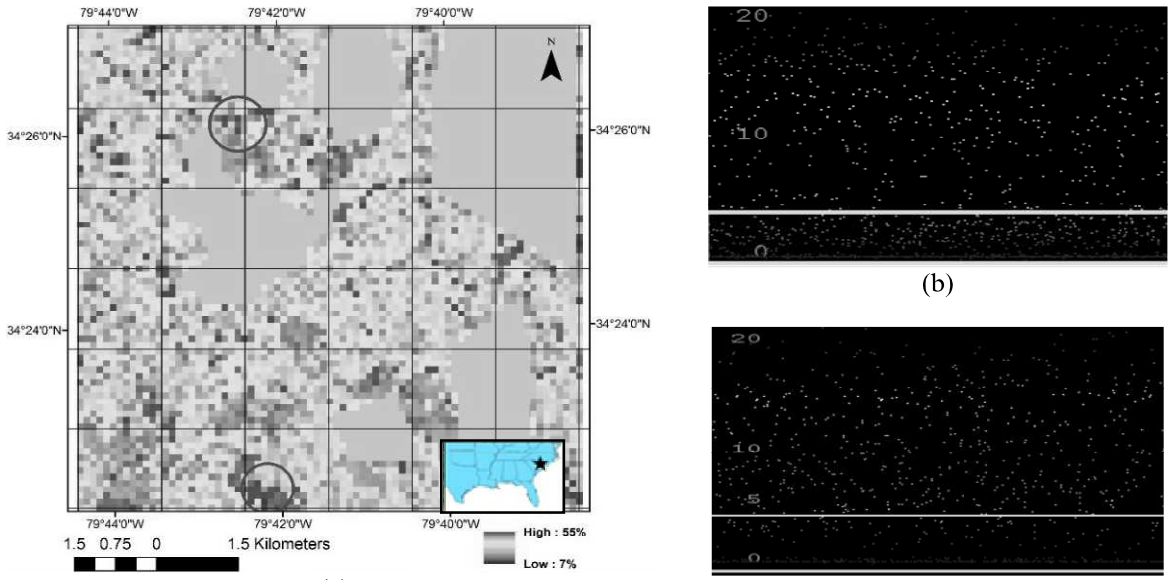


Figure 3 (a): Understory cover estimate map for the specified area, generated using our main model. The dark squares correspond to lidar data tiles (each ~1.5 km across). The light grey areas are non-forested land, where our model is not applicable. The two dark green circles represent two areas of interest. The area on the top has grid cells of estimated high understory cover. A vertical profile of lidar returns of part of this area (with high estimated cover) is shown in figure 3 (b). In this figure, blue represent low returns, while red are high returns. The height above ground is denoted in dark grey numbers (meters). The bold yellow line denotes the approximate 0 to 3m height bin. The other green circle, at the bottom of the figure, marks an area of low estimated understory cover. The vertical profile for this can be seen in figure 3(c). From these figures (3(b) and 3(c)), one can see that there are much more returns from the understory bin in the top (high estimates) area.

Conclusions

This work shows that there is promise in combining data from disparate lidar datasets in one model to generate wall-to-wall maps of understory cover. Although the RMSEs are rather large on a plot level, spatial patterns could be preserved on the landscape level. This work thus points to the potential of estimating the rough spatial distribution of several metrics related the forest 3D structure, which are difficult or almost impossible to get using other remote sensing techniques.

References

[1] Gopalakrishnan et. al. (2013). Efficacy of using heterogeneous lidar datasets in predicting canopy heights over a large region. Proceedings of the SilviLaser conference 2013, Beijing, China.

[2] Westfall, J. A. and Woodall, C. W. (2007). Measurement repeatability of a large-scale inventory of forest fuels. Forest Ecology and Management, 253.1-3: 171-176.

[3] Wing, B. M. et al. (2012). Prediction of understory vegetation cover with airborne lidar in an interior ponderosa pine forest. Remote Sensing of Environment, 124: 730-741.

Assessing understory diversity using a lidar/hyperspectral fusion approach

Beth R. Stein, Valerie A. Thomas, Philip J. Radtke, Randolph H. Wynne

Department of Forest Resources and Environmental Conservation, Virginia Tech

Highlights: We propose a novel lidar/hyperspectral fusion approach to identify canopy gaps and relate their species diversity to spectral reflectance. Although spectral modelling is limited to areas seen by the overhead sensor, regression approaches can predict understory diversity with moderate accuracy.

Key words: *Canopy gaps, understory, biodiversity, lidar, hyperspectral, data fusion*

Introduction

The forest understory, which consists of vascular plants under 1 m in height, provides numerous essential ecosystem services. The understory is important for forest biodiversity, influences the trajectory of forest regeneration, forms linkages with the overstory, and cycles essential nutrients [1]. In addition, the understory is a valuable food source for wildlife, and the diversity of species improves the aesthetic value of the landscape. Consequently, the importance of the understory vegetation leads to a clear need to quantify species diversity. Estimates of tree species richness are generally done with field measurements or remote sensing. However, below the forest canopy, the understory is scarcely visible with air- or spaceborne sensors. As a result, few studies have attempted to study understory plants or estimate diversity using remote sensing [2].

We propose a novel lidar/hyperspectral fusion approach to identify canopy gaps and relate their species diversity to spectral reflectance. NASA Goddard's LiDAR, Hyperspectral, and Thermal (G-LiHT) high-resolution imager offers a unique opportunity to achieve this goal. We hypothesize that gaps in the canopy can provide sensors with a glimpse of small pockets of underlying vegetation. Use of lidar to map large canopy gaps is well-established [3], and several recent studies have successfully modelled species diversity in other vegetation types using hyperspectral remote sensing. These studies include simulated wetlands [4], rainforests in Hawaii [5], and the Peruvian Amazon rainforest [6]. However, whether lidar/hyperspectral fusion can identify small gaps and estimate their species diversity is unknown. Thus, the following question arises: How does the forest understory in canopy gaps differ from the rest of the understory, and what is the effect of the understory composition on reflectance? Our objective is to determine how well a fusion of hyperspectral and lidar technology can describe the forest understory, both within gaps and beyond. To achieve this end, we take a two-tiered approach with a spatial gap assessment followed by a reflectance-based gap assessment. Given the launch of several hyperspectral and lidar satellites in the future, this research is particularly relevant.

Material and methods

The study area consists of a 0.5 km x 1 km region in northern Blackwood Division of Duke Forest. According to Duke University, there are 984 plant species in 459 genera and 134 families in the Duke Forest. Dominant tree species in the mixed hardwood forest include red maple (*Acer rubrum*), green ash (*Fraxinus pennsylvanica*), yellow-poplar (*Liriodendron tulipifera*), and sweetgum (*Liquidambar styraciflua*). During the field campaign, we selected 34 canopy gaps for analysis with a range of sizes (0.32-62.95 m^2). Mean gap area is 13.15 m^2, with a standard deviation of 16.16 m^2. Although overstory species remain similar across the area, gaps were distributed across a topographic and moisture gradient to encompass a large number of species and abundance patterns. To assess gap vegetation, we randomly placed 0.5 m x 0.5 m quadrats at three locations within the gap and three locations outside of the gap: 5m N, 10m N, and 10m E. We identified and counted the number of unique vascular plant species per plot and inside and outside each gap.

G-LiHT and the European Space Agency's HyPlant sensor acquired lidar and hyperspectral data in September and October, 2014. The G-LiHT system includes an airborne scanning lidar instrument with 1 m spatial resolution and a 10 cm footprint at an altitude of 335 m, with 0.3 mrad beam divergence angle [7]. The HyPlant sensor has a spatial resolution of 1 m, spectral resolution of 3-10 nm, and wavelength range of 370-2500 nm. We pre-processed the lidar data by mosaicking the individual datasets, reclassifying the lidar points into ground and non-ground classes, and then using linear interpolation to calculate height above ground. Next, we produced a digital surface model (DSM) and digital elevation model (DEM) from the lidar returns. To calculate the percentage of points hitting the ground, we divided the number of surface points by the total number of points in each pixel. We extracted these percentages at the gap GPS points, and then evaluated the percent distributions for gap and non-gap points. We conducted a t-test to verify a statistical difference in distributions.

We used the density of lidar points below the DSM to locate canopy gaps. Based on field measurements, we evaluated the success of the lidar method at identifying small canopy gaps. Then we conducted a two-tiered approach to examine understory diversity within canopy gaps. First, we examined understory diversity using traditional and spatial statistics. We used a T-test and ANOVA to determine any significant differences in vegetation within the gaps and across the sampling distance. Then, we used spatial statistics (such as Moran's I) to study how understory characteristics change with distance.

The second approach correlates diversity indices and percent vegetation coverage with hyperspectral reflectance. Within each gap, we extracted mean reflectance values from the hyperspectral imagery. Next, we calculated Pearson correlation coefficients between mean gap reflectance and species diversity. Finally, we conducted best subsets multiple linear regression to develop predictive equations for diversity. We used two sets of predictors: (1) reflectance at wavelengths determined by Thenkabail et al. as the optimal combination for maximizing non-redundant information [8], and (2) texture metrics at these same wavelengths. We calculated texture metrics from image reflectance using 3x3 windows with an 8-pixel neigborhood; variables include mean, variance, homogeneity, contrast, dissimilarity, entropy, second moment, and correlation [9]. Our criteria for multiple linear regression model selection included number of variables and Bayesian Information Criterion.

Results

Our method of lidar gap identification successfully found 29 of 34 gaps measured in the field. Gaps were typically located within 3 m of their field GPS location. Problems with gap identification were generally due to missing returns in the lidar data. In the few cases where the lidar algorithm showed gaps as completely closed canopy, the gaps were smaller than 2 m in diameter. Over 500 understory plants were identified to the species. The vegetation comprised 86 different species; the most common species were two non-natives, *Microstegium vimineum* and *Lonicera japonica*, and *Acer rubrum*. ANOVA results indicate significant differences in the number of species between plot locations (F=13.73, p=0.00). Compared to plots outside of gaps, plots inside gaps generally have more species. Figure 1 shows the mean and ranges of number of species by location. However, a matched pairs test between total number of species inside and outside each gap found no significant differences in species diversity (t=-0.74, p=0.47). These results imply that while plots inside gaps tend to have more species, intra-gap variation in species diversity is large enough to cancel differences when plots are pooled by location. Thus, mean gap diversity may not represent a particular plot.

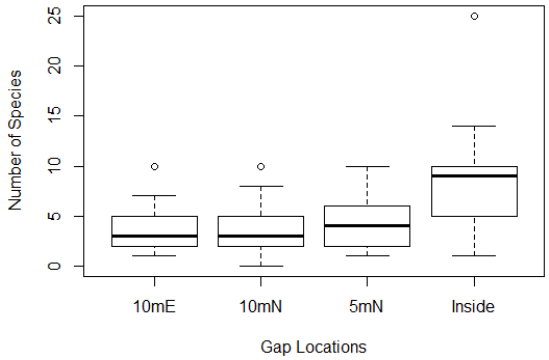


Figure 1: Boxplot depicting mean and ranges of species diversity at different gap locations across 34 plots.

Moran's I between all plots indicated significant spatial autocorrelation in species diversity (observed=0.18, expected=-0.00, p=0.00). A separate Moran's I analysis with only plots inside gaps reinforced these findings that inside-gap plots are spatially autocorrelated (observed=0.38, expected=-0.01, p=0.00). However, the total species diversity inside each gap (as defined by number of different species across all three inside quadrats) is not spatially autocorrelated (observed=-0.04, expected=-0.04, p=0.91). This finding implies that some method is necessary to estimate the number of species within each gap, but then the number found between quadrats can be interpolated.

There were several significant correlations between mean gap reflectance and number of different species ($p<0.05$). Significant wavelengths include 1939-2000 and 2439-2477 nm. The selected multiple linear regression model with 3 variables had a predictive capability of $R^2=0.46$, Adjusted $R^2=0.39$, and RSE=3.71. The intercept and two variables were significant at $p<0.05$; the third variable was significant at $p<0.10$. In addition, the use of texture-based approaches in regression models moderately predicted understory species diversity within gaps ($R^2=0.49$, Adjusted $R^2=0.44$).

Discussion

The lidar algorithm identified the majority of gaps found in the field, and only missed a few of the smaller gaps. Efforts to model understory diversity with hyperspectral data were moderately successful. Our results indicate that species diversity can be estimated using various regression approaches, but approximately half the

variation in diversity remains unexplained. These findings are consistent with some earlier research on species diversity models for other vegetation types [10]. While spectral signatures within a pixel may be linked to a target species or physiological condition (e.g. nutrient, structure), and algorithms exist for spectral unmixing, there is additional complexity to estimate species richness. Greater inter- than intra-specific spectral diversity may be required to further increase model accuracy [11]. However, there are several additional options for improving the relationships between reflectance and plant diversity. We would like to incorporate other related variables such as topography, leaf area index, and lidar structural metrics, which may correlate with productivity, light interception, and overstory diversity. Another possibility is to assess beta diversity, which compares the number of common species between plots using an index such as Sorenson's Similarity Index.

Spatial autocorrelation results were inconsistent across locational classes. The absence of any spatial autocorrelation in the inside-gap diversity measurements is unexpected given the significant spatial autocorrelation of both total and outside gap diversities. We hypothesize that this result may be due to underlying environmental variability. In a future study, we would like to further examine these variables, such as temperature (e.g. G-LiHT thermal data), moisture, and tree species. In addition, the next step is to evaluate whether gap size has any effect on inside-outside gap models or spectral reflectance.

The regression models from this study can produce predictive maps for understory diversity in other gaps nearby. These predictive maps may provide natural resource managers with a starting point for understory inventories, as they identify potential biodiversity "hotspots" and areas with particularly low diversity. While the former may indicate good wildlife habitat or potential conservation areas, the latter are also informative. At many gaps, extremely low species diversity indicated the presence of a dominant invasive species. Knowledge of these locations is critical to map plant invasions and monitor their spatial patterns. Thus, understory diversity estimates at both extremes could be useful for a variety of purposes.

Conclusion

Our study shows that lidar/hyperspectral fusion can be a useful technique for studying understory vegetation diversity within canopy gaps. Although spectral modelling is limited to areas seen by the overhead sensor, regression approaches can provide estimates of species diversity within gaps. While there are many applications for predictive maps generated with these methods, users must be mindful of prediction error.

Acknowledgments

We are grateful to our colleagues in the NASA/ESA Joint Campaign and the Duke Forest staff for their assistance with fieldwork, image acquisition, and data processing. This research was funded by a grant through the USDA National Needs Fellowship, USDA-NIFANNF-2010-03349, the McIntire-Stennis Cooperative Forestry Research Program through the USDA CSREES under Project VA-136614, and the Department of Forest Resources and Environmental Conservation at Virginia Tech.

References

[1] Gilliam, F.S. (2007). The ecological significance of the herbaceous layer in temperate forest ecosystems. Bioscience, 57, 845-858.

[2] Singh, K. K., Davis, A. J., & Meentemeyer, R. K. (2015). Detecting understory plant invasion in urban forests using LiDAR. International Journal of Applied Earth Observation and Geoinformation, 38, 267-279.

[3] Asner, G. P., Kellner, J. R., Kennedy-Bowdoin, T., Knapp, D. E., Anderson, C., & Martin, R. E. (2013). Forest canopy gap distributions in the southern Peruvian Amazon. PloS one, 8, e60875.

[4] Heumann, B.W., Hackett, R.A., & Monfils, A.K. (2015). Testing the spectral diversity hypothesis using spectroscopy data in a simulated wetland community. Ecological Informatics, 25, 29-34.

[5] Carlson, K. M., Asner, G. P., Hughes, R. F., Ostertag, R., & Martin, R. E. (2007). Hyperspectral remote sensing of canopy biodiversity in Hawaiian lowland rainforests. Ecosystems, 10, 536-549.

[6] Féret, J. B., & Asner, G. P. (2014). Mapping tropical forest canopy diversity using high-fidelity imaging spectroscopy. Ecological Applications, 24, 1289-1296.

[7] Cook, B.D., Nelson, R.F., Middleton, E.M., Morton, D.C., McCorkel, J.T., Masek, J.G., Ranson, K.J. Ly, V., & Montesano, P.M. (2013). NASA Goddard's LiDAR, Hyperspectral and Thermal (G-LiHT) Airborne Imager. Remote Sensing, 5, 4045-4066.

[8] Thenkabail, P.S., Gumma, M.K., Teluguntla, P., & Mohammed, I.A. (2014). Hyperspectral remote sensing of vegetation and agricultural crops. Photogrammetric Engineering & Remote Sensing, 80, 697-709.

[9] Haralick, R. M., Shanmugam, K., & Dinstein, I. H. (1973). Textural features for image classification. IEEE Transactions on Systems, Man and Cybernetics, 6, 610-621.

[10] Nagendra, H. (2001). Using remote sensing to assess biodiversity. International Journal of Remote Sensing, 22, 2377-2400.

[11] Asner, G. P., & Martin, R. E. (2008). Airborne spectranomics: mapping canopy chemical and taxonomic diversity in tropical forests. Frontiers in Ecology and the Environment, 7, 269-276.

The use of Multispectral LiDAR for Improved Forest Mapping

Dr. Antoine Cottin[1], Sam Fleming[1], Prof. Iain H. Woodhouse[1,2]

[1]*Carbomap Ltd., Edinburgh, s.fleming@carbomap.com, a.cottin@carbomap.com*
[2]*School of Geosciences, University of Edinburgh, i.h.woodhouse@ed.ac.uk*

Highlights: This project produced the first spectral map of a forest understory using airborne Multispectral LiDAR. This is the first time actual airborne data have been used, as previous work was conducted in the laboratory or with model simulations. This breakthrough paves the way for a range of new LiDAR applications.

Key words*: Multispectral, LiDAR, Forest, Spectral, Structural*

Introduction

Over the last 40 years LiDAR systems have fundamentally changed the world of mapping and surveying, with airborne systems able to cover large areas and remote locations. Since the first introduction of a LiDAR system, there have been many technological developments such as multiple pulses in flight, and full waveform recording. The next major development within the realm of forest mapping is proving to be Multispectral LiDAR[6,7]. This project demonstrates the next steps in this process, generating new methods for extracting key information about forests from cutting edge LiDAR technology.

Materials and Methods

In June 2013 Riegl flew a combination of three LiDAR systems on one airborne platform, over the course of two flights, over Corbin in Virginia, USA. Each laser was a different wavelength, at 532nm, 1064nm, and 1550nm. Similar data was also collected in multiple data collection campaigns in Austria, those these data were not used in this project[1,2]. The goal for these flights was to simulate an integrated Multispectral LiDAR system, providing an opportunity to explore the potential applications for such data. Carbomap received the data from Riegl USA to demonstrate the operational capabilities of Multispectral LiDAR with forestry applications, specifically with regards to the lower canopy layers, extracted through the full-waveform element of the data.

Vertical Spectral Variation

In order to demonstrate the results two approaches were used: the first was to segment a Pseudo Normalised Difference Vegetation Index (pNDVI) into three vertical layers, the first return, mid canopy, and low vegetation and ground. The generation of this map demonstrates that differences in the spectral response can be detected within the vertical profile. The pNVDI was determined using the point clouds of the 1064nm and the 532nm wavelength systems, as the two colour channels for the NDVI equation. This approach first selected one of the point clouds as a reference layer. Each point from this was then considered and the closest point in the other point cloud identified and the point couples tied together, allowing the pNDVI to be calculated.

The second approach was to create a three-channel false colour composite, to further highlight the difference in spectral response in the vertical direction of the forest canopy profile. The same process for tying together two points clouds for the pNDVI was used to create triplets between the three point clouds of different wavelength, from which this dataset was generated. The energy returned from each wavelength LiDAR was then converted into an RGB value; until now this has only previously been demonstrated under laboratory conditions.

Spectral Calibration

Reflectance is a function of the incidence angle between the incident laser beam and the surface normal of the target material. Therefore all the intensity values recorded were corrected to remove the incident angle effect. However, no target surface was used during the data collection, which would have provided a true spectral correction to apply. Therefore there was a need to empirically calibrate the datasets relative to each other.

In order to achieve this we were able to use a dry concrete surface, found in a parking lot in front of a building contained within the dataset. An average multispectral reflectance profile was used to establish the average ratio offset to apply for each of the three wavelengths. This ensured that for the target surface each dataset was able to match the average ratio for this type of material. The distribution of reflectance is specific for each wavelength and they were relatively narrow and similar for each channel. It was also possible to see a distinct pattern in the reflectance of the energy returned from each wavelength LiDAR. This is the approach that was also used for a similar Riegl dataset in Austria[1].

Results

The result of this investigation was the very first understory canopy map from a multispectral LiDAR, as shown in Figure 1 below. The three layers not only demonstrate a difference between the different land cover types, but also spectral differences between the different layers of vegetation.

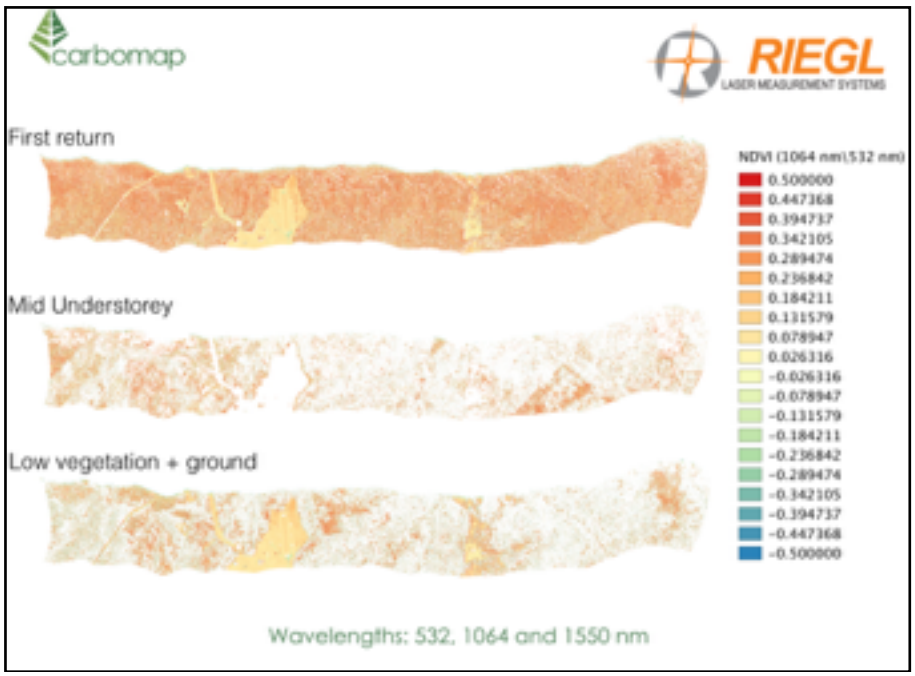


Figure 1: Pseudo-NDVI of three layers within the vertical structure of the forest, demonstrating differences in NDVI throughout the vertical canopy. (Top) First return from the LiDAR instrument, including the ground layer; (Middle) The mid-section of the forest canopy; (Bottom) Low section of the forest canopy and the ground layer.

The very top layer is akin to the data you would receive from a passive multispectral system, in that it contains the data from the very top surface of the scene. This layer demonstrates a full range of pNDVI values, and it is clear to see that the grass in the middle of scene had a lower pNDVI value, 0.1-0.15, than the surrounding areas of forest which had values of 0.3-0.5. Similarly gaps for roads and buildings can be easily determined as their pNDVI value is less than 0. This supports the assertion that Multispectral LiDAR can be used accurately classify land cover types. The middle layer demonstrates that it was possible to extract spectral information about vegetation which would typically be hidden from view by the top level of the canopy. This image also demonstrates that there are horizontal spatial differences within the spectral response of the mid understory level, with some areas having a lower pNDVI than others. Additionally it is also possible to see variation in the density of the mid understory vegetation, of particular note is the section at the bottom centre-right, which can be used to further characterise different sections of forest. The third layer, presents the results from the ground layer and low vegetation. The differences between the grassy ground and the ground surface and low vegetation contained within the forested areas are clearly demonstrated. It is also possible to see spectral variation between different locations within the forest. We can also see that each layer produces a different result from the others, except the grass as this is the same data. However, there was difficulty in extracting blanket coverage for the mid- and low- levels of vegetation, due to the type of LiDAR system used (discrete return).

Discussion

The results demonstrate that it is possible to extract spectral variation within vegetation in a three-dimensional manner. When closer examination of far-right area of the location is examined we are able to see a difference between one particular section and the areas surrounding it. It has a very high pNDVI value for the first return, with a lower value for the bottom layer, yet almost no returns in the mid canopy. Possible reasons for these differences include forest regrowth or clearing, or a different species group. Similarly the ground level vegetation is obviously different from that surrounding it, either as grass or low level shrubby vegetation, with no vegetation in the mid level. It is also possible to see more extreme variation in the pNDVI within the mid- and lower-levels, indicating that it may be possible to differentiate between green leafy vegetation and woody vegetation, which would have a much lower value. The differences in pNDVI make it clear that differentiation of vegetation types should be possible. It was also successful at identifying spatial variation of the response from undercanopy vegetation. This is the first hurdle for demonstrating the potential use of Multispectral LiDAR technology for forest mapping applications. If it were not possible to identify spectral variation within the vertical forest canopy, then this technology would be unable to contribute to the range of potential applications.

There are two key areas for further development and testing, that arose from this project. The first is a more thorough validation of the results. Whilst the spectral calibration conducted was successful in this case, there are clear limitations to such an approach. However, at the time that the data was collected no target material was

used with which to correct, and therefore our approach was an acceptable alternative. The ideal scenario would be with the use of target materials, such as a runway, deliberately covered as part of the data collection with a spectral assessment simultaneously conducted on the ground. Similarly, the collection of spectral data from within the forest could be used to validate differences in spectral variation identified within the LiDAR data.

The other area of limitation for the project was the LiDAR system configuration used was not optimal for mapping vegetation as the wavelength lasers used were 532, 1064, and 1550nm. Previous work indicates that the best wavelengths for forest mapping are 531, 650, 660, and 780nm[3,4,7]. Similarly, the data for this project was collected by flying multiple LiDAR systems on the same platform, with two instruments on a first flight, and a third on a second flight. This setup resulted in overlapping coverage of the point clouds at different wavelength, but with non-coincident points. It was fortunate in this case that whilst the points do not lie on exactly the same place in 3D space the spatial sampling was close enough to provide spectral measurements. NDVI should be independent of the energy loss through the canopy but because separate LiDAR systems were used, and the beams were non-colinear, the laser beams had different paths through the canopy, therefore they may have lost different amounts of energy as they attenuate through the canopy, which would in turn have an effect on NDVI. However, the spatial variation appears to be coherent, meaning we were able to assume there were discrete areas of understory with identical properties, allowing us to accommodate for the non-congruous point clouds. An instrument which is able to provide fully coincident measurements would provide a much more accurate result.

Conclusion

The next steps in the development process is to to fully evaluate Multispectral LiDAR technology for mapping understory spectral response. Although significant research has focussed on this, only now that LiDAR manufacturers are producing operational Multispectral instruments are we able to conduct a real-world evaluation of their effectiveness. As the work presented here demonstrates, spectral variation in the understory is possible to identify. This is significant for forest managers globally as many issues relating to tree health and invasive species are first evident in the undercanopy, which is inaccessible from discrete return and/or single-wavelength LiDAR, or passive remote sensing. A follow-up project is being conducted this summer, in collaboration with Optech, using their Titan system, and Riegl, using the LiDAR “constellation” setup. LiDAR data will be collected in Canada and Austria in May and June 2015, with fieldwork carried out concurrently. The goal is to back-up the findings of this project, in that airborne Multispectral LiDAR can be used for mapping the forest undercanopy. By demonstrating this principle we have begun to open up many new avenues for potential applications. This could include mapping of tree health, or mapping of invasive species; for example by building upon laboratory work which has demonstrated that the spectral signature of Rhododendrons, a particularly problematic invasive species in the UK, can be distinguished from other shrub bushes[5].

Ultimately a true Multispectral LiDAR will provide a point cloud where each point is recorded at multiple wavelengths. To do this, manufacturers will have to make a system where the beams overlap exactly and the returns are measured simultaneously, and consistent calibration across the different wavelengths must be maintained. Combined with a LiDAR configuration optimised for vegetation mapping, spectral information will be available for all aspects of a forest canopy, not just the very top surface. With these exciting advances in the use of Multispectral LiDAR for mapping forest under-canopies, the future is bright for this technology.

Acknowledgements

The authors would like to thank James Van Rens, Riegl USA CEO, for the provision of the data used for this project.

References

1. Briese, C., Pfennigbauer, et al. (2012). Radiometric Calibration of Multi-Wavelength Airborne Laser Scanning Data. ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., I-7, 335-340
2. Briese, C., Pfennigbauer, et al. (2013). Multi-Wavelength airborne laser scanning for archaeological prospection. Proc. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Strasbourg, France, September, 2013
3. Morsdorf, F., Nichol, C., et al. (2008). Modelling Multispectral LiDAR vegetation backscatter - assessing structural and physiological information content. Silvilaser proceeding, Silvilaser, 257-265
4. Morsdorf, F., Nichol, C., et al. (2009). Assessing forest structural and physiological information content of Multispectral LiDAR waveforms by radiative transfer modelling. *Remote Sensing of Environment*, 113, 10, 2152-2163
5. Moussavi, M.S., Abdalati, W., et al. (2014). Applicability of an automatic surface detection approach to micro-pulse photon-counting LiDAR altimetry data: implications for canopy height retrieval from future ICESat-2 data. *International Journal of Remote Sensing*, 35, 13, 5263-5279
6. Taylor, S.L., Hill, R.A., and Edwards, C., (2013). Characterising invasive non-native Rhododendron ponticum spectra signatures with spectroradiometry in the laboratory and field: Potential for remote mapping. ISPRS Journal of Photogrammetry and Remote Sensing, 81, 70-81
7. Wallace, A., Nichol, C., and Woodhouse, I.H. (2012). Recovery of forest canopy parameters by inversion of multispectral LiDAR data. *Remote Sensing*, 4, 2, 509-531
8. Woodhouse, I.H., Nichol, C., et al. (2011). A multispectral Canopy LiDAR Demonstrator Project. *IEEE Geoscience and Remote Sensing Letters*, 8, 5, 839-843

Regeneration detection by 3D segmentation in a temperate forest using airborne full waveform Lidar data

Nina Amiri[1], Wei Yao[1], Marco Heurich[2] & Peter Krzystek[1]

[1]*Department of Geoinformatics, Munich University of Applied Sciences, 80333 Munich, Germany, (n.amiri, wei.yao, peter.krzystek)@hm.edu*
[2] *Bavarian Forest National Park, 94481 Grafenau, Germany, marco.heurich@npv-bw.bayern.de*

Highlights: A novel 3D segmentation method is adapted to detect the regeneration coverage density in the understory level of a temperate forest using airborne full waveform Lidar. The provided segmentation algorithm detects up to 70% of the regeneration correctly in comparison with field measurements.

Key words: *Regeneration, 3D segmentation, normalized cut, Full waveform Lidar.*

Introduction

Forest understory distribution and regeneration coverage are important factors for sustainable forest management. Forest inventory at understory level is previously based on the very limited field sampling data which is then extrapolated across the entire forest area. One of the major disadvantages of this approach is that the limited number of plots often covers less than a few percentage of the total forest and this can influence how representative the data may be for the forest strata [1]. Moreover, the forest understory covers are in the permanent spatial and temporal changes. The analysis of forest regeneration requires accurate and continuously updated field data from real forest sites, which is difficult and time consuming to obtain. There is a growing research trend towards the techniques of "forest understory mapping" which results in detection of regeneration vegetation structure with realistic 3-dimensional (3D) approaches [2], including the potential application of remote sensing technology.

Remote sensing technology can provide objective, cost effective and practical solutions for developing and maintaining automated forest analysis. In particular, Lidar has been widely used in mapping the earth's surface and forest structures. Techniques for single tree identification from Lidar data have been investigated in both overstory and understory levels to identify the important parameters such as DBH, crown volume, Biomass and etc. for mapping purposes [2-3-4]. Airborne full waveform Lidar can more efficiently recover backscattered pulses within the travel path of a laser beam by waveform decomposition and give a higher spatial point density, and even information about vertical structure of understory and reflectional attributes [2-5]. The potential information contained in full waveform Lidar data can be utilized to improve the analysis of the regeneration in understory canopy with poor surface response [6].

In this research we aim to use a clustering-based 3D segmentation in combination with airborne full waveform Lidar data for improving the detection of regeneration cover in the temperate forest (below 5 m tree height). The core of our 3D segmentation method is a normalized cut segmentation in combination with the mean shift clusters and provides an effective way to detect the forest understory coverage density. Our results show how the detection of regeneration across datasets with different properties are achieved using the provided method.

Materials and methods

Materials

Experiments are organized in the Bavarian Forest National Park as a temperate forest during three time periods (2009, 2011 and 2012) which is located in the south-east part of Germany along the border to Czech Republic. From the study area in 2012, 23 circular sample plots with an approximate area size of 500 m², from the 2011 a single plot with the area of 920 m² and from the 2009 a single plot with an approximate area of 2450 m² are selected. These plots contain a mixture of mountain forest, spruce forest and beech forest with various regeneration coverage densities. The coverage of the regeneration in these plots range from 1 to 50 % in deciduous forest type and 1 to 40% in the coniferous. Regeneration reference data for 2012 has been collected for sample plots spread over the study area by ocular assessments. For 2011 and 2009 the reference data have been acquired from TLS (Terrestrial laser scanning) data.

Airborne full waveform Lidar data has been collected by Milan Flug GmbH with the Riegl 680i (Full wave) scanner at leaf-on condition in July 2012, by Riegl 560Q with leaf-off foliage condition in 2011 and in 2009 by

Riegl 560Q with leaf-on condition. The obtained point density is determined by the pulse repetition frequency (PRF), flying height, flying speed and strip overlap [7]. The flying height for the current experiment in 2012 was 650 m leading to an average point density of 30 points/m². In 2011 and 2009 the flight height was 400 m with a point density of 25 points/m². The calibration of the pulse intensity in full waveform Lidar data is achieved by using the intensity of the emitted Gaussian pulse with help of a calibration flight in the airport [2].

Regeneration detection

The basic idea of our study is to use the 3D segmentation algorithm, which was initially designed for single tree detection, to estimate the coverage density of the regeneration. The procedure is defined by the following steps: (i) Decomposition of full waveform data, (ii) local tree maxima filtering, (iii) mean shift clustering, (iv) feature derivation for mean shift clusters, (v) normalized cut segmentation and (vi) classification of the segmentation results based on a tree height threshold [2]. The mean shift clustering algorithm is a nonparametric, feature-space clustering technique which neither requires prior knowledge of the cluster number nor constrains the shape of clusters [2]. The most computationally expensive component of the mean shift clustering is related to the identification of neighbouring points in 3D data space which is defined as the kernel bandwidth (h). This parameter affects the regeneration segmentation procedure, therefore this control parameter has to be set in advance by empirical experiments ($h = 2.4\ m$). Also, the normalized cut segmentation control parameters such as normalized cut threshold and minimum number of clustering points for the current study area are empirically defined.

The results of the 3D segmentation are classified by removing those ones with tree height more than 5 m. These filtered segments contain laser points which correspond to the regeneration trees. The threshold applied to the segmented trees excludes point clouds which belong to dominating trees of the upper layers and are, hence, not be considered as regeneration. The disadvantage of a direct height threshold applied to the point cloud would be a high commission error rate in the regeneration coverage estimation.

Results and discussion

Regeneration detection method was applied to all the selected plots in a batch procedure. Figure 1 summarizes the estimated ratio of regeneration for the entire plots under different coverage density. The results are evaluated by comparing them with the ground reference data of regeneration coverage in real forest sites. It can be seen that in the plots which are covered by a high regeneration density (see the blue polyline in Figure 1), such as plot 12-P14, 12-P8, 12-P7 and 12-P12 of 2012 and the single plot of 2011, we detected more than 60% of the regeneration correctly compared to the ground reference data. The accuracy in detection seems to be correlated to the foliage condition of overstory canopy since the lowest detection accuracies happened in the plots with high upper canopy density and complexity such as plot 12-P4 and 12-P17 (graph is not shown here).

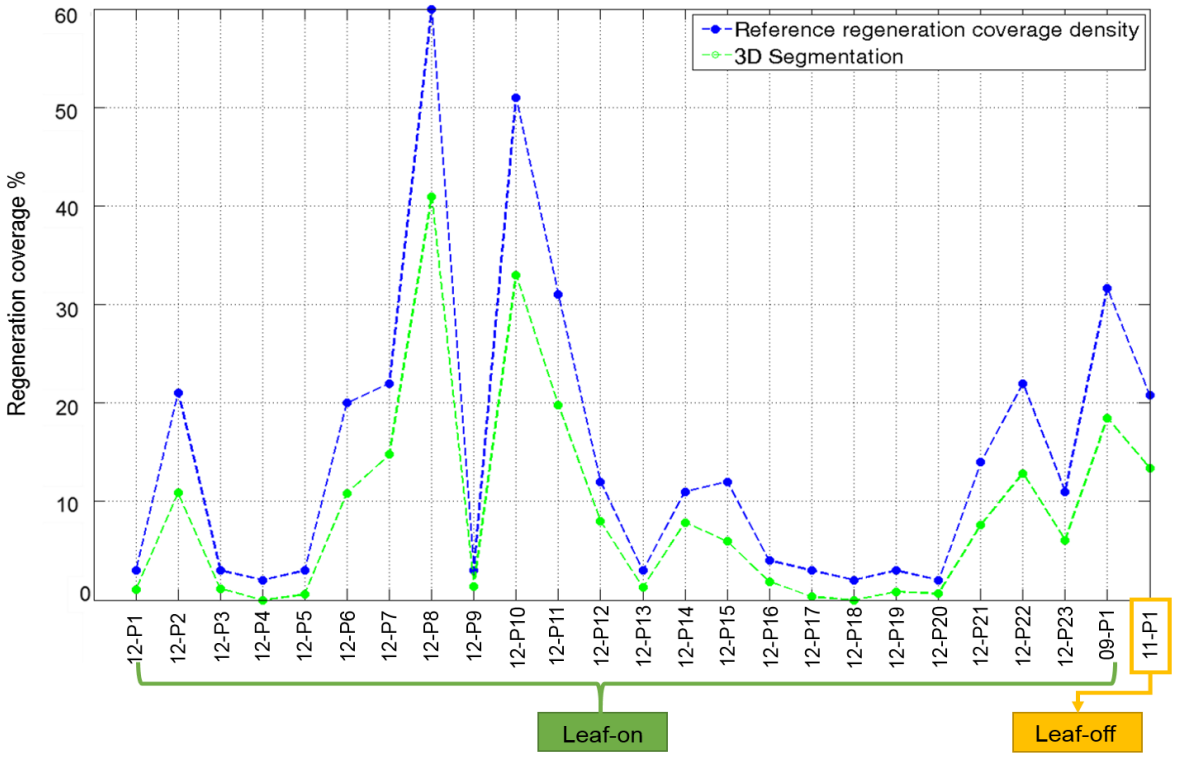


Figure 1: Regeneration coverage estimation by 3D segmentation vs. ground reference. (Note: 12-P represents the plots in 2012 (leaf-on), 09-P the plot in 2009 (leaf-on) and 11-P in 2011 (leaf-off)).

As expected, we can see in Figure 1 the lowest detection rate of ca.1% belonging to the plots with low coverage density of regeneration. It turned out in the 3D visualization of the plots with low regeneration coverage density (graph is not shown here), such as 12-P4, 12-P18 and 12-P5, that these plots had denser tree crowns in the upper canopy layer compared to others. The visualization of these sample plots shows that there are no point clouds at the understory level.

The developed detection method [2] can be seen as an alternative approach with high potentials for understory coverage estimation. Our experiments, which used the plots with various coverage density and characteristics following the previous studies [2-7] confirmed that full waveform laser scanner data have a potential to detect the reflections from the understory canopy cover. By looking into the results of coverage estimation in the forest understory level we demonstrated that it is possible to detect more than 60% of the reference regeneration in plots with low upper canopy coverage. The young trees with the tree height below 5 m in a multi-layered forest have a higher chance to be detected due to the waveform decomposition of full waveform Lidar data. However, the overstory canopy layer in leaf on situation is a main factor to prevent the laser beam from reaching to the understory layer. Therefore, in most of the plots with ca. up to 2% regeneration coverage we did not have enough point clouds to construct a regeneration segment in the algorithm.

The complex canopy and mixed branches reduce the accuracy of the regeneration detection in the segmentation especially in the deciduous forest type such as plots 12-P15 and 12-P17. Therefore, the performance of the segmentation algorithm for regeneration detection is directly related to the forest characteristics and properties such as canopy density and tree species distribution.

Conclusion

The current research deals with the regeneration detection in a mixed and multi-layered temperate forest by applying a 3D segmentation method to airborne full waveform Lidar data. The method can be viewed as a coverage estimation and analysis for regeneration areas. Additionally, our study shows that the detection accuracy still depends on the plot properties such as canopy density. In leaf-on condition, an increase of 10% in the number of trees in the upper canopy layer decreases the detection accuracy of regeneration about 9% (graph is not shown here). Further improvements of method could be achieved by extending the 3D segmentation algorithm with a classification of regeneration into deciduous and coniferous species.

References

[1] Köhl, M., Magnussen, S. S., & Marchetti, M. (2006). Sampling Methods, Remote Sensing and GIS Multiresource Forest Inventory. (1st ed.). Springer Science & Business Media, (Chapter1).

[2] Yao, W., Krzystek, P., & Heurich, M. (2013). Enhanced detection of 3D individual trees in forested areas using airborne full-waveform LiDAR data by combining normalized cuts with spatial density clustering. International Archives of Photogrammetry and Remote Sensing and Spatatial Information Sciences, II–5/W2, 349–354.

[3] Korpela, I., Ørka, H. O., Maltamo, M., Tokola, T., & Hyyppä, J. (2010). LiDAR – Effects of Stand and Tree Parameters , Downsizing of Training Set, Intensity Normalization, and Sensor Type. The journal of Silva Fennica, 44, 319–339.

[4] Heurich, M. (2008). Automatic recognition and measurement of single trees based on data from airborne laser scanning over the richly structured natural forests of the Bavarian Forest National Park. Forest Ecology and Management, 255, 7, 2416–2433.

[5] Stilla, U., Yao, W., & Jutzi, B. (2004). Detection of Weak Laser Pulses by Full Waveform Stacking. International Archives of Photogrammetry and Remote Sensing and Spatatial Information Sciences, 36, 2000, 25–30.

[6] Yao, W., & Stilla, U. (2010). Mutual Enhancement of Weak Laser Pulses for Point Cloud Enrichment Based on Full-Waveform Analysis. IEEE Transactions on Geoscience and Remote Sensing, 48, 9, 3571–3579.

[7] Reitberger, J., Schnörr, C., Krzystek, P., & Stilla, U. (2009). 3D segmentation of single trees exploiting full waveform LIDAR data. ISPRS Journal of Photogrammetry and Remote Sensing, 64, 6, 561–574.

LiDAR-based characterization of understory trees in a complex temperate forest in Ontario, Canada

Karin Y. van Ewijk[1], Paul M. Treitz[1], Neal A. Scott[1]

[1]*Department of Geography, Queen's University, Kingston, Ontario, K7L 3N6, Canada. ekyv@queensu.ca; paul.treitz@queensu.ca; neal.scott@queensu.ca.*

Highlights: Understory information is often lacking in forest resource inventories, yet is essential for biodiversity and habitat assessments. We used *k*-NN imputation and LiDAR metrics to generate understory information for a complex forest. Imputation results were moderately correlated with tree heights, DBH, and basal area. However, composition and density were poorly predicted.

Key words: *Full waveform LiDAR, k-Nearest Neighbor imputation, understory tree species.*

Introduction

The characterization of overstory vegetation is often the primary focus of forest resource inventories (FRI). However, understory vegetation, comprised of seedlings and saplings, suppressed juvenile trees, and shrub species, is critical for the preservation of biodiversity and wildlife habitat and plays an important role in forest ecosystem function. The understory vegetation often constitutes a major component of the plant species diversity in temperate forests. Understory trees also provide essential shelter, nesting, and food resources for many wildlife species and are indicative of habitat availability and quality [1]. As they constitute a new adult cohort they can have a major influence on which tree species are successful and hence are important for evaluating long-term sustainable forest management practices. Key stand-level characteristics of understory vegetation are height and diameter-at-breast height (DBH) distributions, stem density and species composition.

Remote sensing can provide landscape wide vegetation information, however, understory vegetation patterns are often more difficult to detect because of shading and occlusion of the understory vegetation by the upper canopy trees, especially in mature forests with a multilayered or closed canopy [2]. Over the last decade Light Detection and Ranging (LiDAR) has emerged as an important active remote sensing technology that can capture numerous three-dimensional forest characteristics including some of these understory vegetation attributes (e.g., [2, 3-5]). More recently, several studies [6] have indicated that full-waveform (FWF) LiDAR may be able to improve understory vegetation characterization because of its much higher pulse density and the ability to derive additional waveform characteristics (e.g. width, amplitude and intensity of the echo).

Based on this premise, our objective was to evaluate the ability of FWF LiDAR to estimate several understory vegetation characteristics with the purpose of generating stand- and tree-level understory information that could be included in an enhanced FRI (eFRI) and could serve as input data for an empirical forest growth model. In addition, we examined which FWF LiDAR metrics are important to estimate these understory vegetation characteristics The anticipated stand-level understory information derived for an eFRI may allow for biodiversity and habitat assessments while inclusion of understory vegetation information in empirical forest growth models would allow for the forecasting of long-term forest ecosystem dynamics, thereby providing for future biodiversity and habitat assessments.

Materials and Methods

Study area and data collection

Our study site is the Petawawa Research Forest (PRF) (45^{o} 57' N, 77^{o} 34' W) which is situated along the Ottawa River, northwest of Ottawa, Canada and lies within the Great Lakes – St. Lawrence Forest Region. This forest is characterized by tree species such as eastern white pine (*Pinus strobus* L.), red pine (*Pinus resinosa* Ait.), eastern hemlock (*Tsuga canadensis* L.), sugar maple (*Acer saccharum* Marsh) and yellow birch (*Betula alleghaniensis* Britton).

During the summers of 2007, 2009 and 2010, we established 63 circular sampling plots (11.28 m radius; 400 m^2) and sub-plots (3.99 m radius; 50 m^2) in forest stands (average tree height $\geq$ 5 m) across our study area using a stratified random sampling design. In the sub-plots we collected information on tree species, height and DBH for all trees with a DBH < 9 cm, i.e., the understory trees. The centre location of each plot was georeferenced with a Trimble® Geo XT™ GPS unit with external antenna and later differentially corrected to achieve sub-meter accuracy.

FWF LiDAR data were acquired for the entire study area in August 2012 with a Riegl Q680i sensor mounted on a Cessna 172 (Leading Edge Geomatics Ltd., Lincoln, NB). The system was flown at approximately 750 m above ground level (a.g.l.), with a 20^{o} scan half angle, a scan frequency of 77 Hz, and a pulse repetition frequency of 150 kHz. Leading Edge Geomatics Ltd. discretized the FWF data to a nominal sampling density of 48 pulses m^{-2} and provided the data in LAS format classified into ground and vegetation returns.

Characterizing understory trees: k-NN imputation modelling

Spatial predictions of the understory characteristics were performed using a *k*-Nearest Neighbour (*k*-NN) imputation method. Imputation methods are popular for forest attribute estimation because of their ability to simultaneously relate multiple *in-situ* measured forest attributes to ancillary, in this case LiDAR-based, metrics [7]. When using a single neighbour (i.e., 1-NN) approach, it is possible to generate both stand-level and tree-level forest information for un-sampled plots.

We initially selected six understory tree variables for our multivariate *k*-NN imputation model: average and top understory tree height (AHGT and THGT), average understory DBH (ADBH), understory stem density (SD), relative basal area of the dominant understory tree species ($RBA_{DOMSPECIES}$), and dominant understory tree species. We used the Random Forest (RF) proximity metric as our distance metric and used a variable selection procedure to select the optimal LiDAR metrics for each of the response variables. The accuracy of the imputed understory tree characteristics was determined via Pearson's correlation coefficient (r) and the Root Mean Squared Difference (RMSD), a measure of model error for *k*-NN imputation models. Predictor variable importance was also assessed.

Our set of predictor variables for the imputation of understory trees included 23 FWF LiDAR-based metrics which were grouped into statistical, canopy height, canopy complexity and terrain predictors and had a spatial resolution of 20 m. The selection of our set of predictor variables was based on other studies conducted within PRF and studies focused on characterizing the understory vegetation using LiDAR metrics (e.g., [1, 4]). Understory structure and composition is often influenced by the density of the overstory and topography [1]. In addition, [4] found that last return and lower fraction/density LiDAR metrics were strongly correlated with the quantity of young trees. Hence we included LiDAR-based canopy height and complexity metrics, which are reflective of understory, mid-, and upper-canopy structural characteristics, and terrain metrics. The canopy height metrics also included several strata metrics, where each stratum was one meter high and the standard deviation and skewness of the LiDAR returns were calculated within each stratum.

Results and Discussion

Our final multivariate *k*-NN estimates of AHGT, THGT, ADBH and $RBA_{DOMSPECIES}$ were moderately correlated ($0.5 \geq r \leq 0.7$) with our validation data (Table 1). RMSD was relatively small for the tree height and DBH response variables and largest for the $RBA_{DOMSPECIES}$ response variable (Table 1). Both discrete and FWF LiDAR data have been used to characterize understory vegetation but often with less success compared to the characterization of upper canopy properties [2]. Several studies have shown moderate to good correlations for the number and mean height of understory trees in boreal forests using LiDAR-based height metrics [2,3] and have successfully mapped understory shrubs using LiDAR-based terrain and canopy height metrics [1]. However, tree species discrimination based on LiDAR-based height and intensity metrics has been more difficult [2,5]. For example, [5] found that the classification of the understory layer was problematic in complex plots due to occlusion by taller vegetation. Instead, they proposed the use of echo width rather than intensity for the classification of understory species composition or the use of a multi-spectral canopy LiDAR instrument [5]. Most of these studies also concluded that the density of the dominant tree layer influenced the accuracy of their results. Our results are consistent with most of these findings as we had moderate success in estimating understory tree height and DBH and were unsuccessful in discriminating between tree species. However, understory tree stem density in our study site was also poorly predicted which may be explained by the higher density and multilayered structure of the upper canopy in our study area. Our final *k*-NN model therefore excluded understory tree stem density and species composition because of poor model performance. However, both are relevant information within an eFRI and are required if the purpose is to generate an understory tree-level list that can be used as input in a forest growth model.

In terms of predictor variables, the final *k*-NN imputation model was developed using 16 out of 23 FWF LiDAR predictor variables. Each of the imputed response variables was predicted by a different set of LiDAR metrics determined through a variable selection procedure. The four most important predictor variables were all canopy height metrics, i.e., P50, $Stratum2_{sk}$, $Stratum4_{sk}$, and $Stratum6_{sk}$ (Table 1). The canopy complexity metrics, VCI and VDR, were used to predict understory tree height and DBH, and the terrain metric slope was used to predict tree height and $RBA_{DOMSPECIES}$ (Table 1). The choice and influence of LiDAR metrics within our *k*-NN imputation model (Table 1) indicate that metrics such as the height percentiles and lower strata metrics are important in predicting understory tree species characteristics. Similar LiDAR metrics were relevant in analyzing the quantity and vitality of juvenile trees in a boreal forest in Norway [4]. In addition, the establishment and presence of understory trees in a stand is often related to canopy openings, the presence of tall trees for seed sources, and soil factors such as temperature and humidity [4]. In this context it makes sense that canopy complexity and terrain metrics were also included in our *k*-NN model.

In terms of the *k*-NN methodology, the choice and effectiveness of distance metrics is very site specific and it is unclear as to whether predictor variable selection should be performed or not [7]. The use of *k*-NN imputation methods to predict forest inventory attributes does place a very high requirement on the reference data, as these should cover the full range of values of the response variables. This requirement may be difficult to achieve when multiple response variables are included. In our study, the RF *k*-NN approach seemed to give the best imputation results and a reduction of the predictor variable set did improve model results. The moderate understory trees imputation results in our study are likely, in part, a result of having a relatively small dataset to analyze.

Table 1: LiDAR predictor variables and evaluation statistics for the understory tree *k*-NN imputation model (Y_{AHGT}, Y_{THGT}, Y_{ADBH}, $Y_{RBADOMSPECIES} = f(X_{AHGT}, X_{THGT}, X_{ADBH}, X_{RBADOMSPECIES})$). Predictor variables with a high importance are listed in bold.

Understory tree variables	LiDAR predictor variables	r	RMSD
AHGT (m)	**P50, Stratum2$_{skew}$**, Stratum3$_{skew}$, Stratum5$_{sd}$, VCI, VDR, Slope	0.69	0.79
THGT (m)	P30, **P50**, Stratum5$_{sd}$, VDR, slope	0.50	1.05
ADBH (cm)	P01, **P50**, **Stratum4$_{skew}$**, Stratum5$_{skew}$ and $_{sd}$, **Stratum6$_{skew}$**, VCI	0.53	0.85
$RBA_{DOMSPECIES}$ (%)	P05, P30, **P50**, P70, P90, Stratum3$_{skew}$, Stratum5$_{skew}$ and $_{sd}$, PLC, Slope	0.53	20.65

AHGT = average juvenile tree height; THGT = top juvenile tree height; ADBH = average diameter at breast height; $RBA_{DOMSPECIES}$ = relative basal area of the dominant juvenile tree species; Px = x^{st} height percentile; Stratumx_{skew} = skew of non-ground returns > x-1 m and ≤ x m; Stratumx_{sd} = standard deviation of non-ground returns > x-1 m and ≤ x m; VCI = vertical complexity index; VDR = vertical distribution ratio; PLC = plan curvature (across the slope); r = correlation coefficient, RMSD = root mean squared difference

Conclusion

The use of remotely-sensed data to generate understory tree information within the context of an eFRI or as input to forest growth models would benefit landscape-wide assessment of biodiversity, habitat quality and availability. To be relevant for biodiversity and habitat analyses, however, not only are structural attributes required, but information on species composition of the understory trees is necessary. The latter, remains a significant challenge. Accurate information on juvenile and minor tree species located beneath the upper canopy is difficult to obtain as LiDAR returns from the lower layers in the canopy tend to be scarce and noisy [4,6]. FWF LiDAR, with its higher pulse density, may alleviate some of these issues. Our study demonstrates that the approach we used still requires substantial improvements. As a next step, the inclusion of LiDAR-based intensity metrics and other information from FWF LiDAR, such as width, amplitude and intensity of the echo may provide further discrimination of understory attributes.

Acknowledgements

Support for this project was provided by the Ontario Centre of Excellence for Earth and Environmental Technologies, the Natural Sciences and Engineering Research Council (NSERC), and the Premier's Research Excellence Award (PREA). We also gratefully acknowledge the contributions of the Forest Research Partnership (Tembec, Inc., Ontario Ministry of Natural Resources, and Canadian Forest Service). Special thanks go to the Petawawa Research Forest staff for their assistance at the research site and to Murray Woods. We also thank our field assistants: Melissa Fredrigo, Nick Gralewicz, Stephanie Gagliardi, Anne Hagermann, Fraser McLeod and Neil Bodimeade.

References

[1] Martinuzzi, S., Vierling, L.A., Gould, W.A., Falkowski, M. J., Evans, J. S., Hudak, A. T., & Vierling, K. T. (2009). Mapping snags and understory shrubs for a LiDAR-based assessment of wildlife habitat suitability. Remote Sensing of Environment, 113(12), 2533–2546.

[2] Korpela, I., Hovi, A., & Morsdorf, F. (2012). Understory trees in airborne LiDAR data - Selective mapping due to transmission losses and echo-triggering mechanisms. Remote Sensing of Environment, 119(1), 92–104.

[3] Maltamo, M., Eerikäinen, K., Pitkänen, J., Hyyppä, J., & Vehmas, M. (2004). Estimation of timber volume and stem density based on scanning laser altimetry and expected tree size distribution functions. Remote Sensing of Environment, 90(3), 319–330.

[4] Bollandsås, O. M., Hanssen, K. H., Marthiniussen, S., & Næsset, E. (2008). Measures of spatial forest structure derived from airborne laser data are associated with natural regeneration patterns in an uneven-aged spruce forest. Forest Ecology and Management, 255(3-4), 953–961.

[5] Morsdorf, F., Mårell, A., Koetz, B., Cassagne, N., Pimont, F., Rigolot, E., & Allgöwer, B. (2010). Discrimination of vegetation strata in a multi-layered Mediterranean forest ecosystem using height and intensity information derived from airborne laser scanning. Remote Sensing of Environment, 114(7), 1403–1415.

[6] Maltamo, M., Næsset, E., & Vauhkonen, J. (2014). Forestry Applications of Airborne Laser Scanning: Concepts and Case Studies, Managing Forest Ecosystems 27. Dordrecht: Springer Science+Business Media.

[7] Hudak, A. T., Crookston, N. L., Evans, J. S., Hall, D. E., & Falkowski, M. J. (2008). Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. Remote Sensing of Environment, 112(5), 2232–2245.

A benchmark of single tree detection methods using data from alpine forests

Lothar Eysn[1], Jean-Matthieu Monnet[2] and Markus Hollaus[1]

[1]*Vienna University of Technology, Department of Geodesy and Geoinformation, Research Group Photogrammetry, Gußhausstraße 27-29, A-1040 Vienna, Austria (Lothar.Eysn, Markus.Hollaus@geo.tuwien.ac.at)*
[2]*Irstea, UR EMGR Écosystèmes Montagnards, centre de Grenoble, F-38402 Saint-Martin-d'Hères, France (jean-matthieu.monnet@irstea.fr)*

Highlights: This study presents a new open-access dataset of 18 plots from alpine forests of the Alpine Space. Eight detection algorithms were tested and evaluated against forest inventory data using a novel, automated matching procedure. Forest structure remains a key issue limiting tree detection, and algorithms would probably benefit from an adaptive tuning in order to achieve a better trade-off between omission and commission errors.

Key words: *Tree extraction, comparative testing, mountain forest.*

Introduction

Area-based methods provide statistically calibrated maps of forest stand parameters such as growing stock, stem density and stand height, and are used for large-area forest inventory (FI) and long term forest management planning. Meanwhile, in complex alpine forests, single tree information is highly valuable. The spatial distribution of trees and their characteristics are important inputs for silviculture in uneven-aged stands, for growth simulation models or for the evaluation of the forest protection effect against rockfalls.

Numerous tree detection methods have been published in the literature, but benchmarks, i.e. comparison of different methods' performance on various forest stands are scarce. A study of Kaartinen, *et al.* [1] is mainly focused on Scandinavian forests and Vauhkonen, *et al.* [2] used very different forest types. The comparison of different algorithms and their evaluation on various forest types remain difficult, partly because of the absence of an open-access dataset with both, the airborne laser scanning (ALS) data and corresponding FI data. During the NEWFOR project, funded by the European Territorial Cooperation, a dataset of 18 mountain forest plots from 8 areas in 5 countries was constituted. Eight single tree detection algorithms were tested and evaluated with a novel, automated matching method [3]. The method uses a restricted nearest neighbouring approach to automatically link the detection results to the ground truth FI data. This article has three objectives. First, to present the dataset, which is now publicly available [4]. Second, to present the automated evaluation method, and third, to present the benchmark results.

Material and methods

Dataset

The ALS dataset is constituted of eight different surveys. It is somehow representative of the currently available operational data within the alpine countries. Data are heterogeneous as they originate from different sources, who acquired the data for different purposes. Point densities range from 5 to 121 m^{-2}, and four different types of scanners are used. Within these eight surveys, 18 forest plots are available, with stem locations, diameters, heights and species. The total surveyed area is 4.5 ha. Basal area among the plots ranges from 15.5 to 68.1 $m^2.ha^{-1}$, and mean height from 13.7 to 36.7 m. Coniferous proportion is between 23 and 100%. The plots were classified in four forest types: Single-Layered Mixed forest (SL/M, 4), Single-Layered Coniferous forest (SL/C, 5), Multi-Layered Mixed forest (ML/M, 7) and Multi-Layered Coniferous forest (ML/C, 2).

Detection algorithms

For each plot, the ALS data and rasterized terrain models at 0.5 and 1 m resolution were provided to benchmark participants. Participants applied their fully-automated tree detection algorithms in order to output a list with tree coordinates and heights for each plot. In total eight algorithms are tested. Most are based on local maxima (LM) detection using a rasterized canopy height model (CHM) and one relies on 3D point cloud processing. The LM-based methods use different approaches as for example filtering approaches, region growing, usage of multiple derived CHMs in different height layers, watersheds or polynomial fitting. Different settings as for example kernel size or spatial resolution of the input CHMs can be found. The method which is

based on the 3D point cloud uses a segmentation and ellipsoidal clustering approach. The methods are applied to the data without any additional knowledge about the local forest (i.e. species composition).

Performance assessment

The evaluation of the detection results is carried out in a reproducible way by automatically matching them to precise in-situ FI data using a restricted nearest neighbouring approach. The matching algorithm is implemented in the OPALS forestry package [5]. Starting from the highest detected tree the restricted nearest neighbouring reference trees within a defined neighbourhood are detected. Restricted nearest neighbouring means that there are height and neighbourhood criterions which need to be fulfilled to match two trees. Finally, trees with the best neighbourhood and height vote are assigned. Only trees inside an area of interest are considered within the matching to overcome possible detection limitations at the borders of the input ALS data.

Based on the matching procedure, quantitative statistical parameters such as percentages of correctly matched trees and omission and commission errors are derived and presented in graphical representations. The benchmarking results are prepared in different levels of information, starting with investigations based on the method and forest type. Additionally an overall performance of the benchmark is evaluated.

Results and discussion

Evaluation of the matching procedure

A random sample of 699 detected trees was manually interpreted to test the developed matching algorithm. Only 3% were wrongly treated by the matching procedure. The overall accuracy is 97% with a Kappa of 0.94. This procedure enables interpreter-independent and reproducible results in short amount of time. Thus, it is an appropriate way to efficiently compare the detection results of several algorithms and for several test plots.

Detection performance

The vertical structure of the forest has a major impact on the detection results of the different methods (Figure 1). Trees in the lower layers are difficult to detect, even for point cloud-based methods, which results in lower assignment rates in the multi-layered stands. Mixed stands are also more troublesome: more complex crowns in deciduous trees are probably the reason for the higher commission rates. The best matching rate was obtained for single-layered coniferous forests. Dominated trees were challenging for all methods. In general, the tree detection rates show a higher variation than estimated tree heights. The overall performance shows a matching rate of 47%, which is comparable to results of other benchmarks performed in the past [1-2].

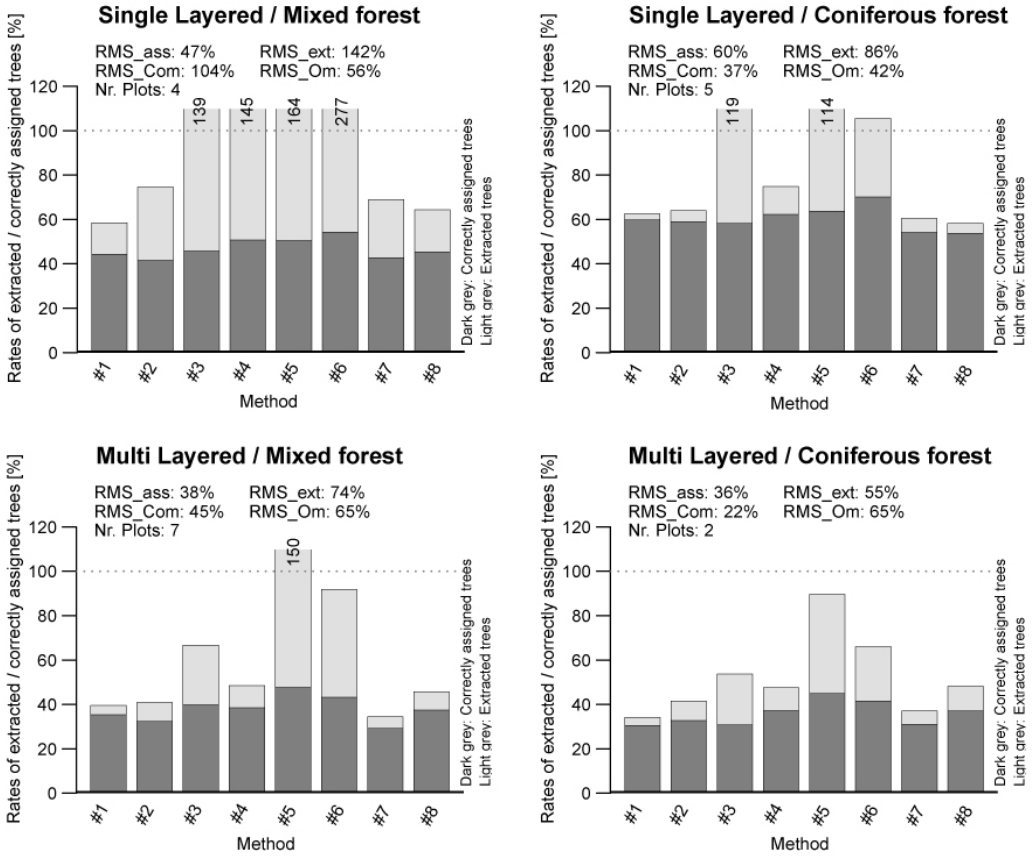


Figure 1: Detection results by forest type.

A method based on local maxima detection within a canopy height model using variable-sized moving windows is rated as the best performing algorithm. A point cloud clustering-based method gained the best results for trees in subdominant layers which is rated as advantage over raster-based methods.

The trade-off between omission and commission errors turns out to be a critical point regarding tree detection. Some methods are probably intrinsically more efficient, but algorithm parameters such as raster resolution, kernel size and horizontal vertical exclusion thresholds have a major impact on detection results. Depending on the forest structure and on the ALS acquisition parameters, the filters required for tree extraction have to be chosen or at least tuned specifically.

Conclusion

The study brings new insight regarding the potential and limits of tree detection with ALS and underlines some key aspects regarding the choice of method when performing single tree detection for the various forest types encountered in alpine regions. It turned out, that complex multi-layered forests are challenging for all tested methods. The best detection results could be obtained for single-layered coniferous forests. In order to improve the detection algorithm available for forest practitioners, it seems important to have datasets, that allow to test the robustness of algorithms on a wide range of forest structures and to design algorithms able to optimize their setting either based on internal (ALS data itself) or external data. Additionally a standardized matching procedure is needed for testing the algorithms in a clear and reproducible way. The presented matching methodology and the published dataset [4] should help to overcome this issue for future studies and applications.

References

[1] Kaartinen, H., Hyyppä, J., Yu, X., Vastaranta, M., Hyyppä, H., Kukko, A., Holopainen, M., Heipke, C., Hirschmugl, M., Morsdorf, F., Næsset, E., Pitkänen, J., Popescu, S., Solberg, S., Wolf, B.M. & Wu, J.-C. (2012). An international comparison of individual tree detection and extraction using airborne laser scanning. Remote Sensing, 4(4), 950-974.

[2] Vauhkonen, J., Ene, L., Gupta, S., Heinzel, J., Holmgren, J., Pitkänen, J., Solberg, S., Wang, Y., Weinacker, H., Hauglin, K.M., Lien, V., Packalén, P., Gobakken, T., Koch, B., Næsset, E., Tokola, T. & Maltamo, M. (2012). Comparative testing of single-tree detection algorithms under different types of forest. Forestry, 85, 27-40.

[3] Eysn, L., Hollaus, M., Lindberg, E., Berger, F., Monnet, J.-M., Dalponte, M., Kobal, M., Pellegrini, M., Lingua, E., Mongus, D. & Pfeifer, N. (2015). A benchmark of Lidar-based single tree detection methods using heterogeneous forest data from the Alpine Space. Forests, 6(5), 1721-1747.

[4] NEWFOR, The NEWFOR single tree detection benchmark dataset. http://www.newfor.net/download-newfor-single-tree-detection-benchmark-dataset/ (last access: March 2015).

[5] Pfeifer, N., Mandlburger, G., Otepka, J. & Karel, W. (2014). OPALS–A framework for airborne laser scanning data analysis. Computers, Environment and Urban Systems, 45, 125-136.

Inventorying the forest: Laser scanning vs close range photogrammetry on a UAV

Udaya Vepakomma[1], Denis Cormier[1]

[1]*FPInnovations, Pointe-Claire, Canada – H9R 3J9*
(udayalakshmi.vepakomma; denis.cormier)@fpinnovations.ca

Highlights: Varying conifer stand structure at the tree level could be well described by point clouds from both image and lidar mounted on a UAV. This has a potential of cost-effective and rapid means of inventorying remote, smaller or scattered areas for tactical and operational management.

Key words: *Unmanned aerial vehicle, multi-sensor, point clouds, tree metrics, canopy*

Introduction

Efficient forest management demands detailed, timely information on forest metrics at several scales – block, stand as well as tree. Time consuming and tedious ground methods have been suitably replaced with remote sensing based assessments, both image and lidar, at fairly large scales [1], albeit not cost efficient when detailed information needs to be gathered in remote, smaller or scattered areas for tactical and operational management. Logistics and low altitude flying to gather high resolution data increases the cost [2].

Recently, imaging and lidar instruments, that have the ability to provide very high temporal and spatial resolution data with similar characteristics as those of full scale platforms, have been deployed on-board small UAVs (Unmanned Aerial Vehicles, drones or remotely piloted aircrafts) [3]. 3D point clouds generated from convergent imagery using 3D vision / modern photogrammetric methods is expected to be dense and as accurate as that of lidar [4]. Due to limited infrastructure requirements, simple logistics, ease of acquiring convergent images or lidar of objects from close range, we presume that UAV platform has potential use for continuous and detailed assessment of small and remote areas of forests.

In this study, we would like to test the feasibility and strength of 3D point clouds generated from both lidar and convergent imagery from a UAV platform in assessing detailed canopy structure of the forest. In particular, here we would investigate tree metrics from both data sources for a softwood stand. Analyses are in progress and only preliminary results are presented at the time of this abstract.

Material and methods

Study area

The study site falls within Petawawa Research Forest (PRF) in Ontario, Canada, that is predominantly white pine mixed with red pine, balsam fir and poplar. Part of the 10.3 ha. selected area had recently been harvested under a shelterwood system for managing white pine leaving an untreated 1 ha. control block of forest. The terrain is rocky but fertile, and is relatively flat.

UAV and data acquisition

The study site was scanned with Yellowscan lidar system and imaged in tandem with Nikon D600 (full-frame) with Zeiss 50 mm lens (RGB) and Tetracam mini MCA-6 (6 multi-spectral channels in 450-1050nm) mounted on Renegade UAV rotor system on 10th Nov., 2014. Due to a very slight shift in mounting of Nikon camera, the area was reflown on 10th Jan., 2015. Yellowscan as well as MCA-6 are independently integrated with internal INS capability and (RTK-ready for lidar) GPS receiver. The 2kg-20W laser scanner has a maximum range of 100 m, can record up to 3 echoes per pulse and provides geopositioned point clouds. The 6 flight lines with a 40% overlap and +/- 50° scan angle were conducted at an altitude of 100m AGL. A total of 414 and 588 images covered the site with Nikon and MCA-6 respectively. Analyses of MCA-6 data are beyond the scope of this paper.

Field data

Three plots of size 20 m X 20 m, 2 in control and 1 in the treated block, were laid. Plot corners were DGPS surveyed (GeoXH; accuracy up to 30 cm). DBH, species, azimuth and the distance from a reference DGPS point were noted for every tree over 8cm DBH. Total tree height (using a Vertex from two vantage points) and DGPS locations of four softwood trees falling in different DBH class were measured within each of 10 m X 10 m

quadrats. Independently, total tree height, DBH and distance between the tree stems of 10 pairs of randomly selected trees (>8 cm DBH) were measured in the cut block.

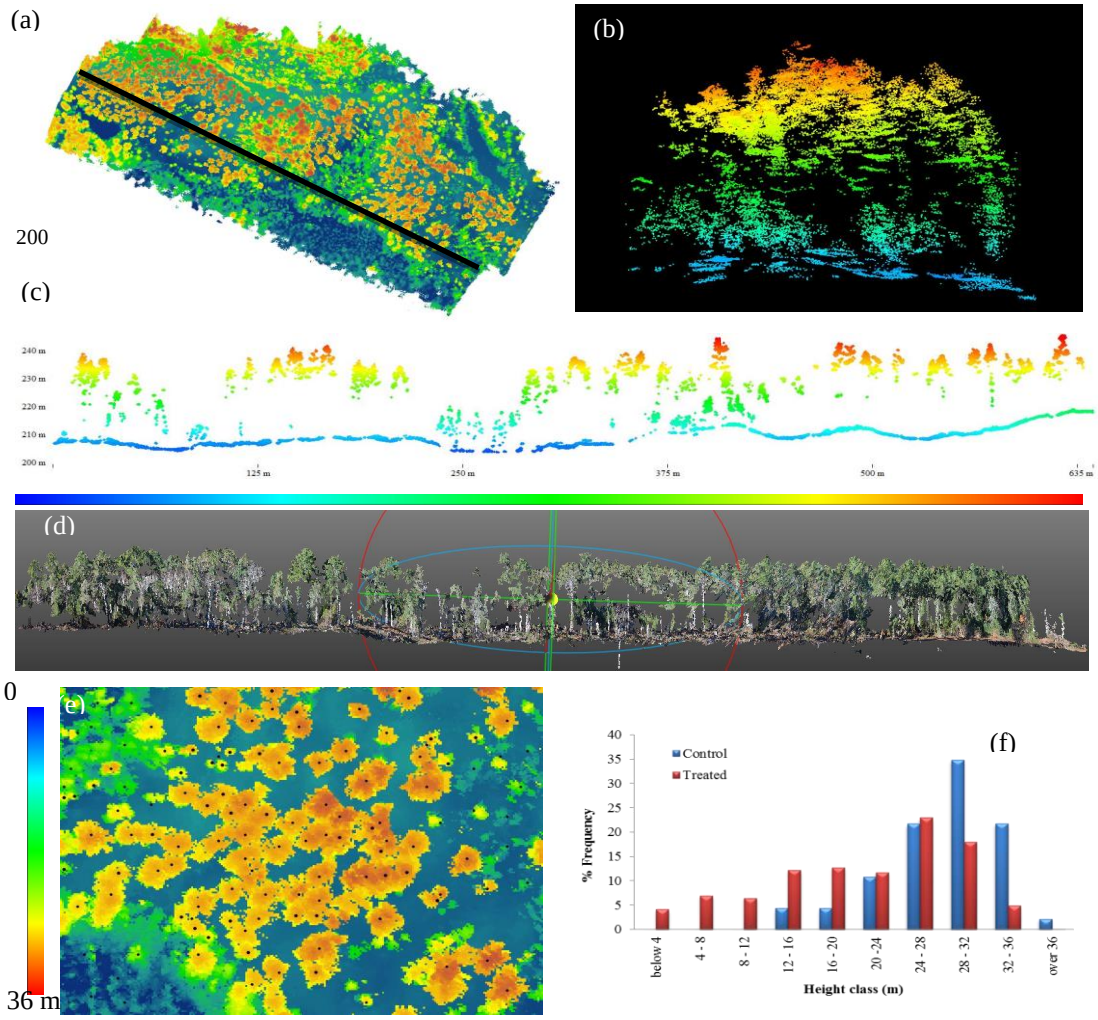


Figure 1: Lidar and image coverage of the 10.3ha. test site. (a) DSM from lidar; (b) Vertical profile from lidar showing emergent and understory growth visible from the dense multi-echo lidar; (c) Vertical profile from lidar along the transect shown in (a); (d) Vertical profile from 3D reconstructed image (e) Single tree detection on lidar CHM (f) Tree height class distribution of the control and treated blocks of the test site

Lidar and Image data processing

Raw lidar data was processed in Global Mapper. Automatic classification into ground and non-ground was manually corrected and validated by inspection. Lidar CHM (canopy height model) was generated by taking the difference in DSM (TIN generated from the highest point for every 5 points of non-ground return with automated optimal grid-spacing) and DEM (TIN generated from the lowest point for every 5 points of ground returns).

All images of Nikon over the study site were used for 3D reconstruction, dense point clouds and finally orthomosaics using Agisoft's Photoscan photogrammetric software. Due to the presence of excess snow over the vegetation causing actual tree height anomalies we used only the November images for this comparison study. The imagery, synchronised GPS position of each image and DGPS positions of reference targets were used as inputs. The processing steps included automatic aero-triangulation, bundle-adjustment, dense 3D point cloud generation and ortho mosaicking. Image DSM is created using the highest point in every 5 points and automatic optimization of grid-spacing. For this study, image-CHM is the difference in image DSM and lidar-DEM.

Single tree height extraction on lidar and image surfaces

Extending the methodology used in [5], the location of individual tree tops was determined based on local maxima filter with a circular non-overlapping (moving) window of a radius of 35 cm (based on field observation) on the Gaussian filtered CHM (twice filtered). Tree height was estimated as difference of the DSM height at the extracted peak and DEM.

Results and discussion

Point cloud – lidar vs image

The structure of individual trees and undergrowth as well as underlying terrain of the dense undisturbed control block (Fig 1b) and the test site in general (1c) is well captured by the lidar point cloud even at the maximum range (100 m) of the lidar system. All return density was 23 hits /m^2 (30 in the control block). Preliminary classification showed that there are sufficient returns from the ground (over 1%) even in highly dense multi-storied canopy to derive a DEM, but a significant number of the first returns (over 40%) came from ground in the treated block. The point cloud extracted from the convergent images with 60% overlap between flight lines (Fig 1d) provided a dense representation of the stand structure. In many places the ground as well as tree stems are also clearly visible. In general, the point density is very high at 900 points /m^2 with about 15% representing the ground in the treated block and 1% to almost none in the control block.

Detection of trees - lidar vs image

Visually, excepting undergrowth, trees of all sizes are better detectable using image-point clouds compared to lidar. Preliminary results of the automated single tree detection algorithm on the CHMs showed that both methods are equally good (Fig. 1(e)), with no omission or commission errors for trees over 7 m, but 5% commission errors on lidar-CHM and 2.7% omission on image-CHM for trees below 1 m. All the 20 field measured random trees were detected on both data sets. It may be noted that the automated single tree algorithm was validated in a different forest [5], its improvement or accuracy assessment has not been attempted in the present study.

Total tree height and height distribution using lidar

Height class distribution (Fig. 1(f)) based on the single tree detection (Fig. 1(e)) realistically represents the control and treated block. Comparison of the randomly measured trees in the field and detected trees on lidar shows a good correspondence within 20 cm error in tree height estimation (Table 1). However, overall tree height is slightly over-estimated by lidar. The difference in minimum tree height could not be explained as trees with smaller than 7 cm DBH were not measured in the field. Since the data was acquired after recent treatment, slash, cut tops and stumps may all have been detected as small trees (below 0.5 m) by the algorithm.

Table 1: Summary of the field and lidar tree measurements.

	Field – DBH (cm)			Field – Tree height (m)			Lidar – Tree height (m)		
	Min	Max	Avg	Min	Max	Avg	Min	Max	Avg
Control	8.0	62.0	18.0	13.3	34.8	24.2	7.0	34.9	26.2
Treated	8.0	68.0	27.0	14.9	36.8	24.0	0.3	36.9	24.5*
Random	12.0	64.0	32.5	7.0	34.4	20.7	7.0	34.6	20.8

**based on trees over 7m height*

Acknowledgements

This work was partly funded through integrated research along the entire forest sector value chain at FPInnovations in partnership with Natural Resources Canada. Technical support for the UAV flights, Lidar and image acquisition was provided by RME Geomatics (Canada) and L'Avion Jaune (France). The authors would also like to acknowledge the Canadian Forestry Service, particularly Peter Arbour, and Petawawa Army Base to have made the area available for our trial and for providing background data.

References

[1] McRoberts, R. E., and Tomppo, E.O.(2007) Remote sensing support for national forest inventories. Remote Sensing of Environment 110.4 (2007): 412-419.

[2] Van Leeuwen, Martin, and Maarten Nieuwenhuis (2010). Retrieval of forest structural parameters using LiDAR remote sensing. European Journal of Forest Research 129.4, 749-770.

[3] Colimna, I., & Molina, P. (2014). Unmanned aerial systems for photogrammetry and remote sensing: a review. ISPRS J. of Remote Sensing, 92, 79-97.

[4] Leberl, F., Irschara, A., Pock, T., Meixner, P., Gruber, M., Scholz, S., & Wiechert, A. (2010). Point Clouds: Lidar versus 3D Vision. Photogrammetric Engineering & Remote Sensing 10 (2010): 76,1123-1134.

[5]Vepakomma, U., Kneeshaw, D., and DeGranpré, L. (2015) Edge influence of natural and anthropogenic linear canopy openings on forest structural patterns investigated using lidar (*Landscape Ecology* – in review).

Individual tree species identification using the multispectral return intensities of the Optech Titan lidar system

Benoît St-Onge[1] and Brindusa Cristina Budei[2]

1. *Department of Geography, University of Quebec at Montreal, Canada*
2. *Institute of Environmental Sciences, University of Quebec at Montreal, Canada*

Highlights: After checking the vertical coregistration of the point clouds of the green, near infrared, and middle infrared point clouds of Optech's Titan multispectral lidar, we have performed a classification of the three channel intensities on a crown-wise basis, and successfully identified tree species (broadleave vs. needleleaf, and genus).

Key words: *Multispectral, lidar, Optech, Titan, species, classification.*

Introduction

The identification of individual tree species is crucial for forest inventory, carbon stocks assessment, as well as for habitat studies. Several remote sensing approaches have been proposed, such as object-based classification of multispectral airborne images, or the use of single tree metrics derived from lidar return heights and intensities [1]. The former approach is complicated by the fact that the radiometric response of trees in airborne images is variable due to the varying sun-object-sensor geometry. The latter approach is somewhat limited because lidar sensors typically provide intensities in a single wavelength, usually in the near infrared. Dual wavelength commercial lidars have started to emerge (e.g. Optech's SHOALS), and multi-channel lidar prototypes were developed (e.g. Carbomap's four wavelength lidar demonstrator). In December of 2014, Optech Incorporated (Toronto, Canada) introduced the Titan multispectral airborne lidar, the first commercial lidar system that operates in three wavelengths (hereafter termed channels 1-3, or C1-C3): middle infrared (1550 nm, C1), near infrared (1064 nm, C2), and green (532 nm, C3). Each of Titan's three lasers fire with a different forward tilt, respectively 3.5° for C1, 0° for C2, and 7° for C3. Both infrared channels have a divergence of 0.35 mrad, but C3 has a divergence of 0.70 mrad. Because, by design, the lasers are not aligned, the respective point clouds do not coincide. Therefore, a given arbitrary small scene element is often hit by just one of the lasers. For this reason, multispectral analysis cannot be applied at the image pixel level in a classical way, nor by analysing the multispectral intensities of a given pulse, but rather by using multispectral statistics on a per object basis, in this particular case, by considering the intensities of all the hits falling within a given crown. This creates a new category of multispectral remote sensing problem, which calls for novel solutions.

Our objective was to explore the potential of individual tree species identification using Optech's Titan multispectral lidar. In this first study, we deliberately did not use the 3D data components in order to focus on the information contents of the multispectral intensity data. However, to ensure that the three independent channels can be used together without coregistration issues, we first checked the vertical adjustment between the three different point clouds. We then used individual crown multispectral intensity statistics to classify tree species into broad groups (hardwood vs. softwood), and into individual genera, using a random forest approach.

Study region and data

We used a dataset graciously provided by Optech. It was generated during a test flight over a 38 ha terrestrial area located in a suburb of the city of Toronto, Canada (79°68'W, 43°47'N). The flight was performed on 2 October 2014, a time of year when the foliage colour of broadleaved trees had not yet changed significantly. The aircraft followed three parallel flight lines at 400 m AGL. Each laser operated at a pulse repetition rate of 200 kHz (600 kHz overall system repetition rate). The pulse footprint at target level at nadir was 14 cm in C1 and C2, and 28 cm in C3. The lidar strips had an average 43 % lateral overlap. This resulted in a 10.6 first returns/m^2 per channel and per strip, 31.8 first returns/m^2 3-channel overall density per strip, and 53.8 first returns/m^2 overall multi-strip density (in overlay zones). The 12 bit intensities were recorded for each of 4 possible returns per pulse, in each channel. The range information of each return was also provided in the dataset. A total of 20 genera of trees were identified in the study area (16 broadleave and 4 needleleaf). The most common broadleaf genera are *Acer* and *Fraxinus*, and the most common needleleaf trees are *Picea* and *Pinus*.

Methods

Inter-channel coregistration

Because we needed to pool together lidar returns generated from different lasers that fire at different angles and that have different divergences (C1=C2<C3), we first checked that there was a good vertical correspondence between the three independent points clouds. We specifically wanted to check if the varying firing angles and divergences influenced the degree of penetration within tree crowns. For this, we have compared the Z values of all pairs of returns separated by a horizontal distance of 2 cm or less. This analysis was done for all inter-channel combinations, and also for same-channel (NIR) pairs of returns originating from two adjacent strips, the latter serving as control. The results were separated per broad land cover class (asphalt, grass, and trees).

Individual tree species classification

The return intensities were corrected for range according to the equation of Korpela et al [2]. Because multispectral signatures cannot be constructed on a pixel basis, or on a pulse basis, we have created signatures by extracting statistics for each individual sample crown. For this purpose, 523 tree crowns were manually delineated on a representation of the lidar digital surface model. The delineation excluded the fringe of the tree where a certain proportion of the pulses travelled unintercepted by the foliage and were reflected by objects located within the crown footprint, but under the crown itself. The tree species were identified based on a field census of the Toronto area tree species [3], and by using Google's Street View™ (in which terrestrial images for this area were acquired in summer of 2014 by Google). They were grouped into broadleaf and needleleaf classes (level 1: 2 classes), and genus (level 2: 8 classes). Measures of central tendency (mean and median), as well as dispersion (standard deviation) per crown and per channel were calculated from the normalized intensities of the first returns only. We then performed a random forest classification of each tree, at each classification level, using combinations of the above statistics. The current study being exploratory, and the test area being rather small, we have not divided the crown sample into training and validation datasets, and we report only the random forest "out-of-bag" error.

Results

Inter-channel vertical coregistration was found to be very good. For asphalt and grass respectively, the inter-channel Z RMSE were 1.8 cm and 1.9–2.7 cm. For grass, the greatest variations were observed between the two infrared channels (C1-C2: 2.7 cm), and the lowest variations for C1-C3 (1.9 – 2.1 cm). As expected, the Z RMSE for tree crowns is much higher than for flatter surfaces: 359 cm to 467 cm. The highest RMSEs were obtained for one of the visible-infrared combination (C1-C3 or C2-C3, depending on flight lines). For the same sites, the C2-C2 comparison used as control also yielded high RMSEs of 364 cm (flight lines 1-2), and 387 cm (flight lines 2-3). Moreover, there does not appear to be any strong Z bias between channels on tree objects, the median of inter-channel Z differences being between -0.2 cm and 3.8 cm.

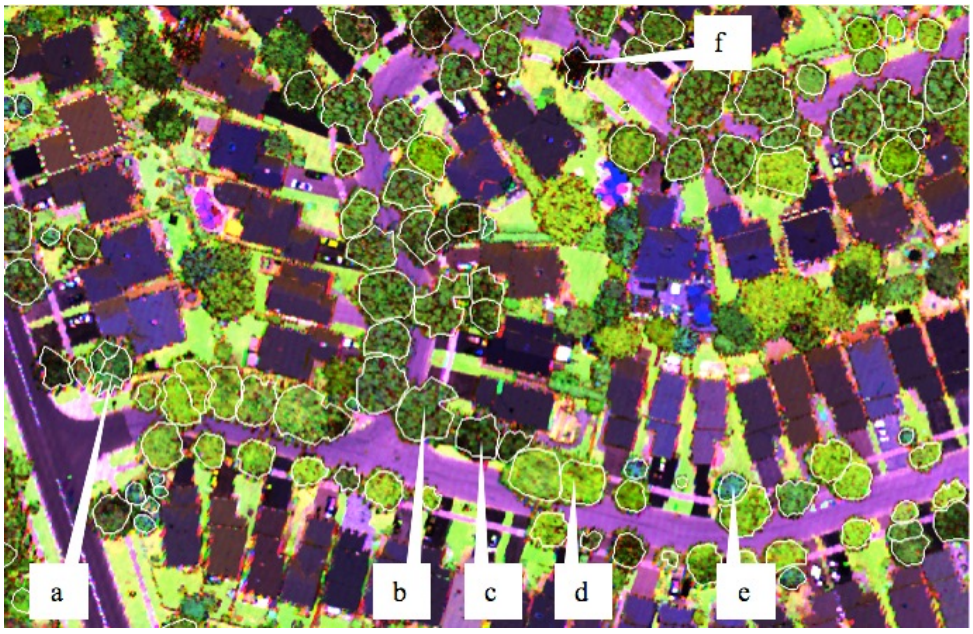


Figure 1: Multispectral image subset created using the 1st return intensities in three channels (crowns : a = *Pinus*, b = *Fraxinus*, c = *Betula*, d = *Acer*, e = *Picea*; crown f is severely defoliated). Sample trees are outlined.

Before presenting the classification results, we study the characteristics of the intensity data by examining a visual representation created by interpolating intensity values and creating a false colour combination (figure 1). We see that brightness and colour vary within individual crowns, but that colour contrasts between trees of different species are often important (see for example crowns labelled b, d and e on figure 1). Trees with severe defoliation appear very dark (see crown labelled f on figure 1). The crowns edges sometimes appear jagged, as grazing pulses may have hit either the crown periphery or the ground underneath, thus creating large variations in the radiometric response. As mentioned previously, these fringes were excluded from the analysis.

Applying a random forest classification based on the crown-wise mean and standard deviation in each channel (6 variables) to the set of sample crowns yielded an "out-of-bag" classification error of 4.59% in the case of the level 1 classification (broadleave vs. needleleaf trees), and of 24.29% in the case of level 2 classification (8 genera). The confusion matrix for the latter level is presented in table 1. Other models (e.g. using the median only, etc.) were less accurate.

Table 1: Confusion matrix of the random forest classification of individual tree at genera level, for 454 trees of the 8 most frequent genera, based on the three channel standardized intensities.

Genus	*Picea*	*Pinus*	*Acer*	*Fraxinus*	*Populus*	*Gleditsia*	*Betula*	*Tilia*	error
Picea	64	2	0	5	0	1	1	0	0.12
Pinus	6	9	0	2	0	0	0	0	0.47
Acer	0	1	181	5	1	6	4	3	0.10
Fraxinus	4	1	13	49	1	2	9	1	0.39
Populus	0	0	7	0	0	0	1	0	1.00
Gleditsia	1	0	10	4	0	22	2	0	0.44
Betula	0	0	3	8	0	0	17	0	0.39
Tilia	0	0	6	1	0	0	0	4	0.64

We see that the classification accuracy varies widely between genera, with *Acer* and *Picea* having the best scores. The overall error figure of 24.29% is influenced by the fact that the species with the best scores were also the ones with the most observations. This error score would rise in a dataset composed of an equal number of crowns per genus.

Discussion and conclusion

The point clouds from the three independent lasers of Optech Titan sensor proved to be well coregistered vertically. The 3-4 m variations in Z for a same XY location observed for tree crowns was expected, as laser pulses coming from different angles might travel between twigs and branches at various depth below the crown convex hull. Wind-induced branch movements can also contribute to this effect. A similar variation in Z in the flightline overlap zones in the same channel (the common 1064 nm lidar band, Titan's C2) was observed, indicating that the variations in wavelength and divergence have minimal impact on variations of penetration between channels. The three channels can therefore be used together, for 3D or multispectral analysis.

Large intensity variations within a given crown were also both expected and observed. This is likely caused in great part by variations of the cross-section area of the foliage intercepting the incident pulses, and to a much lesser extent by variation of reflectance within a single tree. This assertion is supported by the fact that using measures of central tendency and dispersion per crown, we could demonstrate that the overall crown-wise spectral signature can be used to identify species class (broadleaf vs. needle leaf), and to a certain extent, the genus of the trees. Further work is needed to assess the contribution of each spectral channel separately. Although the accuracy is still modest at this level, we expect that by combining the multispectral spectral data to the 3D data (crown shape and vertical distribution of returns), the species identification of single trees will be significantly enhanced compared to using standard single wavelength lidar data.

Acknowledgements

We express our gratitude towards Optech Incorporated (Toronto, Canada) for graciously providing, for the purpose of this study, a multispectral lidar dataset acquired using the Optech Titan sensor.

References

[1] Vauhkonen, J., et al. (2014). Tree species recognition based on airborne laser scanner and complementary data sources. In M. Maltamo, E. Naesset & J. Vauhkonen (Eds.), Forestry Applications of Airborne Laser Scanning (pp. 135-156). Springer.

[2] Korpela, I. & F. Rohrbach (2010). Variation and anisotropy of reflectance of forest trees in radiometrically calibrated airborne line sensor images - Implications to species classification. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences _ISPRS Archives, 38, 342-347

[3] Urban trees, Toronto tree map, http://urbantrees.ca/

Individual tree key parameters extraction based on LiDAR and hyperspectral data

Dan ZHAO[1,2], Yong PANG[2], Zengyuan LI[2], Yuan ZENG[1]

1. Key Laboratory of Digital Earth Science, Institute of Remote Sensing and Digital Earth (RADI), Chinese Academy of Science, Haidian District, Beijing, 100094, China
2. Institute of Forest Resource Information Techniques, Chinese Academy of Forestry, Haidian District, Beijing, 100091, China

Highlights: We use high density airborne LiDAR data and high spatial resolution hyperspectral data to extract key parameters of individual trees, including position, height, crown radius, species, in coniferous mixed forest of northeast China.

Key words: *Individual tree parameters, LiDAR, hyper-spectrum, individual tree isolation*

Introduction

Light Detection and Ranging (LiDAR) and hyperspectral techniques are applied in various fields including forestry by providing high quality 3-dimension vertical information and high spectrum resolution respectively. Isolating individual trees and extracting its key parameters are significant for automatic and semi-automatic forest management. This study introduces an experiment of individual tree isolation and species identification based on high density airborne LiDAR and high spatial resolution hyperspectral data in northeast China. In which a watershed method based on morphological crown control is applied to isolate individual trees and a combined classification method based on individual tree spectrum extraction and optimization is applied to identify individual tree species. Results show the methods can isolate most dominant and subdominant trees, more than half of intermediate trees and some suppressed trees, and efficiently identify dominant and sub-dominant species. The experiment indicated that LiDAR and hyperspectral data have great potential to provide more detail information in individual tree scale for forest management.

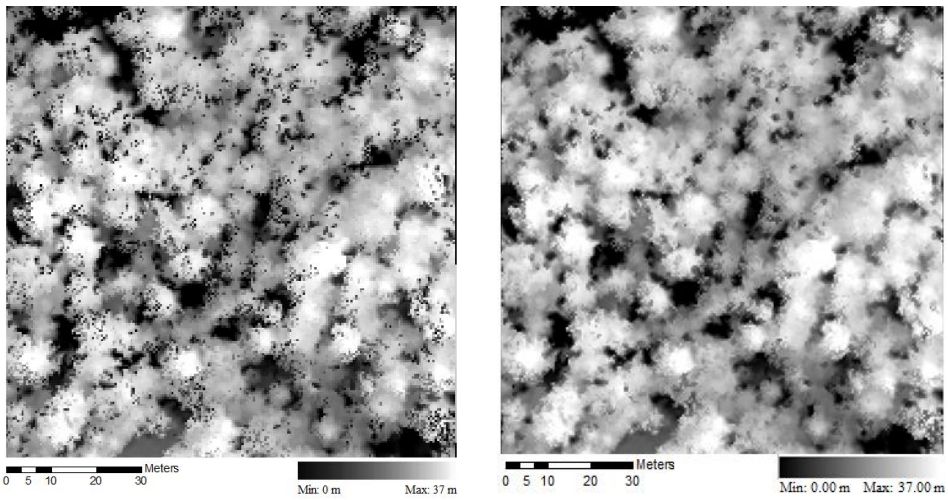
Content

The study area is located in Liangshui National Reserve ($47°10'N, 128°53'E$), Heilongjiang Province, China. The Liangshui National Reserve was established in 1980 and is a national reserve that protects the ecological system of a mixed forest of conifer and broadleaf species. A broadleaved Korean pine mixed forest sample plot in the reserve was selected for a detail field survey for this study. The plot is a square of 300 m × 300 m, which is divided into 900 pieces of 10 m × 10 m small quadrats (Jin et al., 2006). The species, DBH, tree height, and position of all individual trees with DBH greater than 2 cm were measured. Ground spectrum measurement was synchronized to hyperspectral data acquired through flight by FieldSpec 3 Spectroradiometer.

High-density airborne LiDAR data and charge-coupled device (CCD) images were simultaneously acquired in August 2009 by LiteMapper 5600 while the hyperspectral data were also acquired in August 2009 by using a CASI-1500 (Compact Airborne Spectrographic Imager) with a visible and near infrared (VNIR) pushbroom sensor.

The workflow of this study is briefly described as follow.

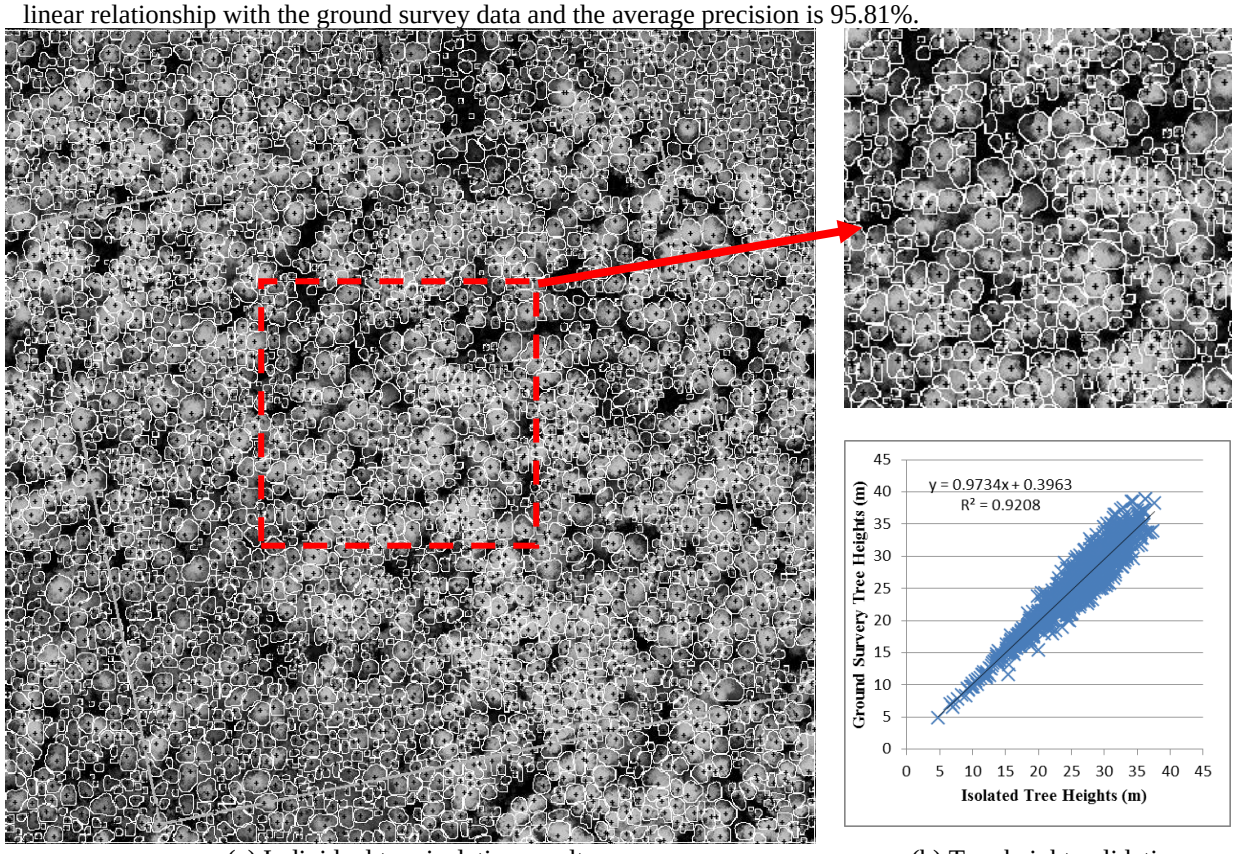
First, canopy height model is derived from LiDAR point cloud and an invalid value filling algorithm (Zhao et al., 2013) is applied to improve its quality. First, the Laplacian operator is applied to an original CHM to determine possible invalid values. Then, the morphological closing operator is applied to recover the crown coverage. By combining the two results, the possible invalid values in CHM can be confirmed and replaced by corresponding values in the median-filtered CHM. This approach can fill the most invalid values well while refraining from overfilling (Figure 1).



(a) Original CHM (b) Invalid value filled CHM

Figure 1: Invalid value filling results.

Second, the morphological crown control-based watershed method (Zhao et al., 2014) is applied to isolate individual trees and obtain their position, tree height and crown radius. Morphological crown control was introduced to determine the crown areas, which were used to limit the watershed results. The local maxima algorithm was adopted to identify potential individual tree positions. Double watershed transformations, in which a reconstruction operation was inserted into the two transformations, were then applied to delineate the tree crowns. Finally, the individual trees were isolated, and their parameters were extracted from the optimized watershed results (Figure 2). The results show that the method can isolate most dominant and co-dominant trees, more than half of the intermediate trees and some suppressed trees. For tree heights, the result has a pretty good linear relationship with the ground survey data and the average precision is 95.81%.

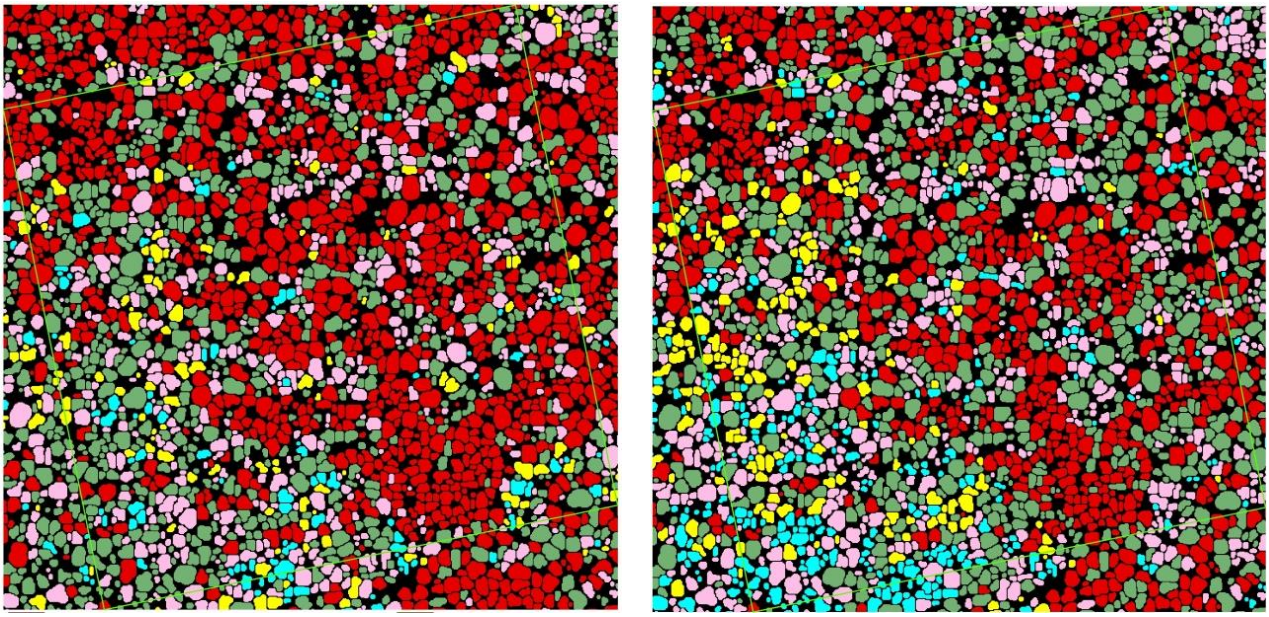


(a) Individual tree isolation results (b) Tree height validation

Figure 2: Individual tree isolation and tree height validation.

Third, a combined classification approach based on individual tree spectrum extraction and optimization is applied to identify individual tree species (Zhao et al., 2015). The approach first isolated individual trees and extracted their parameters from LiDAR data. Then, the spectra of the individual trees were extracted from hyperspectral data and merged to avoid mixed pixel and four-component problems. Finally, two common classification algorithms, namely, SVM and SAM, were applied to classify the individual tree species (Figure 3).

The identification correctness of the dominant species was over 90%, whereas that of the sub-dominant species was over 60%. The overall correctness of SVM and SAM were both over 70%. However, the SVM reflected the species distribution of the experiment area better than SAM.



(a) ITS-C-SVM (b) ITS-C-SAM

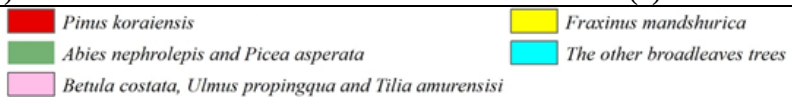


Figure 3: Classification results.

References

[1] Jin G.Z., Xie X.C., Tian Y.Y., & Kim J.H. (2006). The pattern of seed rain in the broadleaved-Korean pine mixed forest of Xiaoxing'an Mountains, China. Journal of Korean Forest Society, 95(5): 621-627.

[2] Zhao D., Pang Y., Li Z.Y., & Sun G.Q. (2013). Filling invalid values in a LiDAR-derived canopy height model with morphological crown control. International Journal of Remote Sensing, 34(13): 4636-4654.

[3] Zhao D., Pang Y., Li Z.Y., & Liu L.J. (2014). Isolating individual trees in a closed coniferous forest using small footprint LiDAR data, International Journal of Remote Sensing, 35(20), 7199-7218.

[4] Zhao D., Pang Y., Liu L.J., Zeng Y., Li Z.Y., & Zhao Y.J. (2015). Individual tree species mapping based on LiDAR and hyperspectral data in Northeast China mixed forest. ISPRS Journal of Remote Sensing and Photogrammetry, Under review.

Tree-centric mapping of forest carbon density from airborne LiDAR and hyperspectral data

David A. Coomes[1], Xiaohao Cai[1,3], Michele Dalponte[1,2], Juheon Lee[1,3], Carola Schönlieb[3]

Author's affiliations. 1 Department of Plant Sciences, University of Cambridge, UK; 2 Department of Sustainable Agro-ecosystems and Bioresources, Fondazione E. Mach, Trento, Italy; 3 Department of Applied Mathematics and Theoretical Physics, University of Cambridge, UK

Highlights: A tree-centric approach to carbon mapping is described, based on measuring the dimensions of individual tree crowns (ITCs) from LiDAR data, and species from hyperspectral data, from which individual-tree carbon stocks are calculated. We should tree-centric approaches deliver highly precise and accurate carbon maps.

Key words: *individual tree crown delineation, graph cut, hyperspectral, carbon mapping.*

Introduction

Forests are a major component in the global carbon cycle, and accurate estimation of forest carbon stocks and fluxes is important in the context of anthropogenic global change. Airborne LiDAR is increasingly recognized as outstanding data source for high-fidelity mapping of carbon stocks at regional scales. Most LiDAR-based approaches to carbon mapping work with summary statistics derived from the point cloud within pixels, which are calibrated with field-plot data. We develop a tree-centric approach to carbon mapping, based on identifying individual tree crowns (ITCs) and species from airborne remote sensing data, from which individual-tree carbon stocks are calculated.

Content

The first analysis we present uses airborne surveys of a 32 km^2 area of the Italian Alps, which is dominated by conifer species. The average LiDAR point density was high, at 48 pts/m^2 and accompanied by hyperspectral data (AISA Eagle II). ITCs we identified from the laser-scanning point cloud using a region-growing algorithm and identifying species from airborne hyperspectral data by machine learning. For each detected tree, we predicted stem diameter from its height and crown-width estimate. From that point on, we use well-established approaches developed for field-based inventories: aboveground biomasses of trees are estimated using published allometries and summed within plots to estimate carbon density. We show this approach is highly reliable: tests in the Italian Alps demonstrated a close relationship between field- and ALS-based estimates of (a) tree biomasses ($r^2 = 0.65$) and (b) carbon stocks ($r^2 = 0.98$) (Figure 1).

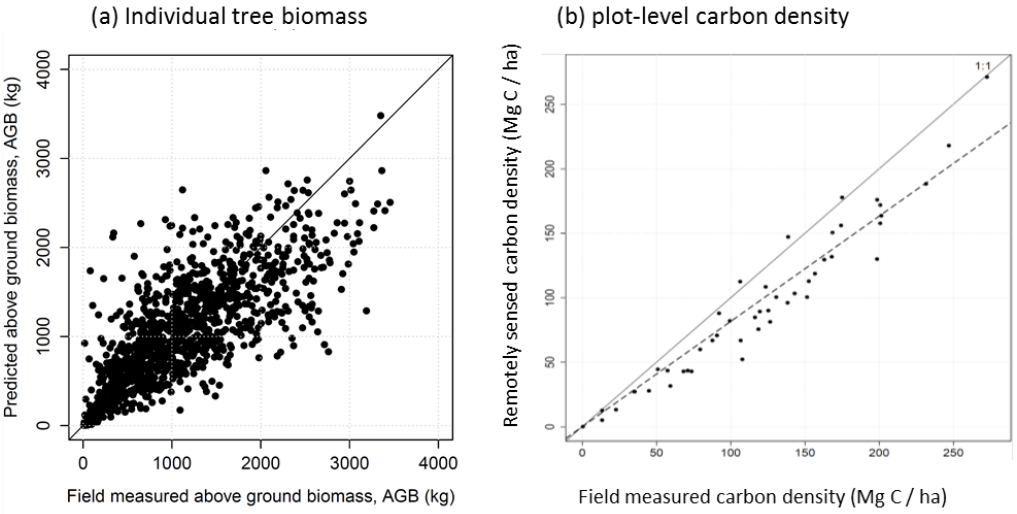


Figure 1: (a) biomasses and (b) carbon densities estimated from airborne imagery using tree-centric methods

Small trees were invisible from the air, meaning that 23% of biomass was not sensed (dotted regression line vs solid 1:1 line in Figure 1), so a correction factor is required to accommodate this effect.

The second illustration of the tree-centric approach uses an airborne survey of lowland tropical forests in Borneo, which is dominated by numerous dipterocarp species. Here the ITCs were identified using a new approach which utilizes the entire point cloud, based on graph cut methods, which allows understory trees to be detected as well as the canopy giants. Despite the complexity of these forests, species detection from hyperspectral imagery (AISA Fenix) was reasonably successful, and tree biomasses estimates were unbiased.

An advantage of the tree-centric approach over existing plot-level methods is that it can produce maps at any scale, and is fundamentally based on field-based inventory methods, making it intuitive and transparent. Airborne laser scanning, hyperspectral sensing and computational power are all advancing rapidly, making it increasingly feasible to use ITC approaches for effective mapping of forest carbon density.

Individual tree-based change detection using repeat pass airborne LiDAR at Teakettle Experimental Forest, California

Laura Duncanson[1], Colin Robertson[2], Bruce Cook[1]

1. *Biospheric Science Laboratory, Goddard Space Flight Center*
2. *Department of Geography, Wilfred Laurier University*

Highlights: We conduct an individual tree-based change detection of a mature conifer forest in California. LiDAR data were collected in the summers of 2008 and 2013. A crown delineation algorithm is applied to quantify individual tree growth and mortality, and trends are explained as a function of topography and tree size.

Key Words: *Crown delineation, biomass, tree height, change detection, carbon accounting*

Introduction

Forest ecosystems play a critical role in global biogeochemical cycling, particularly with respect to carbon, representing the largest terrestrial carbon stock on the planet. Indeed, the flux of carbon into forests through photosynthesis and out of forests through respiration and plant degradation largely controls annual fluxes of atmospheric carbon [1]. Understanding changes to plant growth and disturbance is therefore of the utmost importance for improving carbon cycle models and mitigating climate change through policies such as the U.N.'s Reduced Emissions from Deforestation and Degradation [2].

Repeat-pass LiDAR remote sensing has been used to detect changes in forest carbon stocks at a map or pixel level through analyzing changes in two maps of aboveground biomass [3]. These methods do not directly attribute change to ecological processes at the level of individual trees. In order to better understand changing dynamics in forests it is necessary to assess change at the level of the individual. In this research we conduct change detection at the individual tree level in order to increase our understanding of the processes driving the forest carbon balance in California.

Our research goals are three-fold: First, to determine the drivers of individual tree growth and mortality (topography, tree height, crown area), second, to estimate growth and mortality rates as a function of tree size/demography, and third, to determine whether the forest in question is a net source or sink of carbon.

Methods

Teakettle Experimental Forest is located in the Western Sierra Nevada Mountain range in California. Dominant species include *abies concolor (*white fir), *pinus ponderosa* (ponderosa pine), *abies magnifica* (red fir) and *quercus kelloggii* (California black oak). The elevation range of the site is approximately 1000 m to 2500 m above sea level, with AGBM values averaging 200 Mg ha^{-1} with individual tree values up to 20 Mg $tree^{-1}$.

Lidar data at Teakettle were flown in the summer of 2008 with the University of Florida's OPTECH GEMINI ALTM unit, operating at 100-125 kHz with a maximum 25° scanning angle. Data were flown ~600-750 m above ground, with 50%-75% swath overlap yielding an average return density of approximately 18 returns/m^2.

In 2013, a second LiDAR acquisition was conducted with NASA Goddard's lidar, Hyperspectral and Thermal Imager [4] instrument. G-LiHT uses a 300 kHz multi-stop scanning lidar operating at 1550 nm with a 60° field of view and 10 cm diameter footprint.

Individual tree metrics are derived from the lidar point clouds through a multilayered canopy delineation algorithm. This algorithm has been validated at Teakettle in a previous study [5]. Biomass is estimated for each delineated tree through the application of an empirical relationship relating height to Diamater at Breast Height (DBH) and the subsequent estimation of biomass through the Jenkins et al. equations [6]. Spatial overlap of tree crown polygons is examined for changes in morphology using the spatial-temporal analysis of moving polygons [7] method which characterizing change using an event-

based framework. Events of interest for crown morphology change include expansion, contraction, and height and biomass change.

Results and Discussion

In the five year period between LiDAR campaigns, there are areas of the landscape that exhibit regrowth (clearings where saplings are apparent in 2013), individual tree mortality events (Fig. 1), and individual tree growth and degradation. On average, the crowns in the study area are smaller in 2013 than 2008, with a mean area of 123 m^2 in 2008 reduced to 98 m^2, and a corresponding reduction in average crown radius from 2.89 m to 2.68 m. These are only preliminary results, and our research regarding individual tree growth and mortality is ongoing.

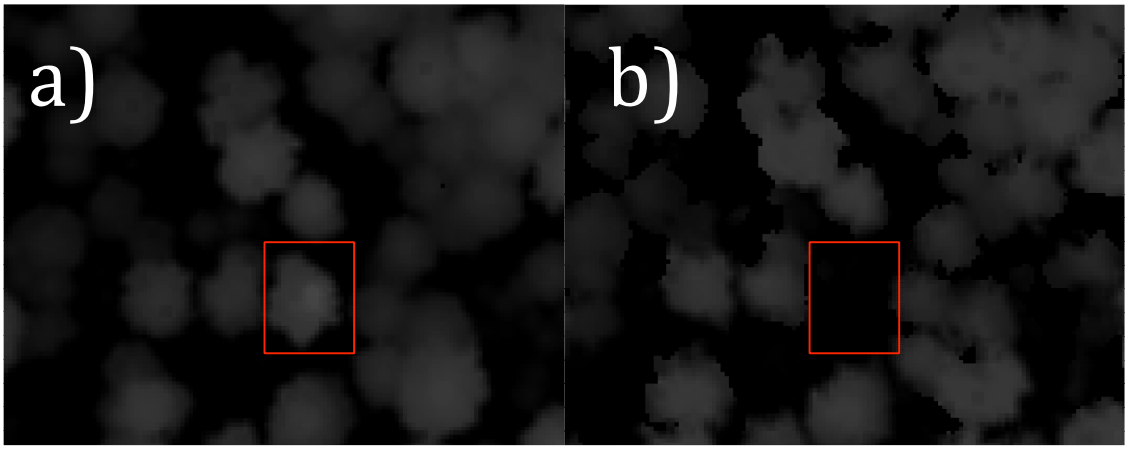


Figure 1. A 59 m tree in 2008 (a) is no longer present in 2013 (b). This single mortality event represents a loss of 11.8 Mg of biomass from the landscape.

References

[1] Keeling, C. D., Bacastow, R. B., Bainbridge, A. E., Ekdahl, C. A., Guenther, P. R., Waterman, L. S., & Chin, J. F. (1976). Atmospheric carbon dioxide variations at Mauna Loa observatory, Hawaii. *Tellus*, *28*(6), 538-551.

[2] Gibbs, H K, Brown, S., Niles, J.O., & Foley, J.A. (2007). Monitoring and estimating tropical forest carbon stocks: making REDD a reality. *Environmental Research Letters*, 2, 045023.

[3] Dubayah, R. O., Sheldon, S. L., Clark, D. B., Hofton, M. A., Blair, J. B., Hurtt, G. C., & Chazdon, R. L. (2010). Estimation of tropical forest height and biomass dynamics using lidar remote sensing at La Selva, Costa Rica. *Journal of Geophysical Research: Biogeosciences (2005–2012)*, *115*(G2).

[4] Cook, B.D., Corp, L.A., Nelson, R.F., Middleton, E.M., Morton, D.C., McCorkel, J.T., Masek, J.G., Ranson, K.J., Vuong, L., & Montesano, P.M. (2013). NASA Goddard's LiDAR, Hyperspectral and Thermal (G-LiHT) Airborne Imager. *Remote Sensing*, 5, 8, 4045-4066.

[5] Duncanson, L. I., Cook, B. D., Hurtt, G. C., & Dubayah, R. O. (2014). An efficient, multi-layered crown delineation algorithm for mapping individual tree structure across multiple ecosystems. *Remote Sensing of Environment.*

[6] Jenkins, J. C., Chojnacky, D. C., Heath, L. S., & Birdsey, R. A. (2003). National-scale biomass estimators for United States tree species. *Forest Science*, *49*(1), 12-35.

[7] Robertson, C., Nelson, T.A., Boots, B. and Wulder, M.A. (2007) STAMP : Spatial – temporal analysis of moving polygons. Journal of Geographical Systems. 9: 207-227.

Lidar-based assessment of forest edge effects across a degraded landscape in the Brazilian Amazon

Maiza N. dos-Santos[1], Ekena R. Pinagé[1], Marcos Longo[1],
Luciana Spinelli-Araujo[1], Marcos Adami[2] , Douglas Morton[1,3] and Michael Keller[1,4]

[1] EMBRAPA Satellite Monitoring, Campinas, São Paulo, Brazil, maizanara@gmail.com, ekenapinage@hotmail.com, mdplongo@gmail.com, luciana.spinelli@embrapa.br
[2] INPE – Amazon Regional Center, Belém, Pará, Brazil, marcos.adami@inpe.br
[3] NASA Goddard Space Flight Center, Greenbelt, Maryland, USA, douglas.morton@nasa.gov
[4] USDA Forest Service, International Institute of Tropical Forestry, San Juan, Puerto Rico, USA, mkeller.co2@gmail.com

Highlights: We quantified edge effects in forest fragments using lidar-based tree height over a degraded landscape in the Brazilian Amazon and found that the distance that characterizes an edge is more variable than previous studies have detected. This edge variability is key to understanding effects on ecosystem services such as changes in carbon stocks.

Key words: *edge effect, forest degradation, airborne lidar, Amazon.*

Introduction

Forest fragmentation divides forest areas into smaller discontinuous fragments, increasing the forest edge area that experiences different environmental conditions from interior areas [1]. According to one study, logging and deforestation generate ~32,000 and 38,000 km of new forest edge each year in the Brazilian Amazon [2]. Nearly 20% of the world's forests are within 100m of agriculture, urban, or other non-forest land uses [3].

Forest edges suffer changes in microclimate, with reduced moisture and increased variability of temperature compared to the forest interiors. In addition, winds cause greater structural damage near the forest ends and affect other ecological processes such as pollination, seed dispersal, nutrient cycling and carbon storage.

The extent of forest edges in the Brazilian Amazon grows each year, caused by deforestation and forest degradation from human activity. However, the extent of edge forests, and the changes in carbon stock resulting from alterations in microclimate and disturbance near forest edges are difficult to quantify. Previous studies have used multispectral remote sensing and geographic information systems (GIS) analysis to quantify fragmentation and the effect of edges [2]. We used lidar (light detection and ranging) measured canopy heights as a proxy variable to quantify forest edge effects across a degraded forest landscape of the Amazon in the Paragominas Municipality in Para State, Brazil. To our knowledge no previous study has used variability in forest height from lidar data to quantify changes in forest structure near edges in tropical forests. Lidar provides data with sub-meter vertical and horizontal accuracy, greatly improving forest structure quantification compared to traditional field studies.

Material and methods

The study area is located in the municipality of Paragominas (3°S, 50°W), in the eastern Brazilian Amazon. Paragominas was a center for timber production in the Brazilian Amazon in the 1980s and 1990s and large areas were deforested for cattle pasture. Today, the landscape is characterized by a range of agricultural uses (ranching, soy, plantation forestry, and natural forest management), with remnant logged, burned, secondary, and unmanaged forests interspersed across the municipality. High density (average of 17.9 points/m^2) lidar data was acquired over 30 transects (200 x 5000m), systematically distributed across a 34 by 180 km area inscribed within the borders of the municipality using an airborne scanning laser system (Optech Orion M-200) in August/September 2013 and June/July 2014.

In our study, forest edge was considered as the interface between a forest and a non-forested land. This definition of edge was applied on the lidar derived Canopy Height Model (CHM), a high resolution (1 meter) raster representation of the tree canopy height. We identified 57 forest/non-forest edges within the 30 transects. Based on the TerraClass land cover maps [4], we classified the forests as either old-growth, forests that were never cleared but may have been burned and/or logged and secondary, forests that regenerated after total clearance for pasture or agriculture. The interface land-uses included annual agriculture, pasture, pasture with regeneration and roads.

We defined a series of 10m wide buffers from each forest edge based on the CHM. The maximum distance from each edge was either half the fragment width or a maximum of 710m as only 26% of the fragments were

wider than this width, generating up to 71 strips. A variable number of 5x5m samples (up to 18, depending on buffer geometry) were randomly allocated within each strip, and the 95th percentile of height within each sample was calculated using Fusion software [6] to represent canopy height. The 5x5m sample window size was chosen as a compromise to fit within the 10m strips and to smooth canopy variability.

In order to test the existence of an edge effect, the set of samples of canopy height in each buffer segment was compared with samples of canopy height in the forest interior. The forest interior was defined for each lidar transect using first an area of old-growth forest classified by TerraClass land cover maps [4], free of perturbation by fire [5] and/or logging. In case no area complied with these characteristics, we selected a zone of secondary forest according to TerraClass maps [4] without fires [5] or logging. We generated 1000 bootstrap realizations based on the samples in each edge width and the interior to test the difference of the means and conservatively considered samples statistically distinguishable only if they were separated with 99% confidence.

Results and discussion

Figure 1 illustrates the extent of edge effects for the 57 edges in the Paragominas study area. The length of each bar shows the maximum distance into the forest where an edge effect (height differences) could be statistically separated with 99% confidence from the interior forest.

Nearly half of the analyzed edges had statistically significant differences in height compared to interior forests throughout the entire forest sample length (confidence interval of 99%). Secondary forests were more likely than old-growth forests to be classified entirely as edge. Forests fragments with a maximum measured transect of less than 200m in old-growth forest and 100m in secondary forest were all classified as edge. The average length of forest analyzed was 355m ± 256m (mean ± standard deviation) and the average edge effect observed was 162m ± 170m (mean ± standard deviation).

Three edges, all occurring in areas classified as old-growth forest, presented no discernible edge effect, meaning that the forest edge did not show difference in height when compared to the interior. Edge effects have a major impact on large trees, which often die off within 300m of the forest edge [7]; we found that approximately 44% of the studied edges were significantly different up to a distance of 320m from the edge.

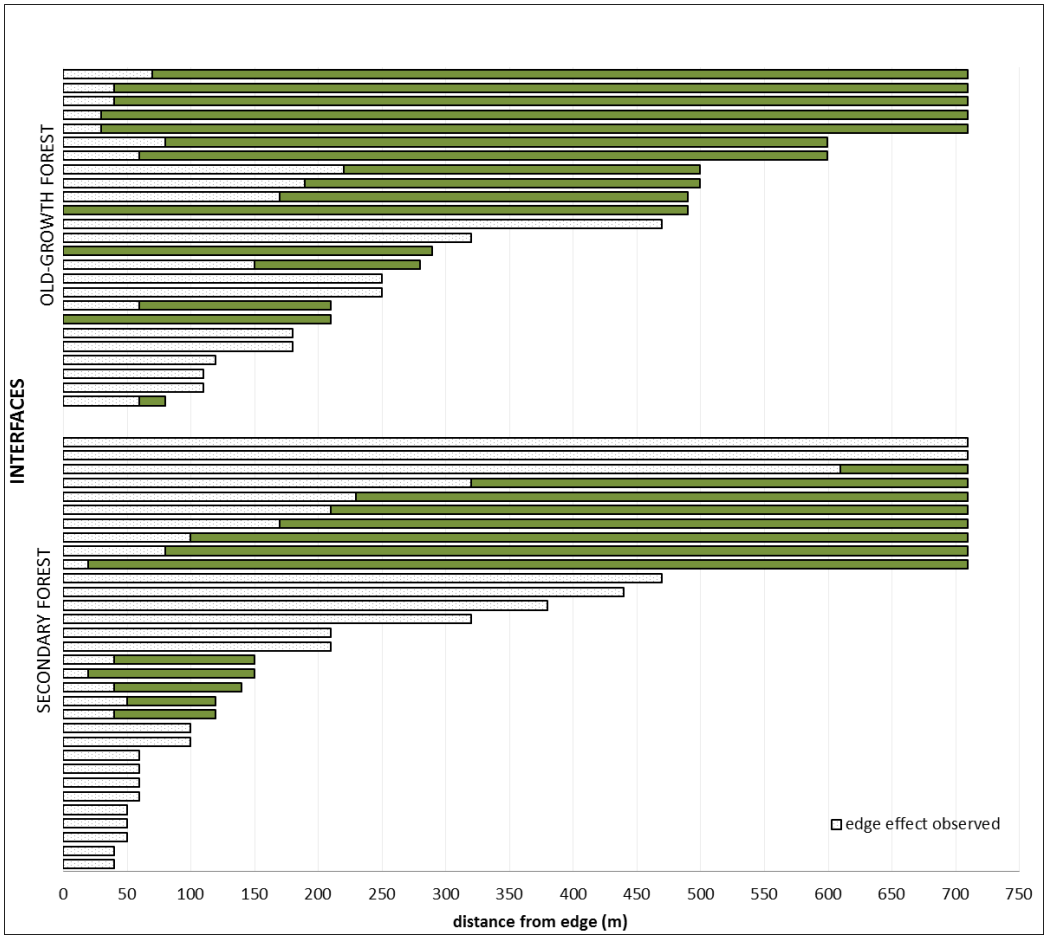


Figure 1: Edge effects observed in a degraded landscape. The white portion of the bar indicates the distance from the forest edge where edge effects were observed (statistically significant differences in forest heights, relative to the interior site at 99% confidence). On the top of the figure are the edges with primary forest/other land-use interface (25 edges) and on the bottom are the edges of secondary forest/other land use (32 edges).

Conclusion

Using tree height to measure edge effects we found that edge effects in a complex degraded forest landscape were highly variable. In general, secondary forests had a greater edge distance than old-growth forests but this difference was not statistically significant. Comparisons to previous studies that examined edge effects in controlled experimental settings [1] suggest that in real landscapes edges have greater variability. Under a mixture of land uses, edges are exposed to a broad range of degradation processes overlapping in space and time. Controlled experiments, while useful for process understanding, may not be suitable for extrapolation across broad areas. Vegetation height is well correlated with biomass; therefore, our results present an avenue for quantifying the importance of edge effects on Amazon region forest carbon stocks.

Acknowledgments

This study was supported by the Sustainable Landscapes Brazil project, collaboration between the Brazilian Agricultural Research Corporation (EMBRAPA) and the US Forest Service with financial support from USAID, and the US Department of State. This research was developed under the project "Land use changes and their interactions with forest degradation processes in the Amazon", supported by CNPq.

References

[1] Laurance, W.F., 2000. Do edge effects occur over large spatial scales? Trends in Ecology and Evolution 15, 134–135.

[2] Broadbent, E. N., Asner, G. P., Keller, M., Knapp, D. E., Oliveira, P. J., & Silva, J. N. (2008). Forest fragmentation and edge effects from deforestation and selective logging in the Brazilian Amazon. Biological conservation, 141(7), 1745-1757.

[3] Haddad, N. M., Brudvig, L. A., Clobert, J., Davies, K. F., Gonzalez, A., Holt, R. D., ... & Townshend, J. R. (2015). Habitat fragmentation and its lasting impact on Earth's ecosystems. Science Advances, 1(2), e1500052.

[4] Almeida, C.A., Pinheiro, T.F., Barbosa, A.M., Abreu, M.R.B.S., Lobo, F.L., Silva, M., Gomes, A.R., Sadeck, L.W.R., Medeiros, L.T.B., Neves, M.F., Silva, L.C.T. and Tamasauskas, P.F.L.F., "Metodologia para mapeamento de vegetação secundária na Amazônia Legal". Instituto Nacional de Pesquisas Espaciais, 32 (2009).

[5] Morton, D.C.; DeFries, R.; Nagol, J.; Souza Jr, C.; Kasischke, E.; Hurtt, G.; Dubayah, R. Mapping canopy damage from understory fires in Amazon forests using annual time series of Landsat and MODIS data. Remote Sensing of Environment, n. 115, p. 1706-1720, 2011.

[6] McGaughey, R. J. FUSION/LDV: Software for LIDAR Data Analysis and Visualization. Seattle, WA: USFS, 2014.175 p.

[7] Laurance, W.F., Delamonica, P., Laurance, S.G., Vasconcelos, H.L., Lovejoy, T.E., 2000. Rainforest fragmentation kills big trees. Nature, 404.

Patterns of nitrogen retention linked to LiDAR derived stand level architecture

Katherine C. Britt[1], Brian D. Strahm[2], Valerie A. Thomas[3]

College of Natural Resources and Environment, Virginia Polytechnic Institute and State University
[1]*kcorrell@vt.edu*, [2]*bstrahm@vt.edu*, [3]*thomasv@vt.edu*

Highlights: Nitrogen cycle dynamics, indicated by $\delta^{15}N$, are tied to stand level architecture rather than other foliar biochemical components. Fusion of spectroscopic data and LiDAR data may more robustly predict these values, providing monitoring capabilities to managers and policy makers.

Key words: *nitrogen isotopes, image fusion, nutrient cycling.*

Introduction

As global environmental change becomes a higher priority for research and policy, a better understanding of the nitrogen cycle is an important key to the focus on the carbon cycle [1]. Increased change to an already dynamic system highlights forests as a key resource providing ecosystem services to the biosphere, atmosphere, and hydrosphere [2]. Anthropogenic nitrogen inputs have doubled [3]. These inputs will have an unknown, complicated effect on multiple ecosystems. Some systems may experience increased growth as nitrogen, a typically limiting nutrient for plant growth, increases in availability. This retention may serve as a buffer for at risk ecosystems that are down stream and may partially offset carbon emissions [2]. Other systems may not retain input nitrogen, allowing through flow to at risk systems that are downstream [2].

Fusing remotely sensed data such as spectroscopic and LiDAR data provides a powerful assessment tool for ecosystem monitoring. The stable nitrogen isotopic ratio, $\delta^{15}N$, can provide information to forest managers about nitrogen cycle dynamics, with different $\delta^{15}N$ signatures for systems that retain or loses nitrogen [4]. The nitrogen cycle dynamic, rather than an instantaneous measure of nutrient content, is indicated by $\delta^{15}N$ [4]. Spectroscopy has been found to effectively predict $\delta^{15}N$ in vegetation [5]. Previous studies have concluded that the ability to predict foliar $\delta^{15}N$ is tied to the ability to spectroscopically predict foliar nitrogen content, suggesting potential common wavelengths that might be significant in both relationships [5]. Previous studies predicting $\delta^{15}N$ have primarily been conducted either at the leaf level or for crops, which typically have less complex structure than do forests. When applying such relationships to the canopy level, especially for heterogeneous forests, structural components both internal to the leaves and of the canopy can affect these relationships [6]. Here, structural data and spectroscopic data is combined to account for complexity inherent to the ecosystem of interest.

This study shows significant relationships between $\delta^{15}N$ and stand level architecture, a suite of metrics describing trees crowns, setting, and surrounding characteristics, to be significantly related. Foliar $\delta^{15}N$ and other leaf level biochemical components were largely unrelated, signifying the importance of providing spectroscopic relationships for $\delta^{15}N$ independently from existing predictive spectroscopy for other leaf biochemical components.

Materials and Methods

Foliar samples and forest inventory analysis was collected from the Blackwood Division of the Duke Forest in Chapel Hill, North Carolina. The forest is a mixed forest consisting of uneven-aged mixed hardwood and even age planted loblolly stands in the Piedmont of North Carolina. Sampling was conducted within a week of a NASA G-LiHT data acquisition. G-LiHT data includes LiDAR, hyperspectral, and thermal data collection.

Foliar samples were collected from major, canopy contributing species. Sampled trees were both tall enough to contribute to the top of the canopy and were large enough in area to contain multiple, unmixed pixels. Samples were selected to include a gradient in age, soil type, and elevation. Foliar samples were processed to determine a suite of macronutrient contents, $\delta^{15}N$, and chlorophyll.

Hyperspectral imagery for the flight requirements of this acquisition has a ground sample distance of 2.1 meters. LiDAR metrics were produced at the same resolution for the study area. Spectral signatures and LiDAR metrics for each tree in the sample set were extracted and averaged from the center of the tree, for a 3x3 set of pixels, and for the tree crown.

Results and Discussion

Previous work indicated that foliar $\delta^{15}N$ and foliar nitrogen content were correlated [5]. Across all species in this study area, the correlation between $\delta^{15}N$ and nitrogen content has a Pearson Correlation Coefficient r=0.147. For loblolly pines, r=0.838, but for hardwoods, r=0.090. Inconsistencies in the patterns of nitrogen content and

$\delta^{15}N$ correlation indicate that there are different underlying mechanisms and meanings for $\delta^{15}N$ and nitrogen content. These differences indicate that there may be a different suite of wavelengths that reflect the physiological processes driving these values.

Foliar $\delta^{15}N$ was correlated with several LiDAR derived stand level architectural metrics. Correlations exist with measures of point distribution such as kurtosis and skewness. Density of returns in the understory and canopy were significant. Understory height was also significant. Measures of the terrain were also significant, potentially as an indicator of nutrient movement through the ecosystem. Many LiDAR metrics are autocorrelated, meaning data reduction and component analysis is important in determining which metrics are most important in predictive modelling of $\delta^{15}N$.

Conclusions

Foliar nitrogen content and foliar $\delta^{15}N$ are decoupled. The different underlying meaning and processes indicate that modelling of $\delta^{15}N$ using spectroscopic data needs to be performed separately from existing spectroscopic nitrogen models. The significant wavelengths for these relationships may be different, as reflectance wavelengths are a function of underlying physiological properties of the foliage.

The significant relationships between foliar $\delta^{15}N$ and stand level architectural metrics indicates that both internal leaf structure and canopy structure and environmental setting are significant. These relationships should be incorporated into predictive modelling of $\delta^{15}N$.

In order to accurately model $\delta^{15}N$ levels at the canopy level, fusion of spectroscopic data and LiDAR data may be the best option for robust modelling. These predictive relationships will allow forest managers and policy makers to rapidly identify and monitor ecosystems that are effective and ineffective at nitrogen retention. Ecosystems that are effective at nitrogen retention can be promoted since they can serve as a buffer for systems that may be at risk of acidification from terrestrial nitrogen inputs. Systems that are ineffective at nitrogen retention can be managed to mitigate terrestrial nitrogen inputs.

References

[1] Law, B. (2013**)**. Nitrogen deposition and forest carbon. Nature, 496(7445), 307-308.

[2] Reay, D. S., Deneter, F., Smith, P., Grace, J., &Feely, R. A. (2008**)**. Global nitrogen deposition and carbon sinks. Nature Geoscience, 1(7), 430-437.

[3] Gruber, N., & Galloway, J. N. (2008**)**. An Earth-system perspective of the global nitrogen cycle. Nature, 451(7176), 293-296.

[4] Pardo, L. H., & Nadelhoffer, K. J. (2010). Isoscapes: Understanding movement, pattern, and process on Earth through isotope mapping. Netherlands: Springer, (Chapter 11).

[5] Wang, L., Okin, G. S., & Macko, S. A. (2010). Isoscapes: Understanding movement, pattern, and process on Earth through isotope mapping. Netherlands: Springer, (Chapter 3).

[1] Knyazikhin, Y, Schull, M. A., Stenberg, P., Mottus, M., Rautiainen, M., Yang, Y., Marshak, A., Carmona, L., Kaufmann, R. K., Lewis, P., Disney, M. I., Vanderbilt, V., David, A. B., Baret, F., Jacquemoud, S., Lyapustine, A., & Myneni, R. B. (2012**)**. Hyperspectral remote sensing of foliar nitrogen content. PNAS, E185-E192.

Linking lidar-derived canopy structure with MODIS black-sky albedo

Lauri Korhonen[1], Miina Rautiainen[1], Pauline Stenberg[1]

[1]*University of Helsinki, Department of Forest Sciences, Helsinki, Finland*

Highlights: A clear negative correlation was observed between MODIS shortwave albedo and lidar-based canopy cover estimates. Canopy cover explained the variation in albedo better than traditional forest attributes. Including mean height and species into the model improved the fit slightly.

Key words: *albedo, canopy structure, airborne lidar, MODIS*

Introduction

Land surface albedo is one of the essential terrestrial climate variables listed by the Global Climate Observing System [1]. It is defined as the fraction of the incident shortwave radiation that is reflected by the Earth's surface, and is directly related to the planetary radiative energy budget [1]. Cuttings, changes in species composition, and other forest management practices may influence the climate through changes in forest albedo [2]. However, although theoretically the albedo should depend quite strongly on the canopy structure [2], very weak relationships between standard forest inventory estimates and albedo have been reported in empirical studies [3].

Airborne lidar data is currently available for fairly large areas, and can provide a consistent description of the canopy structure. Thus, our objective was to examine if lidar-derived metrics of forest structure can explain the variations in satellite-based albedo estimates better than standard forest inventory data. We estimated canopy cover and mean canopy height directly from the lidar, and regressed these values against MODIS black-sky albedo values. For comparison, we tested also canopy cover and volume estimates produced by the Finnish national forest inventory (NFI).

Materials and methods

Our study site is located in the boreal forest of Seitseminen, Finland (61˚54´N, 23˚29´E). The dominant tree species are Norway spruce (*Picea abies* L. *Karst.*) and Scots pine (*Pinus sylvestris* L.). The lidar data was acquired on 2–11 May 2013 as a part for national digital terrain model creation. An area of 21 × 24 km was selected for the study. The pulse density was 0.7 m^{-2} and the maximum scan angle was approximately 15°. Forest canopy cover was estimated as the fraction of pulses that produced echoes > 2 m above ground level. The mean height of the first echoes above this threshold was also calculated. A spatial resolution of 30 m was used with lidar data.

Additional NFI data (© Finnish Forest Research Institute, 2013) were available for the year 2011 and included the volume fractions by tree species and canopy cover. The variables were estimated using multisource k nearest neighbour imputation based on field plot measurements and satellite imagery [4]. The resolution of the NFI maps was 30 m, i.e. similar to the lidar data.

MODIS MCD43A albedo product is a composite of satellite observations from the nearest 16 days, obtained by fitting a bidirectional reflectance distribution function into multiangular data and integrated over the observation angles to obtain estimates of black sky (directional-hemispherical) shortwave albedo which is the most appropriate for climate modelling [1]. Only the pixels with best quality retrievals were accepted. We used the MODIS albedo composite from 21 August 2013, which had the best overall quality for summer 2013.

Lidar and NFI-based forest attribute maps were converted into MODIS sinusoidal projection and averaged into 500 m resolution corresponding to the MODIS pixels. MODIS pixels containing water bodies were omitted, but other land use categories (including some agricultural fields and small settlements) were included because more than 90% of the land area was forest. Finally, 1230 valid MODIS pixels were available for the regression analysis, where lidar and NFI variables were applied to predict the observed albedo.

Results

The results showed a clear negative correlation between lidar-based canopy cover estimates and MODIS black-sky shortwave albedo (R^2 = 0.34, s.e. = 0.0056) (Figure 1). Canopy cover and height were highly correlated (R=0.71), but using the height instead of canopy cover resulted in weaker relationship (R^2 = 0.31) than using canopy cover. Adding both predictors into the model to approximate the forest biomass decreased the standard error only slightly (s.e. = 0.0054, R^2 = 0.38), i.e.canopy cover alone explained most of the variation in the observed albedo.

The NFI variables did not explain variation in albedo as well as the lidar variables. NFI canopy cover and volume estimates produced nearly similar model accuracies (R^2 = 0.20 and 0.19, s.e. = 0.0061 and 0.0062, respectively). However, supplementing the lidar variables with NFI species information improved the model precision. Our best multivariate model included lidar canopy cover, lidar mean height, and NFI pine volume fraction (s.e. = 0.0053, R^2 =0.40).

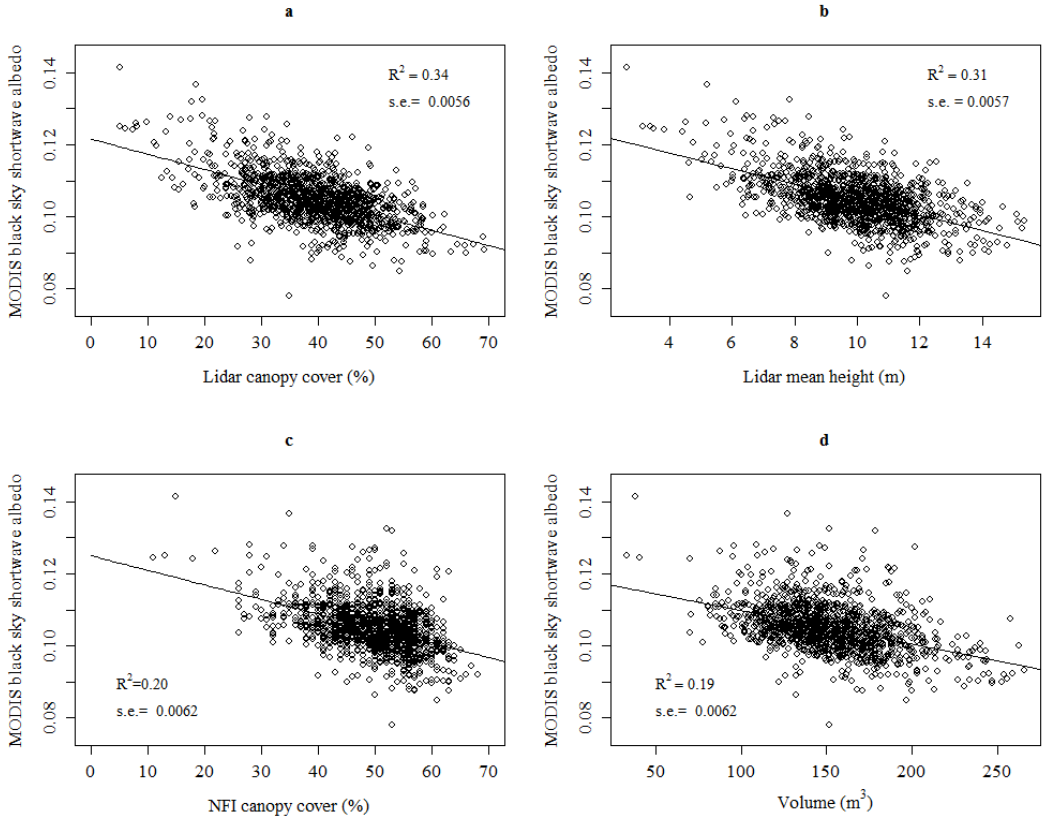


Figure 1: The relationships between MODIS black sky shortwave albedo and lidar-derived canopy cover (a), lidar-derived height (b), NFI canopy cover (c) and NFI total volume (d).

Conclusion

Lidar-derived metrics of forest structure were capable of explaining the variation in black-sky albedo better than standard forest inventory estimates. Among the studied variables, canopy cover seemed to be the main driver of shortwave albedo, although the inclusion of height and species composition into the model improved the model slightly (R^2 increased from 0.34 to 0.40, standard error decreased from 0.0056 to 0.0053). The spectral properties of the forest understory may also have a significant influence on the albedo, but this effect could not be directly quantified with these data. As a conclusion, lidar based canopy structure variables were shown to be useful in characterizing relationships between forest structure and albedo.

Acknowledgements

This research was supported by the Academy of Finland (grants no. 252257 and 135502).

References

[1] Schaaf, C.B. (2009). Assessment of the status of the development of the standards for the terrestrial essential climate variables. T8 albedo and reflectance anisotropy (V12, GTOS63). Rome: NRC, Food and Agriculture Organization of the UN (FAO), pp. 1-20.

[2] Lukeš, P., Stenberg, P., & Rautiainen, M. (2013). Relationship between forest density and albedo in the boreal zone. Ecological Modelling, 261–262, 74–79.

[3] Lukeš, P., Rautiainen, M., Manninen, T., Stenberg, P., & Mõttus, M. (2014). Geographical gradients in boreal forest albedo and structure in Finland. Remote Sensing of Environment, 152, 526–535.

[4] Tomppo, E., & Halme, M. (2004). Using coarse scale forest variables as ancillary information and weighting of variables in k-NN estimation: A genetic algorithm approach. Remote Sensing of Environment, 92, 1–20.

LiDAR-based quantification of carbon emissions from recurrent fires on a tropical peat swamp forest in Central Kalimantan, Indonesia

Kristina Konecny[1,2], Sandra Englhart[2], Uwe Ballhorn[2], Juilson Jubanski[2], Peter Navratil[2], Florian Siegert[1,2]

[1]*GeoBio-Center, Ludwig-Maximilians-University Munich, Germany*
[2]*RSS Remote Sensing Solutions GmbH, Germany*

Highlights: With the use of airborne LiDAR both aboveground biomass and fire-driven peat surface subsidence in relation to fire frequency were quantified within a tropical peat swamp forest in Central Kalimantan, Indonesia. Based on these figures an integrated carbon emission estimation was derived, comprising aboveground and peat carbon emissions.

Key words: *Carbon emissions, aboveground biomass; airborne LiDAR; tropical peatlands*

Introduction

Tropical peatlands play a significant role in the context of global warming, since they not only comprise huge amounts of aboveground biomass (AGB), but are also one of the largest near-surface pools of terrestrial organic carbon [1]. Persistent anthropogenic impacts by logging and drainage diminish their ability to sequester carbon and, as a consequence, enhance the risk of vegetation and peat fires [2]. Since the 1997/98 El Niño-induced drought, where peatland fires in Indonesia alone released an equivalent of 13–40% of the mean annual global carbon emissions from fossil fuels, they have recurred on an almost annual basis [3].

The estimation of the amount of carbon released during wildfires on tropical peat swamp forests is a complex task, since a large number or parameters, both relating to the carbon content of aboveground vegetation and carbon released from peat fires, need to be quantified. Light Detection and Ranging (LiDAR) showed to be a promising technique in quantifying these emissions. While the use of LiDAR is being extensively investigated in regard to AGB estimations, the state of knowledge on carbon emissions from tropical peatland fires is poorly developed with only a few studies reporting the mass of peat consumed during these events [4].

This study is based on approximately 270,000 ha of an extensive LiDAR dataset, acquired in 2011 covering a peatland dominated landscape in Central Kalimantan, Indonesia. The study area underwent industrial logging operations as well as illegal small-scale logging and was affected by frequent severe wildfires, often as a result of the Mega Rice Project (MRP). Our main objectives were to i) quantify AGB for forests which have been disturbed by single or multiple wildfire episodes, ii) quantify fire-driven peat surface subsidence in relation to fire frequency and iii) produce an integrated carbon emission estimation within our study area.

Methods

Acquisition and processing of airborne LiDAR data

The airborne LiDAR dataset was acquired between August and October 2011 with an Optech Orion M200 at a height of 800 m aboveground and with an average point density of 2.8 points/m². Ground and off-ground points were classified by applying a hierarchic robust filtering to the LiDAR point cloud. A Digital Terrain Model (DTM) with 5 m spatial resolution was interpolated from the ground points by a Kriging algorithm. The vertical accuracy of the DTM was assessed based on 441 checkpoints, with the modelled surface being within a 95% confidence interval of the true surface.

Delineation of burned areas

Active fire data from the NOAA AVHRR and Terra/Aqua MODIS sensors were analysed for the determination of years with fire occurrence. For the years 1990, 1997, 2001, 2002, 2003, 2004 and 2006, an automatic object-based classification of Landsat-5 TM and Landsat-7 ETM+ was performed. Recently burned areas (2009 and 2011) were identified by visual interpretation of RapidEye images as well as the LiDAR DTM. Based on spatial intersection, the burned areas were classified according to the number of fires they were affected by.

Estimation of aboveground carbon emissions

Inventory plots with different plot sizes were established within the LiDAR dataset in 2010 and 2011. In forested areas, 38 circular nested plots (radii of 4 m, 14 m and 20 m) were recorded. Trees of a certain diameter

at breast height (DBH) were measured depending on degradation intensity. In regrowth areas, 17 rectangular plots of 20 x 50 m² were established to record all saplings and trees. A combination of allometric models from Hughes et al. [5] for saplings or trees with DBH < 5 cm and Chave et al. [6] for moist tropical forest stands including DBH and tree height (if DBH ≥ 5) was employed to estimate AGB based on the recorded parameters.

An AGB regression model was developed on the basis of a statistical metric from the LiDAR point cloud, i.e. the Centroid Height (CH) of 0.5 m height intervals. CH was calculated by weighting each bin with its relative number of LiDAR points and then related to field inventory AGB. The AGB regression model employed utilized a power function in the lower biomass range and a linear function in the higher biomass range, resulting in a coefficient of determination r^2 = 0.81 and RMSE = 48.3 t/ha. The predictive power of the regression was assessed by randomly leaving out 10% of the reference field inventory plots and running 1,000 iterations of the regression, which resulted in a mean RMSE of 47.4 t/ha.

For each fire frequency class, as well as for the unburned area, mean AGB in t C per ha was determined based on the LiDAR AGB model. The carbon content was specified as 50% of the biomass.

Estimation of peat carbon emissions

Pre-fire elevation in burned areas was modelled by spatial interpolation from reference elevation values in surrounding unburned areas. On that account, we implemented Bézier approximation and applied it to the two test sites of 35,000 ha in total that were suitable for pre-fire peat surface reconstruction. The methodology incorporated a model refinement based on an uncertainty assessment.

Fire-related peat surface subsidence was quantified by defining it as a function of fire frequency. Mean elevation differences between the reconstructed peat surface and the DTM were calculated for each fire frequency class. Standard deviation and 95% confidence interval of the mean were determined for validation purposes.

By combining fire-related peat surface subsidence measurements with average values for bulk density (0.12 g cm^{-3} for the first fire and 0.115 g cm^{-3} for second and subsequent fires) and a carbon content of 55% peat in the study area, as well as a combustion efficiency of 1.0 [7], we estimated average emission factors for the first, second, third and successive fires.

Results

Carbon emissions for different fire frequencies were determined by addition of the estimates for forest and peat fires. Forested unburned areas featured a carbon content of 104 t C ha^{-1}, while in areas affected by one, two, three and four or more fires it was reduced to 22 t C ha^{-1}, 6 t C ha^{-1}, 2 t C ha^{-1} and 1 t C ha^{-1}. Depending on the time interval between two successive fires (and excluding other diminishing factors like e.g. further logging operations), the remaining carbon content is a combination of the trees remained standing after fire and regrowth in the form of ferns and sedges. Since the amount of trees remaining standing after fire is unknown in drained and degraded peat swamp forests, we assumed the carbon released by one fire event to be the difference between C before and after the fire took place (82 t C ha^{-1} from first, 16 t C ha^{-1} from second, 4 t C ha^{-1} from third and 1 t C ha^{-1} from successive fires). Emission factors for peat carbon loss from first through to four and more fires accounted for 113 t C ha^{-1}, 63 t C ha^{-1}, 38 t C ha^{-1} and 13 t C ha^{-1}.

Table 1: Aboveground and peat carbon emissions in the study area.

Fire frequency	Area [ha]	Aboveground carbon emissions [t C]	Peat carbon emissions [t C]	Total carbon emissions [t C]
1	35,913	2,944,866	4,058,169	7,003,035
2	36,206	579,296	2,280,978	2,860,274
3	39,823	159,292	1,513,274	1,672,566
4+	14,175	14,175	184,275	198,450
Total	**126,117**	**3,697,629**	**8,036,696**	**11,734,325**

Discussion and conclusion

Categorizing burned areas by fire frequency showed a clear trend. With increasing number of fire events, remaining biomass and associated aboveground carbon emissions are reduced. Standard deviations were relatively high compared to mean values, which indicate very high variability potentially resulting from fire intensity and time span since the last fire event.

The peat surface subsidence measurement which formed the basis for our peat carbon emission results are in the same order of magnitude as field measurements and consistent with previous airborne LiDAR measurements published by Ballhorn et al. [4]. However, the advancement of the approach presented here is that through applying a spatially explicit model, the complex dependence of fire related peat surface subsidence on fire frequency can be mathematically described and the wide variation in peat surface subsidence within burned areas is better captured.

The analysis of carbon emissions from wildfires in tropical peat swamp forests in our study area showed that there is an increasing share of emissions resulting from peat fires with increasing number of fire events. While carbon emissions from initial fires can nearly equally be attributed to peat fires (58%) and forest fires (42%), the share of peat carbon emissions increases to 80%, 90% and 93% for second, third and successive fire events. The distribution for first fires is influenced by the fact that the peat swamp forest in our study area has been affected by logging and drainage over the past decades, leading to decreased AGB compared to pristine peat swamp forests and hence reduced emissions from the remaining forest affected by fire. Furthermore, it has to be taken into account that it is unknown how many standing trees remain after the first fire and to what extent the forest recovers.

Airborne LiDAR proved to be an effective tool to quantify both aboveground biomass and fire-driven peat surface subsidence in relation to fire frequency. Based on these figures an integrated carbon emission estimation could be derived, comprising aboveground and peat carbon emissions.

Acknowledgements

The research was supported by AusAid through the KFCP (Kalimantan Forest and Carbon Partnership) programme. The PhD of K.K. is financially supported by the Hypatia programme of the Beuth University of Applied Sciences Berlin.

References

[1] Page, S. E., Rieley, J. O. & Banks, C. J. (2011). Global and regional importance of the tropical peatland carbon pool. Glob. Change Biol. 17, 798–818.

[2] Hooijer, A., Page, S., Jauhiainen, J., Lee, W. A., Lu, X. X., Idris, A. &Anshari, G. (2012). Subsidence and carbon loss in drained tropical peatlands. Biogeosciences 9, 1053–1071.

[3] Page, S. E., Siegert, F., Rieley, J. O., Boehm, H.-D. V., Jaya, A. &Limin, S. (2002). The amount of carbon released from peat and forest fires in Indonesia during 1997. Nature 420, 61–65.

[4] Ballhorn, U., Siegert, F., Mason, M. & Limin, S. (2009). Derivation of burn scar depths and estimation of carbon emissions with LIDAR in Indonesian peatlands. Proc. Natl. Acad. Sci. U.S.A. 106, 21213–21218.

[5] Hughes, R.F., Kauffman, J.B., and Jaramillo, V.J. (1999). Biomass, carbon, and nutrient dynamics of secondary forests in a humid tropical region of Mexico. Ecology 80, 1892–1907.

[6] Chave, J., Andalo, C., Brown, S., Cairns, M.A., Chambers, J.Q., Eamus, D., et al. (2005). Tree allometry and improved estimation of carbon stocks and balance in tropical forests. Oecologia 145, 87–99.

[7] Hooijer, A., Page, S., Navratil, P., Vernimmen, R., van der Vat, M., Tansey, K., et al. (2014). Carbon Emissions from Drained and Degraded Peatland in Indonesia and Emission Factors for Measurement, Reporting and Verification (MRV) of Peatland Greenhouse Gas Emissions – a summary of KFCP research results for practitioners. IAFCP, Jakarta, Indonesia

Using full-waveform lidar to characterise urban habitat structure and function

Steven Hancock[1], Karen Anderson[1], Mathias Disney[2], Kevin J. Gaston[1]

1 Environment and Sustainability Institute, University of Exeter, Penryn Campus, Cornwall, UK.
2 Department of Geography, University College London & the National Centre for Earth Observation.

Highlights: A novel method using terrestrial lidar to calibrate airborne lidar processing is presented. Attention is paid to the ability to detect understorey vegetation. Two deconvolution methods are compared and Gold's method was found to be superior. A variable noise threshold is needed to deal with weak signals and multiple scattering effects.

Key words: Waveform lidar, airborne, terrestrial, signal processing, urban vegetation.

Introduction

Urban areas are covering increasing areas of the world's surface and now contain over 50% of the global population. Increasing attention is being paid to the importance of urban wildlife and the impacts on urban greenspace on human wellbeing and provision of ecosystem services. There are well known relationships between urban greenness, the magnitude of the urban heat island, noise pollution and human-wildlife interactions. There is increasing evidence suggesting that interacting with wild animals improves human well-being [1] , and so there is a scientific driver towards understanding and quantifying greenspace connectivity for optimal town planning. Connectivity models predicting how animals move through a landscape can be used for this purpose, but currently there is little consensus on the best approach and little agreement about the types and scale of the input data needed [2].

A key determinant of connectivity is the three-dimensional distribution of vegetation structure and functional type. Airborne laser scanning (ALS) is the most appropriate tool for providing these data due to its ability to rapidly measure large areas accurately and non-destructively. This paper develops methods to measure the 3D structure of urban vegetation from ALS and discusses how they can be used to investigate urban habitat connectivity. Information on the full vertical canopy profile is important as different species inhabit different height bands. Terrestrial laser scanner (TLS) data were used as ground reference data to optimise the waveform signal processing method. Results were compared to the output of the ALS instrument manufacturer's (Leica) discrete return algorithm (captured in parallel to the full-waveform data) as this is a commonly used data type. Particular attention was paid to the capabilities of discrete return and waveform ALS data for describing understorey vegetation.

In order to extract vegetation structure from full-waveform ALS the signal must be denoised, calibrated to a physical intensity, system pulse shape effects removed, attenuation corrected for and then the signal strength converted to vegetation density at each point in the scene. This paper will demonstrate that process and show the results in the context of urban ecological greenspace mapping.

Field site and data

An urban area, including patches of woodland, buildings, parks and gardens in Luton, UK, was surveyed by the NERC-ARSF aeroplane in September 2012. Measurements were made with a Leica ALS50-II full-waveform lidar, Eagle and Hawk imaging spectrometers (covering the visible, near infra-red and short-wave infra-red wavelength regions) and a standard fixed format survey grade digital camera. The Leica ALS50-II is a 1064 nm lidar which records two streams of data, one using Leica's commercially confidential discrete return algorithm to output up to four points per laser shot and the other as full-waveforms sampled every 1 ns. The ground point density was between 0.25 m^{-2} and 5 m^{-2} depending on flight line overlap.

Eight ground plots were chosen to cover the full range of observed ALS waveform shapes. In August 2014 a Riegl VZ-400 TLS was used to collect between two and seven scans at each plot (depending on vegetation density) which were geolocated with the ALS data. An Ocean Optics USB2000 was used to collect ground, bark and leaf reflectance spectra within the plots.

Signal processing

A histogram of all ALS waveform bins was used to determine the background noise levels. This showed that noise was very stable (99% of values lay within 2 digital numbers of the mode), suggesting that it was due to a dark current rather than background illumination and could be removed by thresholding [3]. Previous work

shows that simply summing all waveform bins above noise gives the most accurate measure of target reflectance for this type of lidar [4]. Returns from Luton airport runway were used to determine the system pulse shape and Gold's method [5] or Richardson-Lucy deconvolution [6] used to deconvolve the waveform, and the results of the two methods were compared. Both are iterative re-blurring algorithms for deconvolving the effect of system pulse blurring in the presence of noise. Gold's method is given by:

$$o^{(k+1)} = o^{(k)} \frac{i}{s \otimes o^{(k)}}$$

Where i is the original waveform, s the system pulse, $o^{(k)}$ the k^{th} estimate of the true signal and $\otimes$ the convolution operator. This is iterated over for either a fixed number of iterations (set to a maximum of 2,000) or until the change between consecutive estimates drops below a threshold (in the examples examined the result always converged). Richardson-Lucy deconvolution is very similar but with an extra blurring by the system pulse:

$$o^{(k+1)} = o^{(k)} \left(s \otimes \frac{i}{s \otimes o^{(k)}} \right)$$

This should make it more robust to noise but will take longer to converge and may not reach as sharp a result. These algorithms require several parameters to be set, such as the deconvolution tolerance, smoothing widths and noise threshold. These were tuned against "true waveforms" produced from the TLS data, representing the signal that the ALS would see without the noise and system pulse effects, but still including attenuation.

To produce these "true waveforms", the TLS data were voxelised [7] to the same size as the ALS waveform bins (15cm vertically and 30 cm horizontally) and a silhouette ray tracer [8], taking gap fraction within each voxel into account, was used to create the true waveforms. Non-linear optimisation was used to find the optimum signal processing parameters. The same TLS voxels (without ray tracing) were used as the 'true', reference vegetation profile in order to optimise the ALS attenuation correction and validate the results. Comparison between the raw ALS data (before any processing) and the TLS produced waveforms showed that the combination of using the beam divergence, hit intensity and gap fraction [8] accounted for attenuation, giving good estimates of the canopy up to the highest points within the plots. The accuracy at the canopy top reduced farther from the plot centre (notably falling off 15 m away), but the results within the central 10m were deemed accurate enough. Plots were split into separate optimisation and validation datasets.

Results

Gold's method was found to converge faster and give more accurate results than Richardson-Lucy deconvolution and so we use Gold's method here. It was not possible to set a fixed noise threshold that removed all subterranean noise and captured the understorey vegetation; either the threshold was too high to detect the weak signal from under dense canopies or else was so low that multiple scattering and electronic lag of the detector caused spurious subterranean hits. Therefore a variable noise threshold was needed, tuned to the background noise for each individual laser shot but also increasing through each waveform in order to avoid spurious returns from multiple scattering and electronic lag.

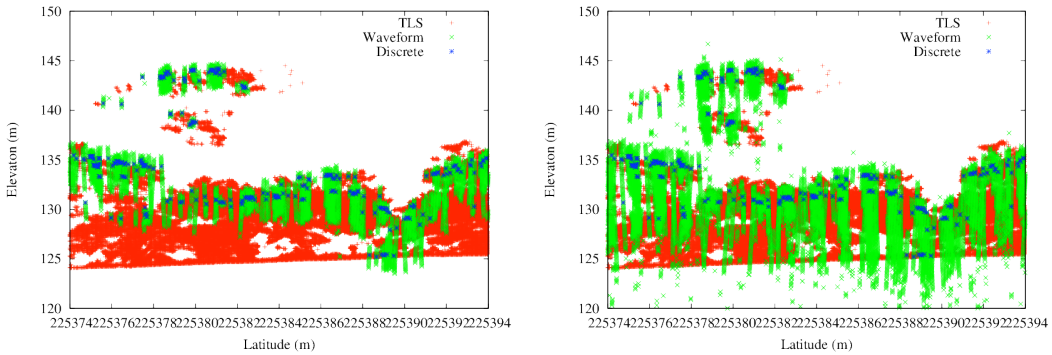


Figure 1: Comparison of TLS, ALS discrete and ALS waveform results with a high (left) and low (right) noise threshold.

Figure 1 shows that the discrete return algorithm gave no subterranean hits but did not give any returns from the understorey, suggesting that it used a high, conservative noise threshold. This would have been exacerbated by the discrete return dead time, of the order of 2 m. The discrete return algorithm will have been optimised for measuring the range to hard targets in order to produce accurate digital elevation models and so it is not surprising that it fails to characterise the complete vegetation canopy. In addition the discrete return data showed a bias in height of the order of an 80 cm underestimate.

Conclusions

It was found that Gold's method outperformed Richardson-Lucy deconvolution in terms of computational expense and RMSE of the retrieved vegetation profile relative to the TLS measurements. A fixed noise threshold was not suitable for all laser shots in a scene as it had to be set too high to measure weak returns from understorey in order to avoid multiple scattering and electronic lag effects. An adaptive noise threshold was required, changing between shots and along a waveform as multiple scattering becomes significant in dense canopies and may require more advanced denoising methods [10]. Further work using radiative transfer modelling will allow a thorough investigation and decoupling of these effects and work towards this using the librat ray tracer [9] is ongoing.

References

[1] Dallimer, M., Irvine, K. N., Skinner, A. M., Davies, Z. G., Rouquette, J. R., Maltby, L. L., ... & Gaston, K. J. (2012). Biodiversity and the feel-good factor: understanding associations between self-reported human well-being and species richness. *BioScience*, *62*(1), 47-55.

[2] Zeller, K. A., McGarigal, K., & Whiteley, A. R. (2012). Estimating landscape resistance to movement: a review. *Landscape Ecology*, *27*(6), 777-797.

[3] Hancock, S., Disney, M., Muller, J. P., Lewis, P., & Foster, M. (2011). A threshold insensitive method for locating the forest canopy top with waveform lidar. *Remote Sensing of Environment*, *115*(12), 3286-3297.

[4] Hancock, S., Armston, J., Li, Z., Gaulton, R., Lewis, P,. Disney, M., Danson, F.M., Strahler, A., Schaaf, C., Anderson, K., Gaston, K.J. (2015). Waveform lidar over vegetation: An evaluation of inversion methods for estimating return energy. In review for Remote Sensing of Environment.

[5] Hancock, S., Lewis, P., Disney, M. I., Foster, M., & Muller, J. P. (2008). Assessing the accuracy of forest height estimation with long pulse waveform lidar through Monte-Carlo ray tracing. *Proc. SilviLaser 2008, Sept. 17–19, Edinburgh, UK*, 199-206.

[6] Wu, J., van Aardt, J., and Asner, G. P., (2011), A comparison of signal deconvolution algorithms based on small-footprint LiDAR waveform
simulation. Geoscience and Remote Sensing, IEEE Transactions on, 49,
2402–2414

[7] Béland, M., Widlowski, J. L., Fournier, R. A., Côté, J. F., & Verstraete, M. M. (2011). Estimating leaf area distribution in savanna trees from terrestrial LiDAR measurements. *Agricultural and Forest Meteorology*, *151*(9), 1252-1266.

[8] Hancock, S., Essery, R., Reid, T., Carle, J., Baxter, R., Rutter, N., & Huntley, B. (2014). Characterising forest gap fraction with terrestrial lidar and photography: An examination of relative limitations. *Agricultural and Forest Meteorology*, *189*, 105-114.

[9] Lewis, P. (1999). Three-dimensional plant modelling for remote sensing simulation studies using the Botanical Plant Modelling System. *Agronomie*, *19*(3-4), 185-210.

[10] Azadbakht, M., Fraser, C. S., Zhang, C., and Leach, J., 2014, Cross-section retrieval from full-waveform LiDAR using sparse solutions, In
Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International , IEEE, pp. 1959–1962.

Tree stem reconstruction from terrestrial laser scanner point cloud using Hough transform and open active contours

Joris Ravaglia[1,2], Alexandra Bac[1], Richard A. Fournier[2]

[1]*LSIS – Université d'Aix-Marseille - France*
[2] *Centre d'Applications et de Recherches en Télédétection (CARTEL), Département de géomatique appliquée, Université de Sherbrooke, Sherbrooke (QC) CANADA*

Highlights: We introduce an innovative methodology for reconstructing tree stems from terrestrial LiDAR data. The method uses a multi-scale combination of an original Hough transform and active contours. Tests were conducted on both simulated data and real LiDAR data. Even though no formal evaluation was completed yet, the procedure seems promising.

Key words: *terrestrial LiDAR, stem reconstruction, Hough transform, open active contours*

Introduction

Tree structural attributes are crucial to forest inventories measurements. Indeed, structural attributes such as diameter at breast height (DBH) or stem taper are used in a wide range of applications. For example, they are employed in forest monitoring, plant growth models and allometric models [1]. However, the measurement of some of these attributes is often limited in the context of non-destructive field work, and their precision can also be limited by measurement tools and operational constraints. Thus, terrestrial laser scanners (TLS) have been introduced as a measuring tool to expand the number and the quality of the structural attributes available in forest inventories. TLS instruments provide a precise 3D point cloud characterizing the surrounding environment of a forest plot from which tree attributes can be estimated in an non-destructive and objective way.

Assuming that tree stem cross-sections are circular, DBH and stem taper can be estimated using pattern recognition algorithms applied to point clouds. Point cloud processing methods such as circle fitting, cylinder fitting or cone fitting are indeed well-suited for reconstructing tree stem [2]. Hough transform is another way of retrieving DBH from point clouds [3]. Traditional Hough transform has shown a stem detection rate above 90% and an average DBH estimation error of 1.7 cm when applied to natural forest environment [3]. The analysis of the Hough space obtained by Hough transform can be a hard task, hence, despite these positive results, shape fitting has become the main method to reconstruct tree stems. However, Hough transform, as a pattern recognition tool, has demonstrated to be adapted for application in complex forest environment.

In this study we aim at developing a novel methodology for tree stem reconstruction in order to estimate DBH and stem taper. This methodology is intended to be able to overcome some limitations of point cloud analysis: occlusion, non-homogeneous sampling density and noise. To do so, we take advantage of the Hough transform qualities and propose an innovative use of this pattern recognition tool.

Material and methods

We reconstruct tree stems as an ordered series of circles in the 3D space with continuous centre locations, orientations and radii, which we refer to as tuboid. We introduce a multi-resolution approach for extracting such series of circles from the TLS point cloud data. At each resolution, in a first step, a Hough space describing potential circles in the data is created. In a second step we employ growing open active contours to extract tuboids from this space.

We use a variant of the classical Hough transform adapted to circle detection based on normal direction at each point of the data to obtain a discrete 4D Hough space (HS). A cell (x,y,z,r) of this space describes the number of data points belonging to the circle c*(x,y,z,r)*. The Hough space is computed according to the following property: the normal direction at a point on a circle passes through the centre of this circle. This property allows creating the corresponding Hough space using fast ray tracing algorithms.

A tuboid in this HS is represented as a curve. Since the elements of the HS represent the number of points belonging to a circle, the elements of highest value tends to be an accurate reconstruction of the data. Thus, we intend to extract curves passing through high value elements. To do so, we use open active contours. We define a curve energy which is minimum when the curve passes through these high value elements with a low curvature in order to reduce noise effect and obtain a smooth taper. Classical active contours need an initialisation of their position. Such a general initialisation along a stem can be a delicate issue. Instead, we initialise small curves at the local maximum of the HS and let the active contours grow along the elements of high value by adding a growing term to the curves.

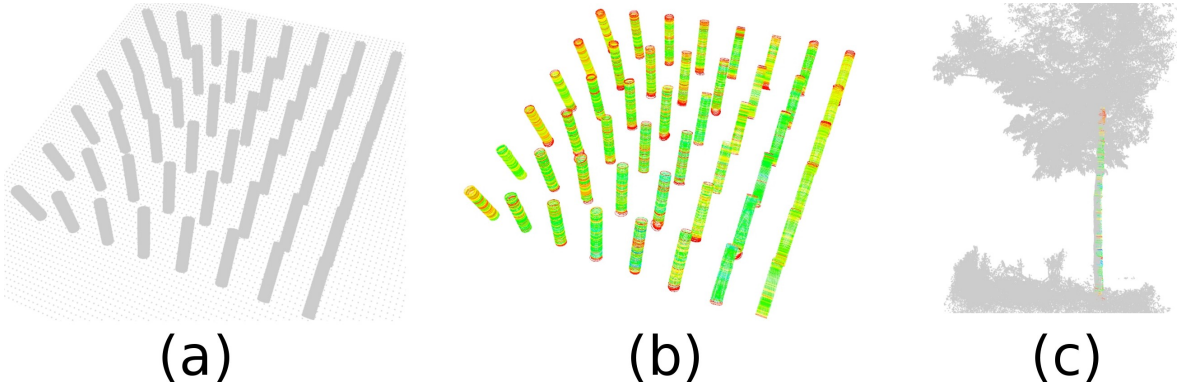


Figure 1 : (a) simulated point cloud : cylinders with increasing orientation angle and noise, (b) reconstruction with Hough transform (circles are coloured according to their value in the HS), (c) laser scanned tree stem and reconstruction as tuboid

The precision of results is related to the Hough space resolution. However, the computation of high resolution Hough spaces for 20 m × 20 m forest plots is time consuming and needs a lot of computer resources. This is why we propose a multi-resolution scheme. Starting from a coarse Hough space, coarse tuboids are extracted. Then, a precise sub-Hough space is computed with finer resolution.

Results

Up to now, the methodology was tested on both simulated and real datasets. The first test dataset is a simulated point cloud containing a set of 48 cylinders with a constant radius (50 cm) with various orientations and increasing noise (from 0% to 10% of the radius). The second dataset is a single-scan acquired point cloud in a real forest environment using the Faro Focus 3D TLS (Fig. 1).Normal directions have been estimated by local plane fitting for both dataset. The 48 cylinders of the first data set were successfully reconstructed with an average error of 1.07 cm. This illustrates the robustness of the HT to noise (Fig. 1). In the second data set from real data, tree stem from the ground to the centre of the crown.

Discussion

The method we introduce is based on a combination of two powerful pattern recognition tools. Our variant of the classic HT needs an estimation of the normal at each points of the cloud. Thus, the results are highly dependent on the quality of the normal estimation procedure. However, this HT has a low algorithmic complexity due to to the availability of the normal information: the number of necessary operations is linearly related with *n,* the number of points in the cloud. Moreover, HS elements contain only local information, and HS analysis with active contours is based on the detection of local high values, making the entire methodology able to deal with non-homogeneous sampling density.

The use of growing open active contours ensures a continuity of location, orientation and radii inside each tuboid following what is observed for tree stem and branches. These growing active contours also avoid the issue of initialisation and the need of *a priori* information about the location of tree stems. Finally, through the minimization of the internal energy of curves and the addition of a growing term to the active contours, missing information in occluded areas may be coherently estimated according to the surrounding non-occluded zones.

Conclusion

We proposed a novel methodology to reconstruct tree stems from LiDAR data collected in forest plots. This methodology aims at overcoming problems related to the use of TLS in forest environment such as occlusion, non-homogeneous sampling densities and noise. By combining active contours with the Hough transform, we take advantage of their respective qualities in an innovative novel way. This treatment differs from the iterative shape fitting procedures since stem reconstruction is embedded in a single object recognition: a tuboid.

References

[1]Schneider, R., et al., Calibrating jack pine allometric relationships with simultaneous regressions. Canadian journal of forest research, 38(10), 2566–2578

[2] Dassot, M., Constant, T., & Fournier, M. (2011). The use of terrestrial LiDAR technology in forest science: application fields, benefits and challenges. Annals of Forest Science, 68(5), 959–974.

[3] Simonse, M., Aschoff, T., Spiecker, H., & Thies, M. (2003). Automatic determination of forest inventory parameters using terrestrial laser scanning. In Proceedings of the scandlaser scientific workshop on airborne laser scanning of forests, 2003, 252-258.

Application of terrestrial LiDAR and modelling of tree branching structure for plant-scaling models in tropical forest trees

Alvaro Ivan Lau Sarmiento[1], Harm Bartholomeus[2], Martin Herold[3], Christopher Martius[4], Yadvinder Malhi[5], Lisa Patrick Bentley[6], Alexander Shenkin[7] & Pasi Raumonen[8]

[1]*Wageningen University and Research Centre, Wageningen, the Netherlands, alvaro.lausarmiento@wur.nl*
[2]*Wageningen University and Research Centre, Wageningen, the Netherlands, harm.bartholomeus@wur.nl*
[3]*Wageningen University and Research Centre, Wageningen, the Netherlands, martin.herold@wur.nl*
[4]*Center for International Forestry Research – CIFOR, Bogor, Indonesia, c.martius@cgiar.org*
[5]*University of Oxford, Oxford, United Kingdom, yadvinder.malhi@ouce.ox.ac.uk*
[6]*University of Oxford, Oxford, United Kingdom, alexander.shenki@ouce.ox.ac.uk*
[7]*University of Oxford, Oxford, United Kingdom, lisa.bentley@ouce.ox.ac.uk*
[8]*Tampere University of Technology, Tampere, Finland, pasi.raumonen@tut.fi*

Highlights: Terrestrial laser scanning, with automated data processing techniques, provides a powerful alternative for estimating tree characteristics. Nevertheless, TLS also offers an unexplored potential where tree architecture plays a major role: forest ecology.

Key words: *terrestrial laser scanning, plant scaling modelling, quantitative structure model.*

Introduction

The whole-organism metabolic rate expresses the rate at which energy and materials are taken up from the environment, transformed in biochemical reactions and allocated for maintenance, growth and reproduction [1]. To relate metabolic rate to size, the theoretical model of West, Brown and Enquist (WBE) proposed the fractal geometry of transport systems as the origin of the allometric scaling laws observed in nature at both, internal (vascular) and external (branching) networks [2].

In the context of the WBE model for plants, a tree's external architecture is defined by branching, furcation numbers, branch diameters and branch lengths. The tree's external branching network is assumed to be symmetrical, self-similar and hierarchical; and values for WBE scaling exponents should not vary throughout branch level and the whole-tree. Therefore, a small piece of the tree is representative of the whole tree and we would be able to use a subset only to estimate whole-tree scaling exponents and predict the whole-tree metabolic rate [2]

Despite its importance, few studies have assessed the external branching patterns from individual trees at branch-level, whole tree-level and analysed data in context of plant-scaling models. Within this scope, TLS has the ability to capture the complex 3D structure of trees, and 3D tree reconstruction models based on TLS point clouds, which allow us to model the shape of the trunk and main branches. [3], [4]. The aim of this study is to assess the feasibility of 3D structure models from TLS pointcloud data to estimate plant scaling parameters from tropical trees.

Materials and Methods

Data were acquired using a RIEGL VZ-400 V-Line 3D© Terrestrial laser scanner [RIEGL Laser Measurement Systems GmbH, Horn, Austria, www.riegl.com], mounted on a survey tripod 1.5 m above ground. The instrument is a full waveform LiDAR, operating at 1500 nm with an angular resolution between 0.0024 and 0.5 degrees, and a laser beam divergence of 0.35 mrad. Since the TLS only collects100 degrees of vertical angle per scan, a full-hemispherical scan was acquired by scanning two times on the same position; one in an upright position (perpendicular to the ground) and one in a tilted position (parallel to the ground).

The dataset was acquired during November 2013. The plot is located in latitude -12.83 degrees and longitude -69.27 degrees with an elevation of 223 masl and is located inside the Tambopata National Reserve, in Madre de Dios region, Peru. The plot is currently managed by the Global Ecosystem Monitoring network (GEM) from the University of Oxford, under the Andes to Amazon Transect Project.

The plot comprises an area of 100 by 100 meters with a regular square sample grid of 20 by 20 meters. Each intersection of the grid was used as a scan position. A total of 72 scan positions (2 scans per intersection, one upright and one tilted) were collected per plot. Individual scans were merged using the RIEGL proprietary software RiScan Pro and manual extraction was done to select individual trees in the plot. For this study 3 trees were manually extracted from the pointcloud data as seen in Figure 1. Finally, the pointcloud of each tree was exported to ASCII format.

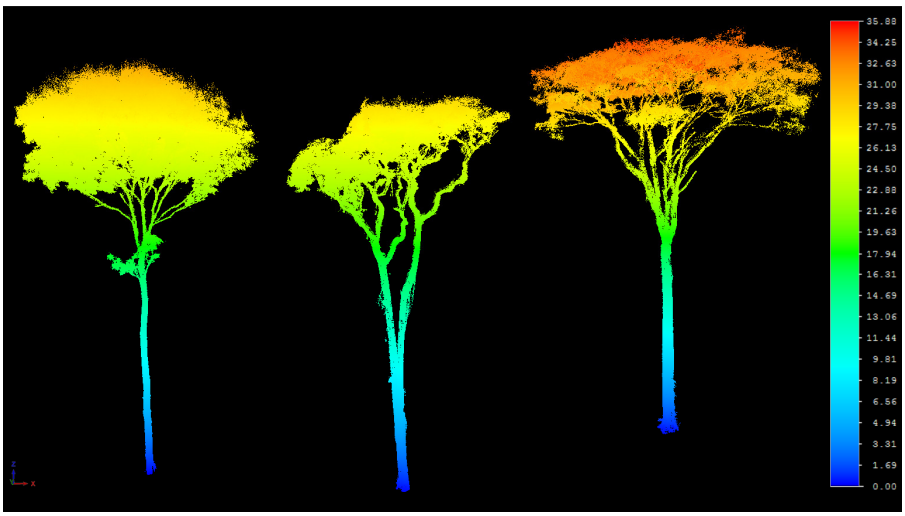


Figure 1: Tree pointcloud of scanned trees (in scale) from GEM plots in Peru; (left) *Tachigali polyphylla*, (centre) *Jacaranda copaia* and (right) *Sclerolobium bracteosum*.

For each tree a quantitative structure model (QSM) [3] was reconstructed using different settings from the tree pointcloud . Visual evaluation of the modelled tree compared to the original pointcloud gave us an insight of the optimal settings in order to determine an accurate evaluation. Then, specific branch parameters, such as branch hierarchy, branch length and radius (lower and upper) were calculated from the QSM. Finally, these parameters were used in the WBE plant-scaling model. This model calculated the following exponents: length ratio scaling, radii ratio scaling and estimated metabolic rate scaling.

Results and Discussion

The theoretical exponent expected from WBE for branch length scaling is 0.3 and for branch radii scaling is 0.5. Across all trees, the calculated branch-level length scaling exponent varied from 0.30 to 0.38 and the calculated branch-level radii scaling exponent ranged from 0.44 to 0.51 (Table 1). The calculated (estimated) metabolic rate scaling exponent was 0.70, 0.78 and 0.81 for tree 01, 02 and 03 respectively (expected to be 0.75 from the WBE model).

Table 1: Scaling exponents for branch length and radii

Tree	Length scaling	Confidence interval	Radii scaling	Confidence interval	Estimated metabolic rate scaling
Tachigali polyphylla	0.38	0.36 – 0.41	0.51	0.49 – 0.54	0.70
Jacaranda copaia	0.38	0.33 – 0.42	0.44	0.40 – 0.49	0.78
Sclerolobium bracteosum	0.30	0.25 – 0.36	0.46	0.40 – 0.52	0.81

Estimations of tree scaling metabolism derived from architecture via TLS scans showed consistent and comparable values to the model predictions for all scaling exponents. Since the scanned trees were different species, these results provide evidence to support the WBE assumption of similarities in branching structure and common set of branching rules across trees. It is possible that these deviations from the WBE predictions were not observed here due to the maturity of the sampled trees. It has been shown that the WBE model predictions apply more closely to trees that are larger and closer to the limits of an infinitely sized network [1].

Bentley et al [2] showed that the scaling of branch radii was less variable than the scaling of branch length. These results support the WBE assumption that energy minimization for water transport leads to minimization of hydrodynamic resistance. This assumption implies that changes in radius lead to bigger changes and more use of energy than changes in length; thus trees would be more able to respond to their environment through changes in branch length compared to branch width. Nevertheless, in the current study we could not find a strong evidence of this variability; length and radii were equally variable. While it might be possible that Bentley's results were driven by the inclusion of a large number of temperate trees, it is more likely that increased variability in length scaling was not observed in our study because the trees studied were large canopy trees. Therefore, these mature, emergent trees did not have high light competition and as such their branching structure could potentially be in close equilibrium with the demands from the surrounding environment. In the future, it would be useful to analyse more trees and their surrounding light environment demands to further explore this variability.

Lastly, these above findings are extremely relevant to field of ecology for the scaling of whole tree carbon and water use. Currently, the only way to extract these branch-level parameters from trees is to painstakingly

measure each branch part by hand. This procedure can take hours and only extracts parameters from part of the tree. Here, by coupling T-LiDAR with pointcloud processing codes we are able to extract branch-level parameters from a much larger percentage of the whole tree (and even in some cases from the whole tree). Thus, this fast and relatively easy way to calculate branch-level scaling exponents has the potential to revolutionize the ability to calculate the scaling of whole tree carbon and water use. That the branching laws follow a theoretically predictable function based on metabolic considerations provides a theory-based validation for this approach. Not only will researchers be able to calculate exponents for trees of much greater sizes than possible by hand, it will be possible to process a greater number of trees in much less time.

Conclusion

Tree scaling metabolism derived from TLS evidenced that (1) length ratio exponent, radii ratio exponent and architecture estimated metabolic rate converge between the tropical trees analysed, (2) there is no strong evidence to support that tropical tress are able to respond to environment through changes in branch length more so than branch width, and (3) length ratio exponent, radii ratio exponent and estimated metabolic rate from the analysed samples are comparable with the predicted values.

Acknowledgements

The present author of this research would like to thank Jose Luis de Tanago, Walter Huarasco Huaraca and the Oxford-GEM team in United Kingdom and Peru for making this possible.

References

[1] V. M. Savage, E. J. Deeds, and W. Fontana, “Sizing up allometric scaling theory.,” *PLoS Comput. Biol.*, vol. 4, no. 9, p. e1000171, Jan. 2008.

[2] L. P. Bentley, J. C. Stegen, V. M. Savage, D. D. Smith, E. I. von Allmen, J. S. Sperry, P. B. Reich, and B. J. Enquist, “An empirical assessment of tree branching networks and implications for plant allometric scaling models,” *Ecol. Lett.*, vol. 16, no. 8, pp. 1069–1078, Aug. 2013.

[3] P. Raumonen, M. Kaasalainen, M. Åkerblom, S. Kaasalainen, H. Kaartinen, M. Vastaranta, M. Holopainen, M. Disney, and P. Lewis, “Fast Automatic Precision Tree Models from Terrestrial Laser Scanner Data,” *Remote Sens.*, vol. 5, no. 2, pp. 491–520, Jan. 2013.

[4] K. Calders, G. Newnham, A. Burt, S. Murphy, P. Raumonen, M. Herold, D. Culvenor, V. Avitabile, M. Disney, J. Armston, and M. Kaasalainen, “Nondestructive estimates of above-ground biomass using terrestrial laser scanning,” *Methods Ecol. Evol.*, vol. 6, pp. 198–208, Nov. 2015.

On the potential of Terrestrial Laser Scanning for revising biomass allometric models of mangrove trees

Adewole, Olagoke[1,2], Christophe Proisy[2], Jean-Baptiste Féret[3], François Fromard[4], Uta Berger[1]

[1]Institute of Forest Growth and Computer Science, TU Dresden, Germany, adewole.olagoke@tu-dresden.de
[2]IRD, UMR AMAP, Montpellier, France, christophe.proisy@ird.fr
[3]IRSTEA, UMR-TETIS, Montpellier Cedex 5, France, jb.feret@teledetection.fr
[4]CNRS; EcoLab; 31062 Toulouse, France, Francois.Fromard@cict.fr

Highlights: A simplified method to compute the volume and biomass of large mangrove trees from TLS data is described. Pixel grids on 2D flattened images of segmented thin sections of trunk yielded biomass via volume estimates while the branches were considered as successive best fitted conical frustums. We proposed new Allometric models using TLS-derived biomass.

Key words: *Terrestrial laser scanning; Stem volume; Biomass allometry; Mangrove; French Guiana*

Introduction

Large trees are keystone components of forests. They contribute significantly to the characteristic structural dynamics and ecosystem functioning. Their presence strongly contributes to ecosystem biomass productivity and may explain variation in biomass distribution across forest landscapes [1]. Accurately determining the biomass of large trees is crucial for reliable forest biomass analysis. Large tree biomass data are however scarce in most tropical forests, and this is also the case in tall mangroves growing in the equatorial region. Even when tree species diversity in mangroves is low compared to the rainforests, mangrove trees of the same species may exhibit distinct growth forms, reflecting adaptation to local coastal environments. This fact may introduce bias when pan-tropical allometric models [2] are considered. Though some allometric models for estimating mangrove tree biomass are available (e.g. [3]), they are deficient in respect of large trees in their calibration. How reliable are these models beyond the DBH range of their calibration cannot be ascertained. Meanwhile, collecting biomass data of large mangrove trees is often a challenging task, considering the tree structure, heavy branches and interlocking canopies. Besides, cutting and weighing of large trees and their components in the field remains an expensive, dangerous and labour intensive task. That necessitates the quest for an alternative approach, and an impetus to consider the prospect of terrestrial laser scanning (TLS) technology.

The TLS recently found relevance in the non-destructive estimation of tree volume and biomass [4-5]. Various approaches have been tested for estimating tree volume and biomass, from geometric shape primitive fittings to voxelization [6]. To our knowledge, the application of TLS in the study of intricately structured mangrove species remains largely unexplored. Here, we addressed a twofold objective: 1) evaluating the performance of a pixel-based analysis of TLS data for accurate estimation of aboveground biomass of large trees, and 2) utilizing the TLS-derived biomass data for extending the validity range of allometric models for *Avicennia germinans* to large trees. To achieve this, we propose a simplified method particularly adapted to mangrove trees with buttresses (irregularities in stem shape) – a common feature of large-diameter mangrove trees. This entails trunk segmentation and pixel-count of thin sections to estimate the trunk volume. We then converted the volume estimates into biomass by integrating a height-dependent wood density model.

Methods

Sample data were obtained from 36 trees of *A. germinans* in French Guiana mangroves. The diameter size of the sample trees ranged from 13.4 cm to 124.5 cm. The TLS data acquisition was performed using a FARO Focus3D X330 laser scanner in a multiple-scanning mode for single or group of 2-5 trees. Post-scan processing of TLS points cloud data, including co-registration of scans, tree isolation and filtering of noises and cleaning background vegetation was achieved with the FARO's SCENE® 5.2 software. We manually separated TLS data into woody trunk and branches while the leaf component was filtered, and the data was exported in the 3D format (Fig. 1A) for further analysis using MATLAB® programming software. We developed a simple trunk volume computation routine that decomposes the 3D trunk shape into successive 10 cm thick sections (TS) which were converted to the 2D flattened images of plan projection (Fig. 1B&C). Omissions and occlusions during TLS acquisitions sometimes resulted in open shapes, which were automatically closed using a simple convex envelope. Then, the area covered by the boundary of the section was divided into a grid of 1cm square pixels. The number of pixels inside the TS projection image was summed up to obtain the surface area of the section, and the section volume was obtained as a product of its surface area and thickness. The height of each section from the ground

level (SH) was recorded and the trunk volume (VTR) was obtained as a total of all constituent sections. We compared the VTR with the geometric volume of successive best fitted conical frustums (Fig. 1D). We obtained branch volume (VBR) as the volume of best fitted conical frustums along the axis plane. Wood density (ρ) at different heights was retrieved, thanks to wood cores sampled at heights varying from 0.3 to 25 m. The biomass estimates of each trunk and branch components were obtained as the sum of the products of each section volume and corresponding ρ, which interpolates the decreasing values along stem axis. We evaluated the TLS-derived biomass by comparing estimates with the reference biomass obtained from the model proposed in [3] and mentioned as the reference model hereafter for sample trees of DBH<42cm. Combining the TLS-derived data with the existing biomass data, we fitted new allometric models of different forms and combination of explanatory variables (ρ, DBH, and height – H) for *A. germinans* trees and the goodness of fit tests performed appropriately.

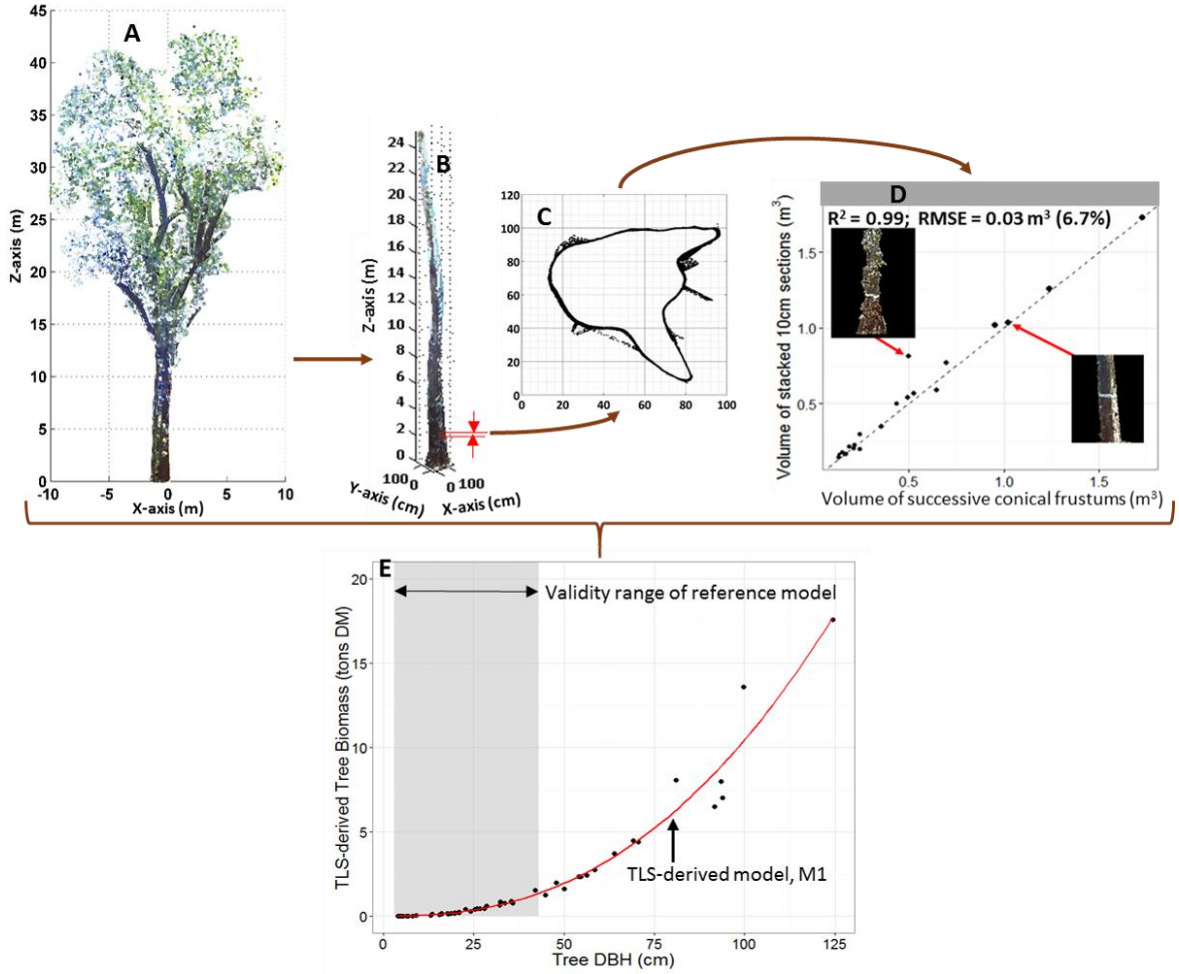


Figure 1: Illustration of tree TLS data, volume and biomass computation method and evaluation tests (*A: Typical view of a large mangrove tree*; *B: 3D Trunk of a tree with DBH=93.5 cm; C: Typical projected plan of 10cm high trunk section extracted at a height of 1.8 m; D: Trunk volumes – pixel-based method versus conical frustums; E: Allometric relationships between DBH and TLS-derived tree biomass; M1: see Table 1 below)*

Results and Discussion

Our approach for wood volume computation yielded a comparable trunk volume estimate to the geometric volume of successive best fitted conical frustums of trees without buttresses. The higher values of trunk volume obtained with our method for some buttressed or irregular shaped large trees reflected its better performance in that respect. Overall, the root mean square error of trunk volume estimate between the two methods reached 6.7% (Fig. 1D. Integrating a decreasing ρ value (height-dependent) into the wood volume yielded a higher accuracy in the estimation of woody biomass of trees. This resulted in a mean deviation of 18.2 kg for biomass of trees within the diameter range of the reference model, against the mean deviation of 47.8 kg returned by the mean ρ value. Overall, we obtained biomass values that correlated to the reference value to 0.99 (RMSE= 13.6%), a comparable accuracy to earlier studies [4-6]. This was convincing to adopt the TLS-derived biomass estimates of large diameter trees. Table 1 provides the summary of the reference model and the new biomass allometric models. All the new models resulted in higher predictive values in comparison to the reference model, but gave a less residual standard error for the biomass estimates of large trees. Though the predicted biomass values are not significantly different for all models, the M1 was adjudged the best model based on the evaluation parameters (Fig. 1E).

Table 1: Summary of allometric models and the associated parameters

Model ID	Model Form*	α	β	RSE (log-scale)	R^2	AIC
Reference Model	$AGB = \alpha.DBH^{\beta}$	0.14	2.44	0.29	0.97	12.3
M1	$AGB = \alpha.DBH^{\beta}$	0.16	2.42	0.23	0.99	-3.86
M2	$AGB = \alpha.\rho DBH^{\beta}$	0.23	2.41	0.23	0.99	-3.82
M3	$AGB = \alpha(DBH^2H)^{\beta}$	0.072	0.91	0.24	0.99	2.76
M4	$AGB = \alpha(\rho DBH^2H)^{\beta}$	0.099	0.91	0.24	0.99	2.57

AGB = Aboveground biomass; H=height; ρ = Wood density; α and β = coefficients; RSE = Residual standard error of estimates; AIC = Akaike Information Criterion, a measure of model's relative quality (the best model takes the lowest AIC); M1-M4= New Models

Concluding remarks

Our experiment confirmed the use of the TLS a suitable approach for measuring the biomass of large mangroves, though some challenges surfaced when applying the TLS to mangrove trees. Generally, buttresses and stem appendages (e.g. aerial roots) are commonly encountered in trees. Our approach, based on a plan projection of thin trunk sections (10 cm high) found relevance in providing accurate volume estimation for such trees over the estimates from the regular circle fittings or DBH measurements. The method yielded satisfactory biomass estimates that allow for extending the allometric models for *A. germinans* to large diameter trees and may hence improve biomass prediction for large trees. The current protocol performed well for near straight trunks and would require orthogonal axis rotation to ensure slicing in orthogonal directions. A seeming caveat is that over-estimation may ensue the TLS-derived biomass estimates should there exist a hollow in large old trees. In the objective of parameterizing mechanistic modelling approach [7] able to simulate changes in mangrove trees allometry, the prerequisite of achieving a full automatic procedure of volume and biomass estimation in natural, scrubby to tall mangrove forests, further refinement of the current TLS data collection and treatment procedure is required. This could be supported by works in tree architecture which are currently almost inexistent in mangroves.

Acknowledgements

This work was achieved in the framework of MANGWATCH project funded by the CNRS research programme 'Incubator for interdisciplinary research projects in French Guiana'. Adewole Olagoke is thankful to the Erasmus Mundus *Forest for Nature programme* (FONASO) for his PhD fellowship and the duo of the French *Centre National de la Recherche Scientifique* (CNRS) and the Technische Universität Dresden Graduate Academy for providing research travel grants. Jean-Baptiste Féret is thankful for the French *Centre National d'Etudes Spatiales* (CNES) post-doctoral fellowship. We also thank the TOSCA program of the CNES for providing us grant for describing 3D mock-up of mangrove trees within the frame of the BIOMASS mission preparation.

References

[1] Slik, J.W.F. et al. (2013). Large trees drive forest aboveground biomass variation in moist lowland forests across the tropics. Global Ecology and Biogeography 22:1261-1271

[2] Chave, J. et al. (2014). Improved allometric models to estimate the aboveground biomass of tropical trees. Global Change Biology 20:3177-3190

[3] Fromard, F., Puig, H., Mougin, E., Marty, G., Betoulle, J.L. & Cadamuro, L. (1998). Structure, above-ground biomass and dynamics of mangrove ecosystems: new data from French Guiana. Oecologia 115:39-53

[4] Calders, K. et al. (2015). Non-destructive estimates of above-ground biomass using terrestrial laser scanning. Methods in Ecology and Evolution 6:198-208

[5] Dassot, M., Colin, A., Santenoise, P., Fournier, M. & Constant, T. (2012). Terrestrial laser scanning for measuring the solid wood volume, including branches, of adult standing trees in the forest environment. Computers and Electronics in Agriculture 89:86-93

[6] Hauglin, M., Astrup, R., Gobakken, T. & Næsset, E. (2013) Estimating single-tree branch biomass of Norway spruce with terrestrial laser scanning using voxel-based and crown dimension features. Scandinavian Journal of Forest Research 28: 456-469

[7] Peters, R., Vovides, A.G., Luna, S., Grüters, U. & Berger, U. (2014). Changes in allometric relations of mangrove trees due to resource availability – A new mechanistic modelling approach. Ecological Modelling 283:53-61

Highly portable lidar mitigates occlusion in tropical forests through high density sampling schemes and novel deployment methods

Author names: Edward Saenz[1], Ian Paynter[1], Francesco Peri[1], Deborah Clark[2], David Clark[2], Leo Campos[2], William Miranda[2], Crystal Schaaf[1]

1. School for the Environment, University of Massachusetts Boston
2. La Selva Biological Station, Costa Rica

Highlights: Highly portable lidar can be deployed at varying heights to capture tropical forest sub-canopy and canopy information while mitigating occlusion. Hybridized reconstruction and voxelization approaches may offer improved biomass estimates from lidar acquisitions.

Key words: Tropical, terrestrial, lidar, occlusion

Introduction

Remote sensing captures important ecosystem characteristics at large spatial scales and meaningful temporal and spatial resolutions. Tropical forests are particularly challenging for airborne and spaceborne remote sensing resources due to overall height, the density and vertical stratification of vegetation, the variable moisture content and frequent cloud cover. Terrestrial lidar has proved a useful validation tool for airborne and spaceborne observations of temperate forests, as well as an increasingly effective independent characterization and measurement tool. However there are limiting logistic challenges to deploying lidar in tropical forests, including the difficultly in accessing and traversing these environments and other conditions which are hazardous to equipment such as frequent rainfall, heat and humidity. Even when deployment can be achieved, the dense understory of tropical forests introduces a large amount of occlusion, limiting the range of lidar pulses and therefore the overall spatial coverage of the lidar scans. In addition, occlusion limits information which can be found above the sub-canopy, including tree heights and crown structures.

Recently developed highly portable lidar can overcome many of the logistic challenges to deployment of terrestrial lidar in tropical forests, and additionally can facilitate mitigation of occlusion both horizontally and vertically.

Methods

The Compact Biomass Lidar (CBL) is a highly portable, rapid-scanning, eye-safe, time-of-flight, 905nm discrete lidar based on the SICK LMS151 and built by the University of Massachusetts Boston in collaboration with Rochester Institute of Technology. The CBL samples at an azimuth and vertical angular resolution of 0.25° and has a maximum range of 40m. In 2014 and 2015 the CBL scanner was deployed in La Selva National Park to acquire scans of long-term Carbono vegetation monitoring plots. In 2015, scans were taken on a 5m resolution grid of the 50m x 100m Carbono A2 and A5 sites (231 scans per site at a height of 1.5m). In addition, scans were taken at 3m, 6m, 9m and 11m height on a 10m x 10m resolution grid for a 20m x 20m sub-plot within the A5 Carbono plot. These variable-height scans were facilitated by mounting the CBL on a portable telescopic radio-mast tripod. Finally, a series of 1.5m scans were taken ad hoc to provide full angular coverage of trees with buttressed root structures within the 20m x 20m variable height sub-plot. The scanner was aligned to magnetic North and levelled prior to each scan to aid the co-location process.

Scans were co-located (relatively positioned in a Cartesian co-ordinate system) by a two-step process. The first step was an unsupervised X,Y and Z positional transformation based on the relative position of the scans according to the three-dimensional sampling scheme. The second step was a supervised, iterative process between sequentially paired scans, starting with the plot center scan as a reference. Calibration targets and GPS are not practical in dense, high-occlusion tropical environments, so co-location remains a challenge. Inertial Movement

Units (IMU) combined with traditional surveying techniques may offer the best avenue for automatic co-location of scans, though the semi-supervised approach employed here was still acceptably efficient.

Quantitative Structure Modelling (QSM) [1] was applied to the individually extracted trees for woody biomass estimation. An automated approach to stem detection previously designed for use in temperate forests was tested for its efficacy when initialized with scans of different heights.

Buttress volume was estimated by mesh formation from Poisson resampled point cloud segments which were extracted by analysis of trunk diameter stabilization, and manual filtering of unrelated ground points between roots. The resulting mesh was reductively facetized to form a closed polygon, the volume of which was then measured.

Information gain with horizontal and vertical view angle was assessed in terms of contributions to tree structure at varying heights; unique view angle analysis; and occlusion mitigation assessed with a combined pulse-tracing and voxelization method. Stabilization of the QSM output with combinations of source scans contributing to the extracted point cloud was also utilized as a metric of information quality.

Thus a hybridized approach to biomass estimation was investigated, utilizing QSM for trees of Diameter at Breast Height (DBH) greater than 10cm; voxelization for a volumetric estimation of vegetation associated with stems under 10cm DBH; and volumized polygons for buttress roots.

La Selva offers a unique tropical forest data set with extensive high quality field measurements of the vegetation in the Carbono plots to serve as validation data. The large amount of lidar data collected at La Selva has also provided the opportunity to explore the information gain from varying horizontal scan densities (10m x 10m, 15m x 15m, 20m x 20m and randomly resampled grid) as well as three dimensional tree reconstructions utilizing view angle contributions across the horizontal and vertical ranges and tree biomass estimation from QSM with different combinations of vertically distributed scans. The multi height deployment also allows a detailed analysis of the sub-canopy occlusion via the contributions to structure from scans above and below (see Figure 1). Finally the availability of detailed scans of buttressed trees facilitates an analysis of buttress volume estimation with respect to view angles, information density and atypical internal structure.

Results

We will present preliminary findings from ongoing view angle analysis, tree reconstruction by QSM, biomass estimation and atypical tree morphology reconstruction efforts.

Discussion

Highly portable and rapid scanning lidar instruments such as the CBL can augment highly capable, but less portable terrestrial lidars in easily accessible environments, and are serving a role in provide robust independent observations in less accessible environments such as tropical forests. The portability offers a positional flexibility that can mitigate occlusion both in terms of offering higher density of scans in the horizontal plane, and alternative view angles within and above the sub-canopy in the vertical plane.

Occlusion analysis is an important measure to quantify the benefits of highly portable lidar and evaluate the information quality of the data. These assessment techniques are rapidly improving in terms of their resolution and computational efficiency. Eventually, occlusion analysis should allow us to analyze information quality on the fly, and form appropriate scanning schemes responsively in the field. Not only would this truly optimize the benefits of portable lidar, but would guarantee the information quality of the deployment. This relies upon appropriate definitions of information quality, which can be determined by sensitivity analyses on results obtained from information-saturated experiments such as these at La Selva.

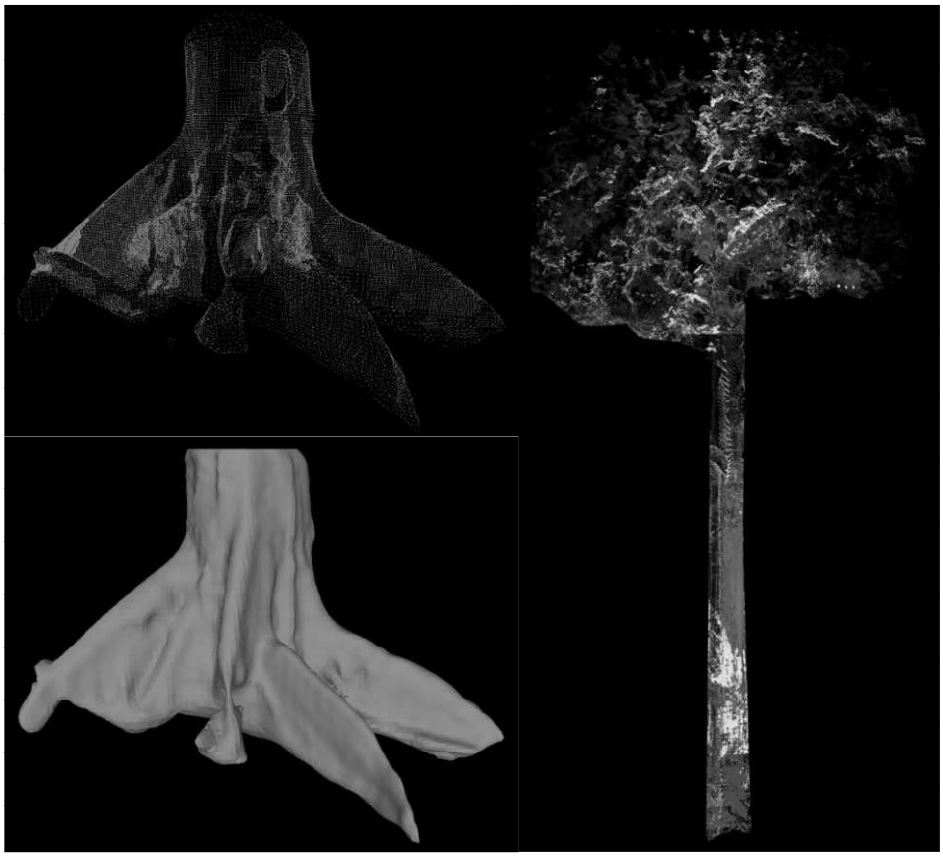


Figure 1: (Anti-clockwise from top-left) Poisson-resampled point cloud extracted from CBL scans of buttress roots of saber tree; surface fitted to k-neighbor mesh of buttress roots; saber tree extracted from 5 CBL scans at varying heights, color of points denotes contribution of individual scans to point cloud

References

[1] Raumonen, P., Kaasalainen, M., Åkerblom, M., Kaasalainen, S., Kaartinen, H., Vastaranta, Holopainen, M., Disney, M., & Lewis, P. (2013). Fast automatic precision tree models from terrestrial laser scanner data. Remote Sensing, 5(2), 491-520.

First examples from the RIEGL VUX-SYS for forestry applications

Mandlburger Gottfried[1], Markus Hollaus[1], Philipp Glira[1], Martin Wieser[1], Milutin Milenković[1], Ursula Riegl[2] and Martin Pfennigbauer[2]

[1]Vienna University of Technology, Department of Geodesy and Geoinformation, Research Group Photogrammetry, Gußhausstraße 27-29, A-1040 Vienna, Austria (gottfried.mandlburger, markus.hollaus, philipp.glira, martin.wieser, milutin.milenkovic@geo.tuwien.ac.at)
[2]RIEGL Research Forschungsgesellschaft mbH, Riedenburgerst. 48, 3580 Horn, Austria (uriegl, mpfennigbauer@riegl.com)

Highlights: Very high point density UAS-based laser scanner point clouds of an alluvial forest acquired with the RIEGL VUX-SYS were analyzed w.r.t. forestry applications. With point densities >1500 points/m² and accuracies <2 cm the study shows that individual stems and branches, understory, lying deadwood, and the terrain are clearly represented in the 3D point clouds.

Keywords: *UAS, LiDAR, forest, single trees, understory, deadwood, terrain roughness*

Introduction

While UAS-based remote sensing using imaging sensors is already widespread, the integration of LiDAR sensors on remotely piloted aerial systems (RPAS) still remains challenging [1]. But the development of compact scanners has opened the possibility to operate laser sensors on UASs for civil applications including forestry [2]. Especially the introduction of UAS-based survey-grade full waveform scanners [3] allows mapping of forested environments in unprecedented detail.

Until now high point density laser scanning data, suitable for modeling individual trees on a stem-branch level, could only be derived from terrestrial laser scans (TLS). For example in [4] TLS data were acquired for two forest stands in extensive field work campaigns. The TLS data were used to model individual trees including the stems and branches. It was shown that the main limiting factor for modeling of entire trees was the reduced visibility of the upper tree parts and the high amount of field work for large area scanning campaigns, especially in topographic complex terrain and dense natural forests. The missing information in the upper canopy could not be complemented by airborne laser scanning data (ALS) with a similar level of detail due to the system configurations (i.e. footprint size and planimetric accuracy ~30 cm, point density ~10 points/m²).

The recent generation of UAS-based laser scanners (ULS) [4] opens new possibilities to model individual trees but also the understory and the terrain in cm-accuracy. Due to the very low altitude, the small footprint size, the large scanner field-of-view (FOV), and the very high effective scan rate the description of environmental scenes can be achieved with very high level of detail and completeness. Especially for forestry applications this data can be used to derive 3D models for individual trees, to describe the understory, and to derive high precision terrain models.

The objectives of this contribution are to demonstrate first examples of high precision UAS-based LiDAR data for vegetation studies and for terrain characterization. In addition to georeferencing issues, the achievable level of detail of the derived data will be shown and discussed.

Material and methods

The study site Neubacher Au (Figure 1b) is a Natura2000 conservation area located at the lower course of the Pielach River (Lower Austria) near the confluence with the Danube River. Within the last two years the area has been monitored with a topo-bathymetric LiDAR system. Due to the complex topography and vegetation structure in the alluvial area the derivation of a high quality Digital Terrain Model (DTM) posed problems using the medium footprint size bathymetric system and advocated for applying UAS-based laser scanning.

On February 26 and 27 the alluvial forest area was captured with the VUX-SYS compact laser scanner mounted on a RiCopter UAS platform under leaf-off conditions. The sensor is equipped with a Global Navigation Satellite System (GNSS) and an Inertial Measurement Unit (IMU) for capturing the flight trajectory. The scanner features an effective measurement rate of 350 kHz with a total FOV of 230°. In addition, two RGB cameras (Sony Alpha 6000) were mounted on the UAS on the first flight day (western part of the study area). Including batteries and all sensor components the payload weighed 15 kg and the take-off mass was just below the maximum of 25 kg, enabling a flying time of 30 minutes.

Data capturing was carried out based on a state of the art ALS flight planning with longitudinal and cross strips. The regular strip distance was 40 m and the flying altitude 50 m above ground. Depending on the sensor-to-target range this resulted in a laser footprint diameter of 1.0-2.5 cm enabling detection of small vegetation

objects. The UAS was autonomously flying the programmed path (based on waypoints) with a speed of 8 m/s. The mission parameters and the large scanner FOV resulted in a mean laser pulse density of 1500 points/m^2 and in multiple strip overlaps. Thus, the vegetation was captured from all sides.

The high strip overlap was additionally used for a thorough calibration of the entire sensor system using the software RiProcess. The georeferencing quality of the 31 strips was further improved in post processing by a rigorous strip adjustment. This was done by simultaneously minimizing the point-to-plane distances of approximately 200000 correspondences. Within the strip adjustment, the acquisition system was fully re-calibrated. This includes the estimation of scanner calibration parameters (e.g. range finder scale error), mounting calibration parameters (i.e. misalignment and lever-arm), and strip-dependent trajectory errors (i.e. GNSS and IMU errors). The adjustment led to a substantial quality improvement of the acquired point clouds with a final relative accuracy of 1.7 cm (defined as 1 sigma of the a posteriori DEM height differences within smooth areas).

Results and Discussion

Figure 1a shows the VUX-SYS consisting of a RIEGL VUX-1 airborne laser scanner including a tightly coupled GNSS/IMU unit, a control unit, and a camera for video downstream. Additionally, a digital RGB camera was mounted on the RiCopter. A perspective view of the dense 3D point cloud colored by reflectance is illustrated in Figure 1c demonstrating the completeness of the 3D data capturing of the forest scene. The seamless data capturing around the trunks and stems of the trees is further highlighted by the horizontal section of the point cloud in Figure 1d. In this figure individual colors are used for the points of each flight strip and, therefore, also demonstrate the remarkable georeferencing quality. Finally, Figure 1e shows a vertical section of the alluvial forest near the terrain. It can be seen that both the terrain and the vegetation is represented with high spatial density in the point cloud. Especially deadwood and accumulated woody debris can be modeled in high detail based on the captured data.

The availability of high density and high precision 3D data for wooded areas opens new possibilities for forestry applications like single tree modeling and terrain characterization under dense forest canopy. As shown in Figure 1c tree stems and branches are represented well in the 3D point cloud and, therefore, geometric tree models can be derived with methods that are in use for 3D tree modeling based on TLS data (e.g.[4]). Such 3D models can support or even partly replace traditional field-measured forest inventories i.e. calculating growing stock and biomass, measuring tree height, characterization of the understory, detection of standing and lying dead wood, etc. Identifying the latter was possible by removing the medium and tall vegetation using standard DTM filtering strategies [5]. The remaining near ground point cloud showed remarkable details of the abundant deadwood of the study area. Moreover, the captured point cloud exhibits only minimal shadowing (cf. Figure 1d) due to the overlaps of the flight strips and the large FOV of the VUX-SYS. In comparison to TLS, ULS is characterized by shorter acquisition times and better representation of the entire vertical forest structure (terrain, understory, and canopy). The capability of acquiring large sample plots or transects also excels ULS as a promising data acquisition method for reference data collection within forest inventories.

Conclusions

In this contribution we showed the first results of a very high resolution 3D data acquisition of a complex alluvial forest area. An average point density of 1500 points/m^2 resulting in a mean point spacing of 2.5 cm was achieved by applying UAS-based laser scanning using the Riegl VUX-SYS sensor system. The device consists of an octocopter carrying the compact survey-grade VUX-1 laser scanner, which is additionally equipped with a GNSS/IMU unit and optional RGB cameras. High georeferencing quality of the flight block consisting of 31 strips could be achieved by applying rigorous strip adjustment based on the data from the highly overlapping strips. Quality control confirmed an accuracy of less than 2 cm.

UAS-based laser scanning allows flexible data acquisition tailored to the needs of the specific study area. For the experiment at hand it turned out that the topography and the vegetation could by captured in high detail and with seamless data coverage. This especially applies for the stems and branches of the trees which are covered with points from all sides. Completeness of the data acquisition resulted from both the dense array of flight strips and the large scanner FOV of 230°. The captured 3D point cloud, therefore, combines the advantages of ALS (homogeneous planimetric data distribution, bird's eye perspective) and TLS (short ranges, 3D vegetation structure). An additional advantage was identified for the canopy area as the slender branches at the top of the trees benefit from the smaller laser footprint compared to the trunk area (canopy: ~1 cm, terrain: ~2.5 cm).

The primary goal of the flight experiment was to create a trustworthy reference DTM capturing all details of the fluvial floodplain topography. Beyond that, analysis of the point cloud in the vegetation area confirmed, that ULS can also serve as reference data acquisition for a variety of forestry applications like biomass estimation, tree height measurements, forest structure assessment, etc. Furthermore, in the open terrain the estimation of surface roughness was identified as a potential field of application for UAS-based laser scanning.

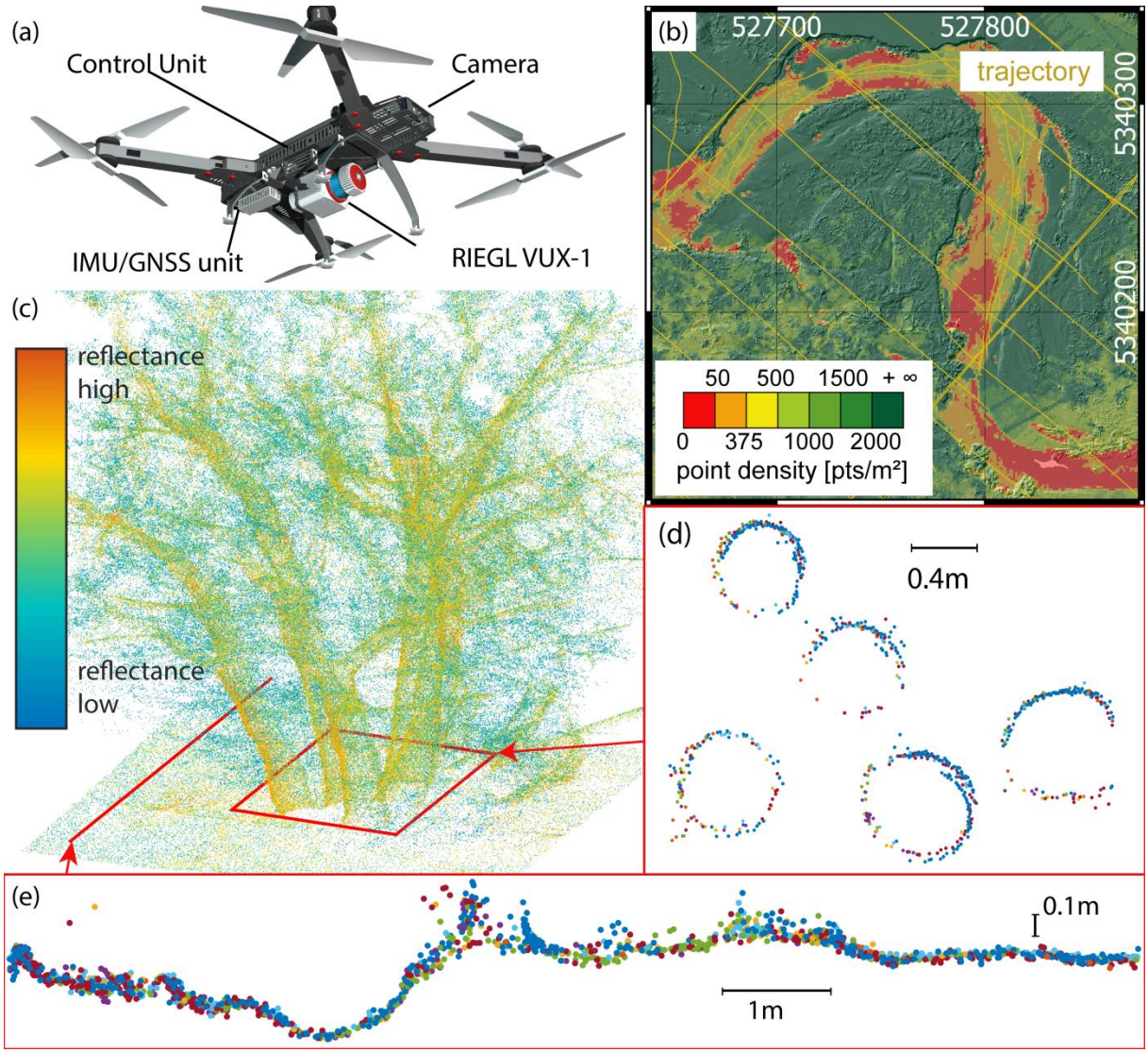


Figure 1: (a) RiCopter/VUX-SYS sensor system, (b) Study area Neubacher Au, DTM shading superimposed with point density map and flight trajectories, spatial reference system: ETRS89/UTM33, (c) 3D LiDAR point cloud colored by reflectance, (d) Horizontal tree stem section at breast height (vertical layer thickness: 10 cm), (e) Terrain profile (width: 10 cm)

Acknowledgements

This work was funded by the Austrian Research Promotion Agency (FFG) COMET-K project "Alpine Airborne Hydromapping – from research to practice" and by the European Commission FP7 project "Advanced SAR".

References

[1] Colomina, I., & Molina, P. (2014). Unmanned aerial systems for photogrammetry and remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensing, 92, 79-97.

[2] Wallace, L. O., Lucieer, A., & Watson, C. S. (2012). Assessing the feasibility of UAV-based lidar for high resolution forest change detection. International Archives of Photogrammetry, Remote Sensing, and Spatial Information Sciences, XXXIX-B7, 499-504.

[3] Amon, P., Rieger, P., Riegl, U., & Pfennigbauer, M. (2014). Introducing a New Class of Survey-Grade Laser Scanning with Unmanned Aerial Systems (UAS). Proceedings of the FIG Congress 2014, Kuala Lumpur, Malaysia.

[4] Eysn, L., Pfeifer, N., Ressl, C., Hollaus, M., Grafl, A., & Morsdorf, F. (2013). A pracitcal approach for extracting tree models in forest environments based on equirectangular projections of terrestrial laser scans. Remote Sensing, 5, 5424-5448.

[5] Pfeifer, N., & Mandlburger, G. (2008). Filtering and DTM Generation. In J. Shan, & C. Toth (Eds.), Topographic Laser Ranging and Scanning: Principles and Processing (pp. 307-333), CRC Press.

Reconstructing forest canopy from the 3D triangulations of airborne laser scanning point data for the visualization and planning of forested landscapes

Jari Vauhkonen

University of Eastern Finland, School of Forest Sciences, jari.vauhkonen@uef.fi
University of Helsinki, Department of Forest Sciences, Helsinki, Finland.

Highlights: A canopy surface modelling approach based on the triangulations of airborne LiDAR data is described and evaluated. The approach preserves the 3D properties for genuine volumetric surface modelling. Realistic visualizations for forest management planning applications are demonstrated.

Key words: *Surface modeling, Delaunay triangulation, Alpha shape, Persistent homology, Simulation*

Introduction

Airborne laser scanning (ALS) has become a routinely operated technique for various forestry analyses [1]. However, operational approaches are most often based on sparse point densities (<1 pulses m^{-2}) and therefore typically limited to area-based analyses of the distributions of height values. Notably, such analyses employ only one of the three dimensions available in the data and ignore the (horizontal) between-point information. Triangulating point data, i.e. subdividing the underlying space of the points into simplices, and subsequent filtering of these triangulations has been proposed as an alternative means of representing the canopy surface [2]. A potential improvement could be obtained by a persistent homology analysis [3] – see the Methods section.

The purpose of this presentation is to describe the triangulation approach for generating reliable 3D models and visualizations of forest canopy. The resulting triangulations, filtrations and canopy volumes were evaluated against field measurements. The filtering approach allows a simulation of the effects of biomass growth and harvests to the canopy, which is demonstrated in a fully automated visualization of a forested landscape.

Material and methods

Study area and data

The study area is a typical boreal managed forest area in eastern Finland (62°31'N, 30°10'E). Altogether 79 field plots of 20×20 to 30×30 m were measured on May–June 2010. Stem volume estimates were calculated using the measured DBH and height in species-specific volume equations.

The ALS data used in the study were acquired on 30 April 2012 with Leica ALS 60 system from a flying height of 2350 m and with a nominal density of 0.79 points m^{-2}. The ground surface was detected and classified and the echo heights were normalized with respect to this surface following a standard TerraScan approach. Only the first echoes (i.e. 'only' and 'first-of-many' of multiple echoes recorded per pulse) extracted employing a ground threshold of 1 m were included in the analyses.

Methods

Opposite to the typical "surface modelling" approaches that are based on 2 to 2.5D images representing counts of height values within pixel or voxel cells, the aim here was to generate genuine 3D models preserving the between-point information in the data. These are obtained by triangulating the ALS point data, i.e. subdividing the underlying space of the points into simplices, which quantify the (empty) space delimited by the points. Reconstructing the canopy volume populated by biomass will thus require filtering to exclude that volume from canopy voids [2]. The degree of filtering was determined by analyzing the inherent algebraic properties of the triangulations such as simplicial homomorphism and topological persistence [3].

Following [2], the point data of each individual plot were triangulated and filtered separately. The filtration corresponded to the family of 3D α-shapes [4], each of which consist of the simplices of the initial triangulation with an empty circumscribing sphere with a squared radius smaller than α. Although an α-shape is defined for every non-negative real α, there are only finitely many different α-complexes and a filtration is obtained by ordering these complexes according to the α [5]. All computations were based on Delaunay triangulations and implemented with C++ and an open-source Computational Geometry Algorithms Library (http://www.cgal.org).

During the filtration, i.e. by increasing the considered filtration parameter α, simplices of a dimension d join together with other d-dimensional simplices to form structures with a higher dimensionality. Particularly, when the value of α allows two points (d=0) to connect, an edge with d=1 between these two points is formed. The edges further join to form facets (d=2) and tetrahedra (d=3). Each simplex can thus be assigned by a value

describing its persistence in the structure formed. The index of persistence was obtained as the absolute difference between the index values of α causing the birth and death of simplices of the given dimension [3], and used as a criterion to determine the degree of filtering and therefore the initial state of the forest.

The numerical evaluation of the approach was based on assessing the correspondence of the total tetrahedral volume of the selected α-complex with the plot-level stem volume measured in the field. To demonstrate the potential of the proposed approach for 3D visualizations of forest canopy, the ALS point data of a forested area of approximately 1000 × 1200 m was divided on a grid with a cell size of 20 × 20 m. Each cell of this grid was separately triangulated and analyzed for its persistent homology. The effect of harvesting biomass from the area was visualized by re-assigning simplices as empty space, i.e. adjusting the value of α downwards.

Results and discussion

The tetrahedral volume derived by analyzing the persistent homology of the triangulations (R^2=0.17 with stem volume) was not as accurate as the reference canopy volume produced by optimization following [2] (R^2=0.65). However, noting that no allometric information on the forest attributes was used in the derivation (unlike in the case of the optimized one), the metric deduced from the properties of the point data does include information on the variation of the stem volume in the area.

Despite the low degree of determination with the field measured stem volume, the positions of the canopy elements derived with the persistent homology approach corresponded well to those of the field-measured tree stems, which showed a considerable potential in producing realistic visualizations of a forested landscape (Figure 1). An inherent property of the approach is that it allows simulating the effects of biomass growth and harvests to the visualizations. The simulated reduction of the canopy volume was found to describe the effects of biomass harvesting realistically (Figure 1). The approach will be further tested in surveys aiming to elicit preference information on the impacts of the harvests to the scenery for designing the management planning of the area.

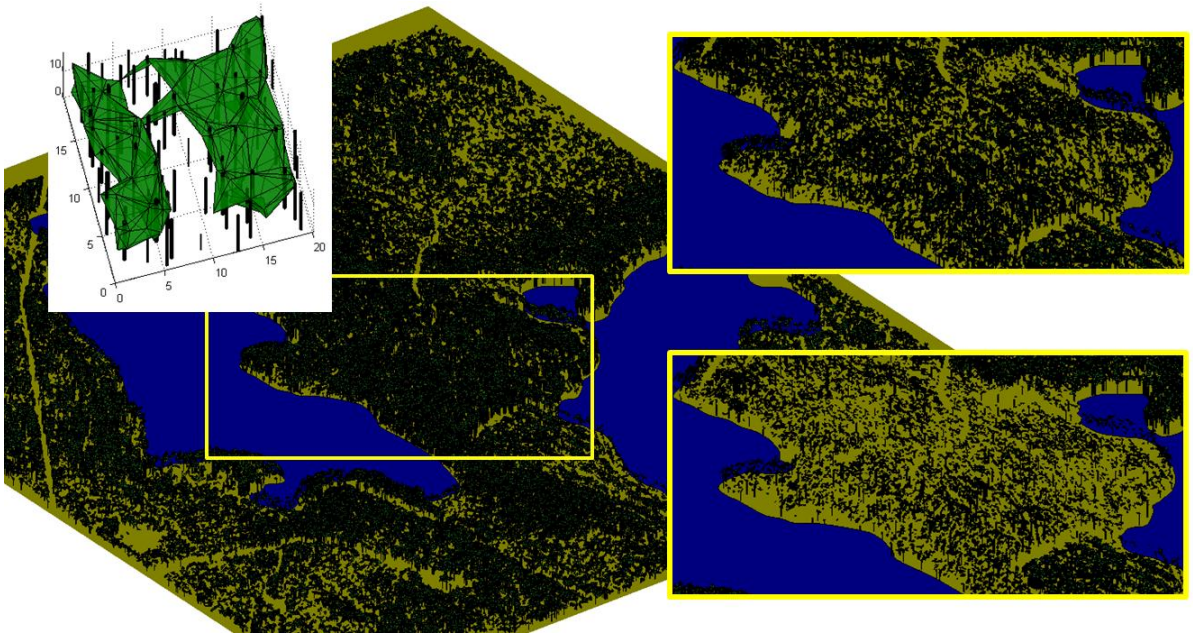


Figure 1: Visualizations of a forested landscape produced by the proposed approach. Upper left: an example of an individual 20× 20 m cell. Right: harvests simulated to the compartment in the centre of the landscape.

References

[1] Maltamo, M., Næsset, E., & Vauhkonen, J. (2014). Forestry applications of airborne laser scanning - concepts and case studies. Managing Forest Ecosystems, 27, The Netherlands: Springer, 464 p.

[2] Vauhkonen, J., Næsset, E., & Gobakken, T. (2014). Deriving airborne laser scanning based computational canopy volume for forest biomass and allometry studies. ISPRS Journal of Photogrammetry and Remote Sensing, 96, 57–66.

[3] Edelsbrunner, H., Letscher, D., & Zomorodian, A. (2002). Topological persistence and simplification. Discrete Computational Geometry, 28, 511–533.

[4] Edelsbrunner, H., & Mücke, E.P. (1994). Three dimensional alpha-shapes. ACM Transactions on Graphics, 13, 43–72.

[5] Delfinado, C.J.A., & Edelsbrunner, H. (1995). An incremental algorithm for Betti numbers of simplicial complexes on the 3-sphere. Computer Aided Geometric Design, 12, 771–784.

From leave scale to tree scale: which structural parameters influence a simulated full-waveform large-footprint LiDAR signal?

Cécile Antin[1], Eloi Grau[2,3], Grégoire Vincent[1], Sylvie Durrieu[3], Marion Jourdan[4], Jean-François Barczi[5], Yves Caraglio[5], Sébastien Griffon[5], Raphaël Pélissier[1]

[1] *IRD, UMR AMAP, F-34000 Montpellier, France*
[2] *CNES, UMR TETIS, F-34000 Montpellier, France*
[3] *IRSTEA, UMR TETIS, F-34000 Montpellier, France*
[4] *CNRS, UMR AMAP, F-34000 Montpellier, France*
[5] *CIRAD, UMR AMAP, F-34000 Montpellier, France*

Highlights: We explored the impact of tree structural parameters on spatial LiDAR signal. We created detailed virtual representations of forest scenes that we subsequently altered structurally and geometrically. We simulated a large-footprint LiDAR signal on the virtual scenes. Alteration of forest scenes impacted the metrics extracted from the backscattered waveforms up to a 136.4% difference.

Key words: *virtual forest scene, plant architecture, radiative transfer, spaceborne LiDAR, simulation, sensitivity analysis*

Introduction

Light Detection And Ranging (LiDAR) technology has the potential to assess forest structural parameters and notably canopy height, wood biomass or volume, and leaf area index [1]. Although these parameters are typically estimated at the stand level, the LiDAR backscattered signal is affected by vegetation structural properties at tree level. For example, leaf clumping and leaf inclination distribution are believed to have a significant effect on the backscattered signal [2], whereas crown archetypes assumptions significantly affect the parameters derived from LiDAR return [3]. Simulation of LiDAR signal on realistic and detailed virtual representations of forest stands can be used to better understand how the LiDAR signal interacts with the different parts of a tree, which is a step towards the development of improved methods for assessing vegetation characteristics from LiDAR data.

The main objective of this work is to explore how vegetation structural parameters at a fine scale (i.e. individual tree level) impact basic metrics extracted from the backscattered waveform of a large footprint LiDAR system. LiDAR waveforms were simulated on a realistic virtual forest scene. A sensitivity analysis was conducted by altering virtual scenes to assess the impact of structural and geometrical simplifications of the tree representation on the simulated LiDAR signal and on few metrics derived from the waveforms. The results of this sensitivity analysis should provide insights into which tree structural parameters (1) precisely impact the LiDAR simulations, and (2) have to be independently estimated for the inversion of a real LiDAR waveform.

Materials and methods

We first applied an original approach to create tridimensional realistic virtual representations of maritime pine plantations by the simulation of tree development in stand (i.e., including inter-tree competition and functional relationships between branch section and supported leaf area). The resulting forest scenes, representative of a set of plots measured in the field, were validated by a comparison with forest inventory data. We then simulated a large-footprint LiDAR signal on the virtual forest scenes and their altered versions in order to test whether the modified parameters impacted the backscattered waveform.

Creation of 3D realistic virtual scenes of maritime pine plantations

The creation of the virtual forest scenes is based on the botanical knowledge of plant architectural development [4] and takes into account plant-to-plant interactions. The scenes were produced with the AmapSim software [5], available from the AMAPstudio software suite for plant architecture modelling [6]. We used forest inventory data and terrestrial LiDAR data from seven plots from the study site of St-Symphorien (France), in the Landes' forest, to calibrate age-dimension distributions of diameter at breast height (dbh), total height, height of crown base, and crown area projection. Additionally, we used published reports to calibrate age-foliage area relationships [7]. The simulations were constrained so as to fit with distribution of tree dimensions observed in the reference plots and reported in the literature.

Sensitivity analysis of a simulated LiDAR signal to vegetation structure

The 3D virtual scenes of maritime pine plantations obtained with AMAPsim were post-processed with Xplo, also available from AMAPstudio, in order to segment virtual trees according to both the nature (leaves vs. wood) and the dimensions (branch section) of their components. The realistic (i.e. non-altered) representations served as references for the sensitivity analysis. They were then altered in three different ways, while preserving tree density, canopy height and LAI of the original scenes. (1) Geometric alterations modifying the shape, size and inclination of leaves in the reference scene were conducted with Xplo. (2) A foliage voxelization was performed with DART by transforming the meshed representation of the foliage into a turbid medium and attributing to series of cubic volumes (voxels) the leaf area density of the reference scene. The leaf inclination angle distribution attributed to each voxel was either assumed to fit a predefined model (spherical, erectophile, or planophile distributions were tested) or was computed with actual angles of individual needles on the reference scene. (3) All tree components (trunks, fine, medium or large branches, and leaves) were alternatively removed to assess their specific contribution to the signal.

A full-waveform large-footprint LiDAR signal was simulated on these representations using DART software for radiative transfer modeling [8]. The simulated laser pulsed at 1064 nm from a distance of 500 km above the ground with an inclination of 2° from the zenithal direction. The simulated footprint measured 20 m of diameter.

Eight metrics were extracted from each simulated waveform (Fig.1): *altMax* (difference between the signal begin and the ground peak), *altPic* (difference between the vegetation peak and the ground peak), *ext* (waveform extent), *VGRP* (ratio between vegetation peak and ground peak), *VGRT* (ratio between vegetation total area and ground total area), *rh15*, *rh50* and *rh85* (distance from the signal begin of the 15, 50 and 85 percentiles of signal intensity). They were then compared across simulations.

Results and discussion

We obtained detailed virtual scenes composed of tridimensional meshed objects with known properties for every mesh, i.e. nature, dimension, orientation, inclination. The virtual trees are botanically realistic as they render the topology and geometry of plant elementary components, and fit the distribution of tree dimensions (diameter at breast height, total height, height of crown base, crown area projection, foliage area) either measured or predicted by allometries for the reference plots. Two virtual forest scenes were selected; corresponding to two inventoried circular plots of 15 m radius of maritime pines aged 19 and 60, respectively.

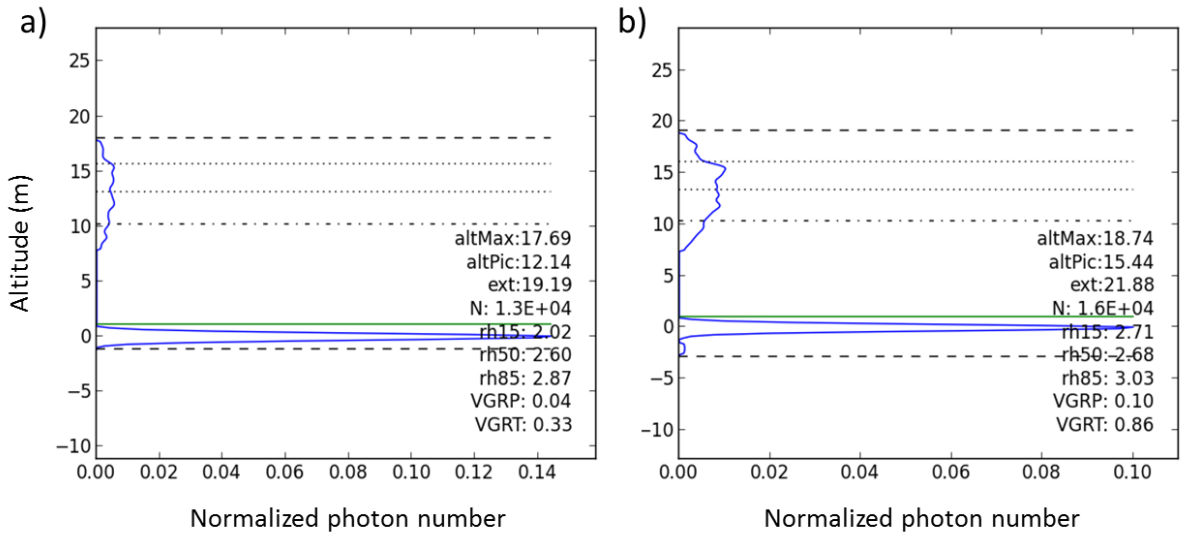


Figure 1: Waveforms and associated LiDAR metrics for a detailed virtual scene (19 years old maritime pines) (a) and for the same scene with a voxelized representation of the foliage with a resolution of 1 m for the voxels (b).

The LiDAR simulations on the scene of maritime pines aged 19 showed that the less sensitive metrics were *altMax* and *ext* (absolute prediction error ≤ 15.26%, see Table 1) whereas *VGRP*, *VGRT* and *rh15* were more sensitive to the alterations performed (absolute prediction error up to respectively 125.00, 136.36 and 131.19%). The geometric alterations had less impact on the metrics than foliage voxelization or components suppression. The foliage voxelization impacted the metrics extracted from the waveform with an effect increasing with the voxel size (25 to 200 cm). For the foliage voxelization with actual predefined leaf inclination angle distribution, the most affected metrics were *VGRP* and *VGRT*. The impact of voxelization was more important for a planophile distribution, intermediate for a spherical distribution and less important for an erectophile distribution (respectively 125, 50 and 25% for *VGRP* and 136.36, 54.55 and 30.30% for *VGRT*). The alternative suppression of tree components showed the predominant role of leaves in the simulated waveform and the minor role of

twigs and trunks. These results are consistent with the conclusions of a similar analysis on simulated small-footprint LiDAR [9].

Table 1: Maximum absolute prediction error (%) for 6 metrics extracted from simulations on altered virtual scenes (only the differences ≥ 5% are reported).

Modifications type (number of simulations)	VGRP	VGRT	altMax	altPic	ext	rh15
Leaf size (4)		6.06		25.95		
Leaf form (2)	25.00	21.21		25.95		8.91
Leaf angle of insertion (9)	25.00	27.27		25.95		8.91
Reflectance of trees (6)		12.12				7.92
Reflectance of soil (6)	75.00	63.64	9.33		8.60	81.19
Tree components suppressions (7)	75.00	90.91	15.26	25.94	14.07	131.19
Foliage voxelization (4)	125.00	136.36		25.95	11.72	34.16

Conclusion

These results indicate that having a precise description of small branches is not of critical importance for the modeling of vegetation-LiDAR interactions, but that a detailed representation of trees performed better than a voxelized scene (turbid medium representation of the foliage) for LiDAR waveform simulations.

The substantial impact of the leaf inclination distribution on the simulated signal suggested that this parameter should be carefully parameterized in simulations and that caution is recommended for the inversion of a real LiDAR waveform. In particular, a spherical distribution hypothesis was not the most appropriate simplification in our study case.

The sensitivity analysis showed that structural and geometrical alterations of detailed tree representations (especially leaves representation) have minor impact on the simulated LiDAR signal. Identifying the right level of simplification of foliage for LiDAR (and other signals) simulation has clear practical implications as it may significantly reduce the computing time in radiative transfer simulations.

Acknowledgements

The authors acknowledge funding from the French CNES/TOSCA program. This work also benefitted from data acquired in the frame of the FORESEE project (ANR-2010-BIOE-008) granted by the French National Research Agency (ANR).

References

[1] van Leeuwen, M., & Nieuwenhuis, M. (2010**)**. Retrieval of forest structural parameters using LiDAR remote sensing. European Journal of Forest Research, 129, 749–770.

[2] Detto, M., Asner, G. P., Muller-Landau , H. C., & Sonnentag, O. (2015**)**. Spatial variability in tropical forest leaf area density from multireturn lidar and modeling. Journal of Geophysical Research: Biogeosiences, doi:10.1002/2014JG002774.

[3] Calders, K., Lewis, P., Disney, M., Verbesselt, J. & Herold, M. (2013). Investigating assumptions of crown archetypes for modelling LiDAR returns. Remote Sensing of Environment, 134, 39-49.

[4] Barthélémy, D., & Caraglio, Y. (2007). Plant architecture: a dynamic, multilevel and comprehensive approach to plant form, structure and ontogeny. Annals of Botany, 99: 375-407.

[5] Barczi, J.-F., Rey, H., Caraglio, Y., de Reffye, P., Barthélémy, D., Dong, Q. X. & Fourcaud, T. (2008**)**. AmapSim: A structural whole-plant simulator based on botanical knowledge and designed to host external functional models. Annals of Botany, 101, 1125–1138.

[6] Griffon, S., & de Coligny, F. (2014). AMAPstudio: An editing and simulation software suite for plants architecture modelling. Ecological Modelling, 290, 3–10.

[7] Shaiek, O., Loustau, D., Trichet, P., Meredieu, C., Bachtobji, B., Garchi, S., & EL Houani, M.H. (2011). Generalized biomass equations for the main aboveground biomass components of maritime pine across contrasting environments. Annals of Forest Science, 68, 443-452.

[8] Gastellu-Etchegorry, J.-P., Yin, T., Lauret, N., Cajgfinger, T., Gregoire, T., Grau, E., Feret, J.-B, Lopes, M., Guilleux, J., Dedieu, G., Malenovský, Z., Cook, B.D., Morton, D., Rubio, J., Durrieu, S., Cazanave, G., Martin, E., & Ristorcelli, T. (2015). Discrete Anisotropic Radiative Transfer (DART 5) for modeling airborne and satellite spectroradiometer and LIDAR acquisitions of natural and urban landscapes. Remote Sensing, 7, 1667-1701.

[9] Romanczyk, P., van Aardt, J., Cawse-Nicholson, K., Kelbe, D., McGlinchy, J. & Krause, K. (2013**)**. Assessing the impact of broadleaf tree structure on airborne full-waveform small-footprint LiDAR signals through simulation. Canadian Journal of Remote Sensing, 39, S60-72.

Iterative and hierarchical forest plot reconstruction from terrestrial laser scanning point clouds for Canopy Radiative Transfer Modelling

M. Bremer [a, b*], K. Schmidtner [a,c], M. Rutzinger [a,c]

[a] Institute of Geography, University of Innsbruck, Innrain 52, 6020 Innsbruck, Austria.
[b] alpS Gmbh, Centre for Climate Change Adaptation Technologies, Grabenweg 68, 6020 Innsbruck, Austria.
[c] Institute for Interdisciplinary Mountain Research, Austrian Academy of Science, Technikerstr. 21a, 6020 Innsbruck, Austria.

Highlights: Branching architectures of forest plots are step-wisely reconstructed ensuring both geometrical and topological completeness and correctness of the included tree models by an iterative and hierarchical processing chain. This procedure is addressing the generation of mesh geometries of trees for a forest plot.

Introduction

Canopy Radiative Transfer Modelling is an important task for understanding the radiation budget of forest plots including direct, diffuse and reflected radiation in different wavelengths [1]. Using detailed 3D tree models supports the understanding of energy fluxes within a forest plot. If the single components of radiation considering vegetation architecture can be identified and quantified in spaceborne optical sensors such as Landsat 8, SPOT 5 and the upcoming Sentinel 2 satellites, mixed pixel effects influenced by different canopy elements (e.g. reflections from leafs and branches) can be understood [2].

The 3D modelling scene should be an ideally detailed representation of a real forest plot for reliable calibration and validation of the simulated radiative transfer. For special applications such as the consideration of changing foliage densities through the yearly life cycle, the foliage representation has to be dynamic over time.

Terrestrial Laser Scanning (TLS) allows a detailed capture of the forest elements as an unorganized point cloud and serves as an adequate technique for input data acquisition at plot level. For the reconstruction of tree architectures a variety of powerful reconstruction methods is available [e.g. 3,4]. However, most approaches are developed only for single tree reconstructions and have to be enhanced and adapted in order to be used operationally [5] on a forest plot. In most of these cases the data is strongly affected by inhomogeneous coverage i.e. containing data gaps. The nature of tree crowns implies a complex mixture of thick main branches and small twigs. The latter adds a significant amount of occlusions even under leaf-off conditions and introduces fine structures confusing the topological quality of the reconstruction results.

In order to achieve both topological and geometrical completeness and correctness of tree architecture models in canopies we enhance classical reconstruction strategies by an iterative and hierarchical workflow. The automatic processing pipeline working on a co-registered and georeferenced TLS point cloud of a plot, consists of classical terrain filtering, extraction of main branches by reflectance and Principle Component Analysis (PCA). These steps are followed up by a step-wise combined Dijkstra region-growing and skeletonisation procedure. By combining the derived tree skeletons with the input point cloud, branch thickness is measured and applied to a pipe-model for each tree. In a last step foliage is modelled by the approach presented by cote et al. 2009 [6].

Methods and Materials

Data sets: Two forest plots of a larch stand (*Larix decidua*) in the Austrian Alps (Stubai, Tyrol) are scanned with a Riegl VZ6000 TLS. The point cloud data consisting of three scans per plot were co-registered into merged point clouds of 42 million and 56 million points.

Stem and branch extraction: For these plots, non-ground points were extracted by a simple ground filtering approach. For a set of trees, points corresponding to the main branches are extracted using the range-corrected reflection intensity values of the laser points. This allows separating small targets such as young twigs from larger targets such as tree stems and the main branches. For automatically finding seed positions of single trees in the filtered dataset, we conducted PCA as suggested by Bremer et al. [7]. Using a 1 m radius for each data point and searching within a 10 cm voxel thinned neighbourhood point cloud, vertical elongated structures are detected as tree stems.

Basic Dijkstra Growing: Each lowest point of the extracted tree stem features is used as a seed point for an iterative Dijkstra Region Growing (DRG) [4] which is the basic strategy to find branch connections presented in this paper. The DRG is carried out on an imaginary point cloud graph that is defined by a certain threshold edge length. In this graph, all *k* neighbours within a threshold distance are connected by an edge. Performing cost path analysis on this graph associates all edge connected points to the seed position with lowest path cost.

If a point shows several paths to multiple seeds, it will be processed multiple times, which leads to a competition of the given graph-path-distances between the seeds [7]. This induces an allocation of all graph-connected points to one of the seeds of the given forest plot.

Iterative Approach: Addressing the issues of input data quality and the nature of canopies, we modified and enhanced the DRG as follows. We use an iterative growing-and-bridging. After one completed DRG cycle on a small edge length graph, the result is locked. In a follow-up DRG cycle the intermediate segmented points are used as a set of seeds again. The point attributes of these seeds keep locked, while edges with an elongated length between segmented points and un-segmented points are considered. This enables bridging gaps in the point cloud and considering additional points for a subsequent growing cycle on the small edge length graph. Keeping the edge length of the main graph small ensures a minimum loss of detail, while the bridging allows a complete processing of all geometries available.

Hierarchical Approach: Further, we introduced a hierarchical constraint for the DRG. In the first set of iterations, only stem points are allowed to be integrated to the graph connections. This is controlled using the difference between the Euclidian distance and the path distance between a data point and its associated tree seed. Allowing only small differences restricts the growing procedure to the monopodium of the tree trunk. In a second set of iterations only data points with high reflection intensities are used, ensuring that only the main branches are used and noise is removed. In the last set of iterations all unconnected data points are added.

Skeletonisation: The topological and geometrical information of the Dijkstra-results is used to convert the stem objects and their superimposed structures into 3D line-drawings and polygonal pipe models. Using the derived graph path distance attribute, the approach splits each Dijkstra-segment into distance bins. Within each distance bin a connected component analysis based on an expected branching distance is performed. For each connected component, skeleton nodes are derived as their centroids. The node 0 is connected to each of its child nodes by evaluating the path information of the DRG. Recursively, these child-nodes become nodes themselves, being connected to their child nodes again. This is done until all nodes are connected and the skeleton graph of the object is completed.

Pipe modelling of branches: To give the reconstructed tree models a stem and branch thickness, a controlled branch diameter derivation from the original point cloud is applied. For each node the associated connected component of points is projected onto a plane, which is perpendicular to the orientation of the edge connecting to its parent node. A minimum circumscribed circle is fitted to the projected point distribution deriving a pipe radius for each node. Evaluating the edge orientations of the skeleton graph and the radii associated to the nodes, orientated polygons, approximating the circular branch cross-sections, are centred at each node. The corresponding vertices of the polygons centred at associated parent and child nodes are connected, leading to the construction of polygonal prism facets for each edge of a skeleton.

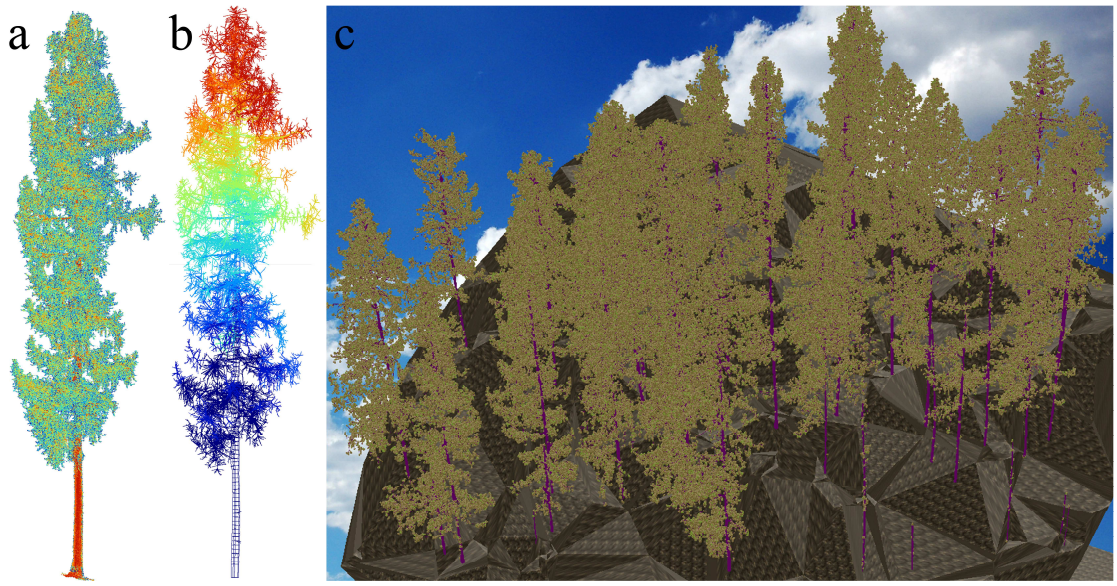


Figure 1a) point cloud of a leaf-off larch tree with small twigs in the outer crown; b) reconstructed vector model of the larch tree; c) a forest plot for 36 single trees visualized in the DART Software Package [1].

Results and Discussion

The quality of the reconstructed geometries is theoretically assessed by testing the workflow on simulated TLS point clouds for digital reference forest models showing varying crown coverage values. The results of the reconstruction are single tree architectures of high geometrical correctness and object completeness for a given forest plot (Figure 1). Looking at the basic methodological set up of the DRG, large edge thresholds lead to complete seed allocation results showing a similarity to Voronoi-regions.

The path reconstructions within these regions show an eccentric or random structure not representing real branch connections within a forest. If the point cloud has a sufficient point density for the description of the present geometry, a DRG process with small edge length thresholds becomes more defined. However, this leads to a variety of missed features in the canopies. As a lot of features especially in the upper crown part are separated from the main structure of a tree by a number of data gaps, using a small edge length graph for the Dijkstra growing will lead to geometrical correct but incomplete path estimations. Introducing the iterative DRG behaviour helps to gain both high completeness and high correctness of the results. Especially for tree species with dominant tree trunks, the hierarchical strategy of the procedure adds enhanced improvements for the reconstructions. In contrast to non-iterative approaches, errors in forest plots with high crown coverage values are subdued. This applies in particular to the main branches of trees. However, for the finer twigs, the quality of the results is linearly decreasing with increasing crown coverage.

Using the described processing pipe-line, the created models carry reliable hierarchical attributes that are used in a procedural modelling environment. Hierarchy information is used for the modelling of leafs and needles depending on twig sizes, hierarchical level, position and light availability within the tree crown as presented by cote et al. [6]. If a certain position within the canopy fulfils the requirements for needle growth, a needle clump template, consisting of single triangles per needle, are inserted. The state of needle growth can be addressed using different template types representing young and mature needles. This fulfils the needs of geometrical correctness and completeness of the branching architectures and dynamic foliage density representations for Canopy Radiative Transfer Modelling. The architectural tree models and leaf and needle geometries are linked to specific reflectance values in the satellite related spectral domains and are directly integrated into the Discrete Anisotopic Radiative Transfer (DART) Software [1] (Fig.1c).

Conclusion

The presented processing pipe line is a fully automatic procedure for the reconstruction of forest geometries. It includes ground filtering, stem detection, single tree identification and single tree model reconstruction. Introducing an iterative and hierarchical behaviour into the DRG procedure allows a high quality reconstruction.

Allowing the processing of an entire forest plot in a comprehensive procedure allows convenient scene compilation for Canopy Radiative Transfer Modelling. As a fully automated method, processing a set of trees in parallel, in future it might be used e.g. for point clouds of other sensors such as mobile laser scanning. This would allow the reconstruction of larger forest plots, fulfilling important needs of Canopy Radiative Transfer Modelling for satellite imagery.

Acknowledgements

This work has been conducted within the project Phenosat-alpha funded by the Austrian Space Agency Program (ASAP) by the Austrian Research Promotion Agency (FFG).

References

[1] Gastellu-Etchegorry, J.-P.; Yin, T.; Lauret, N.; Cajgfinger, T.; Gregoire, T.; Grau, E.; Feret, J.-B.; Lopes, M.; Guilleux, J.; Dedieu, G.; Malenovský, Z.; Cook, B.D.; Morton, D.; Rubio, J.; Durrieu, S.; Cazanave, G.; Martin, E.; Ristorcelli, T., 2015: Discrete Anisotropic Radiative Transfer (DART 5) for Modeling Airborne and Satellite Spectroradiometer and LIDAR Acquisitions of Natural and Urban Landscapes. *Remote Sensing*, 7, 1667-1701.

[2] Koetz, B., Schaepman, M., Morsdorf, F. Bowyer, P., Itten, K. Allgöwer, B. 2004: Radiative transfer modeling within a heterogeneous canopy for estimation of forest fire fuel properties. *Remote Sensing of Environment*, 92, 3, 332 – 344.

[3] Bucksch, A., Lindenbergh, R., 2008. Campino – a skeletonization method for point cloud processing. *ISPRS Journal of Photogrammetry and Remote Sensing*, 63, 1, 115–127.

[4] Livny, Y., Feilong, Y., Olson, M., Chen, B., Zhang, H., El-Sana, J., 2010. Automatic reconstruction of tree skeletal structures from point clouds. *ACM Transactions on Graphic*s, 29, 6, 1511–1518.

[5] Côté, J.F., Fournier, R. A., Frazer, G.W., Niemann, K.O. 2012: A fine-scale architectural model of trees to enhance LiDAR-derived measurements of forest canopy structure. *Agricultural and Forest Meteorology*, 166-167, 72-85.

[6] Cote, J.-F., Widlowski, J.-L., Fournier, R. A., Verstraete, M. M., 2009. The structural and radiative consistency of threedimensional tree reconstructions from terrestrial lidar. *Remote Sensing of the Environment*, 113, 5, 1067–1081.

[7] Bremer, M., Wichmann, V., Rutzinger, M., 2013. Eigenvalue and graph-based object extraction from mobile laser scanning point clouds. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, II-5/W2, 55-60.

Using laser scanning and ray tracing for scaling photosynthesis from leaf to stand level

Martin van Leeuwen[1], Jan van Aardt[1], Nicholas Coops[2], Dave Kelbe[1], Keith Krause[3]

[1] *Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, 54 Lomb Memorial drive, Rochester, 14623, NY, USA*

[2] *Forest Resources Management, University of British Columbia, 2424 Main Mall, V6T 1Z4, BC, Canada*
[3] *National Ecological Observatory Network (NEON), 1685 38th Street, 80301, Boulder, Colorado, USA*

Highlights: We demonstrate ray tracing in virtual coniferous forests and simulate half-hourly time-series of absorbed photosynthetically active radiation (APAR) at the shoot level. We then demonstrate how fine spatio-temporal variation in APAR can be used for scaling leaf-level photosynthesis to the stand and explain 75% variation in eddy-covariance-derived half-hourly GPP.

Key words: *simulation, virtual scene, ray tracing, structure, function, forest*

Introduction

The ability to synoptically describe biophysical characteristics of the Earth's surface across a range of scales is paramount to the application of remote sensing. In forestry, remote sensing has been widely used to capture stand structural characteristics and variables related to the health, vigor, or developmental stage of vegetation. These data typically have been acquired at the stand level and are used by governments and land managers to monitor land-use change and forest growth. Such approaches have served the need to oversee broader landscape management, but a gap remains in the use of remote sensing to explore relationships between fine canopy structure and physiological functions, such as the acclimation of foliage to microclimates, or light competition between individual trees. The application of remote sensing to address research questions at these finer spatial scales is valuable, however, and can be used to calibrate and validate vegetation models that are effective at broader scales.

In this research, we demonstrate techniques for creating virtual stands from combined laser scanning data and tree regeneration models. We simulate light transfer in coniferous canopies, where individual needles and shoots are modeled using a technique known as instantiation. The propagation of light is modeled using stochastic ray tracing in forward mode. Ray tracing provides for radiometric quantities to be estimated with high accuracy and at arbitrary scales or geometric fidelity. Ray tracing has been used in studies to validate canopy-level reflectance and transmission models [1] and to provide antecedent science data for emerging imaging systems [2]. At the finest levels, however, the method becomes computationally intensive, even more so when lighting conditions are dynamic, and this has restricted its use for operational scenarios. Methods have been proposed to derive scattering-related information from ray tracing for use in coarser-level modeling applications [3, 4].

Studies thus far have been restricted to modeling radiometric properties and little effort has been spent on the coupling of ray tracing techniques with plant physiological models, i.e., for the study of structure vs. function relationships. Several precursory studies exist in the literature that have applied ray tracing for these purposes. For example, [5] studied exposure of fruit trees in orchards to hemispherically-downwelling radiation, while [6] investigated light acclimation in peach trees. However, to our knowledge, no attempts in forestry have applied ray tracing for the coupling of photosynthesis-related information at leaf and canopy levels.

Here we present a method to simulate dynamic lighting conditions using forward ray tracing (see Figure 1 below for model detail). The method predicts leaf-level probabilities of light absorption from over a thousand hemispherical directions by individual canopy elements. Once a look-up table of absorption probabilities is computed, dynamic illumination conditions can be simulated in a computationally efficient manner. We subsequently show how fine spatio-temporal variations in photosynthetically active radiation (PAR) can be used for modelling stand-level gross primary productivity (GPP) and for investigating relationships between structure and function.

Methods

A 30 x 30 m plot was selected within a mixed coastal Douglas-fir forest located on Vancouver Island, Canada, near the city of Campbell River. Ground-based laser scans were acquired using the EVI instrument (CSIRO, Australia) and half-hourly GPP along with meteorological records were obtained from an eddy-covariance (EC) tower.

Single EVI scans were used to detect tree stems that were co-registered to a plot-centered coordinate frame. Tree crowns were modelled using tree regeneration software and whorl sizes were manipulated based on a

Voronoi tessellation, using stems as seeds. Foliage was modelled as planar polygons to represent clumping into whorls, and as shoot meshes comprising a woody stem and individual needles. The trees were placed in a scene surrounded by periodic boundaries and absorption probabilities were computed for each canopy element to describe the portion of radiant flux from incident angle Ω absorbed by element *p*. Whorl or shoot level APAR was computed using these probabilities and measured diffuse and direct downwelling radiation.

A leaf-level photosynthesis model was created based on a hyperbolic response curve and down-regulation of photosynthetic capacity and initial quantum yield were modelled as functions of APAR, temperature and humidity. The parameters of this model were optimized as to minimize the difference between EC-derived half-hourly GPP and simulated photosynthetic assimilation integrated over all canopy elements. Finally, the fraction of explained variation in EC-derived GPP was reported for different optimization range restrictions.

Results and Discussion

The use of whorl and shoot level absorption probabilities proved successful in deriving APAR at the respective canopy element level and guaranteed that energy was conserved. The method provides for the simulation of half-hourly APAR at the canopy element level for over 2,000 observations in time, covering the growing season from May 1 to September 17, 2009. Limitations of the method were observed, mostly among the lower canopy strata, and relate to the low sampling density of photons in the Monte Carlo ray tracing simulations, and indicate a need for additional acceleration techniques to improve the proposed method. EC-derived GPP was explained with R^2 varying from 0.66 for a case without down-regulation of photosynthesis, to $R^2 = 0.75$ for a case with modifier functions based on relative humidity, temperature and APAR. Considerable variation in GPP remained and can be attributed, for example, to EC-related errors, inaccuracies in the virtual scene representation and irradiation maps, as well as negligence of the broader set of photosynthetic drivers including soil moisture and nutrient conditions. Estimates of quantum yield and photosynthetic capacity were within previously reported ranges, demonstrating the ability of the proposed method to converge towards realistic optima. Biases in GPP estimates remain, notably around Noon, and this is consistent with earlier findings and may relate to extreme stomatal closure as well as underestimation of daytime soil respiration [7].

Despite the high level of complexity of the presented approach, a number of limitations remain, including the fidelity and realism of the three-dimensional forest scene, and variations in relative humidity and temperature within the canopy space. For example, shading within shoots establishes significant effects on quantum yield and photosynthetic capacity [7] and these effects are neglected when APAR estimates are averaged at the shoot or whorl level. Moreover, artificial canopies were constructed using tree regeneration software, however, higher resolution point clouds essentially provide for a more faithful representation of actual canopy structure. Along with ground-based plant physiological instrumentation, simulations of this nature allow for a tighter coupling of structural and functional aspects in primary productivity models and could provide novel analytic insight derived from first principles.

Conclusion

In this study, a method was presented that provides for rapid updates of the canopy radiation regime, after an initial, computationally intense model initialisation phase. The model relies on the population of a look-up table that stores probabilities of absorbing PAR by individual scene elements and handles the dynamic changes in the canopy PAR that result from solar tracking and changes in atmospheric conditions. Improvements to the model representation, canopy structure, and the inclusion of the transfer of longwave radiation, sensible and latent heat, and *in situ* sensor network data bear great potential to refine model estimates and may lead to a tighter coupling between eddy-covariance stand-level estimates of gross primary productivity and leaf-level observations of photosynthesis.

References

[1] Widlowski, J.-L., Pinty, B., Lopatka, M., Atzberger, C., Buzica, D., Chelle, M., Disney, M., Gastellu-Etchegorry, J.-P., Gerboles, M., Gobron, N., Grau, E., Huang, H., Kallel, A., Kobayashi, H., Lewis, P.E., Qin, W., Schlerf, M., Stuckens, J., & Xie, D. (2013). The fourth radiation transfer model intercomparison (RAMI-IV): Proficiency testing of canopy reflectance models with ISO-13528. Journal of Geophysical Research: Atmospheres, 118, 6869-6890.

[2] Schott, J., Gerace, A., Brown, S., Gartley, M., Montanaro, M., & Reuter, D.C. (2012). Simulation of image performance characteristics of the Landsat data continuity missions (LDCM) thermal infrared sensor (TIRS). Remote Sensing, 4, 2477-2491.

[3] Disney, M.I., Lewis, P., & North, P.R.J. (2000). Monte Carlo ray tracing in optical canopy reflectance modelling. Remote Sensing Reviews, 18, 163–196.

[4] Smolander, S., & Stenberg, P. (2003). A method to account for shoot scale clumping in coniferous canopy reflectance models. Remote Sensing of Environment, 88, 363-373.

[5] Sinoquet, H., Stephan, J., Sonohat, G., Lauri, P.E., & Monney, Ph. (2007). Simple equations to estimate light interception by isolated trees from canopy structure features: assessment with three-dimensional digitized apple trees. New Phytologist, 175, 94-106.
[6] Le Roux, X., Walcroft, A.S., Daudet, F.A., Sinoquet, H., Chaves, M.M., Rodrigues, A., & Osoria, L. (2001). Photosynthetic light acclimation in peach leaves: importance of changes in mass:area ratio, nitrogen concentration, and leaf nitrogen partitioning. Tree Physiology, 21, 377-386.
[7] Ibrom, A., Jarvis, P.G., Clement, R., Morgenstern, K., Oltchev, A., Medlyn, B.E., Wang, Y.P., Wingate, L., Moncrieff, J.B., & Gravenhorst, G. (2006). A comparative analysis of simulated and observed photosynthetic CO_2 uptake in two coniferous forest canopies. Tree Physiology, 26, 845–864.

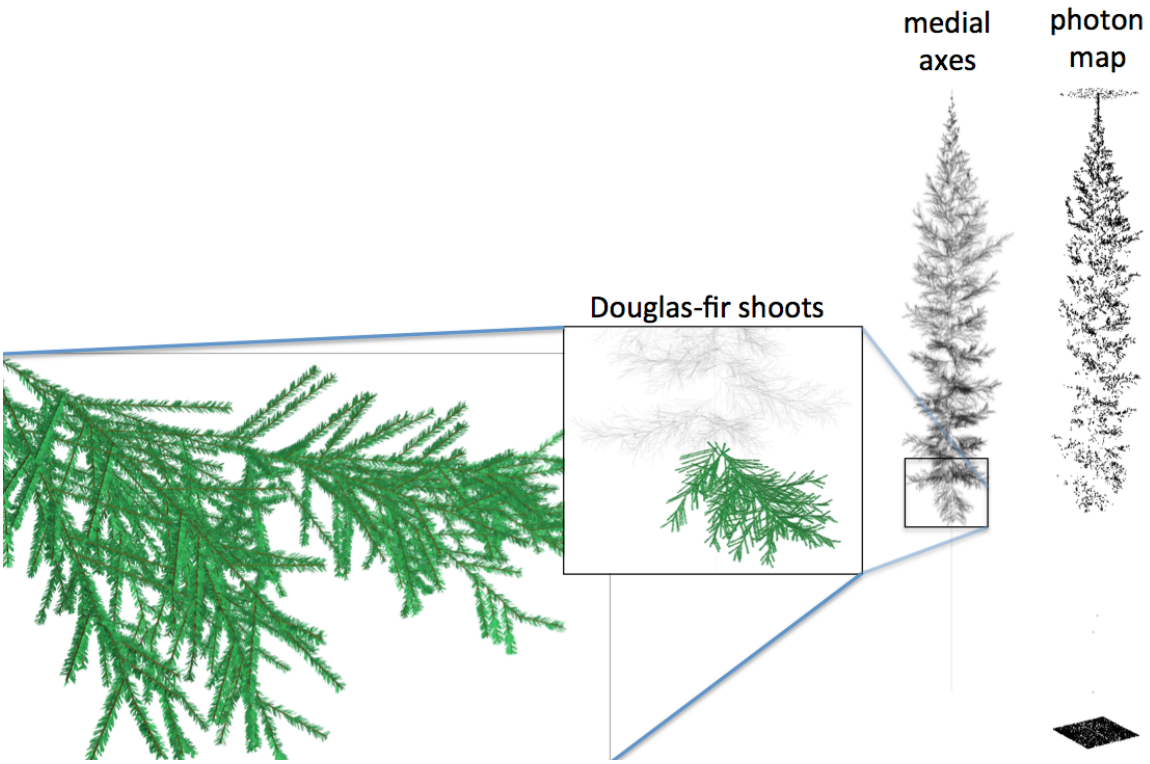


Figure 1. Instantiated conifer shoots and partial results from the ray tracing solution that computes absorption probabilities from directional incident radiation.

Terrestrial LiDAR and 3D tree Quantitative Structure Model for quantification of aboveground biomass loss from selective logging in a tropical rainforest of Peru

Jose Gonzalez de Tanago[1,2,*], Harm Bartholomeus[1], Shijo Joseph[2], Martin Herold[1], Valerio Avitabile[1], Rosa Goodman [2,3], Pasi Raumonen[4], Andrew Burt[5]

Author's affiliations:
(1) Laboratory of Geoinformation and Remote Sensing. Wageningen University (The Netherlands).
(2) Center for International Forestry Research, CIFOR.
(3) Yale School of Forestry and Environmental Studies, New Haven, CT (USA).
(4) Department of Mathematics; Tampere University of Technology (Finland).
(5) Department of Geography; University College London (UK).
** E-mail correspondent author: jose.tanago@wur.nl*

Highlights: We propose a method for quantifying the changes in aboveground biomass (AGB) from selective logging in tropical forests, to estimate carbon emissions in the context of REDD+. The method estimates tree wood volume by reconstructing the 3D tree architecture using terrestrial LiDAR scanner (TLS) data and 3D Quantitative Structure Modelling.

Key words: Terrestrial LiDAR, 3D tree structure modeling, AGB, Tropical Forest, REDD+.

Introduction

Terrestrial LiDAR Scanner (TLS) data has been demonstrated to be a good information source for forest aboveground biomass (AGB) estimation in a non-destructive approach. Tree structural parameters can be measured in a 3D scene, and these measurements can be related to AGB by applying allometric equations [1]. A more direct approach [2], which does not rely on allometric equations, consists of measuring tree wood volume by modelling the 3D tree architecture reconstructed with TLS data [3] and converted to biomass using specific wood density. This novel approach was first applied at the single-tree level in an open Eucalyptus forest in Australia [2] and in a complex tropical rainforest in Gabon [4]. To the best of our knowledge, there are no previous studies applying TLS data and 3D tree architecture modelling to assess AGB losses from selective logging in tropical forest.

The aim of this research is to analyse the performance of a newly proposed method for quantification of AGB in a complex tropical forest environment [2, 3]. The method is then proposed for characterizing the impacts of selective logging for timber exploitation to provide carbon emissions estimates for REDD+. The proposed method provides quantification of the changes in the forest AGB, distinguishing between biomass contained in the trees harvested (removed as timber and left as crop tree residuals) and the AGB lost as collateral damage. We compared the performance of this novel method with the conventional method of applying allometric equations.

Study Area

This study was conducted in the south-western Amazon, in the province of Madre De Dios, Peru, close to the borders with Bolivia and Brazil. This region had relatively little disturbance until the recent creation of the Inter-oceanic highway, connecting the region with the Brazilian and Peruvian road networks.

We worked in Brazil nut (*Bertholletia excelsa*) forest concessions that were also use for harvesting timber. Stem density ranged from 473 to 673 individuals/ha (trees and palms with DBH $\geq$ 10 cm) with a mean of 565 ± 22 stems/ha. Basal area ranged from 16.1 to 32.4 m^2/ha with a mean of 27.2 ± 1.77 m^2/ha. The vertical structure comprised several strata : (i) crowns of emergent trees (35-50 m), (ii) main forest canopy (15-35 m), (iii) understory stratum formed by palms and supressed and juvenile trees (2-15 m), and (iv) lower stratum of shrubs, saplings, and herbaceous plants (0-2 m). Together, the upper two strata covered almost 100% of the surface.

Materials and Methods

Data

Forest inventory data was collected in nine 30×50 m plots, installed around each tree to be harvested. All live trees, palms, lianas, and standing snags with diameter (at 1.3 m or above buttresses; DBH) ≥ 10 cm were inventoried (DBH and species identification) before the harvest. Total height, height of the first branch (stem height), and crown width in two directions (N/S and E/W) were measured on each crop tree. After felling, another inventory was conducted to assess the collateral damage.

Nine trees were harvested. DBH of the cropped trees ranged from 62.7 to 127.6 cm (mean ± standard deviation = 90.0 ± 22.2 cm); total height ranged from 32.0 to 50.5 m (41.3 ± 6.1 m) and estimated AGB ranged from 3.8 to 31.0 Mg (13.2 ± 8.4 Mg). All the harvested trees, except one, were emergent from the canopy.

For validation of individual tree AGB estimates of harvested trees, a reference dataset was produced by first estimating the tree volume from detailed measurements of the stem and all branches (until 10 cm diameter), and converted to biomass using its specific wood basic density.

TLS data was acquired in the nine plots before and after harvest using a multiple discrete returns Riegl VZ 400 scanner (RIEGL Laser Measurement Systems GmbH, Horn, Austria). Eight scans were performed in each plot using an angular nominal resolution of 0.06 degrees. The scan locations were established following a systematic spatial pattern with six locations equidistantly distributed around the edges of the plot and two additional scans locations distributed along the central axis of the plot (aligned with the predicted felling direction).

Methods

Individual TLS scans were co-registered using tie points detection and co-registration with Riscan-Pro software. The average standard deviation allowed for the individual points co-registered was below 5 mm. The Quantitative Structure Model (QSM) 3D tree modelling of Raumonen [3] was used for reconstruction of tree architecture based on TLS data. This model segments the point cloud, models the entire crown branching architecture using a stepwise approach to reconstruct topologically consistent the tree skeleton, and finally performs geometrical model fitting based on cylinders. The model provides automatic calculation of volume of woody structures, which is converted to AGB by multiplying by the specific wood density. Total AGB change per plot was estimated as the sum of AGB of the harvested tree and collateral damage. For calculation of the latter, the 3D QSM was performed for each tree TLS point cloud acquired before and after each tree felling.

The AGB estimations based on forest-inventory approach were computed using the most local specific allometric models by Goodman et al [5] and the most recent pan-tropical models Chave et al [6]. The regional [5], pan-tropical [6] models and the AGB estimations derived from the 3D tree architecture reconstructed with TLS data were compared with the validation AGB reference data (derived from the wood volume measurements of harvested trees). RMSE, bias, and relative error were computed to evaluate each model.

Results and Discussion

The AGB estimates by the optimized TLS-QSM model outperformed the estimates provided by the allometric models tested (Table 1). The gain in the accuracy achieved by the TLS-QSM method in relation to the newest and most enhanced pan-tropical allometric model developed by Chave et al [6] was a 9.8% decrease in CV RMSE. The latter allometric model is most likely to be used in a generic study estimating AGB in tropical rainforest. This represents a similar gain in accuracy as the one reported by Calders et al [2]. Moreover, the bias of the TLS-QSM method was lower than the allometric methods, minimizing the underestimation of AGB of large trees.

Table 1: Summary of crop tree AGB estimate performances by TLS-QSM and allometric models using combinations of DBH (D), wood density (ρ), total height (H), and average crown radius (CR).

Model	RMSE (Mg)	CV RMSE (%)	Bias (Mean error) (Mg)
TLS-QSM	2.46	18.6	-0.37
Goodman et al 2014 (D,ρ,H,CR)	3.70	27.9	-1.31
Chave et al 2014 (D,ρ,H)	3.76	28.4	-2.56

Although the TLS-QSM method performed well compared to allometric equations, it has lower accuracy in a tropical forest (this study) than previously reported for Eucalyptus forests [2]. This is likely due to the lower 3D sampling of top of crown branches because of the high level of occlusion typical of emergent trees in such dense and complex forest. For example, larger errors are obtained when reconstructing crowns than trunks (see Figure 1). The trees in this region have a relatively high proportion of AGB in their branches, especially the very large trees (DBH > 100 cm) [5]. This makes it more difficult to achieve accurate AGB estimates from TLS-QSM tree architecture reconstruction. Nevertheless, the performance of the best TLS-QSM modelling approach outperformed the best allometric equation tested for individual trees, compared to our validation estimates.

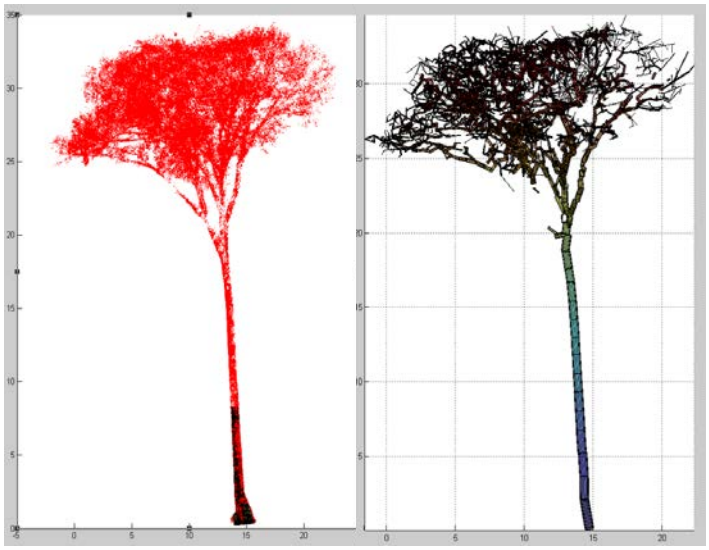


Figure 1: Example of one tree TLS point cloud (left) and the QSM 3D tree modelled output (right).

Conclusions

We found that the proposed method (TLS-QSM) for 3D tree structure reconstruction can be used to estimate AGB of large trees in tropical forests. The TLS-QSM method outperformed the allometric models tested. Moreover, this level of accuracy of the AGB estimation was achieved in very unfavourable conditions for the proposed approach, reconstructing very large and complex tree crowns, within a great vegetation density, and above the main forest canopy.

References

[1] Yao, T., Yang, X., Zhao, F., Wang, Z., Zhang, Q., Jupp, D., Lovell, J., Culvenor, D., Newnham, G., Ni-Meister, W., Schaaf, C., Woodcock, C., Wang, J., Li, X., & Strahler, A.. (2011). Measuring forest structure and biomass in New England forest stands using Echidna ground-based lidar. Remote Sensing of Environment 115: 295-297.

[2] Calders, K., Newnham, G., Burt, A., Murphy, S., Raumonen, P., Herold, M., Culvenor, D., Avitabile, V., Disney, M., Armston, J., & Kaasalainen, M. (2015). Non-destructive estimates of above-ground biomass using terrestrial laser scanning. Methods in Ecology and Evolution 6-2: 198–208.

[3] Raumonen, P., Kaasalainen, M., Åkerblom, M., Kaasalainen, S., Kaartinen, H., Vastaranta, M., Holopainen, M., Disney, M., & Lewis, P. (2013). Fast Automatic Precision Tree Models from Terrestrial Laser Scanner Data. Remote Sensing. 5: 491-520.

[4] Disney, M., Burt, A., Calders, K., Raumonen, P., Gonzalez de Tanago, J., Cuní Sanchez, A., Avitabile, V., Herold, M., Armston, J., Lewis, S. L., Lines, E., & Lewis, P. (2014). New applications of 3D measurement and modelling for quantifying forest structure and biomass. GV2M: Global Vegetation Monitoring and Modeling. INRA, Avignon, France.

[5] Goodman, R. C., Phillips, O. L., & Baker, T. R. (2014). The importance of crown dimensions to improve tropical tree biomass estimates. Ecological Applications, 24: 680-698.

[6] Chave, J., Réjou-Méchain, M., Búrquez, A., Chidumayo, E., Colgan, M.S., Delitti, W.B.C., Duque, A., Eid, T., Fearnside, P.M., Goodman, R.C., Henry, M., Martínez-Yrízar, A., Mugasha, W.A., Muller-Landau, H.C., Mencuccini, M., Nelson, B.W., Ngomanda, A., Nogueira, E.M., Ortiz-Malavassi, E., Pélissier, R., Ploton, P., Ryan, C.M., Saldarriaga, J.G., & Vieilledent, G. (2014). Improved allometric models to estimate the aboveground biomass of tropical trees. Global Change Biology, 20, 3177-3190.

Traceability of essential climate variables through forest stand reconstruction with terrestrial laser scanning

Kim Calders[1,2], Mathias Disney[2], Joanne Nightingale[1], Niall Origo[1], Alexandra Barker[1], Pasi Raumonen[3], Philip Lewis[2], Andrew Burt[2], James Brennan[2], Nigel Fox[1]

[1]*National Physical Laboratory, UK; kim.calders@npl.co.uk*
[2]*University College London, UK*
[3]*Tampere University of Technology, Finland*

Highlights: We develop a processing chain that uses terrestrial LiDAR scans as input data to assess the SI-traceability of various in-situ LAI and fAPAR products via radiative transfer modelling. LiDAR data and tree reconstruction is used to represent the explicit 3D forest structure in radiative transfer models.

Key words: *Terrestrial LiDAR; traceability; LAI; fAPAR; forest stand reconstruction*

Introduction

Climate warming will cause more variable weather, an increased amount of major storms and more extreme weather events. Furthermore, ecological responses to climate change are already visible. The date of leaf emergence in spring is a UK climate indicator and spring activities (e.g. shooting and flowering of vegetation) have occurred progressively earlier since the 1960s [1]. Forest ecosystems cover approximately 4 billion hectares or about 31% of the world's land surface and remote sensing from space is essential to collect global data to underpin climate change research efficiently. Biophysical essential climate variables (ECVs), such as leaf area index (LAI), fraction of absorbed photosynthetically active radiation (fAPAR) or albedo need to be monitored to detect small fluctuations over time.

The validation of spaceborne ECV products is generally based on a direct comparison either with products from different earth observation missions or with in-situ estimates. End-to-end traceability of in-situ measurements and satellite-derived ECVs are difficult. Most in-situ methods do not measure the biophysical quantity directly, but estimate this from indirect measurements using a variety of hypotheses and assumptions. One way to deliver SI-traceable reference measurements for in-situ fAPAR and LAI products is to use radiative transfer modelling and simulations. This allows us to control all aspects of the environment and sensor properties, which would not be possible using measured data. Subsequently, biases resulting from differing assumptions can be quantified.

Radiative transfer models require the forest structure to be quantified. This can either be a simple 2D representation of the forest structure or a full 3D explicit description of the forest stand. Manually recording detailed structural information is laborious and time consuming. A more practical and common method is to use mathematical models to generate virtual trees. This technique was for example used by [2] and based on measurements of relatively few plants, growth rules are defined and conditions are adjusted to match the local environment. Recently, the introduction of terrestrial laser scanning (TLS), also referred to as terrestrial LiDAR, has enabled the possibility of generating plant models based on 3D scans [3-5].

This work will develop a processing chain to use terrestrial LiDAR scans as input data to quantify the SI-traceability of various in-situ LAI and fAPAR sensors via radiative transfer modelling. In this paper we will present the first results from explicitly characterising and reconstructing a deciduous forest stand.

Materials and method

Study area and data collection

A six-hectare study site (200 m x 300 m) is established in Wytham Woods, UK. The forest is dominated by *Acer pseudoplatanus, Fraxinus excelsior* and *Corylus avellana.* TLS and LAI data was collected in June and July 2015, with a leaf-off campaign planned for the winter of 2015. The PAR sensor network will be installed in the summer of 2015 for the duration of the project.

Structural 3D data is collected with a RIEGL VZ-400 terrestrial laser scanner, which records multiple return LiDAR data as well as additional waveform data. The wavelength of the instrument is 1550 nm and the beam divergence is nominally 0.35 mrad. The angular sampling for both zenith and azimuth angle is 0.04°, azimuth range is 0° – 360° and zenith range is 0° - 130°. The complete 6 ha is scanned from 176 scan locations, laid out in a 20 m x 20 m grid (see figure 1).

LAI data is collected through digital hemispherical photography (DHP), Li-Cor LAI-2000 and LAI-2200 and TRAC2 (Tracing Radiation and Architecture of Canopies). We employed the VALERI sampling design for the DHP, LAI-2000 and LAI-2200 measurements (1800 measurements in total) and 100 m transects for TRAC measurements similar to [6].

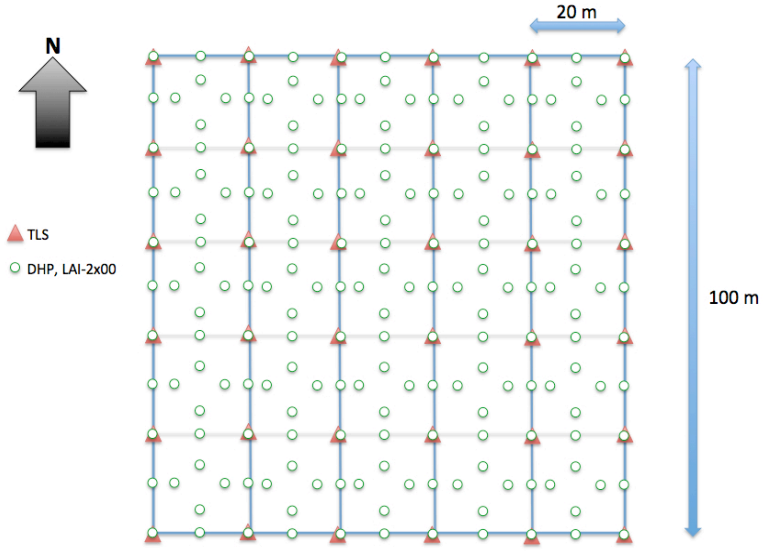


Figure 1: Sampling design for 1 ha

Forest stand reconstruction and radiative transfer modelling

We will use a Monte Carlo ray tracing (MCRT) approach to simulate DHPs, LAI-2000 and LAI-2200, TRAC and PAR measurements. In this study the *librat* MCRT model is used. This model is based on the ararat/drat MCRT model [7] and has been tested in previous studies against observations as well as against other radiative transfer models. This model requires a 3D explicit description of the forest structure, and spectral information about the canopy constituents used to represent the forest structure. We will use TLS data to reconstruct tree models using the quantitative structure model (QSM) approach in [3-4]. These previous studies focused on reconstructing branching structure only, and we will extend this work by adding leaves to the 3D models. Addition of the leaves will be based on the derived light availability, similar to [5]. The TLS point clouds will be filtered using a clustered region growing approach alongside form pruning to exploit the fuzzy appearance of leaf returns due to leaf arrangement, incidence and clumping. This will broadly allow the delineation of wood and leaf returns. We will use the QSM approach to model the tree skeleton and reconstruct the leaves based on the information extracted from the foliage point cloud.

Expected outcomes

This work will give us insight into the SI-traceability of various in-situ LAI and fAPAR products via radiative transfer modelling, stand reconstruction from TLS data and comparison of field measurements with simulations. The actual canopies that are currently being used in projects such as RAMI (radiation transfer model intercomparison) are based on detailed forest inventory measurements. To our knowledge, this study will be the first to reconstruct a large study site directly from TLS data that can be used in a radiative transfer model. Using 3D tree modelling and radiative transfer allows us to control all aspects of the canopy structure and sensor characteristics. Furthermore, the intensive sampling design (figure 1) will allow us to address issues of spatial variance and quantify the effect of different sampling strategies on the inferred ECVs.

References

[1] Walther, G.-R., Post, E., Convey, P., Menzel, A., Parmesan, C. et al. (2002). Ecological responses to recent climate change. Nature, 416, 389–395.

[2] Disney, M., Lewis, P. and Saich, P. (2006). 3D modelling of forest canopy structure for remote sensing simulations in the optical and microwave domains. Remote Sensing of Environment, 100, 114–132.

[3] Calders, K., Newnham, G., Burt, A., Murphy, S., Raumonen, P. et al. (2015). Nondestructive estimates of above-ground biomass using terrestrial laser scanning. Methods Ecol Evol, 6, 198–208.

[4] Raumonen, P., Kaasalainen, M., Åkerblom, M., Kaasalainen, S., Kaartinen, H. et al. (2013). Fast Automatic Precision Tree Models from Terrestrial Laser Scanner Data. Remote Sensing, 5, 491–520.

[5] Côté, J.-F., Fournier, R.A. and Egli, R. (2011). An architectural model of trees to estimate forest structural attributes using terrestrial LiDAR. Environmental Modelling & Software, 26, 761–777.

[6] Leblanc, S.G., Chen, J.M. and Kwong, M. (2002). Tracing Radiation and Architecture of Canopies: TRAC MANUAL Version 2.1.3. Natural Resources Canada.

[7] Lewis, P. (1999). Three-dimensional plant modelling for remote sensing simulation studies using the Botanical Plant Modelling System. Agronomie, 19, 185–210.

On the contribution of dendrometric « rules » to improve accuracy and genericity of ALS models using an area-based approach

L. Saint-André[1,2], J.P. Renaud[3], C.Véga[4], J. Bock[5], S. Durrieu[6], M. Bouvier[6]

1: INRA, Biogéochimie des Ecosystèmes Forestiers, Rue d'Amance, F-54280 Champenoux (st-andre@nancy.inra.fr)
2: CIRAD, UMR ECO&SOLS, F-34000 Montpellier
3: Office National des Forêts, Département RDI, 11 rue Ile de Corse, F-5400 Nancy.
4: IGN, Laboratoire de l'Inventaire Forestier, 11 rue Ile de Corse, F-54000 Nancy
5: Office National des Forêts, Département RDI, 42 quai Charles Roissard, F-73023 Chambéry
6: IRSTEA, Maison de la télédétection, F-34000 Montpellier

Highlights: An approach integrating forest allometric relationships into Lidar models is proposed in a consistent and comprehensive way. This approach, generic and reliable, fits a large diversity of stands and flight conditions (*e.g.* coniferous/broadleaves; mountain/lowland stands), and use only a limited number of Lidar metrics.

Key words: **Forest allometric relationships, relative density index, forest basal area**

Introduction

Lidar remote sensing is a state of the art technology for characterizing forest structure and estimating forest attributes, *e.g.* dominant height, basal area, volume and aboveground biomass. For the last decade, two main approaches have been developed to predict them from Lidar point clouds. The most prevalent one consists in developing models based on Lidar metrics computed at the plot level. This area-based approach is considered to perform well in various acquisition conditions and has already been integrated into operational forest inventory procedures [*e.g.* 3]. However, it is based on empirical models that require field calibration for each Lidar survey [3]. Furthermore, field calibration plots has to be in synchronicity with the Lidar acquisitions in order to avoid bias associated with growth and silvicultural practices (*e.g.* thinning) between field and Lidar campaigns.

The alternative approach is based on individual tree segmentation and characterisation. This strategy provides measurements of tree attributes that are similar to field measurements and could benefit from some established allometries between tree height, crown area or diameters at breast height to refine segmentation. However, such tree segmentation required high point cloud densities (> 5 points/m²) and is computationally demanding. Moreover, its implementation over large areas remains challenging due to computational cost and to the calibration needs to minimize omission and commission errors for a wide range of forest types.

Recent developments have shown that it is possible to improve the area-based approach by using Lidar metrics that match closely different aspects of stand structures. For example, Bouvier *et al.* [1] showed that the use of such metrics improves model genericity, but does not alleviate the need to calibrate models for every Lidar acquisition. In this study, we hypothesize that integrating into area-based approach, constraints arising from allometric rules, well known and widely used by foresters, could further improve model genericity. We developed a model based on height-diameter relationships, and on Reineke's rule of self-thinning [2]. We tested this approach on a large diversity of stands and Lidar acquisition conditions, in order to estimate its robustness.

Material and methods

Proposed methodology

Figure 1 presents the flowchart of the proposed methodology that follows several steps to predict plot basal area (G). First, dominant height (Ho) is estimated at the plot level from Lidar metrics. Inverting standard height – diameter curves yield plot dominant tree diameter (Do) for each plot, which is further refined by using additional Lidar metrics to get unbiased estimates of Do among all stands. Then the maximum stockability in terms of stem density (N) and basal area (Gmax) for a given Do, is assessed based on self-thinning equations based on the

Reineke's relative density index [2, 4, 5]. Next, a model of G ratio (assessing the deviation of the actual G from the maximal possible G value of the plot) is established. At this step, "structural" Lidar metrics are integrated into the model (*e.g.* canopy cover). Finally, back transforming the resulting model yield an estimate of G that only uses a few Lidar metrics.

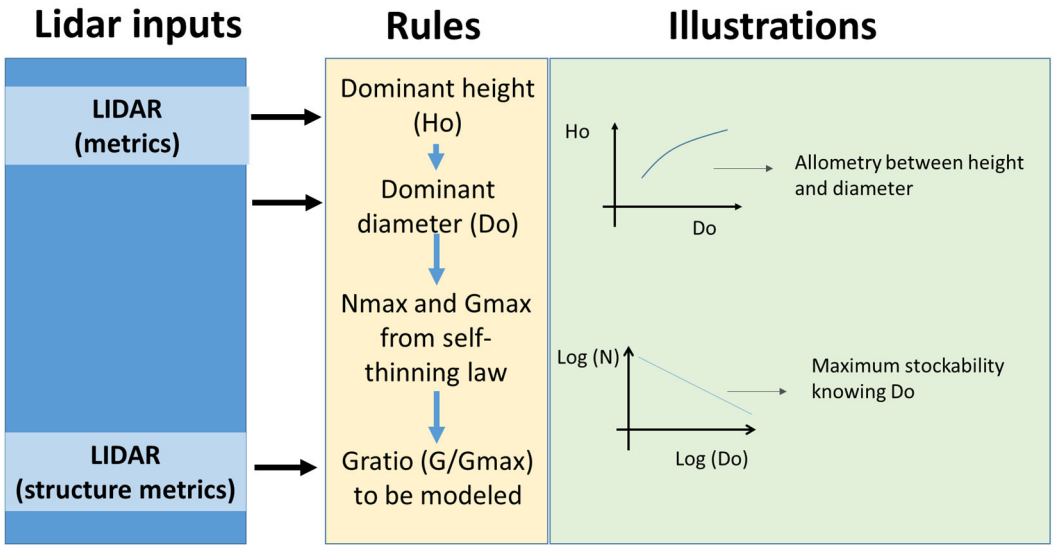


Figure 1. Flowchart of the modelling approach followed, showing the inputs points of the Lidar metrics and the allometric rules used to obtain the final basal area (G) estimations.

Lidar acquisitions over eight coniferous and broadleaves stands were used in this study, covering a large range of Lidar densities (from 2 points/m² to 30 points/m²), topographical conditions (lowland and mountain locations), and both stand structures and composition (*i.e.* regular and irregular stands; monospecific and mixed stands). One broadleaves stand was also flown at 2 vegetation periods (*i.e.* leaf-on and leaf-off). Over 500 plots were used to estimate G using the proposed approach.

Results and discussion

Over all stands examined in this study, the relative residual error (rRMSE) remained under 10% for the prediction of G. The resulting model appeared to be mostly unbiased (Figure 2), suggesting that this approach is robust to a large range of stand and both environmental and Lidar flight conditions. Examination of the residuals showed only a limited number of bias conditions associated to a few species. Globally the proposed model appeared unbiased and robust, providing accurate G estimations and showing neither seasonal, nor stand structure effects (Figure 2).

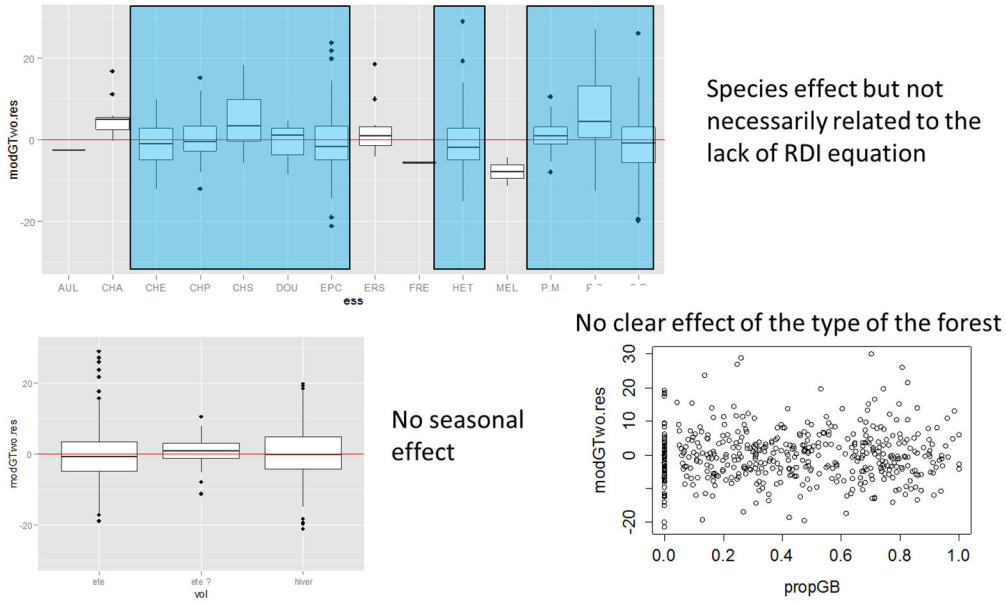


Figure 2. Model residuals as affected by tree species (top), season (bottom left), or forest types (bottom right).

Conclusion

The proposed approach integrates the knowledge developed in forest growth allometry into Lidar models in a consistent and comprehensive way. The work is still ongoing to further improved results and to provide volume, biomass, and stem density estimations along with G estimations.

References

[1] Bouvier, M., S. Durrieu, R.A. Fournier and J.-P. Renaud. 2015. Generalizing predictive models of forest inventory attributes using an area-based approach with airborne LiDAR data. Remote Sensing of Environment 156: 322-34.

[2] Charru, M., I. Seynave, F. Morneau, M. Rivoire and J.-D. Bontemps. 2012. Significant Differences and Curvilinearity in the Self-Thinning Relationships of 11 Temperate Tree Species Assessed from Forest Inventory Data. Annals of Forest Science 69 (2): 195-205.

[3] Næsset, E., T. Gobakken, J. Holmgren, H. Hyyppä, J. Hyyppä, M. Maltamo, M. Nilsson, H. Olsson, A. Persson and U. Söderman. 2004. Laser scanning of forest resources: the nordic experience. Scandinavian Journal of Forest Research 19 (6): 482-99.

[4] Pretzsch, H. 2006. Species-Specific Allometric Scaling under Self-Thinning: Evidence from Long-Term Plots in Forest Stands. Oecologia 146 (4): 572-83.

[5] Pretzsch, H. and T. Mette. 2008. Linking Stand-Level Self-Thinning Allometry to the Tree-Level Leaf Biomass Allometry. Trees 22 (5): 611-22.

Evaluation of an additional LiDAR metric in Forest Inventory

Susana, Gonzalez Aracil[1], David, L, Herries[1], Brian, A, Rawley[2]

[1]*Interpine Group Ltd. (Rotorua, New Zealand) - Susana.Gonzalez@interpine.co.nz*
[1]*Interpine Group Ltd. (Rotorua, New Zealand) – David.Herries@interpine.co.nz*
[2]*Silmetra Ltd. (Tokoroa, New Zealand) - brian.rawley@silmetra.co.nz*

Highlights: LiDAR forest inventory can be improved with an additional LiDAR metric which quantifies point cloud peaks (inferred as tree tops) in the data. This would not necessarily need to be a focused tree count style algorithm, but more of simple peak identification metric. This is because we don't want to "count trees" but just have a metric that represents potential tree tops in a more fractional distribution form than the need to extract crowns themselves. We have experimented with peaks being introduced into our predictor selection model and found this to have the potential to improve estimates of stocking.

Key words: *LiDAR, forest inventory and analysis, canopy height model, peak counts*

Introduction

Last year, 2014, Interpine and Silmetra completed the first large scale commercial forest inventory using LiDAR in New Zealand. This inventory was conducted in the North Island in a forest plantation of 200,000 ha of *Pinus radiata*. The main objective of this project was replacing the traditional stand-level yield inventory for forecasting recoverable volume by log type to a LiDAR forest inventory. This inventory was computed using an imputation model based in k nearest neighbour (kNN) statistical estimation and improved using an additional LiDAR metric, the peak count metric. The peak count is not a tree count because the peaks do not all represent trees and there may be trees for which there is no peak. Suppressed trees are obscured, trees that are close together may be seen as a single peak and there are false positives (e.g. shrubs). Jari Vauhkonen et al. [1] compared six algorithms for single tree detention and his conclusion was not all trees are usually detected, arriving to the result that the algorithms find in average 65% of the number of trees measured in the field. To be useful in a kNN model the peak count need not correspond exactly to a tree count. All that is required is that the tree count and peak count are sufficiently well correlated that the use of peak counts improves the prediction of tree counts and stand density.

Method

Developing the Plot LiDAR Imputation Technique offers a cost-effective alternative to the traditional stand-level inventory used for forecasting forest yield as recoverable volume by log type (measured as log product volume in m^3/ha). K Nearest Neighbour imputation needs two independent sources to produce a model of forecasted log type yield, in our case; 500 field measurement plots of 0.06ha each and airborne LiDAR data of 8 pls/m^2.

The methodology to generate an imputation model was:

- Field data processing: Differential correction of the field reference plots locations collected with a high grade GPS and defining or computing the forest attributes (Age, Top Mean Height, Quadratic mean DBH, Basal Area, Total Stem Volume, Stocking, Pruned length and Grade Mix).
- LiDAR data processing for the reference plots: Extracting the point cloud from the normalized LiDAR data, ground z coordinate replaced as zero and all other points have an elevation that equals their height above or below to the ground, for the circular area on the reference plots and then computing the LiDAR metrics with lascanopy module of LAStools [2].
- Compute the LiDAR metrics for the full estate in a grid cell of 25m by 25m resolution.
- Develop non-LiDAR covariates, such as recorded age and elevation, for both reference plots and grid cells.
- Selection of the LiDAR and non-LiDAR covariates to be used in the imputation model.
- Develop the imputation model.

Plot Imputation is a method where each target pixel is assigned a reference plot, based on similarity in LiDAR and non-LiDAR metrics, and uses a reference plot yield table. The nearest neighbour used in this project was k=3. The LiDAR metrics selected for the model were the 95th and 75th height percentiles, skewness, quadratic average height, canopy cover, planting year, recorded pruned status and three combinations of bincentiles (fractions of points between percentiles of height), height percentiles and intensity:

- The 90th height percentile multiplied by the difference between the 90th bincentile from all canopy returns and the 90th bincentile from first returns.

- The difference between the median intensity from first returns and the median intensity from all canopy returns.
- The difference between the 90th and 80th bincentiles from canopy first returns.

The imputation model was developed using the R package yaImpute with distances calculated using random forests.

Results

The result was good but needed further investigation around the stocking output, one option would have been adding the LiDAR derived tree location, but it's difficult to be sure what you are measuring relative to the tree top, sometimes you are actually measuring the highest point of the tree and another time the highest whorl of the branches. We could extract peak counts from the point cloud if it were not necessary to differentiate between tree tops and highest branches. Instead, we extracted the peaks from a high resolution canopy height model (CHM). Recently a pit-free CHM [3] was developed, this algorithm involves a multi-step, multi-resolution approach that fill gaps between pulses without losing the definition of the tree tops. Following this method the point density of our LiDAR data allowed a final raster of 0.3m by 0.3m resolution to be generated. The CanopyMaxima tool from FUSION [4] provides options to find the individual dominant and co-dominant trees, this algorithm finds the local maxima and is based on the original work by Sorin Popescu [5] where he explains the use of the variable window size on a canopy height model. The window width for our forest of *Pinus radiata* was determined after the measurement of a selection of crowns and heights across a full range of trees, from young to mature, and the coefficients from the following equation where used in CanopyMaxima:

$$Window\ width = 1.3 + 0.2361 * height - 0.002 * height^2$$

Comparing the data from the field tree location (black cross) and the local maxima from FUSION (pink dots) we can see (Figure 1) that the location doesn't match, neither in the number of trees present in the plot, but the average height from the field data and the average of the local maxima dots by plot is quite close.

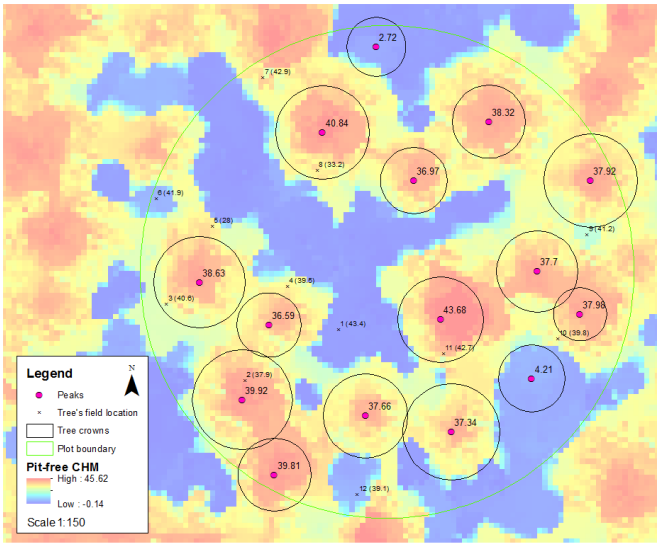


Figure 1: Crown detection of co-dominate and dominate trees on a pit-free CHM.

This additional metric is called peak counts because of uncertainty comparing to actual location and density with the field data. The peak count was computed for the full forest estate and added to the statistic model. The table below (see Table 1) shows the imputation models with and without the peak counts. The peak counts slightly improve the estimate of stocking without having a major impact on other variables. In the table below, rmsdS is scaled root mean square difference and rmsd is unscaled (i.e. in the units of the variable). The generalised root mean square difference was 0.68 for the model without the peak counts and 0.67 for the model with the peak counts.

Table 1: Imputation models with and without the metric peak counts.

Variable	**Model without the metric peak counts**		**Model with the metric peak counts**	
	rmsdS	*rmsd*	*rmsdS*	*rmsd*
Age (years)	0.29	2.5	0.30	2.6
Basal area (m^2/ha)	0.41	6.7	0.41	6.7
Stocking (trees/ha)	0.46	138	0.41	124

Top height (m)	0.16	1.8	0.17	1.8
Total Stem Volume (m^3/ha)	0.32	78.5	0.32	79.6
Pruned length (m/ha)	0.45	343	0.44	339
Quadratic mean DBH (cm)	0.27	3.5	0.27	3.5

Discussion

Previous studies like the one developed by Erik Nasset [6] in Norway proves the positives results using LiDAR data to estimate forest attributes for a large operational inventory. These attributes has been computed as well in Swedish inventories as a normal practices, Johan Holmgren [7] develops prediction models to estimate forests attributes and he points that tree detection methods are usually not efficient if there are few laser measurements on each tree count. On the other hand if the LiDAR metrics, non-LiDAR covariates and the peak counts are added into the imputation the stocking accuracy increase. However, this increment is only 0.05 in the rmsdS for the Stocking variable.

Conclusion

The overall objective for this project was replacing the traditional yield inventory for forecasting recoverable volume by log type to a LiDAR inventory process. Using peak counts as a covariate provided enough improvement in the estimate of stocking to justify further development of this technique.

Acknowledgements

The authors thank Bob McGaughey from USDA Forest Service for the provision of his experience using CanopyMaxima algorithm from FUSION and special thanks to Dr. Martin Isenburg for his assistance generating a pit-free Canopy Height Model using LAStools.

References

[1] Vauhkonen, J., Ene, L., Gupta, S., Heinzel, J., Holmgren, J., Pitkänen, J., Solberg, S., Wang, Y., Weinacker, H., Hauglin, K. M., Lien, V., Packalén, P., Gobakken, T., Koch, B., Næsset, E., Tokola, T. and Maltamo, M. (2012). Comparative testing of single-tree detection algorithms under different types of forest. Forestry. 85(1): 27-40.

[2] rapidlasso GmbH, "LAStools - efficient LiDAR processing software" (version 141218, commercial), obtained from http://rapidlasso.com/LAStools

[3] Khosravipour, A., Skidmore, A. K., Isenburg, M., Wang, T., & Hussin, Y. A. (2013). Development of an algorithm to generate a LiDAR pit-free canopy height model. SilviLaser 2013, October 9-11, 2013 –Beijing, China.

[4] McGaughey, R. J., United States Department of Agriculture, Forest Service, Pacific Northwest Research Station, "FUSION: Providing fast, efficient, and flexible access to LiDAR, IFSAR and terrain datasets" (version 3.42), obtained from http://forsys.cfr.washington.edu/fusion/fusionlatest.html

[5] 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(1-3):71-95.

[6] Nasset, E. (2004). Accuracy of forest inventory using airborne laser-scanning: evaluating the first Nordic full-scale operational project, Scandinavian Journal of Forest Research, 19:6, 554-557.

[7] Holmgren, J. (2004). Prediction of tree height, basal area and stem volume in forest stands using airborne laser scanning, Scandinavian Journal of Forest Research, 19:6, 543-553.

Estimating timber volume with airborne laser scanning using semi-artificial sample plots

Marius Hauglin[1*], Terje Gobakken[1] and Erik Næsset[1]

[1] *Department of Ecology and Natural Resource Management, Norwegian University of Life Sciences.*
[*] *Email: marius.hauglin@nmbu.no*

Highlights: The availability of single-tree field data enables new estimation methods. Data from individual trees were in this study combined to form semi-artificial sample plot data in an area-based forest inventory. Timber volume was estimated and the validation showed comparable results for semi-artificial and field measured plots.

Key words: *Airborne laser scanning, simulation, volume, area-based estimation, single-tree*

Introduction

Field measurements are an important part of an area-based estimation of standing volume with airborne laser scanning (ALS). Field measurements are in boreal forests typically obtained for positioned field plots of 200 – 400 m^2 in size. New approaches to field measurements such as the use of terrestrial laser scanning (TLS) might give accurate data – including position – for single trees, but not necessarily covering all trees at a plot. Depending on a range of factors – such as the scanning time in the case of using TLS – single-tree field data might be obtained at a lower cost than corresponding field plot data. Possible utilisation of single-tree data should therefore be investigated. Breidenbach et al. [1] created artificial sample plots from selected areas within their study area. These artificial plots were created with reference data derived through single tree segmentation of the ALS data. The inclusion of these artificial plots were in that study shown to improve the area-based predictions [1].
In the present study we tested if a combination of single-tree data used to populate semi-artificial plots could be used as part of the modelling procedure in an area-based volume estimation.

Materials and methods

Study area

The study area is located in Aurskog-Høland municipality (59°50' N 11°30' E, 120–390 m above sea level) in the south-eastern part of Norway. The forest type is boreal with Norway spruce (*Picea abies* (L.) Karst.) and Scots pine (*Pinus sylvestris* L.) as the dominant tree species.

Field data

Field data were collected in 2006 – 2008. A total of 631 trees at 40 sites in the study area were accurately positioned and the diameter, height and crown extent in the eight cardinal and inter-cardinal directions were registered. A separate set of 200 field plots of 200 m^2 was also measured, and all trees with a diameter at breast height > 4 cm were calipered. Height were measured on a sample of the trees. Single tree volume were estimated using species specific allometric models, and the single-tree volume for all trees were summed to ground reference plot volume.

The set of 200 field measured plots were randomly split into two subsets; one set of 150 plots for modelling, and 50 for validation.

ALS data

The ALS data were collected in June 2005 with an Optech ALTM 3100 sensor on a fixed-wing aircraft with an average point density on the ground of 0.7 points m^{-2}. Echoes classified as ground were used to construct a triangulated irregular network (TIN) terrain model. The height above ground was calculated for all echoes by subtracting the respective TIN heights from the ellipsoidal heights.

The sensor used in this project records up to four echoes for each laser pulse, but only the first and last returned echoes were used in the current study. For the 631 individual trees ALS data were extracted using the field measured crown extent. All echoes with coordinates within the planar crown extent of a tree were associated with that tree.

Semi-artificial plots

The modelling approach in this study followed the area-based estimation method, as described by e.g. Næsset [2]. The model data were – however – not only field-measured plots, but also semi-artificial plots, created from the field measured single-tree data. Field and ALS-data from a sub-sample of the 631 trees were combined to form each semi-artificial plot. First ALS data from plots within a predefined area known to be without trees were used as a basis. In the present study we used 150 randomly selected locations within a plain field. Each of these plots were associated with a volume of zero. We then added data from a random number of trees to each plot. We selected between one and 40 trees for each plot. The selection of trees to populate a plot was carried out by sampling with replacement from the initial set of 631 trees. Each added trees' field-estimated volume was added to the plots' total volume, and the ALS data for each tree added to the ALS data at the plot. In order to maintain a realistic number of echoes at each semi-artificial plot, the same number of echoes as was introduced by the added tree were removed from the existing plot. Echoes with the lowermost above-ground heights were removed before adding each tree. If for example 30 echoes were associated with the tree being added, the 30 lowermost echoes of the existing semi-artificial plot were removed before the data from the tree were added. This whole procedure resulted in data approximating above-ground echo heights and field data from 150 sample plots of 200 m^2.

ALS-variables were then derived both for the 200 field measured plots, and the 150 semi-artifical plots. Variables similar to those described by Næsset [2] were used and the variables are briefly described in the next section.

Model fitting and validation

We tested the influence of using semi-artificial plots by fitting models to different sets of modelling data. Starting from an initial set of 150 field measured plots we created sets of modelling data by gradually replacing the field measured plots with semi-artificial plots, one-by-one. The final dataset thus consisted of 150 semi-artificial plots. All the ALS derived variables were used as potential explanatory variables in models of the form

$$\begin{aligned}\ln(vol) = {} & \beta_1 \ln(h_{10.f}) + \beta_2 \ln(h_{20.f}) + \ldots + \beta_9 \ln(h_{90.f}) \\ & + \beta_{10} \ln(d_{0.f}) + \beta_{11} \ln(d_{1.f}) + \ldots + \beta_{19} \ln(d_{9.f}) \\ & + \beta_{20} \ln(h_{mean.f}) + \beta_{21} \ln(h_{cv.f}) + \beta_{22} \ln(h_{10.l}) \\ & + \beta_{23} \ln(h_{20.l}) + \ldots + \beta_{30} \ln(h_{90.l}) + \beta_{31} \ln(d_{0.l}) \\ & + \beta_{32} \ln(d_{1.l}) + \ldots + \beta_{40} \ln(d_{9.l}) + \beta_{41} \ln(h_{mean.l}) \\ & + \beta_{42} \ln(h_{cv.l}) + \epsilon\end{aligned} \quad ,$$

where the response variable *vol* is plot volume. The variables h_{10}, h_{20} ,...h_{90} are the height of the 10th,20th,...90th percentile of the canopy laser echo height distribution and d_0 ,d_1,...,d_9 is the ratio of the number of echoes above canopy height bin 0,1,...,9 to the total number of echoes, respectively. h_{mean} is the canopy echo mean height and h_{cv} is the corresponding coefficient of variation. The *.f* and *.l* subscripts denote first and last return, respectively. A stepwise variable selection procedure was carried out in each iteration, reducing the number of variables in the final models. Back-transformation of the values predicted by the log-log models was done with the correction suggested by Sprugel [3].

The models were all validated using the separate set of 50 field-measured plots. Root mean square error (RMSE) from the validation was calculated to compare the models. RMSE is in this study reported in percent of the mean observed volume on the 50 validation plots.

Results

The validation of the models showed a slight increase in RMSE as the proportion of semi-artificial plots in the modelling data increased (Fig 1). The variability in RMSE between relatively similar models – in terms of modelling data – also seems to increase slightly as the proportion of semi-artificial plots increased.

The model fit to data from only field measured plots had an R^2 of 0.85 and in the validation an RMSE of 22.1% . The model fit to only semi-artificial plots had an R^2 of 0.91 and an RMSE of 24.3%.

Discussion

The validation of the models showed that the prediction errors increased as the proportion of semi-artificial plots in the modelling data increased (Fig. 1). The increase in RMSE was however not more than 2 –3 percentage points which in this context must be considered small. The model fit to only semi-artificial plots yielded in this case predictions comparable – in terms of accuracy – to the model fit to data from field measured plots. This

indicates that semi-artificial plots can be used to estimate plot volume, but further research is needed to verify and test this.
The ALS data in the set of individual trees that were used in the present study were extracted using accurate field measured crown extents. This kind of crown measurements are unlikely to be obtained with the same accuracy using e.g. TLS. Further research could therefore investigate the use of semi-artificial plots created from single-tree data obtained through other means.

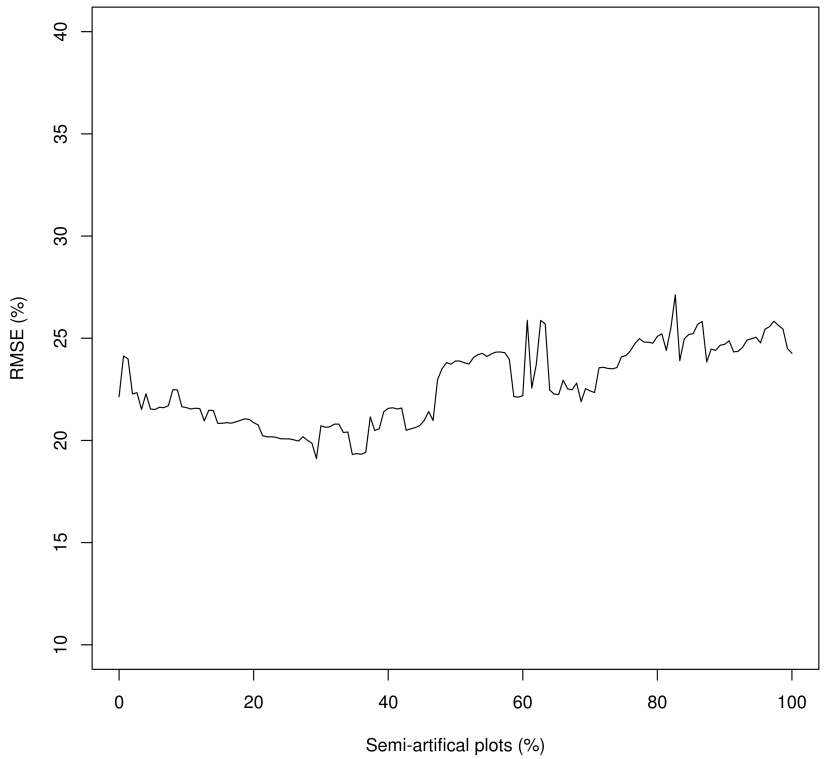


Figure 1: Model validation. RMSE from the validation of the models, with the proportion of semi-artificial plots increasing along the x-axis. At each step one more field-measured plot was replaced with a semi-artificial plot, until the final model was fit to a dataset consisting only of semi-artificial plots.

The data in the present study was not stratified according to site productivity age-class or main tree species, which are commonly adopted stratification criteria in area-based inventories [2]. Stratification could improve the prediction accuracy and further research could investigate the use of semi-artificial plots with stratified data.
Since the semi-artificial plots were created by sampling with replacement from the initial set of individual trees, a large number of plots can be created. This could make the data suited for the use of non-parametric models such as k-nearest neighbour methods. Further research could investigate this as well.
There are several aspects and implications of using artificially created modelling data that were not discussed in this abstract. Thorough consideration and research is needed before this kind of data can be used reliably. It can however be noted that improved knowledge of ALS sensors together with better understanding of the interaction between laser pulses and the trees in the future might allow for creation of artificial model data. This could for example be achieved through a physical simulation of the whole scanning procedure, using ray-tracing. The present study can be viewed as a small step towards the use of artificially created modelling data.

References

[1] Breidenbach J, Næsset E, Gobakken T. Improving k-nearest neighbor predictions in forest inventories by combining high and low density airborne laser scanning data. Remote Sensing of Environment 2012;117:358–65.
[2] Næsset E. Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. Scandinavian Journal of Forest Research 2004;19:164–79.
[3] Sprugel DG. Correcting for Bias in Log-transformed Allometric Equations. Ecology 1983;64:209–2010.

Uncertainty quantification in ALS-based species-specific growing stock volume estimation

Petri Varvia[1], Timo Lähivaara[1], Matti Maltamo[2], Petteri Packalén[2], Timo Tokola[2] & Aku Seppänen[1]

[1]*Department of Applied Physics, University of Eastern Finland*

[2]*School of Forest Sciences, University of Eastern Finland*

Highlights: We propose a Bayesian approach to estimation and uncertainty quantification of species-specific growing stock volume from airborne laser scanning data and aerial imagery. Uncertainty quantification is utilized to identify plots with misidentified dominant tree species. A simpler approach to risk plot identification, usable with conventional area-based methods, is also considered.

Keywords: *uncertainty quantification, Bayesian, forest inventory, airborne laser scanning*

Introduction

Airborne Laser Scanning (ALS), a remote sensing method, has been found to be suitable for obtaining important forest stand attributes such as stem volume (e.g. [4]). The most common approach to ALS-based forest inventory is the area-based approach (e.g. [5]). The area-based approach utilizes statistical relations between the stand attributes, total or species-specific, and metrics calculated from the height distribution of the ALS points in a set of training plots to estimate the stand attributes. For species-specific studies, data from aerial images is usually also used.

The current methods for stand attributes in the area-based approach are usually based on either linear regression or some non-parametric regression method, such as k nearest neighbour (kNN) algorithm [4]. These methods, while generally well performing in the sense of estimation accuracy, do not provide information on the uncertainty of the estimate.

One way to estimate the uncertainty is to formulate the area-based forest inventory problem as a problem of Bayesian statistical inference. The solution of the inference problem is a posterior probability density that describes the uncertainty in the estimate of the parameter of interest. From this probability density, various uncertainty estimates, such as posterior variances or credible intervals (CI) can be readily computed. The procedure also provides point estimates for the stand attributes.

In this study we demonstrate how Bayesian inference approach can be used for estimation of species-specific growing stock volume. We also study if uncertainty quantification can be used to find plots on which the dominant species has a high risk of misidentification. We also consider a simpler, point estimate characteristics based approach to this risk assessment.

Field and remote sensed data

The test area is a typical Finnish managed boreal forest area of around 10,000 hectares located in the municipality of Juuka in Eastern Finland. The dominant tree species in the test area are Scots pine (*Pinus sylvestris* L.) and Norway spruce (*Picea abies* (L.) Karst.), with a minority of deciduous trees, mainly birches (*Betula* spp.). The deciduous trees are considered as a single group. The field measurements were done during the summers of 2005 and 2006. The number of sample plots is 493. For more detailed description of the data, see [6].

The georeferenced ALS data were acquired on 13 July 2005. Nominal sampling density was 0.6 measurements per square meter and the footprint was 60 cm at ground level. Multispectral aerial photographs of the area were taken on 1 September 2005. For details, see again [6]. From the ALS and aerial image data, numerous canopy height percentiles, proportional canopy densities, ALS intensity percentiles, image spectral values, vegetation indices and other metrics were computed. The total number of usable metrics/predictors was 82.

Bayesian inference

The vector of species-specific stem volumes, the variables we are interested in, is denoted by V. The metrics computed from the ALS and aerial image data form the measurement vector z. In Bayesian statistical inference, a prior hypothesis is updated with information gained from measurements (z) using Bayes' rule, see e.g. [1, 3]. In the Bayesian framework both z and V are modeled as random variables. The prior is our a priori information on the stem volumes V (i.e. that volume is never less than zero), encoded in a probability density. The new information on V gained from measurements z is encoded in the so-called likelihood density, which is formulated using a forward model that connects V to z and information (i.e. mean, covariance) on the measurement and modeling errors. As the forward model we used an empirical linear model learned using the set of training plots. The result of the inference is the posterior density $p(V|z)$, which describes the updated information on the target variables.

From the posterior density, point estimates, such as maximum a posteriori (MAP) can be computed. We can also compute credible intervals (similar to confidence intervals, in practice) to get error limits for the point estimates. A kNN-based method was used as a reference

Dominant tree species identification

Dominant tree species is here the species with the largest stem volume on plot. The dominant species identified using the estimated stem volume may be wrong and it would be useful if plots with a high risk of misidentification could be assessed. The posterior density can be used to generate uncertainty indicators that can be used for this. The posterior probability of each species being the dominant one, that is, the probability that the stem volume V_i of the species i is larger than the stem volume of any other tree species j is:

$$P_{\text{dominance},i} = \int_G p(V|z)dV, \quad \text{where } G = \{V \in \mathbb{R}^N | V_i > V_j,\ \forall j \neq i\}, \tag{1}$$

where N is the number of different tree species (or species groups). The dominance probability $P_{\text{dominance}}$ describes uncertainty in the estimated dominant species. If the dominance probability of the estimated dominant tree species is low (for example ca. 0.5), the dominant species identification is uncertain and the estimated dominant species is more likely wrong. Conversely, if the estimated dominant species has very high dominance probability (for example over 0.95), it is likely that the dominant species is correct.

An ad hoc way to find high-risk plots is to examine the plot-level mixedness of the point estimate, whether Bayesian or kNN. If the plot estimate is highly mixed, that is, if it contains fairly equal proportions of multiple tree species, then the identification of the dominant species will be more vulnerable to estimation error.

Results and discussion

An example of posterior marginal densities for the species-specific plot volumes and total plot volume of a single test plot is presented in Figure 1. The figure includes the point estimate values for the variables and 95% credible intervals (shaded area). This example plot contains nearly equal proportions of pine and spruce, with no deciduous trees. Both the Bayesian and kNN-reference overestimate the volume of spruce, with a corresponding underestimation of pine volume. The absence of the deciduous component is estimated more accurately by both methods (although both volume estimates are nonzero). The 95% credible intervals for volumes of pine and spruce are wide. This, in part, tells us that there is considerable uncertainty in the estimated pine-spruce ratio in this example.

Leave-one-out (LOO) cross-validation was used to study point estimation accuracy. The relative RMSE for the Bayesian MAP-estimate was 42.35% for pine volume, 82.34% for spruce, 99.38% for deciduous and 21.26% for total stem volume. Total volume was calculated by summing up the species-specific volumes. For the reference kNN method, the relative RMSEs were 39.01% for pine, 75.16% for spruce, 81.20% for deciduous and 27.64% for total volume. The reference method was slightly more precise in estimating the species-specific volumes.

In the Bayesian case 95% credible intervals were computed for each plot. Ideally, the 95% intervals should contain the true parameter value in 95% of the test cases. The real value was between $92 - 97\%$ for each estimated variable, thus nearly optimal.

The posterior dominance probability method was then used to assess plots with a risk of dominant species misidentification. The results were compared with the kNN estimate mixedness criterion. For each plot in the LOO cross-validation set, the dominance probability of the estimated dominant species (MAP) and plot mixedness for the kNN point estimate was computed. The plots were then sorted for each strategy based on perceived risk: plots with low dominance probability have higher misclassification risk, as have plots with high mixedness. Some percentage of the identified risk plots is then field-checked. By choosing the riskiest plots first, it is hoped to improve dominant species identification accuracy most efficiently.

The base dominant species identification accuracy was 89.7% for Bayesian MAP-estimates and 90.1% for the reference kNN method. If the riskiest 10% of the plots was field-checked, the resulting accuracy was 94.7% for the Bayesian case with dominance probability strategy and 93.9% for the reference (kNN and mixedness strategy). The Bayesian dominance probability strategy is slightly more efficient than the kNN estimate mixedness strategies. However, both strategies work significantly better than field-checking plots randomly.

Conclusions

We presented a Bayesian inference based method for the estimation of stand attributes from ALS data and for uncertainty quantification of this process. The approach is applicable to standard area-based forest inventory scheme, where a set of training plots is field-measured. The approach was thoroughly validated, showing that in the sense of RMSE it performs almost as well as the existing state-of-the-art methods for area-based forest inventory, while also allowing error intervals to be computed for the estimates.

We also studied if uncertainty quantification can be used in finding plots on which the dominant tree species (by volume) is misidentified. A simpler, point estimate mixedness based approach was also considered. We showed that both the Bayesian uncertainty quantification approach and the simpler point estimate based can be used for this kind of risk assessment. In general, the difference between the Bayesian and the point estimate based risk assessment strategy was minor, however, the Bayesian approach has firm grounding on statistical theory. The error bounds and uncertainty analysis provided by our approach would be useful in risk assessment of forest planning.

Acknowledgment

This work was supported in part by the University of Eastern Finland (spearhead project Multiscale geospatial analysis of forest ecosystems) and in part by the Academy of Finland (Finnish Center of Excellence of Inverse

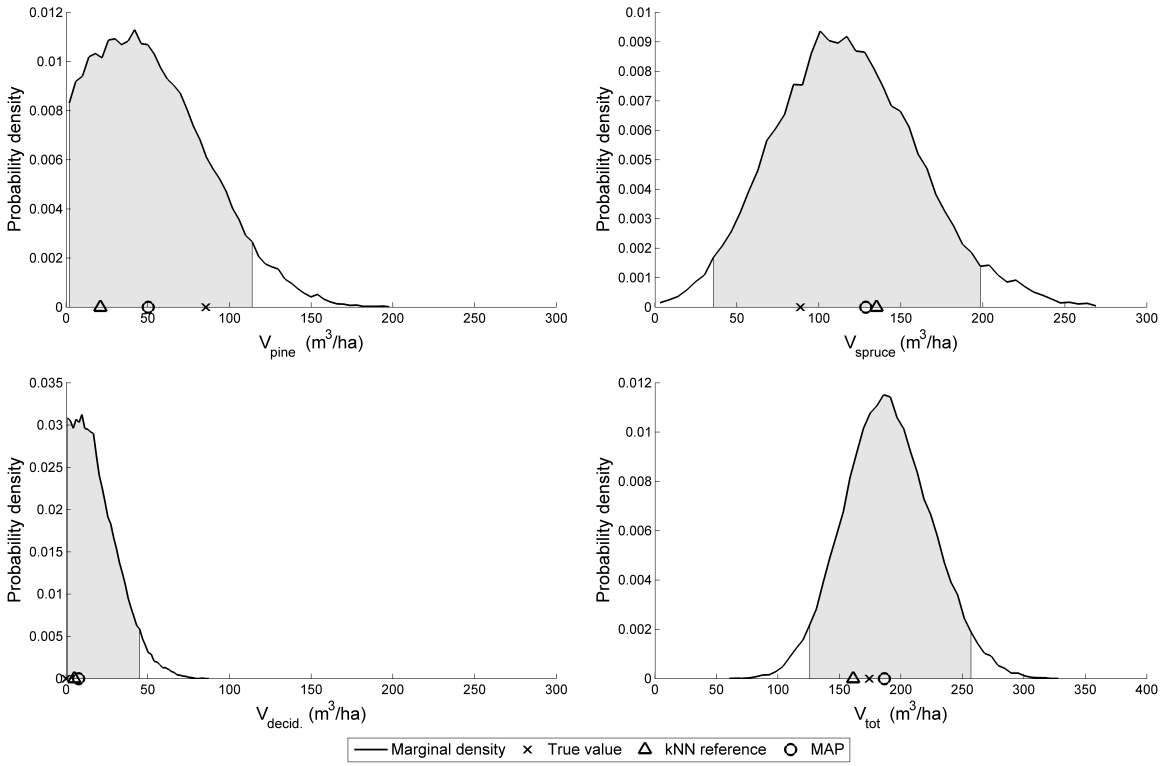


Figure 1: Examples of marginal densities of species-specific and total stem volume of a single test plot. The 95% credible intervals are shaded with gray. Top left: pine, top right: spruce, bottom left: decidous trees, bottom right: total volume.

Problems Research 2012-2017, project number 250215, and projects 270174, 273536 and 257372).

References

[1] Daniela Calvetti and Erkki Somersalo. *An Introduction to Bayesian Scientific Computing: Ten Lectures on Subjective Computing*. Springer, 2007.

[2] H. Haario, M. Laine, A. Mira, and E. Saksman. DRAM: efficient adaptive MCMC. *Statistics and Computing*, 16(4):339–354, 2006.

[3] J. P. Kaipio and E. Somersalo. *Statistical and Computational Inverse Problems*. Springer, New York, 2005.

[4] Matti Maltamo, Erik Næsset, and Jari Vauhkonen, editors. *Forestry Applications of Airborne Laser Scanning*. Springer, 2014.

[5] Erik Næsset. 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, 2002.

[6] Petteri Packalén, Aki Suvanto, Matti Maltamo, et al. A two stage method to estimate species-specific growing stock. *Photogrammetric Engineering and Remote Sensing*, 75(12):1451–1460, 2009.

Influence of sampling design parameters on biomass predictions derived from airborne lidar data

Marc Bouvier[1], Sylvie Durrieu[1], Richard A. Fournier[2], Nathalie Saint-Geours[1], Grégoire Vincent[3], Dominique Guyon[4], Eloi Grau[1,5], Bruno Hérault[6]

[1]*Irstea, UMR TETIS, Maison de la Télédétection, 34196 Montpellier, France*
[2]*Centre d'Applications et de Recherches en Télédétection (CARTEL), Département de géomatique appliquée, Université de Sherbrooke, Sherbrooke, Québec, Canada, J1K 2R1*
[3]*IRD, UMR AMAP, Montpellier, 34000 France*
[4]*INRA, UMR1391 ISPA, 33140 Villenave d'Ornon, France*
[5]*CNES, 31400 Toulouse*
[6]*Cirad, UMR EcoFoG, Kourou, 97300, France*

Highlights: The study provides technical guidelines to optimize sampling design parameters in the context of the implementation of biomass predictive models from airborne lidar data. Parameters under study were those easily controlled when implementing an area-based approach, i.e. lidar pulse density and both field sampling protocols and measurements. Parameters most impacting the model accuracy were identified by conducting a one-factor-at-a-time and a global sensitivity analyses. Number and size of field plot, and plot center position accuracy were identified to be the parameters most impacting model accuracy.

Keywords: **Airborne lidar, Area-based approach, Aboveground Biomass, Model accuracy, Sensitivity analysis, Monte Carlo analysis.**

Introduction

Lidar (Light Detection And Ranging) is used for many forest applications, and, in particular, to support forest inventory. Airborne laser scanning (ALS) coupled with field measurements is an effective approach that can be used to develop predictive models for assessing forest inventory attributes over large areas at a much lower cost than with traditional inventory practices. ALS is now used operationally to enhance existing inventories. With the increased use of ALS in forest applications, good survey design is increasingly important to enhance information content while maximizing cost-effectiveness. However, previous lidar studies have reported considerable variability in aboveground biomass (AGB) prediction accuracy [1]. Numerous parameters may affect the ability to reliably predict forest parameters from ALS data. Prediction accuracy on AGB primarily depends on three groups of parameters: (1) stand complexity, (2) lidar sensor and flight parameters, and (3) field protocols and measurements. These parameters affect prediction quality and consistency. Stand complexity is inherent to the sites under study and cannot be modified; one must cope with it and try to use models that have proven their effectiveness in complex environments. Nevertheless, the two other groups of parameters can be studied and carefully defined in order to maximize the chances of meeting accuracy requirements.

Deciding which lidar sensor and flight parameters are the most suitable when planning and designing a lidar survey involves a trade-off between acquisition cost and result accuracy. Pulse density (in pulses/m^2) is a key parameter in area-based approaches and a major cost driver dictated by ALS system setting and flight configuration [2]. Regarding the effects of pulse density, several studies found only small differences in results when comparing stand attributes predictions from data acquired at different pulse densities varying from 0.8 to 3.2 pulses/m² [3]. However, relevant lidar metrics selected to build predictive models were found to differ significantly with pulse density. It is therefore important to investigate how, and to what extent, pulse density affects stand attribute prediction accuracy.

Field protocols and measurements involve other parameters affecting stand attribute prediction accuracy that need to be optimized as field surveys are time consuming and costly. Field protocol design requires setting many rules regarding the choice of the number and the size of field plots, their spatial distribution, the threshold of the diameter at breast height (DBH, trunk diameter measured at 1.3 m above ground) defining trees to be inventoried, to name but a few. The cost associated with fieldwork is highly dependent on these choices. Gobakken and Næsset [4] examined both the number and size of field plots. They concluded that the optimal configuration is a tradeoff depending on inventory costs and forest structure. Plot size was also shown to influence predictions of AGB as larger plots reduce edge effects [5]. Moreover, higher prediction accuracy has been observed with an increase in plot size [1]. GPS position errors were also found to impact significantly prediction accuracy in Norway forests with a decrease in volume prediction accuracy of 15.8% for position errors up to 5 m [6]. Frazer et al. [5] investigated how plot size and co-registration errors interact to influence

AGB prediction. They found that an increase in circular plot radius from 10 to 25 m reduced the impact of co-registration error and improved AGB prediction accuracy by 13.3%. Thus, volume and AGB predictions from ALS data are highly dependent on the way field surveys were conducted as well as on the way lidar data were acquired. Both these groups of parameters can be, to some extent, set in order to optimize assessment of forest attributes from ALS data.

There is a need for more comprehensive approaches that can quantify the specific impacts of different lidar acquisition parameters, field protocols and measurements on the accuracy of the resulting model. Among the numerous parameters influencing AGB predictive model accuracy, only a few of them can be easily controlled, i.e. lidar pulse density and both field sampling protocols and measurements. Therefore, the goal of this study was to provide technical guidelines to optimize lidar pulse density and field survey protocols in order to implement predictive models of AGB from ALS data. To that aim parameters most impacting the model accuracy were identified by conducting a one-factor-at-a-time and a global sensitivity analyses.

Material and methods

Two study sites were selected: a pine plantation and a tropical forest. The pine site is located in the Landes region in southwestern France (44.40° N, 0.50° W). This forest area is dominated by mono-specific stands of Maritime pine (*Pinus pinaster* Aiton) in even aged plantations. The tropical site is a lowland rainforest located in coastal French Guiana (5.25° N, 52.93° W). Several Lidar data sets are available at several point densities on both sites (0.5 to 4.0 pulses/m² on the first site and 3.5 to 10.0 pulses/m² on the second site).

We investigated the influence of sampling design parameters on AGB prediction accuracy from ALS data using an ABA. Regression models for AGB predictions were developed according to a single method, recently developed by Bouvier et al. [7], so as to focus on the parameters explaining AGB prediction accuracy variability rather than model selection. As impacts are generally inter-dependant and thus difficult to compare, the relative importance of parameters is difficult to assess. The influence of these key parameters was assessed individually by carrying out sensitivity analysis in order to define which parameter has the most influence on AGB prediction accuracy. Thus, changing parameters one at a time (OAT approach), we explored the impact of: pulse density, the number of field plots, field plot size, minimum DBH threshold (DBH_{min}), DBH and H measurement accuracies and field plot position accuracy. As some uncertainties can accumulate and propagate, the influence of field parameters was also investigated in a Global sensitivity analysis (GSA). GSA aims to study how the uncertainty of a model output, i.e. forest attributes, can be apportioned to different sources of uncertainty in its inputs, i.e. survey parameters. It allows for a ranking of the input variables according to their contribution to the output variability. GSA thus helps to identify the key parameters which deserve specific attention. In the first study site GSA was used to study impact of four parameters - i.e., DBH and H measurement accuracies, field plot position accuracy, and allometric equations used- and to compare the relative impacts for four field protocols (two DBH_{min} and two field plot radiuses).

Results

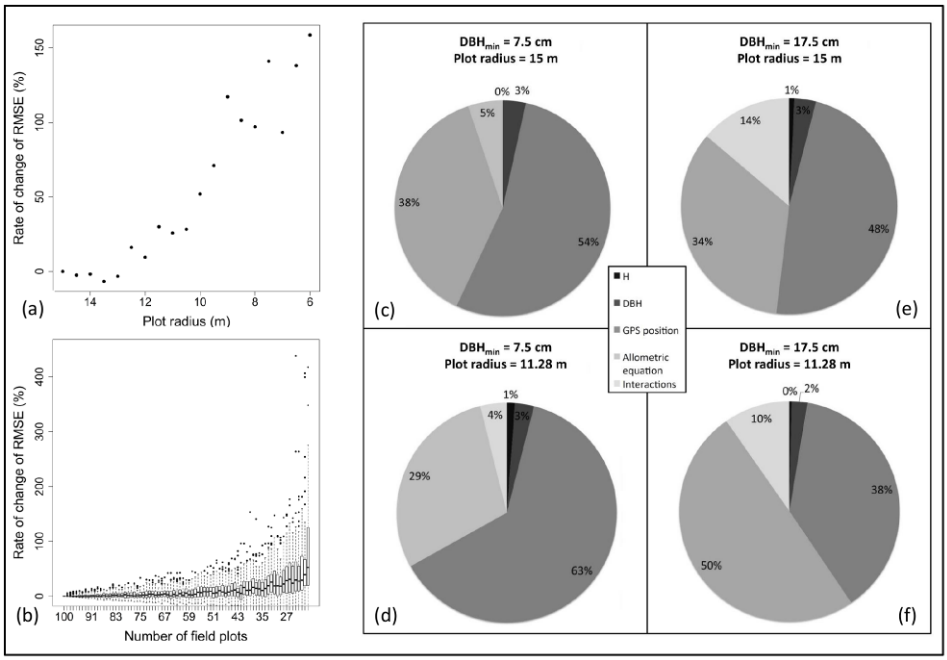


Figure. Results on pine plantation: (a) Rate of change of RMSE, expressed as a percentage of the RMSE obtained with the maximum plot radius for AGB models calibrated and validated using different field plot radius

from 15 to 6 m. (b) Rate of change of RMSE using different field plot radius from 100 to 20 m. Subsets were randomly selected from among all the field plots and selection steps were repeated 100 times. (c) - (f) Part of variance explained by H, DBH and field plot position measurement errors, and allometric equations used to predict AGB. Four protocols have been investigated: (c) 7.5 cm minimum DBH threshold (DBH_{min}) and 15 m field plot radius; (d) 7.5 cm DBH_{min} and 11.28 m field plot radius; (e) 17.5 cm DBH_{min} and 15 m field plot radius; and (f) 17.5 cm DBH_{min} and 11.28 m field plot radius.

A one-factor-at-a-time (figure (a)-(b)) and a global sensitivity analysis (figure (c)-(f)) were applied to identify the parameters most impacting model accuracy. Variability in plot size and number of field plots were observed to significantly impact model performance (figure (a)-(b)). Preliminary results on the tropical forest site confirmed the high impact of plot size and plot number on AGB estimation accuracy. The part of variance attributable to each parameter, and linked to interactions between parameters, varied depending on which of the four inventory protocols was concerned (figure (c)-(f)). Whatever the inventory protocol, the part of variance explained by plot center positioning accuracy (38% - 63%) or the allometric equation used (29% - 50%) were significantly higher than the part of variance explained by the DBH (2% - 3%) and H (0% - 1%) parameters. The part of variance attributable to the interaction between parameters was found to vary between 10 and 14%.

Conclusion

Some recommendations can be drawn from our results for those interested in estimating AGB from a lidar-based ABA in a pine plantation. First, regarding ALS data acquisition, cost savings can be made by reducing pulse density to 0.5 pulse/m² without any major impact on model quality. Even if DBH and H measurement accuracies were shown to contribute to a lesser extent to the prediction error, field measurement costs will still remain high to ensure a good quality model. This is due to requirements regarding field plot number, position accuracy and plot size. Therefore, and despite the relative simplicity of the environment, a minimum of 40 plots is recommended. It is also recommended to inventory field plots at least 13 m in radius when plots contain trees with a DBH equal or greater to 17.5 cm in a temperate forest plantation. In addition, plot positioning efforts should be performed with great care as this parameter has a great impact even on simple and regular stands for which a position accuracy of center plots below 5 m is highly recommended. Different recommendations may be warranted when working in more complex multispecies, multistrata dense tropical forests. The preliminary results will be compared to those obtained in the tropical site. Differences and similarities between the two forest types will enable us to evaluate influence and sensitivity of design parameters with regard to the forest complexity.

References

[1] S. G. Zolkos, S. J. Goetz, and R. Dubayah, « A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing », *Remote Sens. Environ.*, vol. 128, n° 0, p. 289‑298, janv. 2013.

[2] J. Hyyppä, H. Hyyppä, D. Leckie, F. Gougeon, X. Yu, and M. Maltamo, « Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests », *Int. J. Remote Sens.*, vol. 29, n° 5, p. 1339‑1366, 2008.

[3] E. Næsset, « Effects of different sensors, flying altitudes, and pulse repetition frequencies on forest canopy metrics and biophysical stand properties derived from small-footprint airborne laser data », *Remote Sens. Environ.*, vol. 113, n° 1, p. 148‑159, 2009.

[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. For. Res.*, vol. 38, n° 5, p. 1095‑1109, 2008.

[5] G. W. Frazer, S. Magnussen, M. A. Wulder, and K. O. Niemann, « Simulated impact of sample plot size and co-registration error on the accuracy and uncertainty of LiDAR-derived estimates of forest stand biomass », *Remote Sens. Environ.*, vol. 115, n° 2, p. 636‑649, 2011.

[6] T. Gobakken and E. Næsset, « Assessing effects of positioning errors and sample plot size on biophysical stand properties derived from airborne laser scanner data », *Can. J. For. Res.*, vol. 39, n° 5, p. 1036‑1052, 2009.

[7] 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, n° 0, p. 322‑334, janv. 2015.

Edge-tree correction improves the performance of area-based approach with airborne laser scanning

Petteri Packalen[1], Jacob Strunk[2], Juho Pitkänen[3], Hailemariam Temesgen[4] & Matti Maltamo[5]

[1] *petteri.packalen@uef.fi, School of Forest Sciences,University of Eastern Finland, Finland.*
[2] *jacob.strunk@dnr.wa.gov, Washington State Department of Natural Resources, USA.*
[3] *juho.pitkanen@luke.fi, Natural Resources Institute Finland (Luke), Joensuu, Finland.*
[4] *temesgen.hailemariam@oregonstate.edu, College of Forestry, Oregon State University, USA.*
[5] *matti.maltamo@uef.fi, School of Forest Sciences, University of Eastern Finland, Finland.*

Highlights: We describe a new method to improve the correspondence between field and airborne laser scanning (ALS) measurements in an area based approach (ABA). In the proposed solution, predicted tree positions and crown shapes are used to adjust plot and grid cell boundaries. The new method provided lower error rate than conventional ABA.

Key words: *Remote sensing, ALS, area-based approach, edge-tree.*

Introduction

An established practice in forest inventory is to designate a tree as falling within a sample plot if the center of the bole is inside the plot bounds [1]. The same approach is used with area based methods (ABA) to select airborne laser scanning (ALS) points for a plot or grid cell – points are designated as falling within a plot or grid cell if their coordinates fall within the plot or grid cell. However, this approach to selecting ALS points results in edge tree discrepancies with ABA because some of the crowns from "out" trees extend into the plot ALS data (type 1 discrepancy), and crown from "in" trees extend outside of the plot ALS data (type 2 discrepancy). Both types of discrepancies decrease the observed relationship between ALS metrics and plot attributes.

The edge-tree problem in ABA has been remarked. Nelson et al. [2] showed that estimates of biomass and volume were affected by the type 1 discrepancies. Næsset et al. [3] also pointed out the issue with type 1 discrepancies when estimating change in forest biomass from multi-temporal ALS datasets. Mascaro et al. [4] is, to the best of our knowledge, the only article where the edge-tree discrepancy has somehow been addressed. They addressed the discrepancy by spatially mapping trees' carbon content using their crown footprints relative to the plot borders. However, such an approach cannot be used in forest inventory because it does not follow the established definition of whether trees are "in" or "out" of the plot depending upon bole location.

In this study, we present a method to account for edge-tree discrepancy in ABA. The aim is to improve the performance of ABA by considering the predicted crown of edge-trees.

Material

The study area is a boreal managed forest area in eastern Finland. A total of 79 field plots were placed purposively (non-random) to reflect species and size variation for the study area. Scots pine (Pinus sylvestris L.) is the dominant tree species in the study area. This study used field measurements from concentric 6.37 m (127 m^2 area) and 9 m (254 m^2 area) radius circular plots. These data were extracted from larger 20x20 and 30x30 meter square stem-mapped plots. We will refer to these as 127 m^2 and 254 m^2 plots respectively.

ALS data were collected on June 26, 2009 using an Optech ALTM Gemini laser scanning system. The nominal pulse density was approximately 12 pulses per square meter. The test site was scanned from an altitude of approximately 600 m above ground level with a side overlap of 55%. Side overlap of 55% means that each location is covered by at least two flight lines. This reduces the occurrence of shadow areas in which tree foliage obstructs the path of the laser from reaching behind trees. A Digital Terrain Model (DTM) was constructed by first classifying points as ground and non-ground hits and then interpolating a raster using Delaunay triangulation. Heights above ground (dZ) for ALS echoes were calculated by subtracting DTM from the original (above ellipsoid) heights.

Methods

An overview of the method which takes into account edge-tree discrepancy in ABA is given in Figure 1. We will refer to this method as Enhanced ABA (EABA). The idea is to predict tree positions and their crowns and then modify plot and cell areas using predicted tree positions and segmented crowns. The method requires that ALS data is suitable for individual tree delineation (i.e. point density must be higher than in normal ABA). EABA has 5 additional steps beyond typical ABA processing including:

1) individual trees are detected and their crowns are segmented;

2) edge trees (tree crown intersect the plot boundary) are identified;
3) edge trees are labeled as "in" (tree apex is within the plot boundary) or "out" (tree apex is outside the plot boundary);
4) edge trees which are "in" the plot are used to extend the plot boundary (type 2 discrepancy); and lastly
5) edge trees which are "out" determine type 1 discrepancy regions.

For type 1 discrepancies, the dZs are set to zero. This mimic the case where trees outside the plot do not intrude on the plot and therefore ALS points hit the ground. To correct for type 2 discrepancies, the bounds of the sample plot are extended to contain the full tree crowns. By extending plot borders we improve the chance that ALS points for a plot correspond to field measured trees.

Following steps 1 through 5, ALS points are extracted for each modified plot or cell, and the remaining analyses are performed in the typical ABA fashion. Note that although the modified plot area is used when computing predictor variables predictions are still made for the original plot or cell areas. In this study trees were located on the CHM using watershed segmentation [5]. EABA performance was benchmarked against a typical ABA strategy by fitting regression models for the plot volume (V) separately for ABA and EABA using 127 m^2 and 254 m^2 sample plot sizes.

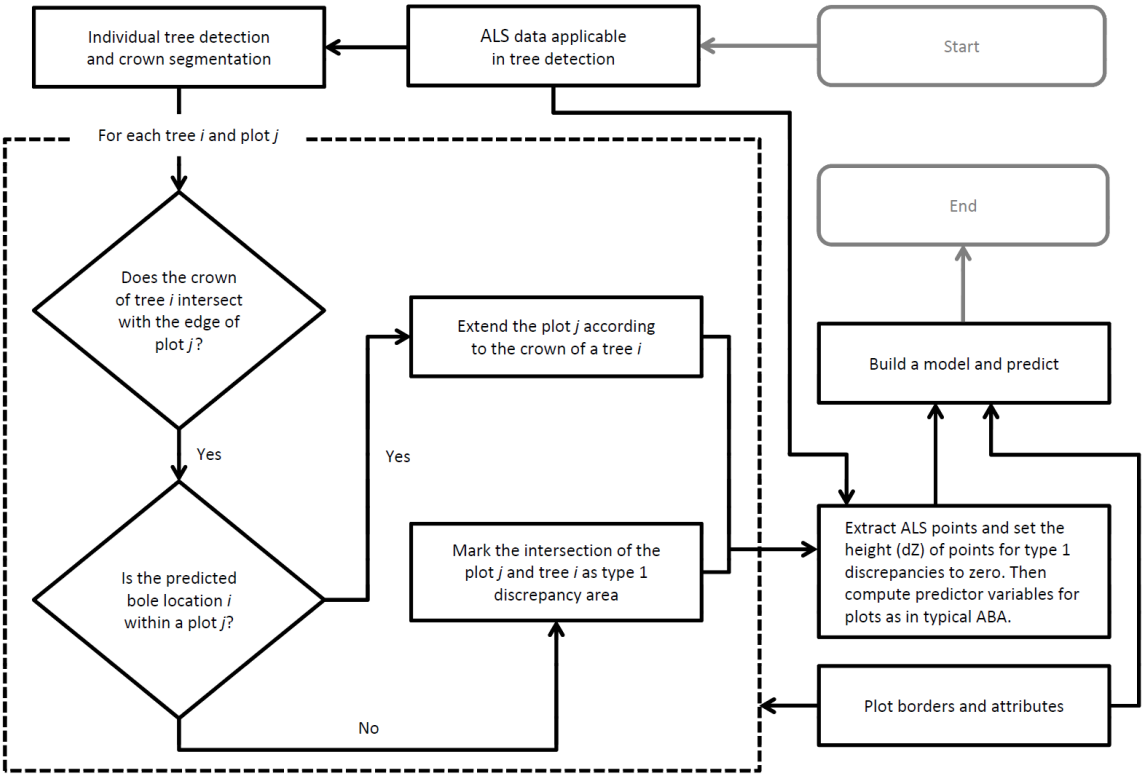


Figure 1: An overview of the EABA method.

Results

RMSE indicate a better fit for the EABA model than for the ABA model (Table 1). With respect to RMSE, the performance improved from 23.16% to 19.11% with 127 m^2 plots and from 19.08% to 16.95% with 254 m^2 plots, so the improvement was 17% with small and 11% with large plots (by absolute means 4.05% and 2.13% respectively). Although the variables selected in the two approaches differed, the choice of variables had minimal impact on our inference as model's RMSEs were nearly identical when we exchanged variables selected for ABA and EABA within the same plot size.

Table 1: Error statistics for ABA and EABA using 127 m^2 and 254 m^2 sample plot sizes.

	Plot size 127 m^2		Plot size 254 m^2	
	ABA	EABA	ABA	EABA
RMSE	23.16 %	19.11%	19.08%	16.95%

Since we consider the observed improvement in RMSE for EABA to be of practical importance, we performed a hypothesis test to determine the probability that the observed difference could arise from random chance. Due to the known difficulty in specifying the sampling distribution of RMSE values, we tested the hypothesis that

RMSE(ABA) ≤ RMSE(EABA) using a bootstrap approach [6]. Following 5000 bootstraps with 127 m^2, plots RMSE(ABA) was less than RMSE(EABA) in 1.46 percent of simulations. Corresponding figure for 254 m^2 plots was 0.62 percent. In other words, predictions are highly statistically significant with "p-values" 0.0062 and 0.0146. These results indicate that the apparent improvement in RMSE for EABA is unlikely to arise due to random chance; we have a high degree of confidence that EABA provides lower error rate than ABA.

Discussion and conclusions

The EABA method we developed provided a lower error rate than a typical ABA approach. The improvement was achieved by detecting and delineating individual trees and including or excluding ALS points based on predicted crown attributes. The EABA method requires higher density ALS data than ABA because improvement is obtained by means of detecting individual trees. This is an important consideration because the cost of ALS data increases with density. However, given historical trends with respect to ALS technology, the advancement of ALS sensor technology means that point densities will most likely increase in future.

Plot size was shown to affect edge errors and ABA performance. As the plot size decreased, the proportion of edge-trees and consequently error due to edge-trees increased. We achieved approximately the same error rate with smaller plots using EABA (19.11%) as with larger plots with typical ABA (19.08%). These results suggest that we could potentially measure 50% fewer trees with the smaller plots using the EABA method and achieve the same performance as we would receive with ABA with the larger plots, whereas with the same plot size EABA provides lower error rate than ABA. Packalen et al. [7] gives an in-depth description of the EABA method.

References

[1] Loetsch, F., Zoehrer, F. & Haller, K.E. (1973). Forest inventory (vol. 2). München. Bern. Wien: BLV – Verlagsges.

[2] Nelson, R., Jimenez, J., Schnell, C.E., Hartshorn, G.S., Gregoire, T.G. & Oderwald, R. (2000). Technical note: Canopy height models and airborne lasers to estimate forest biomass: Two problems. International Journal of Remote Sensing, 21(11), 2153–2162.

[3] Næsset, E., Bollandsås, O.-M., Gobakken, T., Gregoire, T.G. & Ståhl, G. (2013). Model-assisted estimation of change in forest biomass over an 11 year period in a sample survey supported by airborne LiDAR: A case study with post-stratification to provide "activity data". Remote Sensing of Environment, 128, 299–314.

[4] Mascaro, J., Detto, M., Asner, G.P. & Muller-Landau, H.C. (2011). Evaluating uncertainty in mapping forest carbon with airborne LiDAR. Remote Sensing of Environment, 115(12), 3770–3774.

[5] Packalén, P., Pitkänen, J. & Maltamo, M. (2008). Comparison of individual tree detection and canopy height distribution approaches: a case study in Finland. In: Hill, R.A., Rosette, J. & Suárez, J. (eds.). Proceedings of SilviLaser 2008: 8th International Conference on LiDAR Applications in Forest Assessment and Inventory, September 17-19, 2008, Heriot-Watt University, Edinburgh, UK. pp. 22-29.

[6] Efron, B. & Tibshirani, R.J. (1993). An Introduction to the Bootstrap, London: Chapman & Hall.

[7] Packalen, P., Strunk, J.L., Pitkänen, J.A., Temesgen, H. & Maltamo, M. (2015). Edge-Tree Correction for Predicting Forest Inventory Attributes Using Area-Based Approach With Airborne Laser Scanning. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 8(3), 1274-1280.

Low energy LIDARs for biomass applications

J, Lochard (1), J.Costeraste (2), F.Tinto (2), E.Grau (2)(3), S. Durrieu (3), F.Fabre (1)

(1) Airbus Defence and Space, 31 Av. des Cosmonautes, 31 402 Toulouse Cedex, France
(2) CNES, 18 avenue Edouard Belin, 31 401 Toulouse Cedex 9, France
(3) IRSTEA, 361 rue J.F. Breton, BP 5095, 34196 Montpellier Cedex 5, France

Highlights: A new approach for LIDAR altimetry mission for biomass applications (tree height measurement) is explored based on low emitted laser energy at high repetition frequency. Low energy approach drastically reduces the laser induced risks. Altimetry performances meet preliminary science requirements. The proposed instrument design is compatible with a space mission.

Key words: *Canopy height, low energy, high frequency, photon counting*

Introduction

Up to now altimetry LIDAR missions in space have used the full-waveform technique. This technique implies operating energetic lasers (tens of mJ per shot) at low repetition frequency (tens of Hz). Operating lasers at high energy in space comes with risks. Phenomena as the laser induced contamination (LIC) and the laser induced damage (LID) become a real concern at such energy. Eye safety is an additional limitation to high energy as well, essentially in the visible band.

An alternative approach is to operate lasers at lower energy with an increased pulse repetition frequency (PRF). Such option has been adopted for the next LIDAR altimetry mission ICESat-2. Its instrument, ATLAS, operates at a PRF of 10 KHz and at low energy (hundreds of µJ per shot). The detection chain operates then in single photon counting mode (SPC). Lowering the energy by a factor 10^2 to 10^3 with respect to the previous mission represents a major asset to mitigate the laser associated risks for a space mission. Besides lowering the laser associated risks, high repetition frequency LIDARs provide continuous data in contrast with low frequency previous altimetry missions. This continuity of measurements on ground is seen as an asset because of the liberty it brings in data processing and as it expands the information that can be retrieved. It helps for example to tackle a major issue of low repetition frequency LIDAR: the impact of the local slope on the measurement error.

The ATLAS instrument has been designed almost exclusively for hard ground targets with high albedo. Forest monitoring being a secondary goal of the mission, the instrument design is not optimized with that respect. The semi-transparent nature of the forest and the low albedo involved represent major differences with respect to hard targets. Airbus Defence and Space has investigated the viability of such a new low energy approach in SPC regime to estimate forest canopy height from space. This study was funded by CNES. This paper presents the study and its conclusions.

Method and results

The study presented here relies on numerical modelling and simulations. LIDAR data simulations were generated thanks to the radiative transfer model DART [1] for defined reference scenes. Four scenes of interest were studied to cover the two parameters: tree height and canopy density (high/dense, high/sparse, low/dense and low/sparse) (see Figure 1). For each scene, DART provided simulated full-waveforms for a 20 meter diameter spot suitable for forest applications. The low energy approach induces a low level of return signal from the scene, imposing to accumulate successive shot returns. We chose to set an accumulation of 14 successive shots as a baseline to derive the instrument performance. At the considered PRF, the spots move on the ground by 10 m over the corresponding duration. Therefore not only one waveform but a set of waveforms taking into account the displacement of the spot on ground has been provided. Note that each reference scene has been simulated at four wavelengths from UV to IR (0.355 µm, 0.532 µm, 1.06 µm and 1.55 µm).

From the waveforms it was then possible to generate discrete statistically representative low flux events detected by the instrument. Indeed, operating the laser at low energy leads to work with low return flux from the scene (in photon counting regime). The arrival of the incoming photons to the instrument follows the Poisson statistical law. From these discrete pseudo-waveforms one could extract the time duration between the top of canopy echo and the ground echo and derive the tree height in consequence. It is important to note that the data processing algorithm enabling to derive canopy height from the return signal contributes to a large extent to the altimetry accuracy. If future investigations were to be led, this processing aspect should be further addressed. In this study, for the purpose of simplicity of both implementation and interpretation, canopy height was defined based on the cumulated energy of accumulated return signals, with thresholds set to 3% and 97% for canopy top and ground, respectively.

In order to explore the sensitivity of the instrument performance with respect to the signal-to-noise ratio, the link budget was scaled based on the LIDAR product quantity (product of the laser emitted energy and of the instrument collecting surface). The science team had set a preliminary specification on the instrument measurement error to 2 m. By sweeping the lidar product value, we could determine the minimum link budget to meet the error specification.

A first conclusion drawn from these simulations is that the near IR and IR wavelengths are the most suited for the forest mission in the explored wavelength range. The UV and green wavelengths provide degraded performances with respect to NIR and IR. This can be related to the lower albedos of both leaves and ground for the reference scenes. In addition, two reference scenes were not considered for the instrument sizing because the accuracy on the tree height was too poor. The degraded accuracy on these scenes is linked to a vegetation either too sparse (15 % vegetation coverage) or with maximum canopy height too low (4 m).

Therefore the threshold link budget to meet the measurement error specification has been derived at 1.064 µm and 1.55 µm on two reference scenes (high trees for dense and sparse canopies) which were representative of one tropical forest and one temperate forest type, a mangrove and a maritime pine plantation. The tropical scene is driving the design as tropical forests store about 40% of the forest biomass. The resulting proposed design consists of a telescope of 0.8 m diameter and a laser with an energy per shot of 200 µJ at the PRF of 10 KHz. Such characteristics result in an instrument with mass and electrical power consumption budgets of 220 kg and 240 W.

Technical review

In parallel a technical review of both detection subsystem and laser sources was led to identify potential suitable technologies for the mission in both domains. Avalanche photodiodes operated in linear regime despite their moderate maturity appear to be the best candidate for such mission. Such detectors indeed combine high detection efficiency and very low dead time after detection. On the source side, several options exist although the most mature remains the MOPA architecture already chosen for the in-orbit LIDAR missions. The forest spectral response imposes that the LIDAR works in SWIR band. Current common lasers operate at 1.06µm and 1.55µm without any conversion. The better wall-plug efficiency for Nd:YAG at 1.06µm leads us to favour this wavelength over 1.55µm.

Conclusion

This study demonstrates that the use of high frequency and low energy LIDARs for measuring tree height is achievable with a reasonable design and interface budgets compatible with space missions. While maintaining good performances with respect to higher energy LIDARs, operating at low energy enables to greatly reduce the laser induced risks (e.g. LID, LIC, eye safety) which are a major concern for LIDAR space missions. In addition such operating point provides continuous measurements on ground offering more freedom in the data processing and enabling to reduce the altimetry uncertainties on sloped areas.

Acknowledgement

Airbus Defence and Space would like to thank CNES for funding this study and also thanks Christophe Proisy from IRD for providing the mangrove scene.

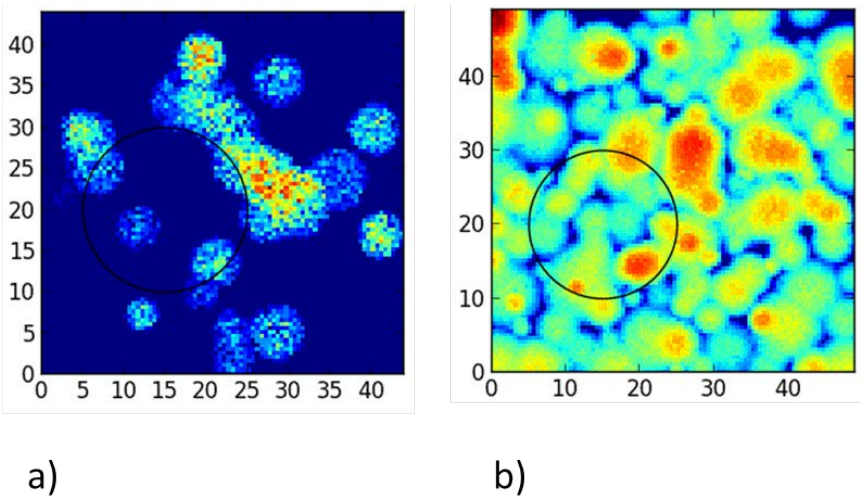


Figure 1: View of two simulated scenes (50m x 50m). The black circle shows the dimension of the LIDAR spot on-ground. A) Landes scenes: high trees, low canopy density. B) Tropical scene: high trees, high canopy density.

References

[1] Grau, E. and Gastellu-Etchegrry, J.P. (2012). Radiative transfer modeling in the Earth Atmosphere system with DART model. Remote Sensing of the Environment, 139, 149-170

Modelling full waveform Lidar data on forest structures at plot level: a sensitivity analysis of forest and sensor main characteristics on full-waveform simulated data.

Eloi Grau[1,2], Sylvie Durrieu[2], Cécile Antin[3], Henri Debise[2], Grégoire Vincent[3], Claudia Lavalley[2], Marc Bouvier[2]

[1]CNES, UMR TETIS, 34196 Montpellier, France
[2]Irstea, UMR TETIS, 34196 Montpellier, France
[3]IRD, UMR AMAP, Montpellier, 34000 France

Highlights: Spatial full-waveform Lidar signal sensitivity to both forest and sensor characteristics was studied using DART Radiative Transfer model. The study focuses on testing various approaches to simulate forest canopies. A sensitivity study of main forest and instrumental parameters on a new set of Lidar metrics is presented.

Key words: *full-waveform, simulation, radiative transfer, sensitivity analysis*

Introduction

Light detection and ranging (Lidar) has become widely used for measuring and monitoring vegetation characteristics. Lidar sensors are mostly deployed using airborne platforms, but the Geoscience Laser Altimeter System (GLAS) instrument of the Ice, Cloud and Land Elevation Satellite (ICESat) space mission has also collected data globally from space [1]. GLAS instrument was mainly designed to measure ice elevation at global scale, thus limiting its performance for forest measurements. Satellite-based Lidar data collection offers consistent, synoptic samples of surface characteristic measurements at a global scale. Preparing a future full-waveform Lidar satellite mission designed for vegetation monitoring requires studying the impact of the instrumental and experimental conditions on the recorded signal.

Simulating Lidar data with a radiative transfer model is a powerful approach to study the link between environmental parameters (forest structure, optical properties and Leaf Area Index -LAI, ground digital elevation model -DEM, and atmosphere properties), sensor characteristics (viewing angle, footprint radius, pulse energy), and recorded waveforms. A 3D radiative transfer model can be used to control every aspect of the experimental and acquisition conditions. In such a radiative transfer model, a forest stand can be represented in a variety of way (e.g. at tree level versus plot level, with turbid medium for leaves versus facets or meshes), depending on the data used to create the simulated stand and the degree of approximation required for the simulated waveforms.

The main goal of this study was twofold. The first objective was to evaluate the impact of the stand representation type on the shape of the waveforms and to identify which stand representations can be used to provide realistic large footprint full-waveforms. To that aim several approaches to represent real forest plots were used as input scenes in the radiative transfer model and the waveforms simulated using the several scenes were compared to real Lidar data acquired on same forest plots. The second objective was to evaluate to what extent using one kind of stand representation can impact the results of a sensitivity analysis carried out in order to study the link between environmental parameters (LAI, optical properties, slope), sensor characteristics (footprint radius, scanning angle) and full waveform signals recorded by either small or large footprint Lidar (i.e. space Lidar). This study should provide insights into which stand representation is needed to conduct reliable sensitivity analyses to help designing future space-borne Lidar missions.

Materials and methods

We simulated 4 Maritime pine plots based on a field campaign conducted in the study site of Saint-Symphorien, France (44.4° N, 0.5° W). This forest area is dominated by mono-specific Maritime pine (*Pinus Pinaster*) in even aged plantations. The campaign included field measurements (diameter at breast height, total height, height of crown base of the trees in 15 m radius plots) and Terrestrial Laser Scanning (TLS). In addition, contemporaneous Aerial Laser Scanning (ALS) data were available.

The virtual forest scenes were created using 3 different approaches: (1) "Lollipop" scenes made up of ellipsoid tree crowns (where leaves are represented with turbid medium) and conical trunks with dimensions based on in-situ measurements and positioned using TLS scans, (2) "TLS" scenes made up of a 3D distribution of turbid voxels characterized by their vegetation density estimated by processing TLS data [2], and (3) "AMAP" scenes , realistic representations of stands in which tree elements are represented by triangles (meshes). These virtual realistic forest scenes were created based on the botanical knowledge of plant architectural

development and takes into account plant-to-plant interactions. The scenes were produced using the AmapSim software [3], part of AMAPstudio, a software suite for plant architecture modelling. Radiative transfer simulation was done using DART (Discrete Anisotropic Radiative Transfer) software [4], allowing an accurate and customizable modelling.

The quality of virtual scenes as regards the simulation of Lidar signal was assessed by comparing in terms of derived Lidar metrics ALS simulations obtained using the same characteristics as those of the actual ALS survey (i.e. pulse density, flight height, beam divergence) to the real ALS data. These Lidar metrics were computed on aggregated Lidar signal at plot level, in order to compare them with Space Lidar Signal (SLS) simulations. Aggregation consists in creating a single waveform corresponding to a vertical large-footprint shot from all the ALS small-footprint waveforms included in the large-footprint. The values of the aggregated waveform bins is calculated by summing the weighted numerical counts stored in the bins of the ALS georeferenced waveforms for each corresponding altitude.

Lidar metrics were computed using a new method: first, the waveform is ortho-rectified correction of the zenith angle of the Lidar), and bins are referenced in order to have the ground level (mean of the Digital Elevation Model - DEM) at height bin zero. Then the waveform is noise filtered, normalized and divided into 2 parts (ground / vegetation). Finally, Lidar metrics are computed considering 3 sets of metrics related to, respectively: (1) the vegetation waveform component, i.e. the part of the waveform corresponding to the signal backscattered by the vegetation (integral of the waveform "tVeg", heights corresponding to 3 percentiles of the vegetation waveform "rh%" and to the maximum peak intensity "altPic"), (2) the ground component, i.e. the part of the waveform corresponding to the signal backscattered by the ground (integral of the waveform "tSol") and (3) the whole waveform (normalization factor "N", extension of the waveform "ext", and vegetation peak / ground peak ratio "VGR").

The sensitivity analysis was conducted by changing the initial characteristics of the scenes (tree LAI, ground reflectance, vegetation reflectance, slope, understorey grass LAI) and Lidar acquisition parameters (scan angle, footprint radius, footprint position), independently, and measuring their impact in terms of Lidar metric changes. Comparison was performed on the three types of scenes.

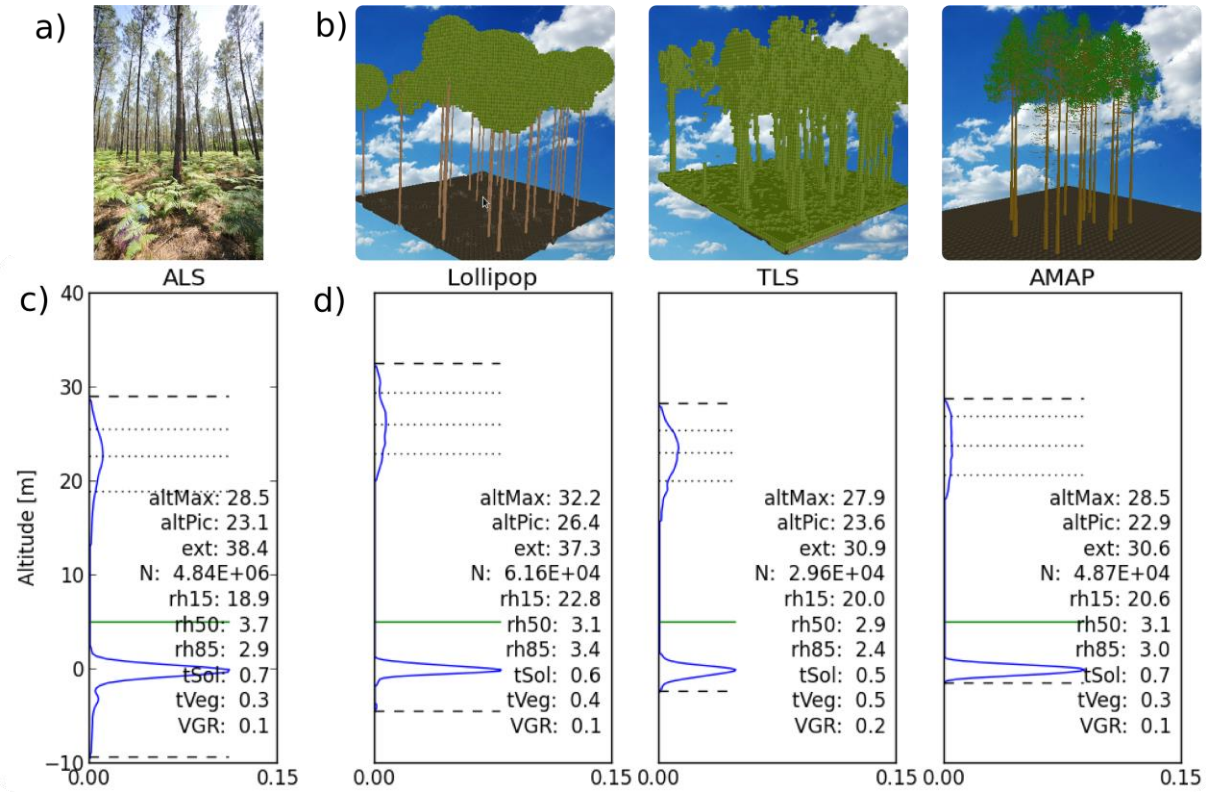


Figure 1: a) Photography of the plot. b) 3D view of the different types of mock-up used in this study for 1 forest plot. c) ALS aggregated measured waveform and associated Lidar metrics, and d) simulated waveforms for the 3 types of virtual scenes and associated Lidar metrics.

Results

Aggregation of actual ALS data allowed a comparison with ALS and SLS simulated data. Firstly, a good agreement was found for each type of virtual scenes and for each plot between Lidar metrics computed from aggregated ALS and SLS when both resulted from simulations. This can be explained by the high number of

ALS shots used to sample each plot. Secondly, the comparison of metrics computed from the waveforms resulting from the aggregation of real ALS data with metrics computed from the simulated SLS waveforms also showed good agreement, provided the virtual scenes are correctly calibrated. Calibration requirements depend on the scene type (low requirements for TLS scenes, whereas for lollipop scenes and even more for AMAP scenes adjustments of some scene characteristics were needed to obtain a good fit between simulated and real data).

Preliminary results of the sensitivity study show various effects of the forest and sensor characteristics on Lidar metrics. Among the environmental parameters, tree LAI and slope were found to be the most impacting parameters. Impact of the reflectance values (ground and vegetation) was not negligible, but affects almost only global metrics (N, VGR). Among the sensor characteristics, the impact of both the footprint radius and the position of the footprint center were limited, due to the relative homogeneity of the forest stand at the study scale. On the contrary, change in scan angle significantly impacted most Lidar metrics.

Conclusion

Different ways to simulate forest canopies used in a radiative transfer model have been tested. Each representation needs different inputs and calibration efforts, and produces virtual scenes with different level of simplification. TLS virtual scenes were the most direct and simplest representations, but they still need to be calibrated in term of total plant area density (PAD), as current voxelization process tends to overestimate PAD and cannot separate woody material from leaves when assessing PAD. Furthermore, TLS acquisitions are useful to create DEM that can be used with any representation of the vegetation. Lollipop scenes required extensive field measurements (position, dbh, crown dimensions and shape) and also produced, despite their apparent simplicity, accurate simulated waveforms. AMAP scenes were the most accurate simulation of the vegetation, but are very costly to produce in terms of field measurements, preparation, calibration, and computation time. The ongoing sensitivity study will further reveal how Lidar derived metrics are affected by vegetation and sensor.

Acknowledgements

The authors acknowledge funding from the French CNES/TOSCA program. This work also benefitted from data acquired in the frame of the FORESEE project (ANR-2010-BIOE-008) granted by the French National Research Agency (ANR).

[1] Abshire, J.B., Sun, X., Riris, H., Sirota, J.M., McGarry, J.F., Palm, S., Yi, D., Liiva, P. (2005). Geoscience Laser Altimeter System (GLAS) on the ICESat mission: on-orbit measurement performance. *Geophysical Research Letters*, 32 (21)

[2] Grau, E., Durrieu, S., Fournier, R., Gastellu-Etchegorry, J.P., Yin, T., Lauret, N., Bouvier M. (2014). Comparing voxelisation methods of 3D terrestrial laser scanning with Radiative Transfer simulation to assess vegetation density. *ForestSat2014*, oral presentation

[3] Barczi, J.-F., Rey, H., Caraglio, Y., de Reffye, P., Barthélémy, D., Dong, Q. X. & Fourcaud, T. (2008). AmapSim: A Structural Whole-Plant Simulator Based on Botanical Knowledge and Designed to Host External Functional Models. *Annals of Botany*, 101, 1125–1138.

[4] Gastellu-Etchegorry, J.-P., Yin, T. et al. (2015). "Discrete Anisotropic Radiative Transfer (DART 5) for Modeling Airborne and Satellite Spectroradiometer and LIDAR Acquisitions of Natural and Urban Landscapes. *Remote Sensing* 7(2): 1667-1701.

Development of an automated vegetation processing algorithm for ICESat-2

Ryan D. Sheridan[1], Sorin C. Popescu[1], and Ross F. Nelson[2]

[1]*LASERS Laboratory, Department of Ecosystem Science and Management, Texas A&M University, 1500 Research Parkway Suite B 217, College Station, TX, 77845, USA, ryan.sheridan@tamu.edu.*
[2] *Code 614.4-Biospheric Sciences Branch NASA/Goddard Space Flight Center, Greenbelt, MD, 20771, USA,* rfn104@gmail.com

Highlights: ICESat-2 is NASA's next generation spaceborne laser altimeter. Photon-counting lidar systems are relatively new, and their ability to produce vegetation measurements remains largely unexplored. For this study we develop a fully automated algorithm for filtering and classifying spaceborne photon-counting lidar data.

Key words: ICESat-2, Algorithm Development, Automated Processing

Introduction

ICESat-2 is NASA's next generation spaceborne laser altimeter, which aims to follow in the footsteps of ICESat. A major difference between ICESat-2 and the previous ICESat mission is the utilization of a photon counting lidar (PCL) system. Such systems are relatively new, and their ability to produce vegetation measurements remains largely unexplored. The overall objective of this study is the development fully automated algorithm for processing spaceborne PCL data.

Materials and Methods

Algorithm development is based on several sources of test data including: (1) data collected by the NASA's Multiple Beam Altimeter Experimental Laser (MABEL) system; and (2) simulated mATLAS data, which is generated from previously collected MABEL data, collected throughout the United States. Our processing algorithm can be divided into two major sections: minimizing noise in raw PCL data and classification of filtered PCL data.

Noise filtering begins by using ancillary data including a global canopy height map [1] and ASTER-derived digital elevation model are utilized in the initial filtering operations. After initial filtering with ancillary datasets is performed A Robust Adaptive Online Repeated Median (ADORE) filter is applied to the remaining to generate envelops for reducing noise remaining after initial ancillary data filtering [2]. The final filtering step is utilizes a running median filter, with the second smallest window size (window size = 5), in an attempt to remove as much of the remaining noise in the dataset as possible.

A fixed-width moving window is utilized to process the filtered data. Photon elevations within each window are utilized to calculate parameters for generating splines to estimate the top of the forest canopy (TOC) and ground surfaces. Parameters used to identify the ground or TOC are varied based on the vertical spread of photons within each moving window. The resulting splines are then buffered ± 1 m and are overlayed on the original, unfiltered data. Photons falling within the TOC buffer are classified as TOC, photons falling with the ground buffer are classified as ground, and photons between the TOC and ground buffers are classified as canopy.

Data from various airborne laser systems, including NASA's Goddard's Lidar, Hyperspectral, and Thermal Airborne Imager (G-LiHT), that are spatially coincident with MABEL and mATLAS data are utilized to assess the accuracy of algorithm results.

Results and Discussion

Early algorithm performance on MABEL data shows lower overall RMSEs for ground than TOC (3.26 m and 9.58 m, respectively). This assessment identified high TOC commission error rates in open areas (i.e. areas with little or no vegetation). In such areas, TOC commission errors were responsible for increased ground omission errors. After adjusting algorithm parameters, visual assessments of subsequent algorithm test runs suggest reductions in open area TOC commission error rates and ground omission rates. Overall, ground identification performed well in areas with dense canopy cover (Figure 1). Similar results were observed when the algorithm was run on mATLAS data. A notable exception is data exhibiting weak beam strength conditions, causing signal

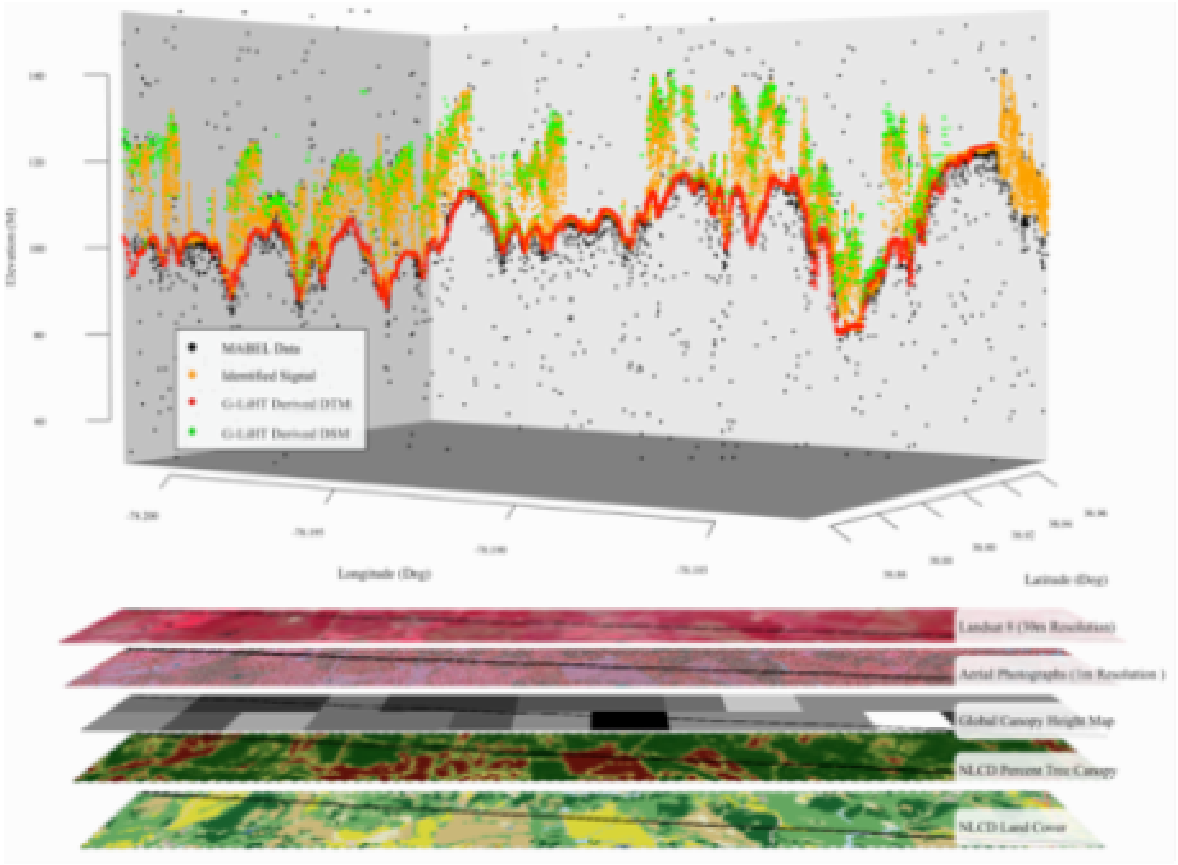


Figure 1: A visual summary of MABEL PCL data, and possible ancillary data sources useful for processing these data.

Conclusion

Our initial algorithm results are promising, and provide a base algorithm for filtering and classifying data collected by spaceborne PCL systems. The classification produced by our algorithm will provide an initial data set for ICESat-2 end users interested in monitoring forests or estimating forest biophysical parameters on a global-scale.

Future work will focuses will focus on: (1) development of methods to identify and flag signal reduction or signal loss due to clouds, haze, smoke, fog, blowing snow, etc.; (2) continued development of methods for automatically varying processing parameters based on conditions present within the data; and (3) the completion of a formal accuracy assessment for mATLAS and MABEL processing results.

References

[1] Simard, M., Pinto, N., Fisher, J.B., Baccini, A., (2011). Mapping forest canopy height globally with spaceborne lidar. J. Geophys. Res. 116, G04021.

[2] Schettlinger, K., Fried, R., Gather, U., (2010). Real-time signal processing by adaptive repeated median filters. International Journal of Adaptive Control and Signal Processing 24, 346–362.

Improved canopy height measurements of single photon lidar (SPL) using a multistage noise filtering method

Hao Tang[1], Anu Swatantran[1], Phil DeCola[2], Terence Barrett[2], Ralph Dubayah[1]

htang@umd.edu
1. University of Maryland, College Park, MD20742, USA
2. Sigma Space Corp. Lanham, MD20706, USA

Highlights: This study developed a multistage noise filtering method to improve current CHM measurements from single photon lidar (SPL) data. Accuracy of the SPL derived CHM was examined by both existing discrete return lidar and ground survey data.

Key words: *single photon lidar, noise filtering, CHM.*

Introduction

Single photon lidar (SPL) is a new technology for rapid mapping of three dimensional forest structure over large geographical areas. It requires only one detected photon of each ranging measurement to generate 3-D topographical maps whereas several thousands are needed in traditional lidar instruments. Advances in SPL technology have significantly improved the speed and coverage of data acquisition while still maintaining a high sampling density. However, data quality of SPL instrument suffers from the high level of background noise accompanied with signal lidar photons, especially over the visible band of solar spectrum. There are thus of considerable interests in developing efficient noise reduction methods and evaluating the data quality collected from SPL, in order to assess its instrumental performance on mapping canopy height model (CHM) and carbon stocks.

Material and methods

The SPL system used in this study is a altitude high-resolution quantum lidar system (HRQLS, pronounced Hercules). HRQLS was flown over the entire Garrett County of Maryland, USA in September, 2013. It produced a 5 × 5 m array of spots on the ground at a nominal above ground level (AGL) of 2.3 km, resulting in a ground-pixel dimension of 0.5 m and a mean point density of 13 per m^2 (including both signal and noise) [1]. Discrete return lidar data was acquired in 2005 during leaf-off condition with an average point density of 1 per m^2. Field data was collected during the summer of 2013 in support of a carbon monitoring system over the entire state of Maryland.

We developed a multistage automated noise filtering process to generate CHM from HRQLS. We then assessed the accuracy of HRQLS CHM with existing discrete return lidar as well as ground survey data from the field campaigns.

Results and discussions

The multistage noise filtering method can efficiently reduce background noise level and significantly improve lidar data quality. About 10% of total photons (mostly noise, from either above canopy or below ground) were filtered out from original point cloud as shown in Figure 1. However, noise photons cannot be fully eradicated, especially over boundary layers between ground (or canopy top) and the air.

Both HRQLS and DRL showed a good agreement with field measurements at plot scale (N=76) despite a slight bias in their measurements (Table 1). The negative bias of DRL is mainly due to the data acquisition over leaf-off period, while the positive bias of HRQLS is most likely introduced by unfiltered noise photons at the boundary between canopy and atmosphere.

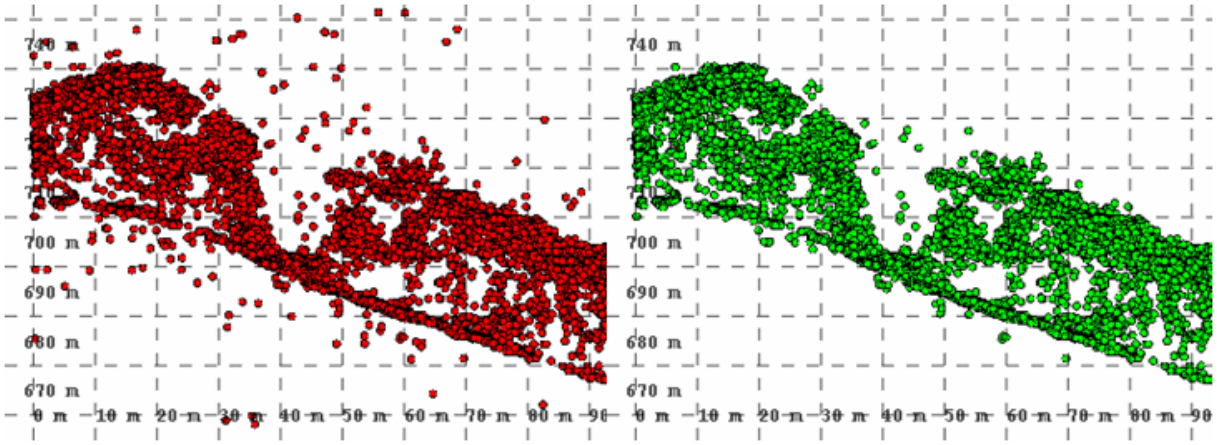


Figure 1: Example of HRQLS point cloud in a 50 m radius plot: left) before filtering; right) after filtering.

Table 1: Canopy height measurement comparisons with field data

Lidar data	R^2	Bias (m)	RMSE (m)
HRQLS	0.60	1.2	5.42
DRL	0.63	-0.26	5.37

Conclusions

This study develops an efficient noise reduction method to improve the data quality of single photon lidar. This method allows the generation of CHM from HRQLS at high resolution. Its further comparison with discrete return lidar and field data suggests high density SPL can provide accurate measurements of canopy height. Further exploration of this study can greatly improve interpretation of current SPL data, and will help inform data acquisition strategy of future SPL instruments as well (e.g. ICESat-2).

References

[1] Degnan, J. & Field, C. T.. (2010**)**. Moderate to high altitude, single photon sensitive, 3D imaging lidars, Proc. SPIE, 9114, 91140H-91140H-11.

Extraction of DTM and Top of Canopy from Photon-counting LiDAR in Preparation for the ICESat-2 Mission

Mahmoud Awadallah[1], A. Lynn Abbott[1], Randolph H. Wynne[2] and Ahmed Ghanem[3]

[1] Bradley Department of Electrical and Computer Engineering, Virginia Tech
Blacksburg, Virginia, USA, {mawadala,abbott}@vt.edu

[2] Department of Forest Resources and Environmental Conservation
Virginia Tech, Blacksburg, VA, USA, wynne@vt.edu

[3] Electrical Engineering Department, Suez Canal University
Ismailia, Egypt, amghanem@gmail.com

Highlights: This paper introduces a multistep algorithm for extracting a Digital Terrain Model (DTM) and top of canopy from photon-counting profiling LiDAR data. The approach emphasizes the identification of bare-ground areas and vegetated areas, to support estimation of canopy height. This work is motivated by the need to process noisy data from ICESat-2, which is scheduled for launch in 2017.

Keywords: *photon-counting LiDAR, DTM, DSM, active contours.*

Introduction

Concerns about global climate change have increased the importance of finding efficient and operational methods to quantify the terrestrial carbon stocks at a global scale. One vital component of the global carbon cycle is the exchange of carbon between various vegetation types (mainly forests) and the atmosphere. Global estimates of the above-ground dry biomass for different types of land cover are critically important, because they can be mapped to estimates of standing carbon [1]. In order to perform these global estimates at a reasonable cost, NASA launched the Ice, Cloud and Land Elevation Satellite (ICESat, 2003-2009). This mission led to important findings related to ecology and cryospheric sciences [2]. ICESat acquired elevation data globally using the Geoscience Laser Altimeter System (GLAS) sensor, a laser altimeter for which elevation estimates are based on the analysis of the waveform returns [3]. The follow-on mission will be ICESat-2, which is currently scheduled for launch in 2017. Unlike the first ICESat mission, ICESat-2 will employ a multibeam system that will collect elevation data using a relatively new photon-counting technology. This system, known as Advanced Topographic Laser Altimeter System (ATLAS), yields clouds of discrete points, each resulting from the return of an individual green ($\lambda = 532$ nm) photon [4].

The change in technology from the GLAS waveform system to the ATLAS photon-counting system will introduce interesting challenges in the detection of ground and vegetation. This paper introduces an algorithm for DTM and top of canopy surface extraction from ATLAS-like elevation data. Of particular interest for this paper is the detection of the ground level underneath a dense canopy. This problem is made difficult by the presence of substantial levels of ambient noise, which depends heavily on the sun angle.

Dataset and Study Area

In order to facilitate algorithm development, the ICESat-2 team collected high-altitude data using the Multiple Altimeter Beam Experimental Lidar (MABEL) [4]. Data from MABEL have been used to simulate photon point clouds for different atmospheric conditions and surface types. Recently, MABEL data have been extended to create the mATLAS dataset, which simulates more accurately ICESat-2 attributes such as larger footprint size. We have obtained experimental results using 17 granules (approximately 60 seconds each) of mATLAS altimetry data. These simulations are distributed geographically along the east and west coasts of North America.

Technical Approach

The proposed multistep algorithm starts with a joint noise reduction technique followed by a non-parametric regression method called locally weighted scatterplot smoothing (LOWESS) which produces a curve that fits the point cloud along the whole transect. This curve is used to distinguish between bare ground and vegetated areas, which are treated differently in the subsequent steps. The algorithm then uses an image processing technique called active contour models [5] to find initial estimates for DTM and top of canopy surfaces. Finally, the

algorithm uses the LOWESS curve to adjust ground surface estimation in the bare ground areas. The following paragraphs will explain the steps of the algorithm in more detail.

Mapping point clouds into a 2D image. The point clouds are first mapped into a 2D grid based on the number of points inside a window of size $w \times w$ pixels. The window is moved through the point cloud with a step size t, and the points inside the window are counted. A 2D map is created with values representing the number of points in each window. The generated image has the size (n, m) where $n = ((\max(x) - \min(x) - w) / t) + 1$ and $m = (\max(y) - \min(y) - w) / t) + 1$, where $\max(x)$, $\max(y)$, $\min(x)$, and $\min(y)$ are the maximum and the minimum values of x and y coordinates in the given point cloud.

Noise reduction using slice-based horizontal projections. This technique is based primarily on the assumption that the point density for the signal return is more than the density of noise points. It was employed in [6] as a technique for active contour initialization. The image is divided into a set of vertical slices and the horizontal projection for each slice is calculated. Then the locations with high density gradient are found as they indicate, roughly, the location of the signal photons within the slice. The area between the highest two successive gradient magnitudes is considered to be most likely to include the signal. This area is kept and all other areas are removed. The drawback of this method is that it may not be able to detect signals for steep terrain (Figure 1).

Bayesian noise reduction. Another method is proposed to remove noise from noisy point clouds using Bayesian decision theory based on the assumption of the different point densities between signal and noise. This method assumes that the noise and signal densities follow two different Poisson distributions $P(X = k) =$ ——, where k is the number of points observed in a specific window w, and λ is the average number of points in the given window. The problem of noise removal in this context can be characterized as a binary classification problem assuming that class ω_1 represents noise and class ω_2 represents signal. The Bayesian decision theory can be used to find a threshold that minimizes the probability of error using the likelihood ratio ——, where and represent the average point densities for classes ω_1 and ω_2, respectively. Each image cell is classified a signal cell if the likelihood ratio is greater than 1; otherwise this cell is considered noise. The main drawback of this method is that it removes parts of signal ground photons if it is low density (Figure 1c).

Noise reduction using the Wiener filter. The third noise reduction method is applying an adaptive Wiener filter to the image. This filter uses a pixelwise adaptive Wiener method based on statistics estimated from a local neighborhood of each pixel. The main drawback of this method is that it cannot remove portions of the very dense noise parts in the image (Figure 1d).

The joint noise reduction method. The proposed joint noise reduction method is based on combining the results of the three previously mentioned noise reduction techniques. A majority voting for every pixel is used to determine whether it represents noise or signal. Figure 1 shows the result of each noise reduction method on one mATLAS transect (Figure 1f).

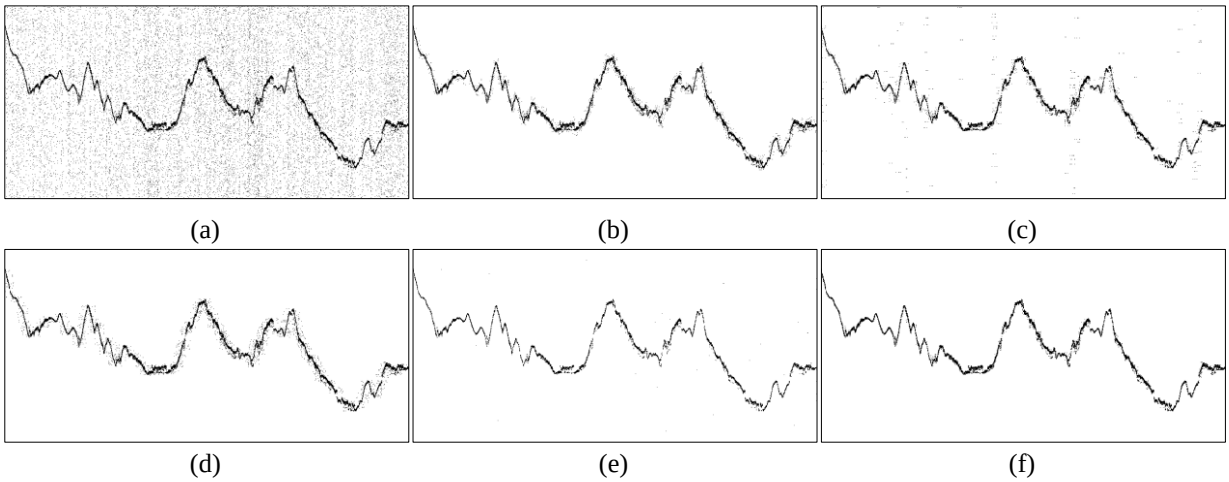


Figure 1: (a) An example of a noisy point cloud. (b) Result of the horizontal projection noise reduction, (c) Bayesian noise reduction, (d) Wiener filter noise reduction, (e) morphological noise reduction, and (f) the hybrid denoising result.

Identifying bare ground and vegetated areas. In order to find the ground surface accurately, a method of identifying bare ground areas should be used in order to treat them differently from the vegetated areas. The algorithm uses the locally weighted scatterplot smoothing (LOWESS) regression method in order produce a curve that fits the point cloud of the whole transect (Figure 2a). The perpendicular distance between the fitted LOWESS curve and each labelled signal photon is calculated, then the transect is divided into a number of 200-photon sections. If the mean absolute deviation of calculated distances in a section is less than a threshold, this area is marked as a bare ground area, and otherwise this area is marked as vegetated area (Figure 2b).

Finding initial surface estimations using active contour models. The next step is to apply an image processing technique called Active Contours Without Edges (ACWE) [5]. This method was proposed in the context of surface finding in [6]. The result of this step is the initial estimation of ground and top of canopy surfaces. This initial estimate requires further processing to be adjusted to final estimated surfaces.

Estimating the final ground and top of canopy surfaces. For each 200-photon section, a final decision is made for the surface finding algorithm according to the type of this section. If this section was considered to be bare ground, the LOWESS curve is used to classify ground photons by selecting the points within a distance of ±2 meters from the LOWESS curve. Otherwise (i.e., in vegetated area), the ground surface is adjusted for this section and the adjacent sections using spline curve fitting. The top of canopy is adjusted afterwards using the highest 95th percentile of points from the estimated ground surface (Figure 2c).

Experiments

In order to assess the accuracy of the algorithm, it was applied on 17 mATLAS transects that have become available recently. Qualitative assessment shows good estimation for both ground and top of canopy surfaces. We have not yet obtained a quantitative comparison with NASA-Goddard's LiDAR, Hyperspectral, and Thermal Imager (G-LiHT) due to the problems in geolocations between the two datasets.

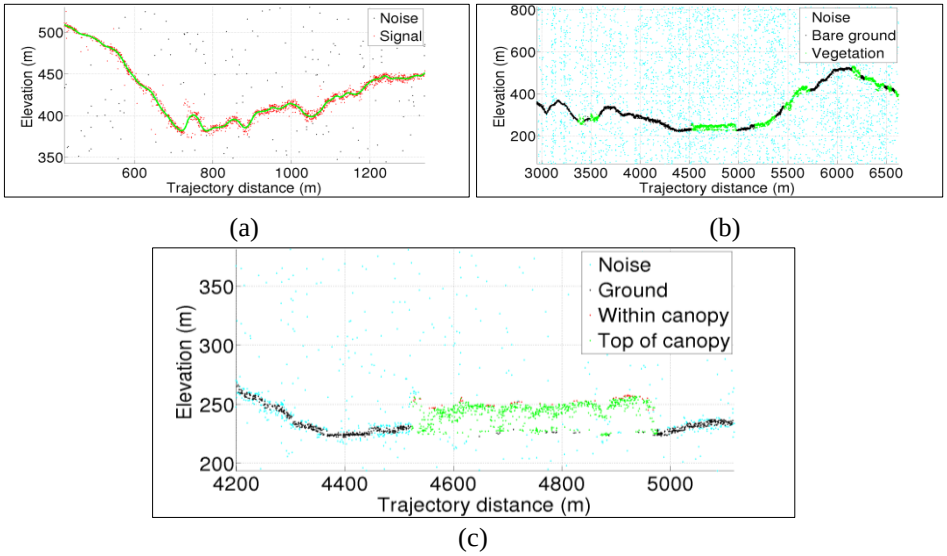


Figure 2: An example of the algorithm's intermediate steps. (a) The LOWESS curve. (b) Area classification (bare-ground vs. vegetation). (c) Close-up view of final results.

Conclusion

This paper introduces a novel multistep algorithm to extract DTM and top of canopy surfaces from photon-counting profiling LiDAR. Experimental results shows good surface estimation results despite the heavy and variable noise levels. The ground truth of this dataset is still in its development phase by the ICESat-2 team. Once it is available we will perform a quantitative assessment of the algorithm accuracy in terms of RMSE of estimated tree hights.

References

[1] R. Nelson, M. A. Valenti, A. Short, and C. Keller (2003). A multiple resource inventory of Delaware using airborne laser data. BioScience, 53(10):981-992.

[2] U. C. Herzfeld, B. W. McDonald, B. F. Wallins, T. A. Neumann, T. Markus, A. Brenner, and C. Field (2014). Algorithm for detection of ground and canopy cover in micropulse photon-counting lidar altimeter data in preparation for the ICESat-2 mission. IEEE Transactions on Geoscience and Remote Sensing, 52(4): 2109–2125.

[3] B. E. Schutz, H. J. Zwally, C. A. Shuman, D. Hancock, and J. P. DiMarzio (2005). Overview of the ICESat mission. Geophysical Research Letters, 32:L21S01.

[4] NASA ICESat-2. http://icesat.gsfc.nasa.gov/icesat2/, 2014. [Online; accessed April 2015].

[5] T. F. Chan and L. A. Vese (2001). Active contours without edges. IEEE Transactions on Image Processing, 10(2):266-277.

[6] M. Awadallah, S. Ghannam, A. L. Abbott, and A. Ghanem (2013). Active contour models for extracting ground and forest canopy curves from discrete laser altimeter data. In Proceedings: 13th International Conference on LiDAR Applications for Assessing Forest Ecosystems, pp. 129-136.

Canopy height and plant area index changes in a temperate forest between 2010–2014 using airborne laser scanning

Fabian D Schneider, Reik Leiterer, Michael E Schaepman, Felix Morsdorf

Remote Sensing Laboratories, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland.

Highlights: Changes in canopy height and plant area index (PAI) in a temperate mixed forest were assessed between 2010 and 2014 using airborne laser scanning. Patterns of canopy height change could be identified and related to forest management and tree growth. PAI changes followed no clear patterns and need further investigation.

Key words: *airborne laser scanning, forest change, canopy height, LAI, PAI, full-waveform.*

Introduction

Airborne laser scanning (ALS) is increasingly being used for forestry applications as well as ecosystem monitoring [1]. Its ability to cover large areas in a spatially continuous way offers unique possibilities to characterize canopy architecture. The fast acquisition and repeatability over time allow for monitoring changes in forested ecosystems, for instance due to forest management, forest degradation, logging, or natural changes like tree growth and tree mortality. Two of the most widely used variables to describe canopy architecture of forests are canopy height and leaf- or plant area index (PAI) [2]. Both are essential variables linked to biomass, ecosystem productivity and biodiversity. In particular, the change in canopy height can be assigned to tree growth, a major ecosystem function and important measure for timber production and increase in biomass. [3]

In this study, we assess changes in canopy height over four years in a temperate mixed forest and compared them to changes in plant area index derived from full-waveform airborne laser scanning data. We expect to detect patterns of forest management (selective logging, clear cutting) and change due to tree growth. However, it was not possible to fully link the canopy height change to PAI change (describing the density of leaves and wood). PAI values might be more heavily affected by varying acquisition geometry and point cloud processing, therefore falsely indicating change patterns due to technology and not ecology. The goal of this paper is to assess the change patterns and to attribute potential change sources: forest management, natural change, data acquisition and quality.

Study area & Data

The study area is a temperate mixed forest at Laegern, Switzerland, covering an area of roughly 2 x 2 km. The site is centered at 8.36° lon, 47.48° lat at an altitude of 680 m a.s.l. Deciduous trees are predominant (mainly *Fagus sylvatica, Fraxinus excelsior, Acer pseudoplatanus*) with patches of needle trees (mainly *Picea abies, Abies alba*). The terrain is characterized by a ridge, spanning from east to west, and a gradient of steep (>45°) to less pronounced (0-10°) slopes north and south of the ridge. The forest is semi-natural, since it is partly managed by selective cutting and natural regeneration [4].

Airborne laser scanning data were acquired on August 1, 2010 using a Riegl LMS-Q680i sensor. Data was recorded at a nominal height of 500 m above ground, resulting in a footprint size of approximately 0.25 m. In 2014, Laegern data was acquired between June 19 and July 25 as part of a larger flight campaign. The same sensor was flown (LMS-Q680i) at a nominal height of 700 m above ground, resulting in a footprint size of approximately 35 cm.

Methods

The processing of the canopy height model (CHM) and PAI values was done for both 2010 and 2014 ALS full-waveform data using the same methodology, described in detail below. Local spatial averaging was applied to reduce the influence of representation errors using a circular averaging filter with a radius of 2 pixels. Differences were subsequently calculated by subtracting the 2010 CHM and PAI values from the 2014 values respectively.

The digital terrain model (DTM) was derived from ALS ground returns, which were extracted using an adaptive multi-scale algorithm, filtered and interpolated to a 1x1 m DTM applying ordinary kriging as described in [5]. For each 1x1 m grid cell, the highest point above DTM was used to calculate the CHM.

The PAI was calculated for 2x2 m grid cells following [5]:

$$PAI = c \cdot ln\left(\frac{1 \cdot t_1 + \frac{1}{2} \cdot t_2 + \ldots + \frac{1}{n} \cdot t_n}{1 \cdot g_1 + \frac{1}{2} \cdot g_2 + \ldots + \frac{1}{n} \cdot g_n}\right)$$

where c is a calibration factor determined from in-situ PAI measurements, t_1, t_2,…, t_n are the total number of echoes for each echo type weighted by the total number of returns (*1, 2, …, n* returns) and g_1, g_2,…, g_n are the corresponding number of ground echoes respectively. The echoes for each grid cell were determined by selecting the canopy echoes first, which lie within the 2x2 m grid and more than 4 m above ground. The pulses of the canopy echoes were then tracked to the ground to find the corresponding ground echoes below 4 m above ground, no matter if they lie inside or outside the current grid cell. This leads to a reduced impact of scan angle and scan geometry compared to [5], where total number of echoes and ground echoes were selected strictly within the vertical column of each 2x2 m grid cell. Moreover, the point cloud was restricted to a maximum of 100 pulses per 2x2 m grid cell to reduce bias from extremely high pulse densities in densely vegetated areas.

Results & Discussion

Figure 1 shows the difference images of 2014 CHM and PAI minus 2010 CHM and PAI respectively. The CHM difference map shows clear patterns. The larger dark blue patches show areas of clear cutting, often connected to forest roads. There are also many areas of selective cutting, where single trees were taken out. In total, roughly 48'000 m^2 of forested area was clear-cut in the time between 2010 and 2014. Besides these clear negative change values, there are many patches showing strong increase in canopy height. Especially in the center region, many areas with more than 2 m increase in canopy height due to tree growth were identified. These areas can be assigned to fast growing rejuvenating forest. In 2010, the mean canopy height of these patches was 10.6 m ± 5.8 m.

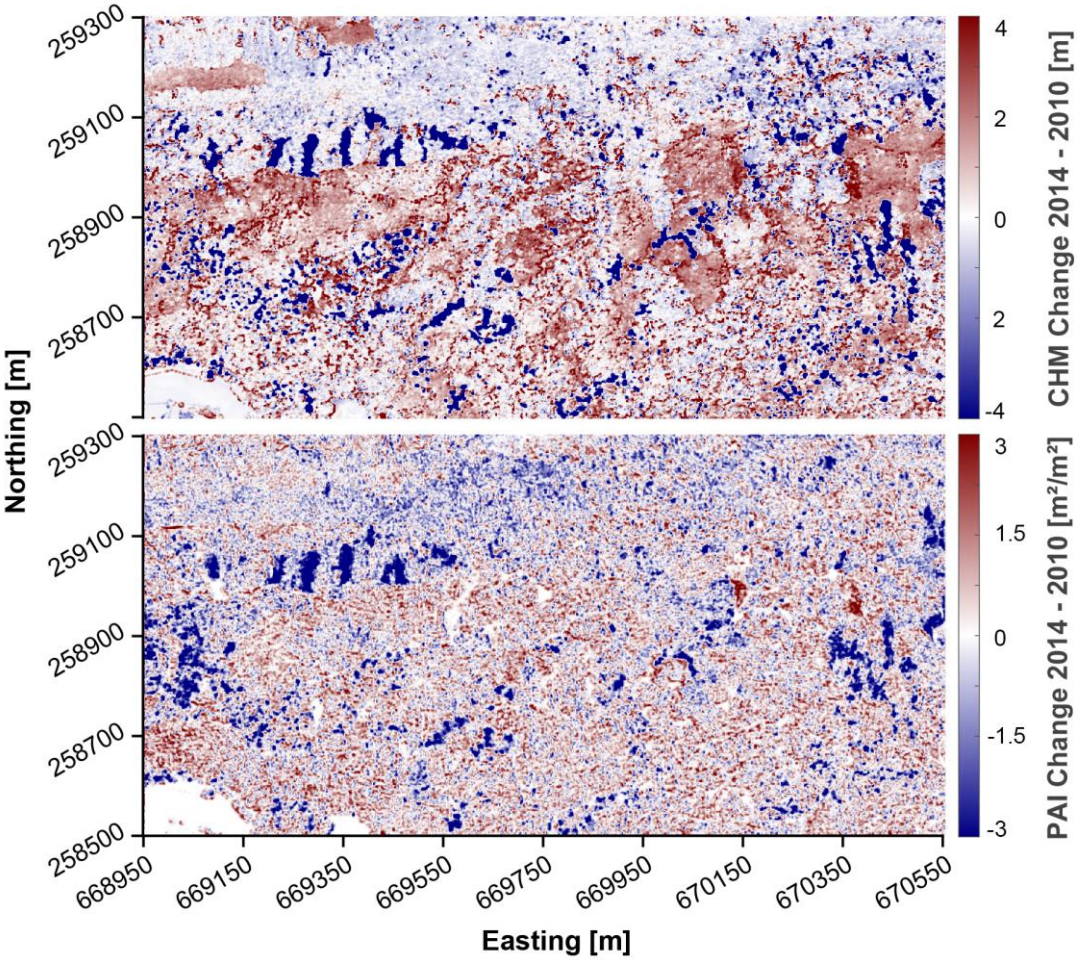


Figure 1: Changes in CHM and PAI between 2010 and 2014 of a 1600 x 800 m subset of the study area (Swiss coordinates CH1903 LV95).

There might be minor local CHM change due to small differences in geo-location of the ALS data. Both datasets have a positional accuracy of < 0.15 m in vertical and < 0.5 m in horizontal direction, as estimated from terrestrially surveyed rooftops. There was no additional co-registration performed between the two datasets. Another reason for the local CHM differences might be representation errors, where the same 3D objects are scanned from two different angles. This can cause CHM differences, which are not linked with the positional accuracy of the datasets. Local spatial averaging helped to reduce these effects, which resulted in a clearer change pattern (Figure 1).

There is a slight underestimation of canopy height in the 2014 dataset compared to 2010. In 2014, the canopy top was probably not captured as reliably as in 2010 due to lower pulse density and higher flight altitude. As reported in [6], an increase in tree height underestimation with increasing flight altitude can be caused by missing the tree tops due to lower pulse density or lower energy being reflected from the canopy surface. Differences in the DTM can represent another source of error, especially in areas of dense understory vegetation. Within the Laegern study area, the mean difference between the DTM of 2010 and 2014 is 0.3 m ± 0.2 m.

The PAI difference map in Figure 1 is more difficult to interpret. The areas of clear cutting and selective cutting can be seen as well. However, all other parts are rather noisy. Some patches with an increase in canopy height also show an increase in PAI, but not all patterns of CHM and PAI change are consistent. In some parts, PAI might already be saturated and not showing an increase in PAI where there is still an increase in CHM observed. When comparing PAI values derived from different acquisitions, representation errors might have a stronger influence than for CHM comparisons. The two datasets were acquired with a different scan geometry: One was flown in east-west direction, whereas the other was flown in north-south direction. So depending on the local observation angle, different parts of the crowns and the ground were observed.

There is no obvious pattern related to pulse density variation, but the pulse density can influence the PAI retrieval in very densely forested areas with few or no ground returns.

Conclusion

We applied CHM and PAI retrieval algorithms on ALS full-waveform data of 2010 and 2014 to compare changes in canopy height and density. Possible sources of the main canopy height changes are forest management (clear cutting, selective cutting) and tree growth (young fast growing trees). Representation errors, differences in data acquisition and geo-correction may lead to a slightly noisy pattern and potential underestimation of canopy height in 2014 compared to 2010. Except for the selective cutting, there are no clear trends in plant area index change. The PAI change patterns are noisier and partly inconsistent with CHM change. Further studies are needed to investigate the influence of pulse density and flight geometry on the PAI retrieval.

Acknowledgements

This study has been supported by the University of Zurich Research Priority Program on ‘Global Change and Biodiversity’ (URPP GCB).

References

[1] Maltamo, M., Næsset, E., & Vaukhonen, J. (2014). Forestry Applications of Airborne Laser Scanning: Concepts and Case Studies. (Managing Forest Ecosystems, 27). Dordrecht: Springer.

[2] van Leeuwen, M., & Nieuwenhuis, M. (2010). Retrieval of forest structural parameters using LiDAR remote sensing. European Journal of Forest Research, 129, 749-770.

[3] Ishii, H.T., Tanabe, S.I., & Hiura, T. (2004). Exploring the relationships among canopy structure, stand productivity, and biodiversity of temperate forest ecosystems. Forest Science, 50, 342-355.

[4] Eugster, W., Zeyer, K., Zeeman, M., Michna, P., Zingg, A., Buchmann, N., & Emmenegger, L. (2007). Methodological study of nitrous oxide eddy covariance measurements using quantum cascade laser spectrometery over a Swiss forest. Biogeosciences, 4, 927-939.

[5] Schneider, F.D., Leiterer, R., Morsdorf, F., Gastellu-Etchegorry, J.-P., Lauret, N., Pfeifer, N., & Schaepman, M.E. (2014). Simulating imaging spectrometer data: 3D forest modeling based on LiDAR and in situ data. Remote Sensing of Environment, 152, 235-250.

[6] Morsdorf, F., Frey, O., Meier, E., Itten, K.I., & Allgöwer, B. (2008). Assessment of the influence of flying altitude and scan angle on biophysical vegetation products derived from airborne laser scanning. International Journal of Remote Sensing, 29, 1387-1406.

Site index assessment based on multi-temporal ALS data

Markus Hollaus[1], Lothar Eysn[1], Bernhard Maier[2] and Norbert Pfeifer[1]

[1]*Vienna University of Technology, Department of Geodesy and Geoinformation, Research Group Photogrammetry, Gußhausstraße 27-29, A-1040 Vienna, Austria (Markus.Hollaus, Lothar.Eysn, Norbert.Pfeifer@geo.tuwien.ac.at*
[2]*Stand Montafon Forstfonds, Montafonerstraße 21, 6780 Schruns, Austria (bernhard.maier@stand-montafon.at)*

Highlights: Based on multi-temporal ALS data the site index of forest stands was assessed for a mountainous study area in Austria. The applied approach is based on derived tree top heights. The promising results show good agreements with forest inventory data and provide new insights in forest productivity analyses.

Key words: *Forest growth; site condition; site class; LiDAR; productivity index; top height.*

Introduction

For estimating the site index of forest stands different approaches can be found in literature. The most important factors affecting the site index are tree species, tree growth, soil properties as well as climatic and topographic conditions. For operational site index estimation data from forest inventories (FI) represents an essential data source. Even though the measurements within FI sample plots are accurate, a high degree of uncertainty is still available in the individual input parameters. For example, exact soil maps with high spatial resolution are often not available for forested areas. Furthermore the delineation of forest stands has often a background in historic forest management activities and the delineated forest stands are actually characterized by varying growth conditions.

The increasing availability of 3D data derived from Airborne Laserscanning (ALS) and image matching opens new opportunities for site index estimation for large areas. Especially the availability of multi-temporal 3D information of the forest canopy allows the estimation of area-wide forest growth, which is directly connected to the site index.

The objective of this study is to analyse the potential of multi-temporal ALS data for site index estimation. The study is done for a mountainous test site in the Western part of Austria, where two ALS data sets are available from 2004 and 2011. In addition to the ALS data FI data are available which are used for validation.

Study area and data

The study area is located in the southern part of the federal state of Vorarlberg, located in the western part of Austria. The elevations range between 800 m above sea level (a.s.l.) in the valleys to a maximum of 2 900 m a.s.l., whereas the timberline is at about 1 950 m a.s.l. The available forests are dominated by Norway spruce (*Picea abies*) and are managed by the local forest administration “Stand Montafon Forstfonds”. Further details of the study area can be found in Hollaus et al. [1].

The FI data from two acquisition times (2002 & 2011) are provided from the “Stand Montafon Forstfonds” and consist of permanent angle gauge sampling plots distributed in a regular grid of 350 m. The statistical analyses as well as the site index calculations based on the FI data were done by experts from the Institute of Forest Growth from the University of Natural Resources and Life Sciences, Vienna. Due to spatial inaccuracies of the plot positions in relation to the ALS data the FI data are co-registered to the ALS data using the approach described in Dorigo et al. [2].

For the study area georeferenced ALS data from 2004 and 2011 are available. The data were provided by the local office of Surveying and Geoinformation. The data were acquired mainly under snow-free and leaf-off conditions during several flight-campaigns using different ALS scanners (i.e. Airborne Laser Terrain Mapper systems ALTM 1225 and ALTM 2050 for the acquisitions in 2004 and a Trimble Harrier 56 and an Optech ALTM Gemini system for the acquisitions in 2011). Further details on the sensor characteristics can be found in Hollaus et al. [3].

Methods

From the georeferenced 3D point clouds digital surface models (DSMs) are calculated based on a land cover dependent approach described in Hollaus et al. [4]. This method uses the strengths of different algorithms for generating the final DSM by using surface roughness information to combine two DSMs, which are calculated based on the highest echo within a raster cell and on moving least squares interpolation with a plane as

functional model. Based on the calculated DSMs a quality check is done by calculating the differences of the DSMs for smooth surfaces, which have not changed their heights between the acquisition times (e.g. open land). As for site index estimation the forest growth is an essential quantity, the observed DSM differences in the range of ~0.2 m are minimized by applying a least square matching (LSM) of the DSMs. The LSM was done for each tile (2.5 x 2.5 km) separately. The derived 3D shift is applied to the ALS data from the first acquisition.

At the Stand Montafon the operational site index estimation is based on the FI data using the DGZ_{100} values. The following formula is used for calculating DGZ_{100}:

$$DGZ_{100} = a_1 * ho + a_2 * ih_{10} + a_3 * ho^2 + a_4 * ih_{10}{}^2 + a_5 * ho * ih_{10}$$

where DGZ_{100} is the mean total volume increment at the age of 100 years, *ho* is the top height in meter from the first acquisition and ih_{10} is the height growth of the top height in ten years in meter. The coefficients are provided by experts from the Institute of Forest Growth from the University of Natural Resources and Life Sciences, Vienna.

In this study the *ho* and the ih_{10} values are extracted from the multi-temporal ALS data. Common approaches to assess the *ho* values from ALS data are either to use a height percentile of the canopy height model or to identify individual tree heights and to average the heights of the 100 tallest tree per hectare. For the current study, the single tree approach is selected mainly due to the fact, that harvested trees can be easily considered. A further reason for this decision is, that the ALS point densities are different in the two ALS data sets, which would lead to different representations of the canopy surface and consequently to biased *ho* estimations. Therefore, single trees are extracted in the new ALS data set and for each individual its corresponding tree is identified in the old ALS data set. For the single tree detection, a local maxima filter with a radius of 1.5 m is applied to the normalized point cloud. For each pair of identified trees, the heights and their changes are extracted from the normalized ALS point clouds. For each tree, the *ho* and the ih_{10} values are calculated by averaging the extracted tree heights of the 100 highest trees in a one hectare circular surrounding. Then, the DGZ_{100} values are calculated based on the above formula. Finally, the calculated DGZ_{100} values for individual trees are averaged for forest stand polygons.

The derived results are quantitatively validated against the available FI data by comparing the estimated *ho* as well as DGZ_{100} values for the individual FI plots. Furthermore, a qualitative validation will be performed by local experts based on the averaged DGZ_{100} for the available forest stand polygons.

Results and Discussion

The geometric quality analyses of the ALS derived DSMs have shown that an improvement of the georeferencing accuracy is required. Based on the adjustments of the old DSMs by applying 3D shifts, the height differences for stable objects could be minimized to less than 0.10 m. Consequently the height changes in the two ALS data sets can therefore be related to forest growth with a high degree of reliability.

The applied single tree detection algorithm leads to satisfying results, because the forest is dominated by Norway spruce, which is known from literature to be well suited for single tree detection. Furthermore, only dominant trees are required for calculating the top height and therefore missing trees in the understory do not influence the calculation of *ho*. The identification of tree pairs from the dominant canopy layer in the multi-temporal ALS data works properly, although the point densities vary between 4 and 20 pts/m². As the single tree detection and the tree height estimation was done in the normalized point cloud no smoothing effects due to surface interpolation distort the extracted tree heights. Nevertheless, several factors can still affect the estimated tree height accuracy, which need to be investigated in more detail in further analyses.

In Figure 1a individual trees are shown, that could be found in both ALS data sets. The colours show the height differences, which represents the forest growth between 2005 and 2011. It can clearly be seen, that with increasing elevation, the forest growth decreases. Furthermore, different topographic conditions (i.e. aspect, slope, etc.), which are important factors influencing the forest productivity, are reflected in the extracted tree growth values. In comparison to the derived site index values derived from the estimated *ho* and ih_{10} values (Figure 1b), the individual tree growth values are characterized by a higher variability even for a homogenous stand where a homogenous site index can be assumed. In figure 1c, the DGZ_{100} values are shown for forest stands.

The quantitative validation of the estimated *ho* values shows a good average agreement with the in-situ measured heights of the sample trees with the 50% mark in the basal area distribution (=central stem), but have a relatively high standard deviation. This can be explained by the fact, that for operational applications and due to constraints of the available FI data it is assumed, that the *ho* is appropriately represented by the height of the central stem. The qualitative validation of the assessed DGZ_{100} values will be done by local foresters in the next months after the snow smelting period.

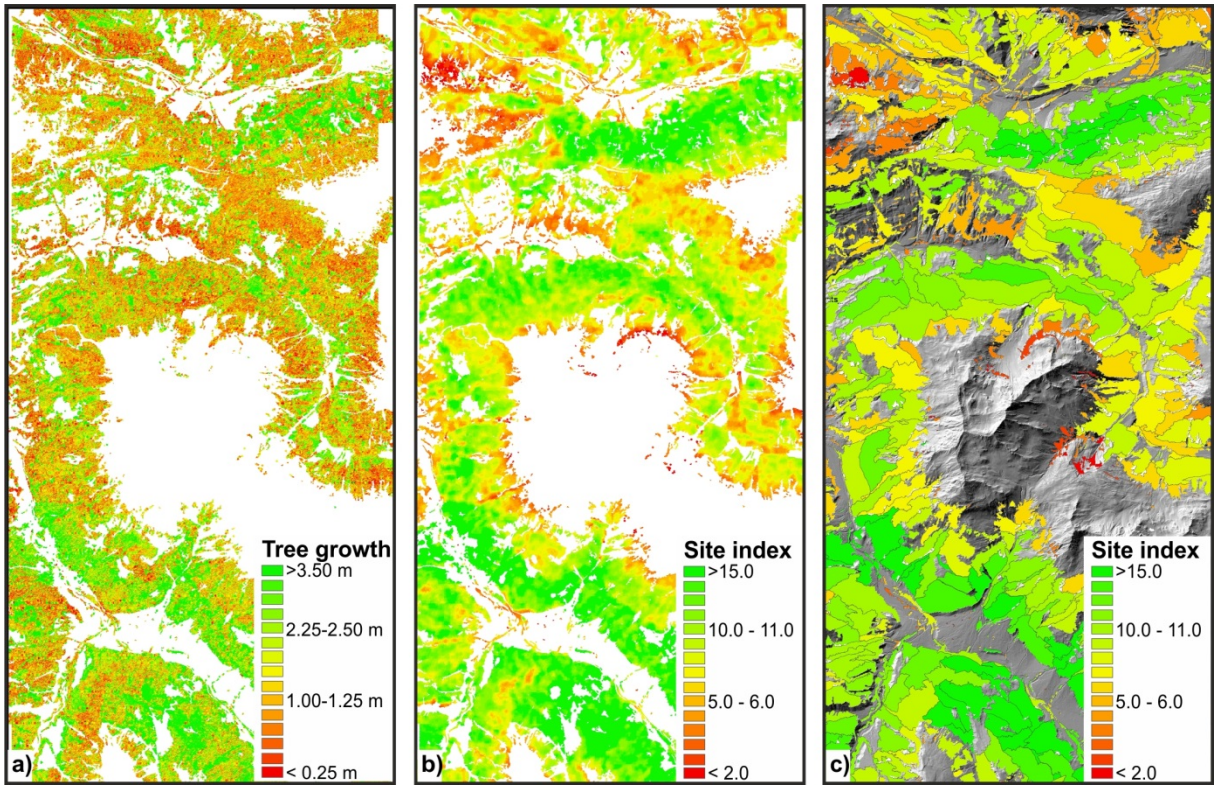


Figure 1: a) Individual trees that could be identified in both ALS data sets colorized by individual tree height differences, b) individual trees colorized by calculated site index values based on the estimated top height values (*ho* & ih_{10}) and c) forest stands colorized by the mean site index; in the background a DTM shading is shown.

Conclusion

This study shows the high potential of multi-epoch ALS data for site index assessment. It could be shown, that advanced georeferencing approaches are required to ensure, that height changes of the ALS data can be related to tree growth, which is a fundamental input information for site index estimation. As the available ALS data sets were acquired by different ALS sensors with varying properties (i.e. flying height, incident angles, point densities, acquisition time), individual tree crown shapes will be represented in the ALS data differently. Furthermore, harvesting of individual trees took place between the two ALS acquisition times which has to be considered. The selected tree based approach, where only clearly identifiable individual trees are considered, can reduce several of these influences. The first results of this study are promising and provide a good overview regarding the site index of the entire study area.

Acknowledgements

This work is founded from the European Community's Seventh Framework Programme (FP7/2007–2013) under grant agreement No. 606971, the Advanced_SAR project. The ALS data are provided by the local office of Survey and Geoinformation of the federal state Vorarlberg, Austria.

References

1. Hollaus, M.; Wagner, W.; Schadauer, K.; Maier, B.; Gabler, K. Growing stock estimation for alpine forests in austria: A robust lidar-based approach. *Canadian Journal of Forest Research* **2009**, *39*, 1387-1400.
2. Dorigo, W.; Hollaus, M.; Schadauer, K.; Wagner, W. An application-oriented automated approach for co-registration of forest inventory and airborne lasescanning data. *International Journal of Remote Sensing* **2010**, *31*, 1133-1153.
3. Hollaus, M.; Eysn, L.; Karel, W.; Pfeifer, N. In *Growing stock change estimation using airborne laser scanning data*, 13th International Conference on LiDAR Applications for Assessing Forest Ecosystems (SilviLaser 2013), Beijing, China, 2013-10-09 - 2013-10-11, 2013; Beijing, China, p 8.
4. Hollaus, M.; Mandlburger, G.; Pfeifer, N.; Mücke, W. Land cover dependent derivation of digital surface models from airborne laser scanning data. In *ISPRS Commission III Symposium PCV 2010 -- Photogrammetric Computer Vision and Image Analysis*, Paris, 2010; Vol. IAPRS Volume XXXVIII Part 3A, pp 221-226.

Towards automated characterization of horizontal and vertical forest structure using multi-seasonal airborne laser scanning

Reik Leiterer, Michael E. Schaepman, Felix Morsdorf

Remote Sensing Laboratories, University of Zurich, Winterthurerstrasse 190, CH-8057, Zurich, Switzerland.

Highlights: We present a method to characterize the vertical layering of forests in space and time based on vertical echo distributions from airborne laser scanning. We further demonstrate successful scaling from local to regional areas, including assessment of transferability, robustness and operational use of the method.

Key words: *Canopy, Large-scale, Automatic, Forestry, Operational, ALS*

Introduction

Forests cover approximately one third of Earth's total land area and are one of the most biologically diverse terrestrial ecosystems on Earth [1]. Moreover, they form habitats maintaining the majority of terrestrial biodiversity and provide a wide range of valuable ecosystem goods and services, including food, timber, and climate moderation. Understanding and monitoring forest ecosystems and their underlying processes allows the projection of biogeochemical and -physical cycles under e.g. changing climate conditions, and is essential for supporting forest management, conservation biology and ecological restoration [2]. The forest structure is considered a particularly crucial constituent of forest ecosystems functioning and processes as it e.g. influences the energy fluxes between the atmosphere and forests, serves as an indicator to determine forest stand resistance to disturbances, enables the estimation of the conservation potential for biodiversity, or allows the identification of recruitment limitations [3]. Airborne laser scanning (ALS) systems are suitable for providing not only horizontal information about the forest structure, but also detailed vertical information based on the physical measurement principles of active sensing and canopy gap fraction [4]. Forest structure information includes geometric variables such as variation in the canopy height and canopy volume, as well as biophysical variables such as fractional canopy cover [5]. Existing approaches for forest structure characterization often require a large amount of prior information at a pre-defined spatial scale and/or rely on manual processing steps, which again require prior information about stand characteristics. Therefore, most of these approaches are limited in their transferability to other sites due to necessary local calibration of the applied models, and they tend not to be directly comparable.

To overcome these limitations, we developed an automated and transferable grid-based approach to provide quantitative descriptions of canopy structure at different scales, where retrievals are based solely on ALS data [6]. Previously developed and tested on a small patch of ALS data (800 ha) from 2010, we transferred the developed method to a much larger scale (180'000 ha) using a different ALS data set in order to prove the transferability, robustness and near-operational use of the developed approach. The area is subject to different forest management practices (from natural situations to highly intensive regimes with silvicultural interventions) and contains a variety of forest types.

Study area and data

The ALS data were acquired in 2014 for the Kanton Aargau, Switzerland, under defoliated (or leaf-off) and foliated (or leaf-on) conditions, covering an area of approximately 1800 km2 (approx. 20 and 40 echoes/m2 for forested areas, respectively). A three-dimensional point cloud was obtained composed of planimetric coordinates (x and y), ellipsoidal heights (z) and the echo type, as well as the physical characteristics of the reflecting object (echo amplitude and width). The topographically corrected point cloud was derived by extracting ground returns from the point cloud and calculating the digital terrain model using an ordinary kriging method [6]. The height aboveground was then calculated for each echo of the point cloud by subtracting the interpolated DTM value at the corresponding horizontal echo location.

Forest inventory data were collected with forest structure information, such as tree species, tree height, forest cover or canopy stratification (e.g., number of canopy layers) as well as information about ground cover and understory vegetation [7]. The forest inventory data were used for validation only.

Methods

One of the essential features in the applied method is the histogram of the vertical echo heights (percentage of echoes per vertical bin) within a given horizontal grid cell. This histogram can be interpreted as a kind of synthetic waveform reflecting various forest structure features.

We derived the histograms both, on a small scale (1x1m) to take account of small-scale variations in the forest structure and on a coarser scale (20x20m) according to the size of the forest inventory plots. The histograms were calculated for the leaf-off and the leaf-on data independently as well as for the combined data. For the histogram based on the combined data, we calculated the following ALS metrics on the coarse scale: i) $canopy_{var}$ - the small-scale variation of the histograms within the coarse grid cell; and ii) $canopy_{lay}$ – the amount and the extent of vertical canopy layers, whereby a canopy layer was considered to have a minimum vertical extent of ≥3m and the spacing between two layers needs to be ≥3m, following the forest inventory specifications. $Canopy_{lay}$ was classified afterwards according to the amount of layers into '1-layered', '2-layered' and 'multi-layered' canopies.

Using the leaf-off and the leaf-on histograms, we calculated $canopy_{diff}$ – the differences in the percentage of points in the top canopy layer, whereby the extent of the top canopy layer was indicated by $canopy_{lay}$. Based on the leaf-on data only, we calculated $canopy_{cov}$ – the ratio between echoes from the forest ground/ forest floor and echoes from the forest canopy above, as a proxy for the percentage forest cover. $Canopy_{var}$ and $canopy_{diff}$ were afterwards classified into significant variations/differences and non-significant variations/differences using the *p*-value with a significance level set to $p \leq 0.05$. Significant changes in $canopy_{var}$ means a high variability in the vertical structure on a small-scale and thus a less good representation of the vertical structure in the coarse-scale histograms in contrast to a homogenous vertical structure, indicates by non-significant variations. Significant differences in $canopy_{diff}$ indicate deciduous vegetation, whereas non-significant differences indicate evergreen vegetation. Table 1 gives an overview about the resulting forest structure classes.

Table 1: Characterization of forest structure according the calculated ALS metrics.

$canopy_{diff}$	$canopy_{lay}$	$canopy_{var}$	$canopy_{cov}$
deciduous	1-layered 2-layered multi-layered	high variability of vertical structure vs. homogeneous vertical structure	continuous classification (1-100%)
evergreen	1-layered 2-layered multi-layered		

Each of the coarse-scale grid cells therefore contain information about the vertical structure and the seasonal variations of the canopy as well as information about the spatial variability of the canopy structure within the coarse-scale grid cell. For the practical application, we finally segmented the structural classes with eCognition into structural homogenous patches, which are comparable to the conventional stands used for forest management in the region.

Results

The validation shows promising results for the determined forest structure classes, particularly in terms of the seasonal ($canopy_{diff}$) and horizontal variation ($canopy_{var}$) in the vertical canopy structure. The overall accuracy for the $canopy_{var}$-based classification was 83%, whereby in average the classification performs better for the evergreen classes. The classification into deciduous and evergreen vegetation resulted in a high overall accuracy of 91%. The estimated forest cover $canopy_{cov}$ shows with $r^2 = 0.76$ a good correlation to the field data. For the vertical stratification of the canopy ($canopy_{lay}$) we reached an overall accuracy of 69%, whereby the lowest accuracy was determined for the multi-layered deciduous forest and the highest accuracies for the 1- and 2-layered evergreen forests. For more developed forest stands (mean tree heights > 30 m), the different forest management strategies are partially reflected in the specific forest structure. In addition, the effect of different forest types (i.e. mixed stands or pure stands) correlates well with $canopy_{var}$ and $canopy_{lay}$. These results nicely agree with the results we achieved with the data from 2010 on the small area and show the transferability of the developed method [6].

Discussion and conclusion

The proposed forest canopy structure characterization improves existing structure classification approaches as it adds a more detailed description of the vertical stratification of the forest canopy and their horizontal variability. For example, the consideration of the small scale variations of vertical canopy structure is an important indicator in terms of biodiversity estimations or habitat assessments. An advantage of the proposed method is the robust development and we showed, that it can easily be adapted and harmonized to different scales as well as applied to different ALS data sets with varying data characteristics.

The validation shows satisfying results, particularly in terms of the seasonal and horizontal variety of the vertical canopy structure. In general, the transition zones between forested and non-forested areas as well as between the specific forest types are problematic, as it is always difficult to determine borders of discrete classes while looking at a continuous, natural feature space. However, the accuracy assessment of $canopy_{lay}$ turned out to be very difficult. The subjective visual evaluation of the canopy stratification by the forestry experts includes a source of error and thus cannot be regarded as an error-free reference, but rather as source for a type of cross-comparison. Additionally, the canopy stratification approaches used in the forest inventory are more related to the composition of different tree development stages and less focused on the actual vertical foliage distribution. For example, received echoes from the vegetation in the lower canopy parts can be either from the forest floor, forest succession or caused by low branches of old growth trees. In terms of the ALS based canopy stratification, we are not able to differentiate between those different sources even if it would be addressed in different ways based on methods of forest inventory or forest management.

We conclude that our method can substantially improve the robustness and reliability of ALS based forest structure characterization and enables an efficient monitoring of forest structure. The proposed approach meets forest management requirements in terms of accuracy, spatial resolution and information content, and has the potential to be applied in a semi-automated fashion in the near future. The next steps will be to transfer the method to a variety of forests (boreal/tropical) to ensure unambiguous use of the application. Further, we plan to investigate the relationship and usability of the derived forest structure classes with different established forest ecosystem goods and services, such as biodiversity, forest stand resistance to disturbances and stand productivity.

References

[1] Pan, Y., Birdsey, R.A., Phillips, O.L., & B. Robert (2013): The Structure, Distribution, and Biomass of the World's Forests. Annual Review of Ecology, Evolution, and Systematics, 44, 593-622.

[2] Bonan, G.B. (2008). Forests and climate change: Forcings, feedbacks, and the climate benefits of forests. Science, 320 (5882), 1444-1449.

[3] Shugart, H.H., Saatchi, S., & F.G. Hall (2010). Importance of structure and its measurement in quantifying function of forest ecosystems. Journal of Geophysical Research, 115 (4), 1-16.

[4] Wulder, M.A., J.C. White, R.F. Nelson, E. Næsset, H.O. Ørka, N.C. Coops, T. Hilker, C.W. Bater, & T. Gobakken (2012). LiDAR sampling for large-area forest characterization: A review. Remote Sensing of Environment, 121, 196-209.

[5] Morsdorf, F., Mårell, A., Koetz, B., Cassagne, N., Pimont, F., Rigolot, E. & B. Allgöwer (2010). Discrimination of vegetation strata in a multi-layered Mediterranean forest ecosystem using height and intensity information derived from airborne laser scanning. Remote Sensing of Environment, 114, 14031415.

[6] Leiterer, R., Morsdorf, F., Torabzadeh, H., Schaepman, M.E., Muecke, W., Pfeifer, N., & M. Hollaus (2012). A voxel-based approach for canopy structure characterization using full-waveform airborne laser scanning. International Geoscience and Remote Sensing Symposium (IGARSS), art. no. 6350691, 3399-3402.

[7] Schneider, F.D., Leiterer, R., Morsdorf, F., Gastellu-Etchegorry, J.-P., Lauret, N., Pfeifer, N., & M.E. Schaepman (2014). Simulating imaging spectrometer data: 3D forest modeling based on LiDAR and in situ data. Remote Sensing of Environment, 152, 235-250.

A Bayesian hierarchical model for spatio-temporal prediction and uncertainty assessment using repeat LiDAR acquisitions for the Kenai Peninsula, AK, USA

Chad Babcock[1], Hans-Erik Andersen[2], Andrew O. Finley[3], and Bruce D. Cook[4]

[1]*School of Environmental and Forest Sciences, University of Washington, Seattle, WA, USA. babcoc76@uw.edu*
[2]*USDA Forest Service Pacific Northwest Research Station, University of Washington, Seattle, WA, USA. handersen@fs.fed.us*
[3]*Department of Forestry, Michigan State University, East Lansing, MI, USA. finleya@msu.edu*
[4]*Code 618, Biospheric Sciences Branch, NASA/Goddard Space Flight Center, Greenbelt, MD, USA. bruce.cook@nasa.gov*

Highlights: Models leveraging repeat LiDAR and field collection campaigns may be one possible mechanism to monitor carbon flux in remote forested regions. Here, we look to the spatio-temporally data-rich Kenai Peninsula in Alaska, USA to examine the potential for Bayesian spatio-temporal mapping of terrestrial forest carbon storage and uncertainty.

Key words: *multi-temporal LiDAR, carbon, prediction uncertainty, temporal misalignment*

Introduction and motivations

Models leveraging repeat LiDAR and field collection campaigns may be one possible mechanism to monitor carbon flux in remote forested regions. Hopkinson et al. showed that it is possible to detect growth using repeated LiDAR collections for a pine plantation in Ontario, Canada[1]. Plot-level field measures of forest height increment paralleled corresponding changes in LiDAR height, indicating that it may be possible to track forest growth by monitoring change in LiDAR metrics from year to year. Yu et al. constructed difference metrics by subtracting LiDAR variables derived from repeat LiDAR collections[2]. Using these as predictor variables in a linear regression, they showed that forest height and volume increment can be predicted moderately well in a boreal forest in Finland. Hudak et al. developed biomass maps for two LiDAR acquisitions over a mixed conifer forest in Idaho, USA and showed that, even when LiDAR point densities differed dramatically between the datasets, it was possible to estimate change in biomass by subtracting the two predicted maps[3].

Considering the demonstrated ability of multi-temporal LiDAR to estimate forest growth, it is not surprising that there is great interest in developing forest carbon monitoring strategies that rely on repeated LiDAR acquisitions for remote areas. Allowing for sparser field campaigns, LiDAR stands to make monitoring forest carbon cheaper and more efficient than field-only sampling procedures. There are issues concerning the few procedures currently proposed to assess growth, and subsequently carbon flux, using multi-temporal LiDAR though. Most linear regression approaches implicitly assume that remote sensing and field data are collected in the same season, which is problematic for most large-area field inventory campaigns. To increase spatial field sampling coverage without incurring extra cost, typically, only portions of the network of permanent sample plots are remeasured each year. Since the LiDAR and field data need to be coincident for proper calibration, the researcher is forced to discard all inventory data from other years or otherwise assume temporal misalignments to be negligible. We need methods capable of using temporally disjointed data appropriately. Further, subtracting maps of separately predicted forest biomass does not allow for prediction uncertainty to be properly carried through to the estimation of carbon flux. Without an accurate and useful assessment of carbon flux uncertainty, we have no way of understanding if predicted values are reliable.

Here, we look to the spatio-temporally data-rich Kenai Peninsula in Alaska, USA to examine the potential for Bayesian spatio-temporal mapping of terrestrial forest carbon storage and uncertainty. The modeling framework explored here can predict forest carbon through space and time, while formally propagating uncertainty through to prediction. Bayesian spatio-temporal models are flexible frameworks allowing for forest growth processes to be formally integrated into the model. By incorporating a mechanism for growth---using temporally repeated field and LiDAR data---we can more fully exploit the information-rich field inventory network to improve prediction accuracy.

Motivating dataset description

LiDAR data for the Kenai Peninsula has been collected on four different occasions---spatially coincident LiDAR strip samples in 2004, 09 and 14, along with a wall-to-wall collection in 2008 (Table 1). There were 436 inventory locations measured at least twice between 2002 and 2014 (Figure 1). Plot locations exhibit a clustered configuration of up to 4 plots within a cluster (Figure 1, inset). Inventory data was collected according to the US Forest Service Forest Inventory Analysis plot design[4]. LiDAR information was acquired at least once over most of the inventory plots with many having LiDAR collected during 2, 3 or 4 different campaigns.

Table 1: LiDAR acquisitions

Acquisition year	Strip sampling / wall to wall	Strip spacing, strip width	Average point density
2004	strip sampling	10 km, 300 m	7.88 points/m^2
2008	wall to wall	NA	1.81 points/m^2
2009	strip sampling	10 km, 300 m	4.13 points/m^2
2014	strip sampling	10 km, 300 m	7.38 points/m^2

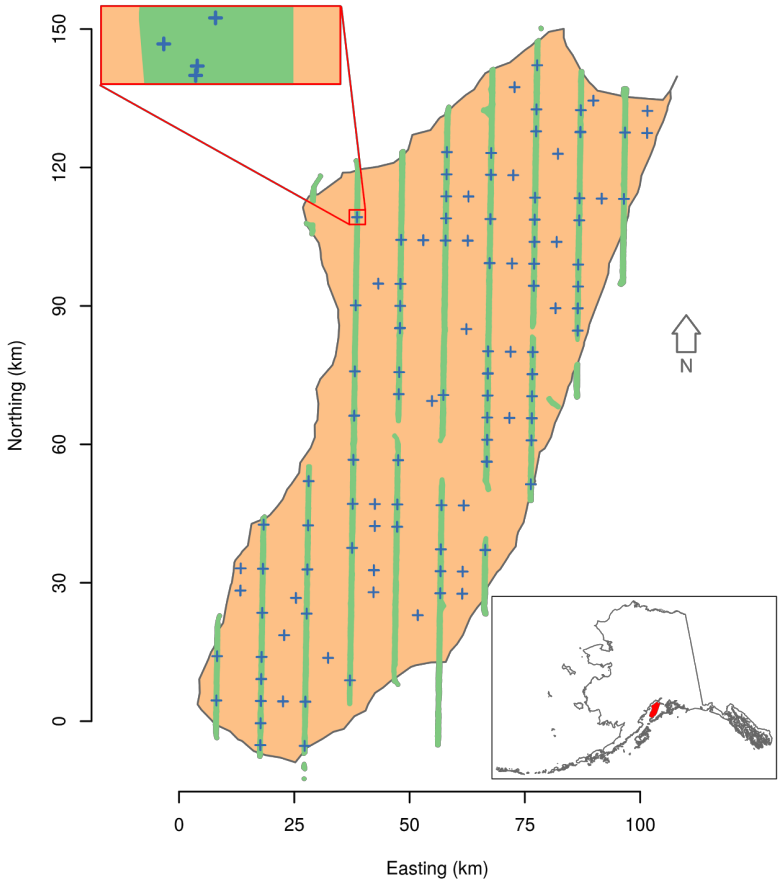


Figure 1: Kenai Peninsula study area. The orange region represents the coverage area for the 2008 wall-to-wall LiDAR dataset. The green areas represent the 2014 LiDAR strip datase coveraget. The 2004 and 2009 LiDAR datasets closely (but not completely) spatially coincide with the 2014 LiDAR dataset. The purple cross hairs locate the inventory plot locations (FUZZED). The inset map in the top left area of the figure zooms in on one cluster of plot locations. The Alaska state boundary map in the lower shows the location of the Kenai Peninsula in red.

Model framework

Bayesian hierarchical spatio-temporal modeling frameworks offer a useful solution to the problems of temporally misaligned data and uncertainty assessment [5]. Here, we explore this class of models to develop a unified statistical framework capable of coupling the temporally misaligned and repeated measures of field inventory and LiDAR data for the Kenai Peninsula. This framework is able to predict forest carbon through space and time, while formally propagating uncertainty through to prediction. Bayesian spatio-temporal models are flexible frameworks allowing for forest growth processes to be formally integrated into the model. By incorporating a mechanism for growth—using temporally repeated field and LiDAR data—we can more fully exploit the information-rich field inventory network to improve prediction accuracy[6]. These frameworks also provide access to spatially and temporally explicit posterior predictive distributions useful for summarizing uncertainty. Because predictions of forest carbon for each year result from the same model in a spatio-temporal framework, it is possible to probabilistically assess uncertainty of carbon flux by summarizing posterior predictive distributions appropriately.

Concluding Remarks

Results from this research will impact how forests are inventoried. It is too expensive to monitor terrestrial carbon flux using field-only sampling and estimation strategies and currently proposed model-based techniques leveraging LiDAR lack the ability to properly utilize temporally misaligned data---we need new and innovative methods to track forest carbon dynamics in remote regions. Bayesian hierarchical spatio-temporal modeling frameworks offer a solution to these shortcomings and, further, easily allow for formal predictive error assessment. which is useful decision making about the certainty of our estimates and about when and where to collect future field and LiDAR data to best improve prediction accuracy.

References

[1] Hopkinson, C., Chasmer, L., & Hall, R. (2008). The uncertainty in conifer plantation growth prediction from multi-temporal lidar datasets. Remote Sensing of Environment, 112, 1168–1180.

[2] Yu, X., Hyyppä, Kaartinen, H., Maltamo, M., & Hyyppä, H. (2008). Obtaining plotwise mean height and volume growth in boreal forests using multi-temporal laser surveys and various change detection techniques. International Journal of Remote Sensing, 29, 1367–1386.

[3] Hudak, A.T., Strand, E.K., Vierling, L.A., Byrne, J.C., Eitel, J.U., Martinuzzi, S., & Falkowski, M.J. (2012). Quantifying aboveground forest carbon pools and fluxes from repeat lidar surveys. Remote Sensing of Environment, 123, 25–40.

[4] Bechtold, W.A., & Patterson, P.L. (2005). The enhanced forest inventory and analysis program: national sampling design and estimation procedures. US Department of Agriculture Forest Service, Southern Research Station Asheville, North Carolina.

[5] Cressie, N., & Wikle, C.K. (2011). Statistics for spatio-temporal data. John Wiley & Sons.

[6] Babcock, C., Finley, A., Cook, B., Weiskittel, A., & Woodall, C. (In Review). Modeling forest biomass and growth: Coupling long-term inventory and lidar data. Remote Sensing of Environment, URL: http://blue.for.msu.edu/BFCW2015.pdf.

Multi-temporal analysis of the retention-trees groups structure using terrestrial laser scanner

Ulysse Rémillard, Robert Schneider

Chaire de recherche sur la forêt habitée, Département de biologie, chimie et géographie, Université du Québec à Rimouski (UQAR), 300 allée des Ursulines, Rimouski (QC), G5L 3A1, CANADA

Highlights: This study demonstrates that terrestrial laser scanner data acquisition allows forest inventory to be enhance by quantifying space occupation using landscape ecology patch metrics at the plot scale and allows the development to be monitored.

Key words: *multi-temporal, terrestrial laser scanner, forest dynamics, windthrow*

Introduction

According to the current Quebec Sustainable Forestry Act [1], new silvicultural approaches must be developed [2]. One such adaptation is variable tree retention (VTR) during clear-cut harvesting. As its objectives, VTR leaves retention trees in the cut-blocks that should survive for the next 40 to 50 years, in order to introduce some structural complexity in the future stand [3]. The first operational trials of VTR in eastern Quebec, Canada, used a grouped-tree method, where 2 groups of unharvested trees covering 200-400 m^2 are left standing per hectare of harvested forest [4]. In order to minimize windthrow, forest managers have tried to establish the groups in parts of the pre-existing stand that exhibit high structural complexity [5]. Quantifying such structural complexity and the development of these metrics over time is not easily accomplish with normal forest inventory data.

Retention group attributes and their changes in time were characterized with multi-temporal terrestrial laser scanner (TLS) data in order: (i) to evaluate the TLS-derived metrics, which can be used to study forest dynamics; and (ii) to identify which attributes are correlated with lower windthrow damage. The TLS-derived metrics were analyzed using landscape ecology approaches [6], as these offer theory and methods that can help to understand spatial patterns. At the landscape level, several studies have used airborne laser scanner (ALS) information to characterize forest structure and pattern using landscape ecology metrics [7]. These can also be used in studying retention groups, by assuming that they are small landscapes where subgroup or individual trees form patches.

Material and methods

A total of 12 retention-tree groups covering areas of 300 m^2 to 880 m^2 from 3 sites were followed for first 3 consecutive years following harvesting (i.e., 2011, 2012 and 2013). Each group consisted of a permanent sample plot that was established in 2011, with trees being numbered and measured (e.g., diameter at breast height and height) in years 1 and 2. This inventory was used to estimate the change in basal area over the 1st year.

In 2011, retention-tree groups were scanned using a Leica ScanStation 2. A Faro LS Focus 3D was used in subsequent years. A total of 3 to 9 scans were used to scan each group, depending upon group size. To reduce the point cloud size in terms of information volume to process, and because of diverse point cloud properties (number and location of scans, resolution and quality), point density was reduced, retaining 1 point per 5 mm-resolution voxel.

Using Computree software, a digital terrain model (DTM) was built with a 50 cm grid. The digital surface model (DSM) was constructed from the highest points in the grid. The difference between the DSM and the DTM represents the height model (DHM). DHM values were grouped into 4 classes corresponding to tree dominance class. This was used to obtain a mosaic where dominant and co-dominant trees are considered as patches. These patches are then used to quantify the basic landscape ecology metrics.

Class heights (Fig. 1b) were analyzed using FRAGSTATS software to estimate canopy change and quantify the effects of mortality on the canopy structure.

- Canopy density (CD), which was used to measure windfall severity, was calculated as the percentage of the area occupied by tree crowns in horizontal projection, with the values for each class representing the vertical layering.
- Patch area (PA), the number of patches (n), and mean patch area (MPA) were used to estimate spatial distribution heterogeneity.
- Connectivity (C) of a class, which was calculated as the relative size of the larger patch, was used to quantify mortality. Tree death will induce a greater connectivity index value if it is concentrated within part of the group rather than tree mortality being spread across the whole group.

- The Division index (DIVI) is based upon the cumulative patch area distribution and is interpreted as the probability that two randomly chosen pixels in the landscape are not situated within the same patch of the corresponding patch type. It is calculated with the dominant and co-dominant height class ($DIVI_4$), and yields information on mature cover distribution. Canopy roughness, which is measured with the Division Index for all height classes (DIVI), yields overall height class heterogeneity or fragmentation in the retention-tree groups.

$$DIVI = \left[1 - \sum_{j=1}^{n} \left(\frac{a_{ij}}{A}\right)^2\right]$$

where a_{ij} = area (m^2) of patch ij
n = number of patches
A = total landscape area (m^2).

- Canopy heterogeneity was quantified with the Rumple Index, calculated as the ratio between the DSM area and the DTM area.

Changes in the connectivity index, canopy roughness (DIVI), and the Rumple Index measure the effect of windthrow within the plot in terms of canopy heterogeneity.

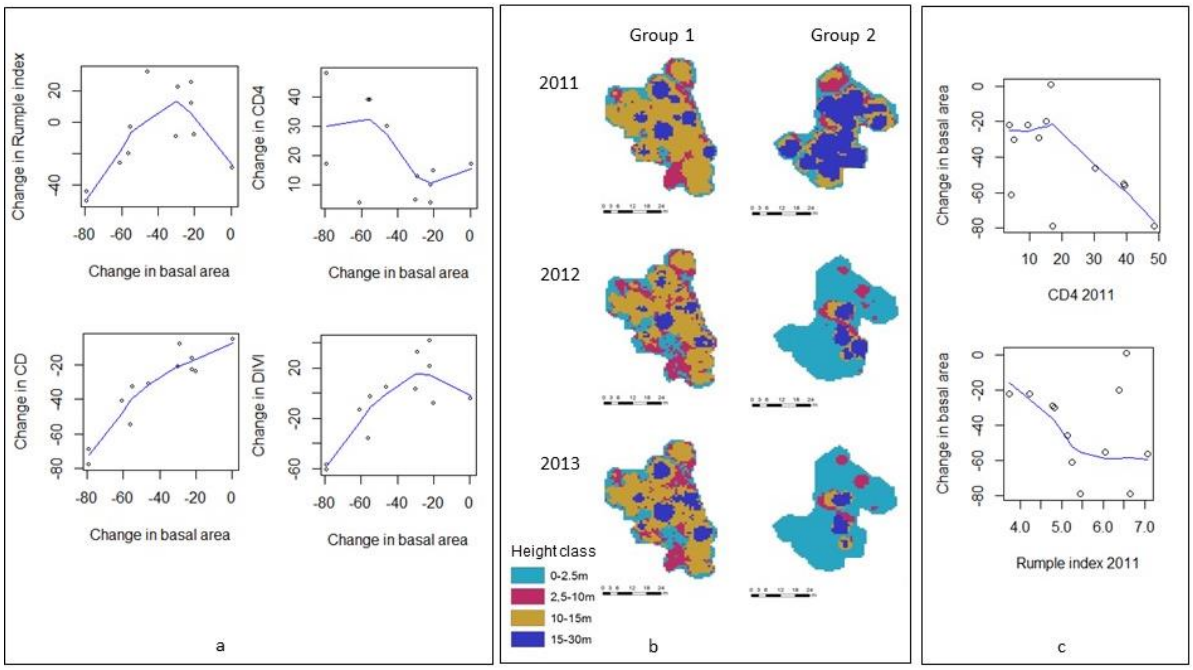


Figure 1: Changes in TLS-derived metrics (Rumple Index, canopy density of all classes, canopy density of the dominants and co-dominants, and division index) with windfall severity (change in basal area) represented by a local regression model (LOWESS) (a); height class dynamics for 2 groups of trees (b); relationship between windthrow damage and canopy density of the dominant-co-dominant class and Rumple Index measured immediately following harvesting (c).

Results

Variation in windthrow was very high, ranging from no loss to about 80 % of the basal area (Fig. 1). A very strong relationship was found between the change in basal area and the change in canopy density, where the reduction in CD was found to be more pronounced for groups with high windfall (Fig. 1a). The Rumple and division indices of all height classes show similar trends (Fig. 1a). The canopy density of the dominant-codominant height class (CD_4) exhibited an inverse trend, where the groups with the most windthrow had a positive increase in CD_4 (Fig. 1a). Group heterogeneity is proportional to windfall damage. Thus, as tree loss increases, it is normal to observe large changes in metrics that are associated with heterogeneity (e.g., CD, DIVI, Rumple Index). Patch area (not illustrated) also changes with windfall damage, where patch area, especially patch area of the dominant-codominant height class, is inversely proportional to the observed loss in basal area.

The relationship between windthrow damage and initial TLS-based metrics is not clear (Fig. 1c). The strongest ones were observed with the Rumple Index and canopy density of the dominant-co-dominant (CD_4) metrics. Windfall was more important for groups with high Rumple Indices as well as with high CD_4 values.

Conclusion

Our results show a high degree of variability in structural complexity during forest development, which coincides with Kane et al.'s [6] study based on ALS data. Windfall is also a complex phenomenon. Our efforts concentrated on determining which TLS-derived metrics could be used (i) to follow forest dynamics and (ii) to help predict the stability of variable harvesting group-trees. The use of traditional landscape ecology metrics at a finer scale is promising, as we demonstrated that they can be estimated from TLS data and used to follow windthrow over a broad range of disturbance intensities. However, the studied TLS metrics are not very useful in predicting group-tree stability.

Our results will help further integrate TLS data into forest management decisions, especially in the cases where there are important shifts in management paradigms.

TLS appears as a promising tool for improving forest inventories by acquiring data that better describe forest ecosystems, and their temporal and geographic changes. Besides monitoring variable retention-tree harvesting, we can imagine, for example, the potential of using TLS and landscape metrics at the plot- and stand-scale to describe and to quantify the differences between managed and natural forests in order for forest managers to quantify how their ecosystem-based management meets the expected objectives.

References

[1] Lois et règlements du Québec (L.R.Q.), 2010. Loi sur l'aménagement durable du territoire forestier, n.d., RLRQ c A-18.1.

[2] Barrette, M., Leblanc, M., Thiffault, N., Paquette, A., Lavoie, L., Bélanger, L., Bujold, F., Côté, L., Lamoureux, J., Schneider, R., Tremblay, J.-P., Côté, S., Boucher, Y., Deshaies, M.-È., 2014. Issues and solutions for intensive plantation silviculture in a context of ecosystem management. Forestry Chronicle 90, 732–762.

[3] Franklin, J. F., Berg, D. R., Thornburgh, D. A., & Tappeiner, J. C., 1997. Alternative silvicultural approaches to timber harvesting: variable retention harvest systems. Creating a forestry for the 21st century: the science of ecosystem management. Island Press, Washington, DC, pp. 111-139.

[4] Leblanc, M., Pouliot, B., 2011. La coupe avec protection de la régénération et des sols avec rétention de bouquets - Fondements et exécution opérationnelle, Québec, gouvernement du Québec, ministère des Ressources naturelles et de la Faune, 9 p.

[5] Lavoie, S., Ruel, J.-C., Bergeron, Y., Harvey, B.D., 2012. Windthrow after group and dispersed tree retention in eastern Canada. Forest Ecology and Management 269, 158–167.

[6] Kane, V.R., McGaughey, R.J., Bakker, J.D., Gersonde, R.F., Lutz, J.A., Franklin, J.F., 2010. Comparisons between field- and LiDAR-based measures of stand structural complexity. Can. J. For. Res. 40, 761–773.

[7] Maier, B., Tiede, D., Dorren, L., 2008. Characterising mountain forest structure using landscape metrics on LiDAR-based canopy surface models, in: Blaschke, T., Lang, S., Hay, G.J. (Eds.), Object-Based Image Analysis, Lecture Notes in Geoinformation and Cartography. Springer Berlin Heidelberg, pp. 625–643.

Development of a scaling technique of ALS-derived aboveground biomass using high resolution satellite data in tropical seasonal forest

Yasumasa Hirata[1], Naoyuki Furuya[1], Pak Chealy[2], Leng Chivin[2], Sokh Heng[2], Ma Vuthy[2], Tetsuji Ota[3], Tsuyoshi Kajisa[4], Nobuya Mizoue[3]

1. Forestry and Forest Products Research Institute, Tsukuba, 305-8687, Japan
2. Forestry Administration, Phnom Penh, Cambodia
3. Kyusyu University, Fukuoka, 819-0395, Japan
4. Kagoshima University, Kagoshima, 890-8580, Japan

Highlights: We develop a scaling technique of ALS-derived aboveground biomass using high resolution satellite data in tropical seasonal forest. First, we investigated the relationship between canopy heights derived from ALS and aboveground biomass from field survey. Then, we modelled the relationship between reflectance properties in satellite data and aboveground biomass from ALS.

Key words: *REDD+, above-ground biomass, tropical seasonal forest, airborne laser scanner (ALS), high resolution satellite, object-based classification*

Introduction

REDD+ (Reducing emissions from deforestation and forest degradation in developing countries; and the role of conservation, sustainable management of forests and enhancement of forest carbon stocks in developing countries) is identified as one of the most effective means to reduce GHG emission in the post-Kyoto climate change negotiation. A reliable and credible system of measurement, reporting and verification (MRV) of forest carbon changes is a cornerstone of any national REDD+. An MRV system should follow the international requirements and also be adapted to the country's specific conditions, e.g. vegetation, economy, culture, institution and/or the deforestation/forest degradation drivers. A REDD+ participating country should prepare such a forest carbon change MRV system prior to a full implementation of REDD+ so as to promptly estimate the carbon budget and its historical trend for the reference level.

Precise and efficient forest carbon monitoring technologies are indispensable for an MRV system. Innovative remote sensing technologies with ground measurement should be able to promote forest carbon monitoring with higher precision and efficiency while following the international requirement. It is thus important for scientists to develop possible carbon monitoring methods utilizing such remote sensing technologies and to evaluate their precision and efficiency so as to implement them in a future carbon MRV.

This study aims to develop a scaling technique of ALS-derived aboveground biomass using high resolution satellite data in tropical seasonal forest.

Materials and Methods

Study area and field data collection

Study area is located in Kompong Thom Province, Cambodia. In this study area, evergreen forest, deciduous forest, mixed forest and regeneration can be seen. We established 35 sample plots (30×30 m or 50×50 m) and determined the coordinates at the centre of each plot by using a global positioning system (GPS) receiver so that we could superimpose the plots on high-resolution satellite images. We measured the diameters at breast height of all standing trees in the sample plots and the aboveground biomass in each plot was calculated.

Remotely sensed data

The ALS data were acquired in January 2012. The flight altitude of the helicopter above the ground was about 500 m, and the average flight speed was approximately 25 m/s. The repetition frequency of laser pulses was 100 kHz. The maximum scan angle (off nadir) was 12°. The sampling density of the laser scanning was 18.7 laser pulses/m^2. Total area of data acquisition by air survey was about 25 km^2.

The DEM and DSM of the study area were generated from the ALS data with a 50-cm cell size. The DEM was generated from the last returns of the laser pulses, and the DSM was generated by assigning the highest first-return value of the laser pulses for each cell as the cell's value. A digital canopy height model (DCHM), which describes the canopy height above the ground level, was calculated by subtracting the DEM from the DSM.

QuickBird panchromatic and multi-spectral data, which has 0.61 m and 2.44 m ground resolution at nadir, were acquired on 1 February 2010. The data were geo-registered to the Universal Transverse Mercator (UTM)

coordinate system (WGS 84, zone 51) with 0.7 m and 2.8 m ground resolution using the nearest neighbour method for resampling to maintain original reflectance properties.

Data analysis

Mean canopy height for each plot was calculated from DCHM derived from the airborne laser scanner data. Then, we investigated the relationship between mean DCHM and above-ground biomass estimated measured in our field survey of the 35 sample plots. We plotted the mean DCHM derived from ALS (x) against aboveground biomass (y) of the 35 sample plots, and we used the results to estimate the parameters p and q in the following Equation.

$$y = px^q \quad (1)$$

Segmentation in the object-oriented approach is characterized by the integrated use of spectral and textural information in comparison to the single use of spectral information in conventional pixel-based classification. We investigated the effect of the parameter concerning heterogeneity in segmentation in the object-oriented approach, on the number and average size of generated objects. The segmentation of high-resolution satellite data was conducted using eCognition software for object-oriented classification. Generated objects were used for this analysis. The average and standard deviation in digital number of panchromatic data and that of the four bands of multispectral data for these objects were calculated. The average and standard deviation of NDVI were also used for the analysis.

Then, we conducted multi-regression analysis using aboveground biomass derived from ALS data by object and properties of high resolution satellite data by object to scale up the result of estimation of aboveground biomass derived from ALS.

$$\text{Biomass}=\beta+\Sigma\alpha_i\ (\text{band } i_{\_\text{mean}})+\alpha 5\ (\text{pan}_{\_\text{mean}})+\alpha_6\ (\text{NDVI}_{\_\text{mean}}) \quad (2)$$

Results and Discussion

Aboveground biomass estimation from ALS

The relationship between mean DCHM from ALS data and aboveground biomass from field survey was shown in Figure 1. From Figure 1, we estimated aboveground biomass using the following Equation.

$$y = 3.8406x^{1.414} \quad (R^2=0.80) \quad (3)$$

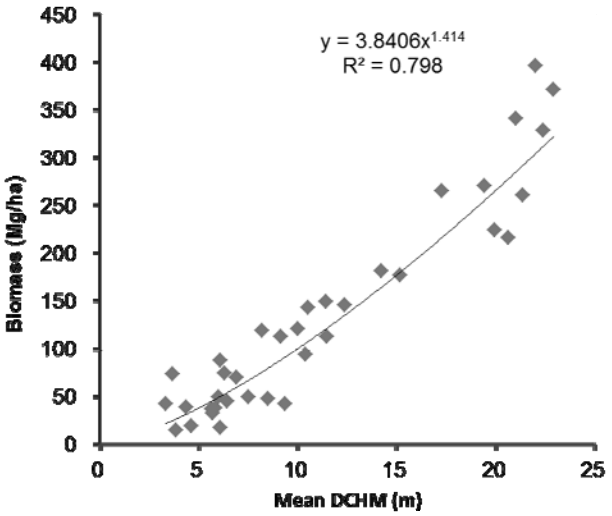


Figure 1. The relationship between mean DCHM from ALS data and aboveground biomass from field survey

Scale-up of aboveground biomass estimation

Segmentation for high resolution satellite data was conducted and objects for domain with similarity of canopy structure were generated (Figure 2). The result of multi-regression analysis using aboveground biomass derived from ALS data by object and properties of high resolution satellite data by object (i.e. coefficients of Equation (2)) was shown in Table 1. Using the model of Equation (2), we estimated aboveground biomass for each object using the properties of reflectance in high resolution satellite data and mapped them (Figure 3).

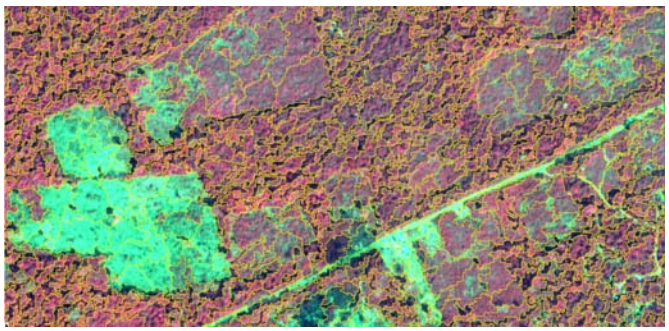


Figure 2. Segmentation for high resolution satellite data

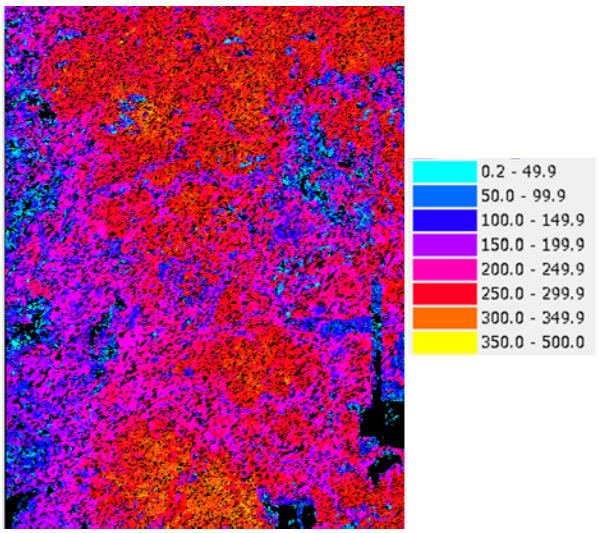


Figure 3. Aboveground biomass mapping using ALS data and high resolution satellite data

Table 1. Coefficients of Equation (2)

β	-221.72
α_1	4.84
α_2	-4.91
α_3	2.57
α_4	0.53
α_5	-0.71
α_6	4.95

Conclusions

In UNFCCC Decision, a combination of remote sensing and ground-based inventory is recommended to establish a national forest monitoring system. However, there are many issues for ground-based inventory in developing countries (accessibility and road condition, weather, ownership, etc.). High precision remote sensing make it possible to estimate emission factor, namely, forest carbon stock per unit area by forest type, instead of ground-based inventory.

Data assimilation in forest inventory: First empirical results using ALS data

Mattias Nyström, Nils Lindgren, Jörgen Wallerman, Anton Grafström, Anders Muszta, Kenneth Nyström, Göran Ståhl, Håkan Olsson
Swedish University of Agricultural Sciences (SLU), Department of forest resource management, SE 901 83 Umeå, Sweden; e-mail: firstname.lastname@slu.se

Highlights: A first data assimilation case study using a time series of ALS for updating forest stand data is presented. Forest stand data are predicted from each ALS acquisition. Kalman filtering and growth models are then used to combine each new ALS based prediction with forecasts from the previous data acquisition.

Key words: *data assimilation; ALS; LiDAR; forest inventory*

Introduction

Data assimilation is a technique that offers great potential for combining all new sources of data of relevance for forest estimates [1]. The success of data assimilation in other areas, such as meteorology, is well documented [2]. However, in order to realize the potentials in the context of forestry the statistical methods need to be adapted to this field of application. In brief, a system for data assimilation in forestry should be based on a geographical model of the forest in which forest data are forecasted using growth models. All new data from remote sensing and from measurements in field, should then be used to adjust the forecasted forest information to the extent motivated by the accuracy of the new data compared to the accuracy of the forecasts.

The objective of this study was to present first empirical results of the application of data assimilation to forest stand data using remote sensing data from airborne laser scanning (ALS). In a previous study made by our research group, the results were based on theoretical assumptions [1]. In the present study we applied data assimilation to predictions based on empirical data from ALS data from six occasions obtained over an 8 year period (2003–2011). The target variables were Lorey's mean height, basal area and growing stock volume. The results from data assimilation were compared with two established methods; prediction using ALS data from the most recent time point or forecasting of ALS predictions using growth models.

Material and Methods

Study area

The study was carried out at the forest estate Remningstorp in south-western Sweden (lat. 13° 37' N; long. 58° 28' E). The forest is dominated by Norway spruce (*Picea abies*) and Scots Pine (*Pinus sylvestris*), with some deciduous forest of mainly birch (*Betula spp.).*

Remote sensing data

Airborne laser scanning (ALS) data from six years: 2003, 2004, 2007, 2008, 2010 and 2011 were used for the assimilation. The scannings were acquired in the autumn under leaf-on conditions except for 2007 and 2011 which were acquired in the spring under leaf-off conditions. Metrics were extracted with LAStools using a height threshold of 2 m [3]. The sample plot area was used for extracting metrics when developing the prediction models and raster cells of 18 m × 18 m when applying the prediction models for the evaluation plots.

Field reference data

The field reference training data set used to predict the three target variables consisted of two field campaigns. The first field campaign was year 2004 and 2005 and collected sample plots with 10 m radius in a regular pattern with 40 m distance between the plot centers. A subset of these plots, in total 258, was used as field reference for the ALS data acquired in 2003, 2004, 2007 and 2008. The field data was for- or back-casted to obtain as good temporal match as possible. This was done using the Heureka forestry decision support system [4].

The second field reference training data set was collected in year 2010 and consisted of 263 sample plots with 10 m radius in a regular pattern with 200 m distance between the plot centers. This reference data was used for the two latest ALS acquisitions.

Prediction of target variables

Prediction of the target variables, Lorey's mean height, basal area and growing stock volume, was done using the area-based method [5]. To predict Lorey's mean height, the 95th height percentile was used. In the prediction model for basal area and growing stock volume, the vegetation ratio and the quadratic average of the returns

were used. These variables were selected with support of best subset regression to find two predictors describing height and density. The selected metrics were forced to be the same for all acquisitions.

Growth models

Growth models were developed to estimate the growth between one assimilation and the date of the next ALS acquisition. These functions were developed through regression analysis and used permanent plots from the Swedish National Forest Inventory [6].

Evaluation plots

The plots used to evaluate data assimilation consisted of ten 40 m radius sample plots in normally developed stands without cuttings in the time period between 2003 and 2011. All plots were however not covered by all laser scanning's and the number of actual scannings used in the assimilation of an evaluation plot varied between two and six, with an average of 3.7. Validation data for year 2011 were obtained by calipering all trees at these evaluation plots and a sample of trees were height measured.

Data assimilation

Existing information about a forest area is forecasted using a model that provides an estimate at the time of the next data acquisition and an estimate of the precision of the forecasted information. Thus, the precision of the forecasted information can be compared with the precision of the new information. In the assimilation step, the two sources of information are combined through weights that are inversely proportional to their variances. The combined estimate is then forecasted to the time point of the next data acquisition, etc.

In this study we applied the extended Kalman filter [7] for the data assimilation. Only one variable at the time was addressed. The other variables used in the growth models were site index, tree species composition and age, which were known from field surveys and assumed known without errors.

Modeling the development over time of the target variable is a core part of the data assimilation system. Using the extended Kalman filter, the assimilated variable, $\hat{x}_t$, can be calculated at time point t as

$$\hat{x}_t = (1 - K_t)\tilde{x}_t + K_t\tilde{z}_t \quad (1)$$

where K_t is the Kalman gain calculated as $K_t = \frac{\tilde{p}_t^2}{\tilde{p}_t^2 + r_t^2}$ where $\tilde{p}_t^2$ is the variance of the forecasted value, $\tilde{x}_t$, and r_t^2 is the variance of the prediction from ALS data, $\tilde{z}_t$. Further, the variance of the assimilated variable, $\hat{x}_t$, can be calculated as $p_t^2 = (1 - K_t)\tilde{p}_t^2$.

The assimilation was conducted for each raster element (18 m × 18 m). For validation on the evaluation plots, the mean value was calculated of the raster elements with center point within each evaluation plot's boundary.

Results

Data assimilation was compared to the two established methods: prediction using ALS data from the most recent time point and forecasting of ALS predictions from the first time point. Field measured evaluation plots consisting of 40 m radius sample plots were used as ground truth. Figure 1 shows the deviation from the field measured value for each of the methods and the ten evaluation plots, for the variable basal area. Table 1 shows the RMSE of the deviation from the field measurements for all three estimated variables. It can be seen that the mean deviations are smaller using data assimilation compared to forecasting the value from the first ALS data prediction. The prediction using the most recent ALS data resulted in lower RMSE for Lorey's mean height though.

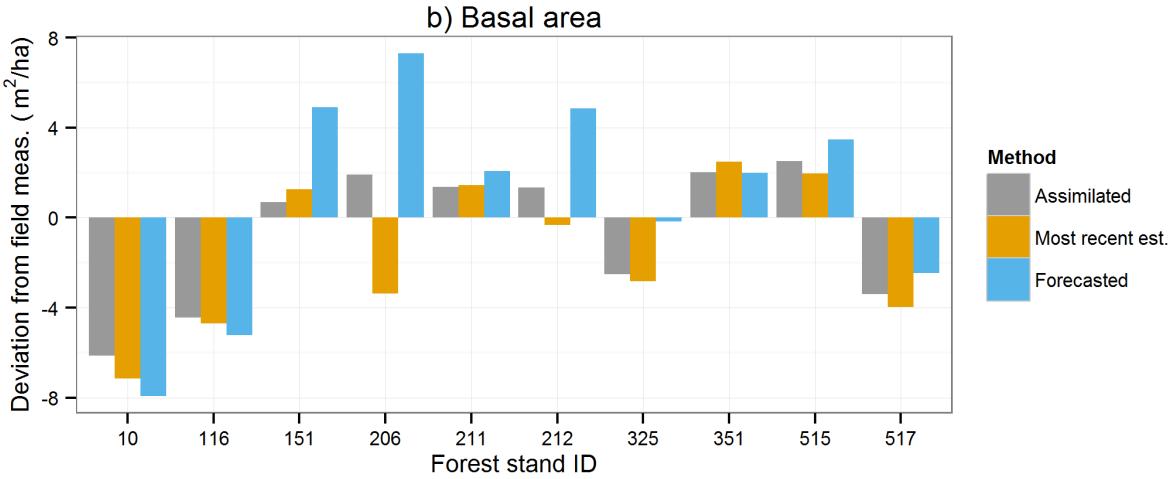


Figure 1: Deviation of basal area calculated as each methods value 2010 minus the field measured value 2010 for the ten evaluation plots.

Table 1: RMSE of the ten assimilated plots. In parenthesis the relative RMSE. The comparison is to the field measured values for the plots.

Target variable	Unit	Assimilated	Last ALS	Forecast from first ALS
Lorey's mean height	m	1.7 (8.5%)	1.6 (8.0%)	2.0 (9.9%)
Basal area	m^2/ha	3.1 (9.5%)	3.5 (10.8%)	4.7 (14.5%)
Stem volume	m^3/ha	41.0 (13.3%)	44.1 (14.3%)	63.5 (20.6%)

Discussion

In this study the potential of data assimilation is verified. The strength of data assimilation will probably first be seen when combining data from several remote sensing techniques. For example if data first is acquired using ALS and the next data acquisition with a cheaper technique and probably lower accuracy, we will be able to update the high accuracy acquisition from the ALS with new data. As we only use ALS data in this early study, there might be temporal autocorrelations between the estimation errors. Methods to compensate for this will be developed in future studies.

The first acquisition is from year 2003 and the last from 2011. This means that it is a rather short time period for forestry. In an operational case we would have a model that continuously is updated when new data become available and probably span over much longer time than eight years.

Further research is needed to investigate at which level the assimilation should be performed. It could be more convenient to assimilate directly on the whole plot instead of splitting the plot into raster cells.

In this study we relied on only two field training data sets and forecasting or back-casting was done to the years' of the ALS acquisitions. The use of only two field surveys is probably the most limiting factor which might explain why the assimilation results is only marginally better than the use of only the latest time-point. In a best case we would have field data from every year where remote sensing data is available. The starting point for the assimilation is an estimate using the first ALS data. Another possibility would be to start with the last known state for each stand, for example from the forest stand register.

Conclusions

This study presents among the first results of a data assimilation of forest stand variables. The results verify the potential of data assimilation, but the study also identifies several practical obstacles that must be addressed before data assimilation can be successfully applied in practice.

Acknowledgement

This study was financed by the Forest Society of Sweden and the Kempe foundations.

References

[1] Ehlers, S.; Grafström, A.; Nyström, K.; Olsson, H.; Ståhl, G. (2013). Data assimilation in stand-level forest inventories. Can. J. For. Res., 43, 1104–1113.

[2] Rabier, F. (2005). Overview of global data assimilation developments in numerical weather-prediction centres. Q. J. R. Meteorol. Soc., 131, 3215–3233.

[3] Nilsson, M. (1996). Estimation of tree heights and stand volume using an airborne lidar system. Remote Sens. Environ., 56, 1–7.

[4] Wikström, P.; Edenius, L.; Elfving, B.; Eriksson, L. O.; Lämås, T.; Sonesson, J.; Öhman, K.; Wallerman, J.; Waller, C.; Klintebäck, F. (2011). The Heureka forestry decision support system: An overview. Math. Comput. For. Nat. Sci., 3, 87–94.

[5] Næsset, E. (2002). Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Remote Sens. Environ., 80, 88–99.

[6] Fridman, J.; Holm, S.; Nilsson, M.; Nilsson, P.; Ringvall, A. H.; Ståhl, G. (2014). Adapting National Forest Inventories to changing requirements - The case of the Swedish National Forest Inventory at the turn of the 20th century. Silva Fenn., 48, 1–29.

[7] Welch, G.; Bishop, G. (2006). An introduction to the Kalman filter; Chapel Hill, NC, USA.

Estimating forest site productivity – approaches based on airborne laser scanning, auxiliary data and Landsat time series

Piotr Tompalski[1]*, Nicholas C. Coops[1], Joanne C. White[2], Michael A. Wulder[2]

[1] *Faculty of Forestry, University of British Columbia, 2424 Main Mall, Vancouver, BC, V6T 1Z4, Canada.*
[2] *Canadian Forest Service (Pacific Forestry Centre), Natural Resources Canada, 506 West Burnside Road, Victoria, BC, V8Z 1M5, Canada*
** corresponding author: piotr.tompalski@ubc.ca*

Highlights: We demonstrate how Airborne Laser Scanning data can be used to estimate forest site productivity. Two approaches are presented: first, we enhance site index predictions with ALS-derived dominant height; second, we develop a productivity model with additional information on stand age derived from Landsat time series.

Key words: *forest site productivity, airborne laser scanning, chronosequence, Landsat.*

Introduction

Sustainable forest management requires accurate information on a range of forest stand attributes. This information, collected during forest inventories, is crucial for evaluating current and projected conditions of a forest, as well as being critical for assessing the consequences of management decisions. Among all possible forest attributes, the potential of a site to produce biomass remains one the most important as this information is fundamental for decisions regarding optimal species composition, rotation age, allowable cut and most importantly, to forecast future timber yield.

The most common approach for measuring forest site productivity is site index (SI), which describes a forest stand as a relation between tree age and height [1]. Site index has conventionally been constructed by modelling the height growth of a subset of the largest trees at a given location and for a given species, according to the concept that stand dominant height is correlated with stand volume. The height of the stand at 50 years (in Canada) is then used as an index of potential site productivity allowing estimates to be compared across sites and across stands of different ages.

Two pieces of information are then needed to estimate forest site productivity: stand dominant height and age. Stand dominant height can be determined using Airborne Laser Scanning (ALS) data. Three-dimensional point clouds acquired with ALS are increasingly used as an information source for forest inventories, and provide the capacity to generate accurate estimates of volume, basal area, and stand height. From all the forest stand attributes possible to estimate with ALS data, height is the attribute estimated with the greatest accuracy.

In managed forests, information about stand age can be obtained from forest inventory records. Alternatively, stand age can be estimated with time series of satellite imagery. With the first mission launched in 1972, the available archive of Landsat data represents more than 40 years of observations, and since 1982, has been acquired at 30 m spatial resolution with consistent spectral resolutions across multiple satellite missions. Time series analyses of annual Landsat imagery enables detection and mapping of stand replacing disturbances and their attributes such as spatial location, extent, and date [2]. Time since disturbance (TSD) can then be used as a proxy for estimating stand age.

In this paper, we demonstrate two methods of how ALS point clouds can be used to estimate forest site productivity. Both methods are based on dominant height estimates derived with ALS data and age estimates obtained from forest inventory or derived with time series of Landsat imagery. We first use existing productivity models and ALS-based dominant height as input data to derive site index estimates. Second, we demonstrate how integrated stand dominant height information from ALS point clouds and stand age derived from time series of Landsat images can be used to develop a forest site productivity model. In both cases, we compare the productivity estimates with inventory data and discuss the strengths and weaknesses of the presented methods.

Methods

The study area was located on northern Vancouver Island, British Columbia, Canada. Forest inventory data were used to provide reference information for forest stand age, dominant height, and site productivity (site index). According to the forest inventory data, the average age of stands was 144 years (σ = 127 years).

ALS point clouds were acquired in 2012 with average first return point density of 11.6 points/m^2. Dominant height was estimated by dividing a stand into 10 x 10 m cells and computing maximum height for each of the cells using the normalized point cloud. The final dominant height estimate is calculated as the weighted mean of all the maximum heights, with the number of non-ground returns within each cell used as the weight.

Information on stand age was extracted from two sources: forest inventory and using an approach based on the Vegetation Change Tracker (VCT) [3]. In the second case stand-replacing disturbance events were detected and attributed to the correct disturbance year. From the disturbance date, we calculated the time since disturbance (TSD) and used this as a proxy for stand age.

We first investigated the outcomes of using ALS-based dominant height information by comparing it with dominant height from forest inventory. We also assessed the impact of the revised site indices on future stand merchantable volume projections. In this case stand age information was provided by forest inventory data [4].

In the second case we developed a site productivity model using a guide curve approach [5]. In this approach, a curve is first fit to the height-age measurements and confidence intervals are calculated. The site productivity model is derived by calculating the proportional distance between the given height and age values and the modelled guide curve. The reference age was set to the maximum year in the chronosequence, which was limited by the Landsat Thematic Mapper (TM) archive (32 years in our case). The model was validated with 120 randomly selected plots. For each plot, we extracted the reference age, dominant height, and site index value from the forest inventory data, and dominant height and age as estimated from the ALS and Landsat disturbance map, respectively. We then compared the reference and predicted values by calculating the bias, RMSE, and correlation coefficient [6].

Results

Stand dominant heights directly estimated from ALS data were compared to stand heights modelled using height-age curves and forest inventory data. Overall, we observed an absolute mean difference of 3.5 m (predicted-observed), and relative mean difference of 25.6%; Similar trends were observed across all species, with western red cedar having the largest, significant difference (ΔH =5.4 m, p = 0.000) and Sitka spruce having the smallest, significant difference (ΔH = 1.7 m, p = 0.023) [4].

The developed productivity model (Figure 1) had a following form:

$$SP = 11.53 - 6.81 * \frac{h_G - h_T}{h_G - h_L} \quad (1)$$

$$h_G = 109.06 * (1 - e^{-0.0041 * age})^{1.0705} \quad (2)$$

$$h_L = 44.72 * (1 - e^{-0.0041 * age})^{1.0705} \quad (3)$$

where SP is the site productivity, h_T is stand dominant height, age is stand age.

Model validation indicated that when the remotely sensed chronosequence data were used, an overestimate of 0.70 m (10.1%), with RMSE of 5.55 (53.02%). The correlation coefficient was 0.63 and the Wilcoxon test indicated that the remote sensing-based estimates of site productivity were significantly different than those derived with the forest inventory reference data (α=0.05) [6].

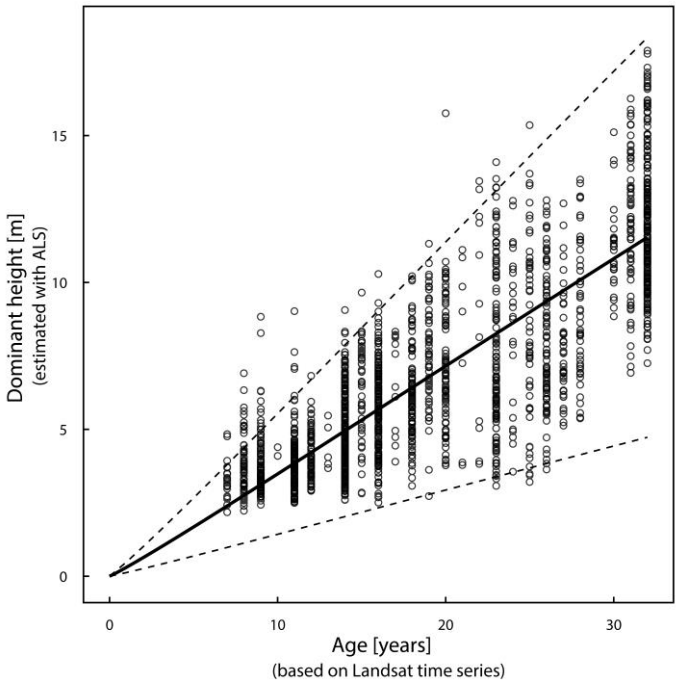


Figure 1. Guide curve fitted in the height-age measurements. 95% confidence intervals presented as dashed lines.

Discussion

In this research, we demonstrated a method of using ALS point clouds to improve existing estimates of forest site productivity and to develop productivity models fully based on remote sensing data.

First, we used ALS data, in concert with age information from a forest inventory, to investigate the robustness of existing site index estimates in the inventory. The site index estimates in the inventory were derived from height-age curves, which were developed using ground sampling. Although the height-age curves themselves are developed using a network of ground samples, the curves are applied in the inventory using attributes that are photo-interpreted (i.e., species, age, height). Our analyses indicated that the current site index estimates in the forest inventory consistently underestimate site potential (by 3.5 m on average) and this finding is in keeping with what others have reported in the literature [7].

Second, we integrated a time series of Landsat imagery and ALS point clouds to provide an indicator of forest site productivity. By integrating information extracted from these two sources of spatial data we established a chronosequence of sample units from which we obtained a good approximation of stand growth. We used the chronosequence to develop a model of site productivity, demonstrating that remotely sensed data sources have potential to provide such information in areas that may be lacking detailed forest inventory information.

We conclude that there is utility in both presented approaches: ALS data can be used to augment the estimates of site productivity by providing measures of dominant height, as well as to develop productivity models, which in the presented case were based exclusively on remotely sensed data sources. Improved information on site productivity is important to support a range of information needs, including forest growth and yield predictions, understanding of climate change impacts, aiding in the estimation of post-disturbance recovery, and enabling spatially explicit carbon budget models.

Acknowledgements

This research was supported by the Canadian Wood Fibre Centre (CWFC) of the Canadian Forest Service, Natural Resources Canada, with valuable in-kind support from BC Timber Sales and Western Forest Products. Support was also provided by a Natural Sciences and Engineering Research Council of Canada (NSERC) grant to Nicholas Coops.

References

[1] Skovsgaard, J. P.; Vanclay, J. K. (2008). Forest site productivity: a review of the evolution of dendrometric concepts for even-aged stands. Forestry, 81, 13–31.

[2] Hermosilla, T.; Wulder, M.; White, J. C.; Coops, N.; Hobart, G. W. (2015). An integrated Landsat time series protocol for change detection and generation of annual gap-free surface reflectance composites. Remote Sensing of Environment, 158, 220–234.

[3] Huang, C.; Goward, S. N.; Masek, J. G.; Thomas, N.; Zhu, Z.; Vogelmann, J. E. (2010). An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. Remote Sensing of Environment, 114, 183–198.

[4] Tompalski, P.; Coops, N. C.; White, J. C.; Wulder, M. A. (2015). Augmenting site index estimation with airborne laser scanning data. Forest Science. http://dx.doi.org/10.5849/forsci.14-175

[5] Edminster, C. B.; Mathiasen, R. L.; Olsen, W. K. (1991). A Method for Constructing Site Index Curves from Height-Age Measurements Applied to Douglas-fir in the Southwest. Research Note RM-510, USDA Forest Service, 6.

[6] Tompalski, P.; Coops, N. C.; White, J. C.; Wulder, M. A.; Pickell, P. D. (2015). Estimating forest site productivity using airborne laser scanning data and Landsat time series. Can. J. Remote Sensing, (in press).

[7] Wulder, M. A.; White, J. C.; Stinson, G.; Hilker, T.; Kurz, W. A.; Coops, N. C.; St-Onge, B.; Trofymow, J. A. T. (2010). Implications of differing input data sources and approaches upon forest carbon stock estimation. Environmental Monitoring and Assessment, 166, 543–61.

Characterizing changes in structural complexity following fire in the western Canadian boreal using a combination of Landsat time-series and airborne lidar data

Douglas K. Bolton[1], Nicholas C. Coops[1], Michael A. Wulder[2]

[1]*Integrated Remote Sensing Studio, Department of Forest Resources Management, Faculty of Forestry, University of British Columbia, 2424 Main Mall, Vancouver, British Columbia, V6T 1Z4, Canada.*

[2]*Canadian Forest Service (Pacific Forestry Centre), Natural Resources Canada, 506 West Burnside Road, Victoria, British Columbia, V8Z 1M5, Canada*

Highlights: We detect high-severity fires using Landsat time-series data and assess structural response to these fires using airborne lidar transects. By sampling > 600 burned patches at various times since fire, we construct a 25-year chronosequence of structural development and assess various lidar metrics that inform on stand structural complexity.

Key words: *Disturbance, Forest succession, Structural development, Structural complexity*

Introduction

Fire is a dominant driver of forest structure and carbon dynamics in Canada's boreal forests. As the northern boreal of Canada is not subject to routine, area-wide, forest inventory, we lack a strong characterization of the impacts of fire on forest structure over this extensive forested area. As access to northern forests in Canada is limited, assessing post-fire structure through field measurements alone is challenging and expensive. Landsat data has been used extensively to detect and describe forest disturbances (e.g, [1]), allowing the timing and severity of boreal fires to be assessed over large areas. However, without information on three-dimensional structure, the role of Landsat for assessing structural development following disturbance is limited. Recent attempts to assess structural development following disturbance have relied on a fusion of Landsat data and Light Detection and Ranging (lidar) data (e.g, [2]). Specifically, Landsat time-series data is used to detect disturbances and lidar data is used to assess post-disturbance structure.

In this study, we detect high-severity fires across 40 million ha of boreal forests using Landsat time-series data (1985-2010), and observe the structural response to fires using transects of airborne lidar data. We focus our attention on lidar metrics that describe stand structural complexity, and discuss the information that each complexity metric provides during early stand development.

Methods

Data and Processing

Thirteen Landsat scenes in the Boreal Shield West ecozone of Canada were selected to detect high-severity fires from 1985-2010 (Figure 1). All available summer imagery from 1984-2013 with less than 60% cloud cover was downloaded and processed for each scene (> 2000 images in total). High severity fires were detected by applying a pre-defined threshold of the differenced Normalized Burn Ratio (dNBR) developed by Hall et al. [3] for boreal fires (see [4] for details).

In the summer of 2010, 34 transects of airborne lidar data were collected across the Canadian boreal by the Canadian Forest Service, in collaboration with the Canadian Consortium for Lidar Environmental Applications Research (C-CLEAR) and the Applied Geomatics Research Group [5]. These transects, which stretched from the Yukon to Newfoundland, totalled 25,000 km in length and were collected using an Optech ALTM 3100 discrete return sensor. The average pulse density was 2.8 returns/m^2 and the minimum swath width was 400 m. Approximately 4,500 km of these transects fall within the Boreal Shield West ecozone.

Calculation of lidar metrics

We calculated lidar metrics for each burned patch that intersected the lidar transects. We assessed structural complexity for each burned patch using four metrics (Table 1). First, we calculated rumple, which describes the roughness of the canopy surface and is calculated by dividing the area of the canopy surface by the area of the ground surface [6]. Second, we calculated the coefficient of variation (CV) of return height, which relates to the variability in structural elements within a canopy (i.e., the more variability in structural elements, the higher the CV, [7]). Third, we calculated the skewness of lidar returns, which could relate to the distribution of tree sizes within a stand (i.e.,high skewness would suggest un-even distribution of tree sizes). Finally, we

calculated the kurtosis of lidar returns, which could indicate if vegetation is dense at uniform height or spread over a wide vertical range.

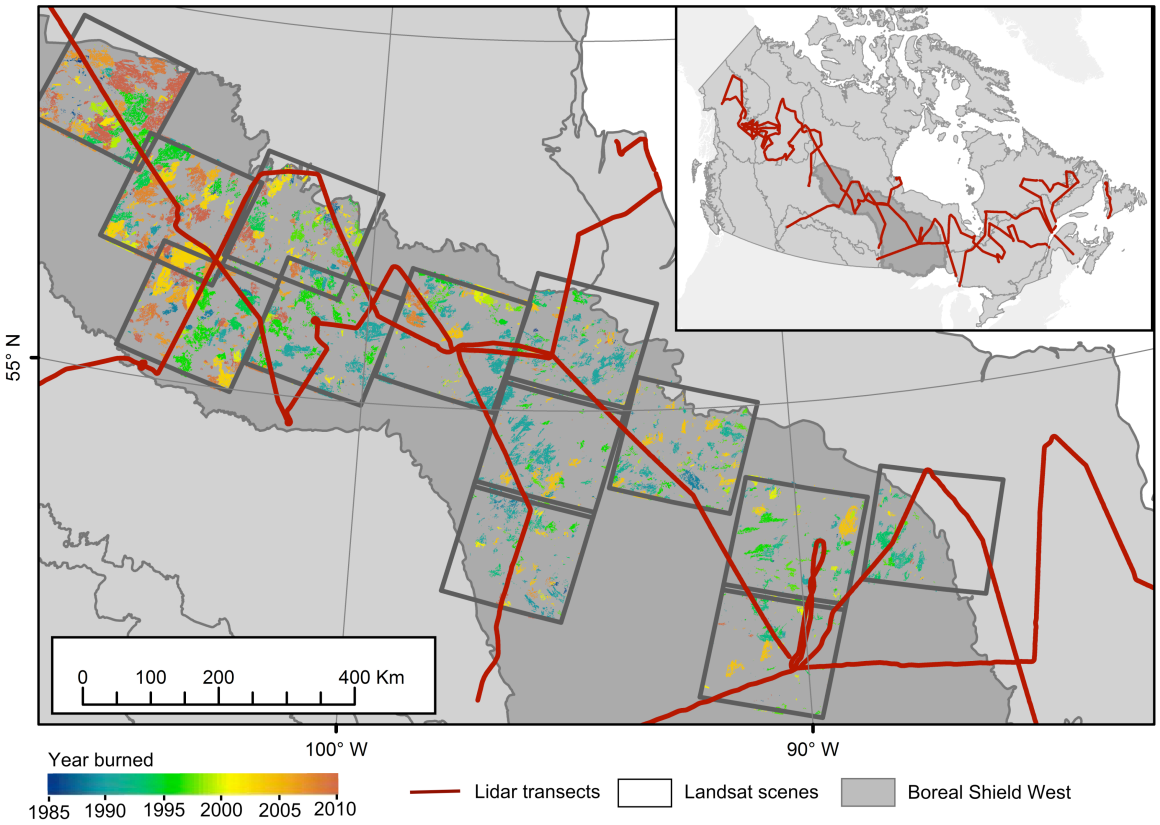


Figure 1: Fires detected across 13 Landsat scenes in the Boreal Shield West ecozone of Canada from 1985 to 2010. These scenes were selected to maximize overlap with airborne lidar transects collected in 2010.

Stratification of burned patches

To assess differences in structural development following fire based on pre-disturbance conditions, a classification tree was developed to stratify burned patches into dense (canopy cover > 50%), open (20 – 50%), or sparse (0 – 20%) forest based on pre-fire Landsat imagery. The classification tree was trained on estimates of canopy cover (percentage first returns > 2m) from the lidar transects. Once the classification tree was developed on circa 2010 Landsat imagery, it was applied to pre-fire imagery for each burned patch to stratify patches into dense, open, or sparse forest (see [4] for details). Due to the small number of sparse patches sampled, this analysis focuses on dense and open patches only. The classification was also applied to forest patches with no record of burning, allowing for a comparison between burned and unburned structure for the dense and open forest classes.

In total, lidar metrics were calculated for 234 dense and 372 open patches that burned between 1985 and 2010, and 629 dense and 957 open patches that had no record of burning.

Table 1:Common lidar metrics used to assess stand structural complexity

Lidar metric	Description
Rumple [6]	Roughness of the canopy surface
Coefficient of Variation	Variability in return height
Skewness	Symmetry of vertical distribution
Kurtosis	Peakedness of the vertical distribution

Results

The lidar metrics employed to assess structural complexity displayed markedly different trends as a function of years since fire (YSF). Rumple remained similar to patches with no record of burning in the first 10 YSF, but was significantly lower ($p < 0.001$) for patches at 10 – 25 YSF for both dense and open forests. Alternatively, CV was significantly lower ($p < 0.001$) in the first five YSF compared to unburned patches, but

increased to 10 – 15 YSF for both dense and open patches. After peaking at 10 – 15 YSF, the CV displayed a decreasing trend to 20 – 25 YSF. By 20 – 25 YSF, the CV was again significantly lower ($p < 0.001$) than in patches with no record of burning. Skewness and kurtosis displayed similar patterns after fire. Both metrics remained similar to unburned patches in the first ten YSF, and increased sharply for dense patches at 10 – 15 YSF. Skewness and kurtosis also increased for open patches, but the increase was more gradual than in dense forest patches. Skewness and kurtosis remained significantly higher for dense and open patches at 20 – 25 YSF compared to unburned forest patches.

Discussion

The area of burned forests sampled in this analysis (> 13,000 ha), and the range of structural responses observed, provides systematic and quantitative means to assess variability in structural response to fire across the Canadian boreal. Here, we have compared and contrasted a number of lidar metrics that inform on stand structural complexity. The varying patterns observed for each metric as a function of time since fire highlight the unique structural characteristics that each metric captures. Rumple remained high in the first ten YSF due to the large contrast in height between residual structures (i.e., surviving trees or snags) and the open areas created by fire. Rumple decreased sharply at 10-15 YSF as regenerating trees began to fill the open canopy space, leading to a young, homogenous stand structure. The results of CV were highly influenced by changes in the mean return height through time (i.e., CV is calculated as the standard deviation of return height divided by the mean of return height). Specifically, the mean height remained high in the first ten YSF, as lidar metrics were sensitive to residual canopy structures. Mean height decreased at 10 – 15 YSF as most returns were from the young, regenerating vegetation at this time. The CV was therefore low in the first ten YSF, when mean height remained high, and increased when mean height decreased at 10 -15 YSF. A slow increase in mean height from 10 -25 YSF led to the decreasing trend in CV. The finding that CV was significantly lower at 20 -25 YSF compared to unburned patches also suggests a young, homogenous stand structure. Alternatively, the sharp increase in skewness at 10 - 15 YSF highlights the contrast between young, regenerating trees and older surviving trees and snags. The sharp increase for kurtosis at 10 -15 YSF suggests that the majority of returns from the young, regenerating vegetation are from a uniform height.

While each metric provides unique information on stand complexity, the metrics tell a similar story. The first ten years following boreal fire are characterized by residual canopy structures, as regenerating trees have not refilled the canopy. At 10-15 YSF, lidar metrics become sensitive to new vegetation growth, which is characterized by young, homogenous vegetation. While these young stands are characterized by more homogenous structure than stands with no record of burning, high skewness suggests that residual structures remain in many of these stands. Through this analysis, we provide an efficient approach for characterizing structural response to disturbance using lidar structural metrics in combination with metrics derived from Landsat time-series data. The approach applied can improve our characterization of the impact of fires on boreal carbon budgets as well as inform management activities that aim to emulate natural disturbance events. Through this analysis, we provide an efficient approach for characterizing structural response to disturbance using lidar structural metrics in combination with metrics derived from Landsat time-series data.

References

[1]Kennedy, R.E., Yang, Z., & Cohen, W.B. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series : 1 . LandTrendr — Temporal segmentation algorithms. Remote Sensing of Environment 114, 2897–2910.

[2] Kane, V.R., North, M.P., Lutz, J.A., Churchill, D.J., Roberts, S.L., Smith, D.F., McGaughey, R.J., Kane, J.T., & Brooks, M.L. (2014). Assessing fire effects on forest spatial structure using a fusion of Landsat and airborne LiDAR data in Yosemite National Park. Remote Sensing of Environment 151, 89–101.

[3] Hall, R.J., Freeburn, J.T., de Groot, W.J., Pritchard, J.M., Lynham, T.J., & Landry, R. (2008). Remote sensing of burn severity: experience from western Canada boreal fires. International Journal of Wildland Fire 17, 476–489.

[4] Bolton, D.K., Coops, N.C., & Wulder, M.A. (2015). Characterizing residual structure and forest recovery following high-severity fire in the western boreal of Canada using Landsat time-series and airborne lidar data. Remote Sensing of Environment. In Press

[5] Wulder, M.A., White, J.C., Bater, C.W., Coops, N.C., Hopkinson, C., & Chen, G. (2012). Lidar plots — a new large-area data collection option: context, concepts, and case study. Canadian Journal of Remote Sensing 38, 600–618.

[6] Kane, V.R., McGaughey, R.J., Bakker, J.D., Gersonde, R.F., Lutz, J.A., & Franklin, J.F. (2010). Comparisons between field- and LiDAR-based measures of stand structural complexity. Canadian Journal of Forest Research 40, 761–773.

[7] Zimble, D.A., Evans, D.L., Carlson, G.C., Parker, R.C., Grado, S.C., & Gerard, P.D. (2003). Characterizing vertical forest structure using small-footprint airborne lidar. Remote Sensing of Environment 87, 171–182.

Spatiotemporal Patterns in Burn Severity and Post-Fire Recovery in Interior Alaska Using Multi-Sensor Airborne Data, Forest Inventory Plots, and Satellite Products

Bruce Cook[1], Douglas Morton[1], Hans Andersen[2], Robert Pattison[3], Andrew Finley[4], Lawrence Corp[5], Megan Kress[4], Chad Babcock[6], and Ross Nelson[1]

[1] *NASA Goddard Space Flight Center, Biospheric Sciences Laboratory, Greenbelt, MD, USA*
[2] *USDA Forest Service, Pacific Northwest Research Station, Seattle, WA, USA*
[3] *USDA Forest Service, Anchorage Forestry Sciences Laboratory, Anchorage, AK, USA*
[4] *Michigan State University, Department of Forestry and Geography, East Lansing, MI, USA*
[5] *Science Systems and Applications, Inc., Lanham, MD, USA*
[6] *University of Washington, School of Environmental and Forest Sciences, Seattle, WA, USA*

Highlights: Data from multi-sensor airborne imaging, forest inventory plots, and satellite estimates of burn rates and severity were used to characterize the spatial heterogeneity of post-fire vegetation succession in interior Alaska, where responses may vary due to fire severity, landscape position and environmental gradients.

Key words: *G-LiHT, boreal forest, fire severity, regrowth rates, data fusion, CWD*

Introduction

Wildfires are the most common forest disturbance in interior Alaska. Recent research suggests that fire severity has increased in interior Alaska in response to climate warming, and more intense fires may alter post-fire succession to favor deciduous stands rather than black spruce [1,2]. These disturbance-driven changes, if consistent at the regional scale, could accelerate expected shifts in vegetation cover and composition from climate warming alone based on the widespread impacts from fire in boreal forests.

Airborne remote sensing platforms offer a unique perspective on the spatial and temporal variability in fire effects, especially in large and remote regions like interior Alaska. The effect of fire on plant community composition and forest regrowth rates reflects a range of interacting processes: pre-fire fuels, synoptic climate conditions, fire history, and environmental factors such as topography, landscape position, soil properties, and depth to surface water. To assess post-fire vegetation succession in interior Alaska across a broad range of environmental gradients and spatial scales, we analyzed a combination of ground, airborne, and satellite data sources based on the time since fire. We used data from a 2014 joint NASA-USFS field campaign in the Tanana Valley Inventory Unit, AK, which included a systematic sampling of approximately 1 million ha using Goddard's LiDAR, Hyperspectral, and Thermal Airborne Imager (G-LiHT; http://gliht.gsfc.nasa.gov)[3]. These fine-resolution (1 m), multi-sensor image data allowed us to characterize the spatial heterogeneity of vegetation structure and composition for fires between 1950-2013. This space-for-time sampling with G-LiHT data was used to evaluate post-fire responses as a function of environmental variables, burn rates, and severity estimates from satellite observations [4-6].

Study Area and Data Analysis

Interior Alaska accounts for approximately 15% of all USA forestland, approximately 450,000 km^2 of boreal forest types. Most of the region is inaccessible from the existing road network, which limits ground-based forest inventories and research studies. Combining ground, airborne and space-based observations allows us to extend our understanding of post-fire spatiotemporal processes to larger regions where vegetation responses vary in response to landscape position and environmental gradients.

During 2014, the USFS funded a pilot study to support the development of a lidar-assisted forest inventory approach for interior Alaska, which supported the collection of approximately 100 new forest inventory plots and a systematic sample of the Tanana Inventory Unit (120,000 km^2) using NASA Goddard's LiDAR, Hyperspectral, and Thermal (G-LiHT) airborne imager [3]. Co-registered G-LiHT data provide a unique combination of high spatial (1 m) and spectral resolution data from LiDAR (vegetation structure), imaging spectrometer (vegetation cover, composition, productivity), and thermal sensors (edaphic conditions, including inundated and frozen soils). The complementary nature of these remote sensing observations offers the opportunity to develop new, multi-sensor fusion approaches to quantify changes in composition, cover, and productivity for this critical region. Individual trees and tree clusters were delineated by a simple watershed algorithm, followed by object based classification using G-LiHT canopy heights and at-sensor reflectance products. Dead and live trees were differentiated using a greenness vegetation index threshold, and vegetation

types were classified using Spectral Angle Mapper (SAM). Reference spectra for different vegetation types were obtained from inventory plots, stem map areas, and visual interpretation of fine-resolution aerial photographs (along-track photos from G-LiHT's 24 MP digital camera).

Wildfires burned more than 22% of forests in the Tanana since 1984. We used burned area perimeters and severity estimates in the Monitoring Trends in Burned Severity (MTBS) database [4], which is derived from pre- and post-fire Landsat image data beginning in 1984, and MODIS burned area (MCD64A1)[5] and burn rate [6] products for fires between 2000 and 2014. We will discuss relationships between recent burn rates and severity estimates and standing dead biomass, and spatial variations in vegetation composition and recovery due to environmental conditions and fire frequency.

Conclusions

Combining ground, airborne and space-based observations allows us to extend our understanding of post-fire spatiotemporal processes to larger regions where vegetation responses may vary due to fire severity, landscape position and environmental gradients. Fine-resolution airborne lidar, imaging spectrometer and thermal data is critical to separating live and dead trees; characterizing edaphic conditions (e.g., inundated and frozen soils); and separating spectral signals from trees, underlying shrubs and ground cover in boreal forests (Figure 1).

Figure 1: Both lidar and passive optical image data are needed to separate live and dead trees; classify vegetation types; and characterize fine-scale heterogeneity in boreal forest structure. Field plots and G-LiHT aerial photographs (shown here) were used to interpret G-LiHT lidar and imaging spectrometer data and assess post-fire plant succession across the Tanana Valley Inventory Unit, Alaska, USA.

References

[1] Kasischke, E. S., Verbyla, D. L., Rupp, T. S., McGuire, A. D., Murphy, K. A., Jandt, R., Barnes, J. L., Hoy, E. E., Duffy, P. A., Calef, M., & Turetsky, M. R. (2010) Alaska's changing fire regime - implications for the vulnerability of its boreal forests. Canadian Journal of Forest Research, 40, 1313- 1324.

[2] Johnstone, J. F., Hollingsworth, T. N., Chapin, III, F. S., & Mack, M. C. (2010) Changes in fire regime break the legacy lock on successional trajectories in Alaskan boreal forest. Global Change Biology, 16, 1281–1295.

[3] Cook, B. D., Corp, L. A., Nelson, R. F., Middleton, E. M., Morton, D. C., McCorkel, J. T., Masek, J. G., Ranson, K. J., & Ly, V. (2013) NASA Goddard's Lidar, Hyperspectral, and Thermal (G-LiHT) airborne imager. Remote Sensing, 5, 4045-4066.

[4] Eidenshink, J., Schwind, B., Brewer, K., Zhu, Z.-L., Quayle, B., & Howard, S. (2007) A project for monitoring trends in burn severity. Fire Ecology Special Issue, 3, 3–21.

[5] Giglio, L., Randerson, J. T., & van der Werf, G. R. (2013) Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4). Journal of Geophysical Research-Biogeosciences, 118, 317-328.

[6] Veraverbeke, S., Rogers, B. M., & Randerson, J. T. (2014) Daily burned area and carbon emissions from boreal fires in Alaska. Biogeosciences Discussions, 11, 17579-17629.

Estimation of tree stem attributes using ground based and airborne laser scanning

Johan Holmgren, Kenneth Olofsson, Mattias Nyström, Håkan Olsson

Swedish University of Agricultural Sciences (SLU), Department of Forest Resource Management, 901 83 Umeå, Sweden; E-mail: johan.holmgren@slu.se

Highlights: Multi-scan TLS data were geometrically fitted into a global coordinate system using matching of canopy height models derived from TLS and ALS measurements. Stem data from TLS data were then estimated and used as training for wall-to-wall predictions with explanatory variables derived on a tree crown level from ALS data.

Key words: *geo-referencing, stem diameter, TLS, ALS, imputation.*

Introduction

The airborne laser scanning technology now makes it possible to produce high precision tree stem estimates for large areas. However, manual field measurements are still used as reference data to translate airborne measurements of tree crowns to estimates of tree stem attributes. The objective of this study was to validate an automatic method for which tree stem estimates from airborne laser scanner (ALS) data are trained by ground based laser scanning (TLS).

Material and Methods

Study area

The study area was located in hemi-boreal forest at the Remningstorp estate in the south-west of Sweden (lat. 58° N, long. 13° E). The dominant tree species were Norway spruce (*Picea Abies*) and Scots pine (*Pinus Sylvestris*) and there were only small areas with birch (*Betula* spp.) and other deciduous trees.

ALS data

The test site was laser scanned with Riegl LMS Q680i during leaf on conditions in September 2014. The measurement density was on average 121 points per square meter. The field of view was 60 degrees but overlaps made it possible to only include data with a maximum scan angle of ±20 degrees. The flight altitude was approximately 440 m above ground level.

TLS data

The TLS data were collected in 40 forest stands. In each forest stand were 16 scanner stations placed in a regular grid with approximately 20 m internode distance to make sure that a large proportion of trees were scanned in a multi-scan-station setup from several directions. The scanner was a Trimble TX8 and the resolution was set to 11 mm at 30 m distance. The amount of data from only one of the 40 stands was approximately 1.5 billion laser reflections.

Geo-referencing of TLS data using canopy height models

The common problem with disturbed Global Navigation Satellite System (GNSS) signals when geo-referencing ground based laser scanning data obtained below dense tree canopies was solved using the following steps: (1) TLS data from all scan stations were rectified in a common local coordinate system using automatic detection of spheres placed at the positions of nearby scan stations, (2) an approximate position of each scan station was measured using GNSS, (3) a first translation and rotation of the TLS data to the global coordinate system based on GNSS data was performed with rigid body fitting using Singular Value Decomposition (SVD), (4) Canopy Height Models (CHM) were created from both the transformed TLS data and from the high resolution ALS data, (5) the CHM from TLS data was rotated and translated to find the best transformation by matching with the corresponding CHM from ALS data and calculation of cross correlation between the two rasters, and finally (6) the TLS returns were transformed using the best transformation found in the previous step according to the calculations of cross correlations.

Tree detection and stem estimation from TLS data

The TLS data were first normalized using a Digital Elevation Model (DEM) and then filtered in order to separate data from tree stems and from other objects, e.g. branches, using indicators of flatness in a local

neighbourhood. The next step was to create segments of individual tree stems or parts of tree stems and this step was followed by using a novel and robust algorithm for estimation of cylinders. The properties of the estimated cylinders were in a final step used to decide if separate cylinders should belong to the same tree stem.

Segmentation of tree crowns from ALS data

The segmentation of tree crowns from ALS data was based on a geometric tree crown model and consisted of four steps: (1) creating a height model (i.e., a raster), (2) creating a correlation surface, (3) segmentation of the correlation surface, and (4) merging segments. Details of the algorithm are described in [1] but a different function was in this study used to decide if a segment should be merged with a neighbour segment. Several metrics were derived from laser data within each tree crown segment to describe tree crown size and shape.

Imputation of tree stems

Tree crowns were imputed at a tree crown level using the segments derived from ALS data. The MSN method [2] implemented in the yaImpute package as part of the R project was used for the imputations. The sums of the variables derived from TLS, from zero, one, or several tree stems that had been linked to a segment were used as the dependent variables. The independent ALS variables used for the MSN imputation were height percentiles, average of heights, standard deviation of heights, crown diameter, and transformations of these variables. The crown diameter was calculated by counting the number of raster cells of the segment and calculating the diameter of a circle with the same area.

Validation

The tree stem diameters were measured 1.3 m above ground level (DBH) in two directions for all trees with a DBH $\geq$ 40 mm and a location within a 40 m distance from the plot centre. A device for ultrasonic based trilateration was integrated in the calliper and used to measure tree positions relatively 16 subplots and the callipered trees were automatically linked to TLS derived tree stem positions with global coordinates using an already developed algorithm [3]. There were two levels of results: (1) direct measurements of tree stem attributes using the TLS data, and (2) tree stem attributes estimated with imputation based on a combination of ground based and airborne laser scanning. If a tree was detected closer than 0.5 m from a scan station this tree was removed because it was probably a detection of one of the spheres used for registration of the point clouds into a common reference system. The estimations of stem attributes from imputations will be validated using leave-one-stand-out cross validation. Only results from TLS based tree stem estimations in one forest stand are presented here due to limited space and time for computations, however, further results will be presented at the conference.

Results

A distinct correlation peak appeared in the correlation images produced from the method used for matching the CHMs from TLS and ALS. The high resolution TLS data from 16 scan stations could be used to produce a CHM that was similar to the CHM obtained from ALS data.

The Geo-referencing of TLS data together with data matching of spatial tree pattern from TLS and manual field measurements made it possible to link TLS based estimated to manual measurements of individual trees without using a total station and without high precision GPS measurements. A high proportion of correctly linked trees could be verified by comparison of manual measured and TLS based estimates of DBH values. In total five outliers, with a TLS measured DBH > 150 mm and manual measured DBH < 100 mm, influenced the results to a high degree (Figure 1). If all linked trees in one stand were included, the RMSE for DBH estimates was 29 mm (15%) and the bias was 8 mm. If only the five outliers were excluded, the same RMSE was 11 mm (5%) and the bias was 5 mm. This tree level validation also included trees outside a 40 m radius plot.

The distributions of TLS detected trees and callipered trees were compared for a circular area with a radius of 40 m from the plot centre, an area of approximately 0.5 ha. For this area were 320 trees detected using TLS data and 260 trees were manually callipered. The histograms with a 20 mm bin size were compared. One error index, calculated as the sum of absolute deviations of observed and estimated number of trees in bins divided by total number of trees, was 0.28 (number of trees). Another error index, calculated as the sum of relative absolute differences in the bins, multiplied by 0.5 to only allow values in the range zero to one, was 0.12 (shape).

Tree crown segments derived from ALS data were linked to TLS derived tree stems to create training data in order to predict wall-to-wall stem data for any area on the test site. Predictions based on imputation techniques are usually improved if the size of the training data base is increased. There is a processing chain in place but a large number of hours for computing are needed. The plan is to use the High Performance Computing Centre North (HPC2N, https://www.hpc2n.umu.se/) in order to create the reference data that will be used for the imputation and then present the results at the conference.

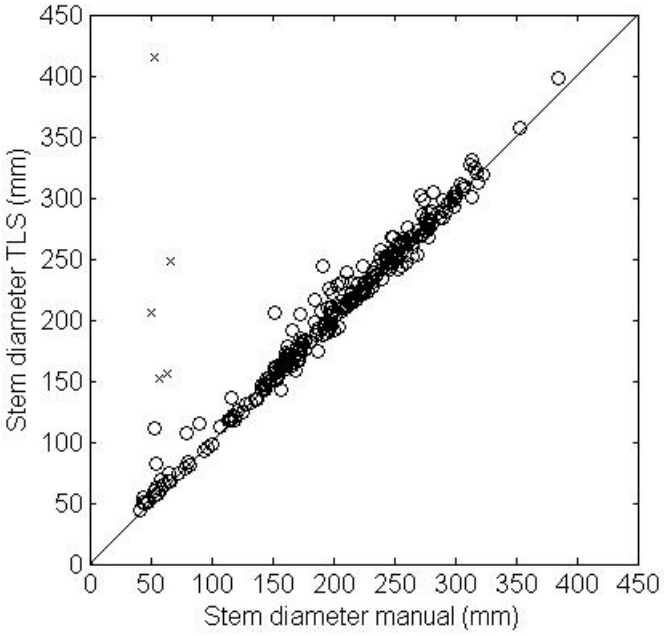


Figure 1: Stem diameter (DBH) estimated from TLS compared with 277 trees that were manually measured with a calliper and automatically linked to TLS, the x are outliers (<100 mm calliper and >150 mm TLS).

Discussion

The similarities between the CHMs from TLS and ALS inspired to the method of matching CHMs in order to complete the geo-referencing of TLS data. A distinct correlation peak indicated that this was a reliable method that works also if the GNSS positions have offsets of several meters. This is a method that could have a high practical value if ground based laser scanning systems are to be used for forest inventory because no long observation times are needed to receive high precision GNSS based positions. In fact in many forest stands an accurate position is very difficult to obtain due to multi-paths in the canopy. Instead of using matching of synthetic images created from tree locations as in [3] were CHMs from TLS and ALS matched to each other.

A new method for segmentation of tree stems and for estimation of tree stem diameter performed well compared with an earlier developed algorithm [4] for the same type of forest, especially for trees with small diameters (40-200 mm). However, an overestimation of stem diameters could be observed. An earlier developed algorithm [3] was used to automatically link TLS trees to manually measured trees. This algorithm was further developed to improve its performance. The validation of stem diameter estimations in one stand only included few outliers that could be because of incorrectly linked trees. The overestimation of stem diameter could be a result of branches although the algorithm in general successfully filtered out data from branches.

The TLS based estimates of stem position and stem diameter were robust and are therefore probably useful as reference data for the ALS based imputation of tree stem data. Imputation results with reference data for individual trees derived from TLS data in 40 stands will be presented at the conference.

Acknowledgements

This study was financed by Nils och Dorthi Troëdsson research foundation and ÅForsk foundation. The acquisitions of data were financed by Hildur and Sven Wingquist foundation for forest research.

References

[1] Holmgren, J. & Lindberg, E. 2013. Tree crown segmentation based on a geometric tree crown model for prediction of forest variables. Canadian Journal of Remote Sensing 39, 86-98.

[2] Moeur, M. & Stage, A. R. 1995. Most similar neighbor - an improved sampling inference procedure for natural-resource planning. Forest Science 41, 337-359.

[3] Olofsson, K., Lindberg, E. & Holmgren, J. 2008. A Method for Linking Field-Surveyed and Aerial-Detected Single Trees Using Cross Correlation of Position Images and the Optimization of Weighted Tree List Graphs. In SilviLaser 2008, 8th international conference on LiDAR applications in forest assessment and inventory, Edinburgh, U.K., 17-19 September 2008; Hill R.A., Rosette, J. & Suárez, J. (Eds.).

[4] Olofsson, K., Holmgren, J. & Olsson, H. 2014. Tree Stem and Height Measurements using Terrestrial Laser Scanning and the RANSAC Algorithm. Remote Sensing 6, 4323-4344.

Biochemical properties from dual-wavelength laser scanning: Deriving equivalent water thickness at leaf to canopy scales.

Rachel Gaulton[1], Steven Hancock[2], F. Mark Danson[3], Magdalena Smigaj[1].

1. School of Civil Engineering and Geosciences, Newcastle University, UK.
email:rachel.gaulton@ncl.ac.uk.
2. Environment and Sustainability Institute, University of Exeter, UK.
3. *School of Environment and Life Sciences, University of Salford, UK.*

Highlights: Improved measurements of canopy biochemistry can allow better monitoring of tree health. The SALCA instrument is able to estimate equivalent water thickness of individual leaves from dual-wavelength laser reflectance. This paper presents the results of a controlled drought experiment and field-based tests for retrieving moisture content at forest canopy scales.

Key words: *terrestrial laser scanning, canopy moisture, tree health, multispectral LiDAR.*

Introduction

Tree health is an area of growing concern internationally. In many countries, alterations in climate are likely to cause significant increases in temperature and increased frequency and severity of extreme droughts. Recent increases in tree mortality caused by drought suggest that climatic factors may already be resulting in forest die-back in some regions (e.g. southern parts of Europe) [1]. In addition, an increased rate of global occurrence and spread of non-native tree pathogens and pests is occurring and climate changes are leading to an alteration in host and pathogen distribution and host susceptibility [2]. Early detection is vital in reducing spread of infections and methods to detect early symptoms of disease (e.g. water stress, defoliation) and physiological stress will be of major benefit to forest management. The lack of adequate data on forest health status globally has been identified as a key information gap in understanding climate change risks for forests [1].

Estimation of leaf and canopy biochemical properties (e.g. leaf pigment and water concentrations) from remote sensing can provide a means to monitor tree health over extensive areas and at frequent intervals. Leaf moisture content, often measured as Equivalent Water Thickness (EWT), is an essential early indicator of forest drought stress, infection by tree diseases and forest pests and is of key importance as an input to models of forest fire susceptibility, ignition and propagation. Whilst a range of existing physiological approaches can be used to measure vegetation stress and EWT in the field, the scope of such approaches for examining temporal and spatial heterogeneity and for long-term monitoring is limited. A number of spectral indices based on near infrared (NIR) and shortwave infrared (SWIR) wavelengths have been developed that strongly relate to EWT and these have been widely applied to satellite data. However, accurate and appropriate validation measurements are crucial if such data are to be fully utilised. Measurements of subtle variations in biochemical parameters and plant health using passive optical sensors are also complicated by the dependence of reflection on illumination and viewing geometry, shadowing effects and the presence of other objects within the pixel (e.g. woody material, understorey vegetation, bare soil). Understorey and canopy response cannot be easily separated, even when passive optical data is fused with LiDAR.

Active reflectance measurements from laser scanners have the potential to overcome these issues, whilst simultaneously providing 3-D information on forest structure and health. Experimental dual-wavelength or multispectral systems have been shown to be capable of assessing crop foliar nitrogen levels [3] and monitoring changes in needle Chlorophyll content over time [4]. This paper examines the ability of a dual-wavelength laser scanner (the Salford Advanced Laser Canopy Analyser, SALCA) to detect changes in leaf and canopy water content due to drought stress and disease. This is tested through a series of experimental and field-based studies from leaf through to canopy scales. The scanner technology and calibration is first described, followed by experimental tests of the instrument at leaf then small canopy scales. Finally applications in real forest canopies are considered.

The Salford Advanced Laser Canopy Analyser

The SALCA instrument is a dual-wavelength terrestrial laser scanner, operating at NIR (1063 nm) and SWIR (1545nm) wavelengths. The system records the full-waveform of backscattered energy in 15cm range bins, with a maximum range of 105m and acquires data over a full hemisphere above the instrument, generating 9.6 million waveforms in a full resolution (1.05 mrad angular resolution) scan. Full details of the instrument can be found in [5]. Estimates of return energy in each wavelength can be extracted from the waveform for individual targets.

For large solid targets, the averaged peak intensity measured provides a sufficient estimate for calibration to apparent reflectance [6]. However, for typically diffuse returns from vegetation canopies, the sum of digital numbers recorded over all bins comprising the return provides a more accurate and consistent measure of return energy [7] and is used here for canopy-level processing. Return energy is then calibrated to apparent reflectance in each wavelength, based on the range of the return and the deployment of calibration boards with panels covering a range of reflectance in the scan, to account for variations in laser output energy (largely due to temperature variations).

Experimental leaf and canopy scale measurements

In laboratory experiments in which individual leaves of three different species were dried and repeatedly scanned over a period of time, a significant linear relationship was obtained ($R^2 = 0.8$, RMSE = 0.0069 g cm^{-2}) between a normalised difference index (NDI) of the two SALCA wavelengths (1063 nm and 1545 nm) and the EWT of the leaves [6]. SWIR reflectance increased significantly as the leaves dried. A small increase in NIR (1063nm) reflectance was also observed from dry leaves. Normalised Difference Index from SALCA also varied significantly at a within-leaf scale, suggesting fine-scale spatial variability in EWT and / or leaf structure [6].

These results indicate potential for retrieval of EWT from active dual-wavelength laser scanner measurements. However, to apply the method at the canopy scale introduces a number of challenges. Woody material within the laser footprint is likely to influence the return intensity when working at a canopy scale, whilst the ability to detect biochemical properties of partially occluded leaves (partial laser returns) has not been experimentally tested. To investigate the potential to measure EWT at canopy scales, a controlled 'dry-down' experiment was conducted in June to July 2013. 15 potted small-leaved lime (*Tilia cordata*) trees (approximately 4m tall) were exposed to varying degrees of drought stress over a 1 month period. The pots of two groups were sealed, whilst control groups were regularly watered. Trees were scanned with SALCA several times per week and leaf spectral reflectance, stomatal conductance, Chlorophyll content and soil moisture were measured at the time of laser scans. A subset of trees were used to provide destructive samples of leaf EWT throughout the experiment. Return energy was retrieved from waveform data and calibrated to apparent reflectance as described above.

From field spectroscopy measurements of leaves, the normalised difference index (NDI) equivalent to that calculated by SALCA ($(\rho_{1063}-\rho_{1545})/(\rho_{1063}+\rho_{1545})$, where ρ is apparent reflectance) stayed relatively constant or increased for the control group trees (mean value ranging from 0.29 at the start to 0.34 at the end) but decreased for trees subjected to drought stress (from 0.33 to 0.18 for drying group 1 and 0.3 to 0.18 for drying group 2) over the course of the experiment. The extent of variability in the spectral index between leaves also increased as drought stress increased, suggesting leaves within the same canopy dried at different rates. Processing of SALCA data is ongoing to obtain relationships with EWT and to examine the extent of variability in the SALCA NDI within the canopies and these final results will be presented at the conference. As with laboratory measurements, the 1545nm apparent reflectance can be seen to increase markedly in the drying groups by the end of the experiment compared to the watered control, whilst the 1063nm reflectance increases to a lesser extent (Figure 1).

Forest canopy scale measurements

In real forest canopies, occlusion will play a greater role and differences in range and scanning angle across the scans will be greater, introducing increased variation in return energy, whilst differences in biochemistry (e.g. EWT) within and between canopies may be more subtle than under experimental conditions. Many UK forests are also coniferous plantations, and the presence of needles rather than broad-leaves will increase the number of partial returns from objects only partly occupying the laser footprint. As yet the ability of dual-wavelength or multispectral terrestrial laser scanners to detect variation in leaf biochemistry within mature forest canopies in field conditions is untested. To examine the potential to detect variations in EWT in forest canopies, four SALCA scans were acquired in forest plots in the Queen Elizabeth II Forest Park, Aberfoyle, Scotland during summer 2014. The scans took place in four Scot's pine (Pinus sylvestris) and Lodgepole pine (*Pinus contorta*) stands with contrasting severities of *Dothistroma septosporum* infection, a fungal infection of which one symptom is decreases in needle moisture content. Trees were visually surveyed to record disease levels and limited needle samples were analysed for EWT and spectral reflectance (measured with an ASD FieldSpecPro field spectrometer). This paper will present the preliminary results of this forest-canopy-scale analysis, comparing the calibrated canopy NDI to observed disease levels.

Conclusions

Active reflectance measurements from a dual-wavelength terrestrial laser scanner allow measurement of equivalent water thickness of individual leaves in laboratory conditions. At canopy scales, a number of complicating factors, such as the presence of occlusion, partial hits and mixed returns from woody material and foliage may complicate the retrieval of biochemical parameters from laser return energy. Despite these factors, distinct changes in apparent reflectance can be detected in trees subjected to drought stress under experimental

conditions. This paper presents initial results of this experiment and of trials of dual-wavelength laser scanning for detection of stress resulting from disease in forest canopies measured in field conditions.

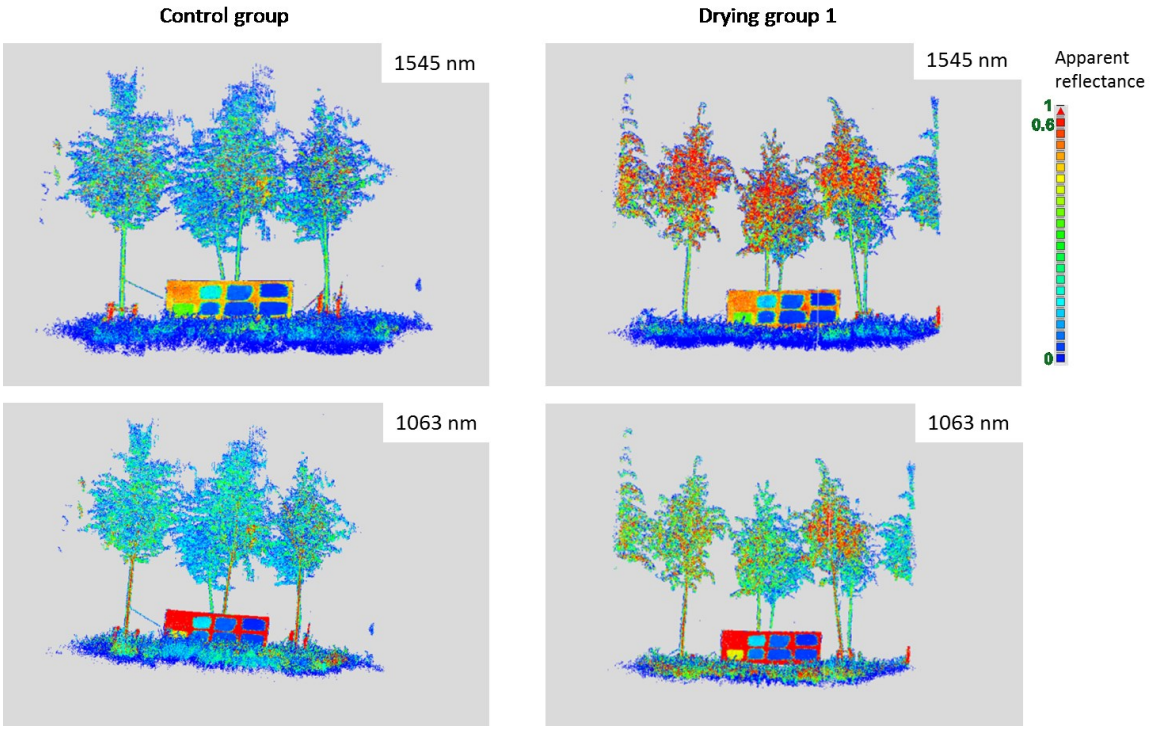


Figure 1: Comparison of apparent reflectance in the control group (watered) and drying group 1 (drought stressed) on day 21 of the experiment for 1545 nm wavelength and the 1063 nm wavelength of SALCA.

Acknowledgements

The research presented was funded by a Natural Environment Research Council grant (NE/K000071/1) and a Royal Society Research grant (RG130211). Field data for Aberfoyle was collected under a NERC PhD studentship (NH/104540921). Thanks are also due to Ali Jahangiri, Elias Berra and Maria Peppa for assistance with field data collection.

References

[1] Allen, C.D. *et al*. (2010). A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. Forest Ecology and Management, 259, 660-684.

[2] Sturrock, R.N., Frankel, S.J., Brown, A.V., Hennon, P.E., Kliejunas, J.T., Lewis, K.J., Worrall, J.J. & Woods, A.J. (2011). Climate change and forest diseases. Plant Pathology, 60, 133-149.

[3] Eitel, J.U.H., Magney, T.S., Vierling, L.A. & Dittmar, G. (2014). Assessment of crop foliar nitrogen using a novel dual-wavelength laser system and implications for conducting laser-based plant physiology. ISPRS Journal of Photogrammetry and Remote Sensing, 97, 229-240.

[4] Hakala, T., Nevalinen, O., Kaasalainen, S. & Mäkipää, R. (2015). Technical Note: Multispectral lidar time series of pine canopy chlorophyll content. Biogeosciences, 12, 1629-1634.

[5] Danson, F.M., Gaulton, R., Armitage, R.P., Disney, M., Gunawan, O., Lewis, P., Pearson, G. & Ramirez, A.F. (2014). Developing a dual-wavelength full-waveform terrestrial laser scanner to characterize forest canopy structure. Agricultural and Forest Meteorology, 198-199, 7-14.

[6] Gaulton, R., Danson, F.M., Ramirez, F.A. & Gunawan, O. (2013). The potential of dual-wavelength lser scanning for estimating vegetation moisture content. Remote Sensing of Environment, 132, 32-39.

[7] Hancock, S., Armston, J., Li, Z., Gaulton, R., Lewis, P., Disney, M., Danson, F.M., Strahler, A., Schaaf, C., Anderson, K. & Gaston, K.J. (2015). Waveform lidar over vegetation: An evaluation of inversion methods for estimating return energy. Remote Sensing of Environment, In Press.

Estimating leaf bulk density distribution in a tree canopy using calibrated t-LiDAR indices

François Pimont[1], Jean-Luc Dupuy[1], Eric Rigolot[1], Vincent Prat[1], Alexandre Piboule[2]

[1]*INRA, UR 629, Domaine de Saint Paul, Agroparc F-84914 Avignon Cedex 9, France.*
[2]*ONF, Département RDI-Pôle Nancy 11 rue de l'île de Corse 54000 Nancy, France.*

Highlights: We propose a method to estimate 3D leaf bulk density distribution, based on the calibration of density indices derived from T-LiDAR over small spherical volumes in which vegetation was sampled. The LiDAR-based bulk density distribution compared well to the inventory-based profiles in the upper part of the canopy.

Key words: *Forest fuel; Quercus pubescens; Mixed pixel; Filtering; Wavelength; Leaf/Wood.*

Introduction

Leaf biomass distribution is a key factor for modeling energy and carbon fluxes in forest canopies and for assessing fire behaviour. The most accurate method for estimating the biomass of a forest canopy is to physically sample it, but this is not feasible on large samples of trees. At forest plot scale, a common method, hereinafter referred to as the inventory-based approach, is to combine a stem inventory and allometric equations for biomass and its vertical cumulative distribution in order to estimate leaf load and bulk density profile [1]. However, the inventory method is time-consuming, so that remote sensing techniques have long been used in forest management and research, especially for estimating leaf area index. These techniques include terrestrial LiDAR (Light Detection And Ranging), referred to hereinafter as TLS, which emerged as a promising tool for leaf biomass or surface area estimation because the distribution of returns can be used to estimate density indices in voxels [2, 3], based on the probability of light transmission. The indices can be specifically computed for the leaves since the intensity can be used to distinguish leaf from wood returns [3, 4]. These direct measurements have been used successfully at tree scale but require voxelization of the scanned scene at a relatively high resolution [4] and multiple scans for a single tree. This is time-consuming in terms of measurements and post-processing, and this may be considered to be a limitation for operational use.

Herein, we propose a method to estimate the spatial distribution of leaf bulk density from TLS scans. It is based on the development of indices similar to those used at tree scale in earlier direct estimation studies. Here, the indices are computed in spherical volumes that are larger than the small voxels employed in earlier studies, and a field calibration is used to provide a less time-consuming but still robust estimation of the distribution at plot scale.

Material and methods

Plot description and inventory-based method

Four 12 m diameter contrasted plots were selected in a *Quercus pubescens* forest in the South-East of France. Maximum heights varied between 8 to 12 m and basal areas between 18 and 40 $m^2\,ha^{-1}$. A stem inventory was carried out and 10 trees of various diameters at breast height were felled and cut in 1-m vertical sections. Leaves and roundwoods were collected, oven-dried and weighted. Data was used to fit an allometric equation for leaf biomass and a cumulative distribution of biomass. The combination of these equations and the stem inventory in each plot were used to estimate leaf bulk density profiles in each plot (inventory-based method).

Calibration volumes

In each plot, ten polystyrene balls (diameter 0.1 m) were placed at different locations in the canopy. The balls were used to mark out the center of the Calibration Volumes (CV), which are virtual spherical volumes bounded by the 0.7 m diameter sphere that has the same center as the polystyrene target. Once the TLS scans had been performed on a plot, the leaves inside the calibration volume were collected. The leaves were then oven-dried at 60°C for 24 hours and weighted.

TLS instruments and measurements

A FOCUS 3D 120 S (FARO Technologies Inc., Lake Mary, USA) TLS instrument was used in this study. It is a phase-shift LiDAR that emits modulated laser beam pulses at 905 nm in a 305° (vertical) x 180° (horizontal) view window. The FOCUS 3D has two material filters on its scanner: "Clear Sky" and "Clear Contour", that filters "sky" points and mixed pixels. Five scans were performed on each plot from the center and from the four summits of the square inscribed in the circular plot. Resolution angle was 0.036° (43822090 pixels per scan) and measurement speed "x4" (488000 pixels s^{-1}, which is 2^4 = 16 fold minimum scanning speed), as recommended for outdoor scanning. Six spherical artificial targets were placed in the scene.

Biomass indices in spherical volumes

Several density indices were computed, according to the number of returns from a spherical volume (N_i), the number of beams emitted by the LiDAR in the cone emerging from the scanner and tangent to the spherical volume (N_t) and the number of returns intercepted by vegetation between the scanner and the spherical volume (N_b). The fraction of point filtered by clear contour f_{cc} and the G function to account for leaf orientation were also incorporated to the index I_3:

$$I_3 = \frac{N_i\left(1+f_{cc}\right)}{N_t - N_b\left(1+f_{cc}\right)}\frac{1}{2G(\theta)}$$

The density indices were fitted with the bulk density measured on the field. In order to predict leaf bulk density, N_i should account for leaf returns only. However, probably because of the wavelength of our LiDAR scan, it was not possible to distinguish leaf from wood returns on the basis of the return intensity value, as it was the case in other studies [3, 4], so that N_i account for all returns.

Model application

The calibrated indices were then used to estimate the bulk density at all points of a grid in each plot. The grid cell was 0.7 m. For each grid point, the scan that has the best "view point" on this grid point (defined as the scan with the highest Ni) was selected for estimation of biomass.

The 3D distribution of estimated bulk density was integrated horizontally to assess vertical bulk density profile that could be compared to the ones obtained with the inventory-based method.

Results

The calibration process entailed to establish that the bulk density in a 0.7 m diameter sphere can be estimated by 0.0727 I_3 (kg m^{-3}) with a determination coefficient R^2=0.73. I_3 performed better than simpler indices, because it accounts for both mixed pixel filter and sky point filter. The results also demonstrated that the leaf orientation was of secondary importance with this plant type.

The model was applied to the four plots. LiDAR-based vertical profiles of bulk density compared well to inventory-based profiles in the upper part of the canopies (Fig. 1). The canopy heights and the levels of the peak bulk density were well captured. However, results were less satisfactory in the lower part of the canopy where the LiDAR-based method overestimates bulk densities, because the model interprets wood and trunk returned echoes as if they were foliage.

Discussion

Method performance

The calibration process employed yielded encouraging results. The determination coefficient was relatively high for index I_3, thus justifying the assumptions made for the processing of noise, occlusion, foliage orientation and proportionality between leaf bulk density and index even without leaf and wood separation. The calibration was obtained over a broad range of bulk density (0.039 - 0.31 kg m^{-3}).

Determination coefficient during the calibration process and agreement between LiDAR and inventory-based bulk density profiles would have probably been much better if it was possible to restrict N_i to leaf returns only.

Benefits and drawbacks of the calibration approach

Leaf area density at plant scale has in the past been estimated directly by several authors [3, 4], but these methods require to compute statistical indices on small voxels, so that a high resolution of scanner and a high scan number are required. The present approach computes statistics on much larger volume (200 times bigger), so that the point density required computing statistics is much lower. It could also, potentially, be applicable to conifers, unlike the direct method which struggles to cope with needle morphology (very thin, aggregated needles oriented at shoot scale). The cost of the calibration procedure is a limitation, especially if the calibration coefficients show poor stability. However, we expect these coefficients to be relatively stable as they depend

only on the distribution of foliage element at the spherical volume scale, that should not change much among plots. For a given species (or group of species with similar morphologies), we can expect variations to be limited and potentially predictable from foliage characteristics, such as SLA.

Conclusion

The method presented here is promising and has potential to become an efficient, operational methodology to estimate leaf distribution. It could be used in combination with airborne and space-borne remote sensing (that often requires ground measurements for calibration), for monitoring of ecosystem, or to provide data for process-based models (functioning or fire).

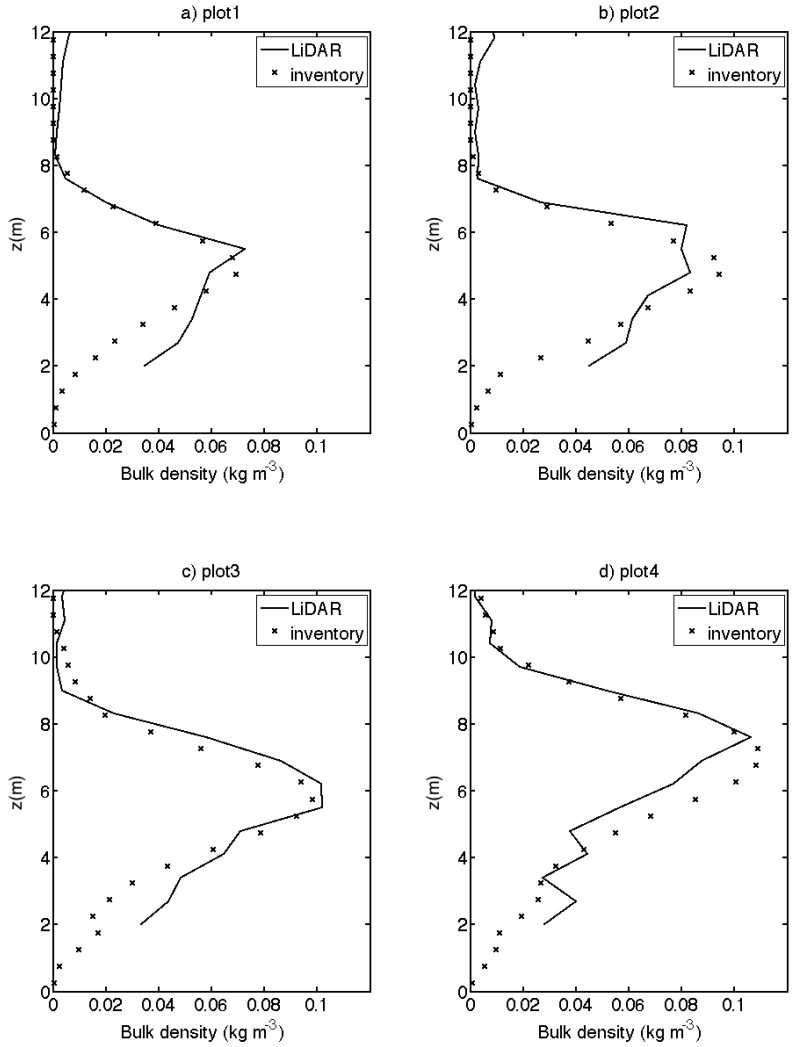


Figure 1: Leaf bulk density profiles estimated from TLS- and inventory-based methods

References

[1] Baldwin, V.C.Jr., Peterson, K.D., Burkhart, H.E., Amatais, R.L. & Dougherty, P.M. (1997). Equation for estimating loblolly pine branch and foliage weight and surface area distributions. *Canadian Journal of Forest Research 27*, 918–927.

[2] Durrieu, S. Allouis, T., Fournier, R.A., Véga, C., Albrech, L. (2008). Spatial quantification of vegetation density from terrestrial laser scanner data for characterization of 3D Forest structure at plot level. In SilviLaser 2008, Edinburgh, UK, 17-19 September 2008.

[3] Béland, M., Widlowski, J.-L., Fournier, R.A., Côté, J.-F., Verstraete, M. (2011). Estimating leaf area distribution in savanna trees from terrestrial LIDAR measurements. *Agricultural and Forest Meteorology 151*, 1252-1266.

[4] Béland, M., Baldocchi, D.D., Widlowski, J.-L., Fournier, R.A., Verstraete, M.M. (2014). On seing the wood from the leaves and the rôle of voxel size in determining leaf area distribution of forests with terrestrial LiDAR. *Agricultural and Forest Meterorology 184*, 82-97.

Changes in canopy and crown structure with stand composition

Martin-Ducup O.[1], Schneider R.[1] and Fournier R.[2]

[1] *Chaire de recherche sur la forêt habitée, Département de biologie, chimie et géographie, Université du Québec à Rimouski (UQAR), 300 allée des Ursulines, Rimouski (QC), G5L 3A1, CANADA*
[2] *Centre d'Applications et de Recherches en Télédétection (CARTEL) Université de Sherbrooke, 2500, boul. De l'Université, Sherbrooke (QC), J1K 2R1, CANADA*

Highlights: Terrestrial laser scanning (TLS) data was used to quantify sugar maple crown metrics and local environment structure indices. It allows quantifying how crown plasticity, potential tree light interception and local environment structure are influenced by stand composition (pure vs mixed).

Key words: *TLS; Competition; Crown structure; Plasticity; Sugar maple.*

Introduction

The forest canopy structure can be defined by the position, size and shapes of each tree in a stand [1]. It is necessary to study the variations in individual tree crowns in order to understand canopy structure and its dynamics. The architecture of a tree at a given time is the result of a complex interaction between intrinsic (e.g. genotype) and external constraints (e.g. the environment) [2]. In fact, the most important environmental factor that influences tree crown structure in forests is the competition for light. This competition is usually quantified using the surrounding trees' crown dimension, size (height and diameter) and species [3]. TLS data could be used as a proxy to quantify the local environment. For instance it can provide metrics related to the shade that a given tree is subject through the woody and leafy material arrangement. The local environment varies widely according to species-specific crown shape and light acquisition strategies, and should thus change between pure and mixed stands.

Our objectives were to quantify with TLS data (i) the differences in the local environments of sugar maple trees in pure and mixed stands, (ii) the variability of crown structure between these two types of stands and (iii) the relationships between the local environment and crown structure.

Material and methods

A total of 36 co-dominant sugar maples (*Acer saccharum* Mill.) and their local environment were scanned (4 scans per tree) using a Faro Focus 3D TLS in Eastern Canada. Half of those trees were scanned in pure stands and the other half in mixed stands. We used voxelized clouds (1 point per non-empty voxel) instead of the raw point clouds in order to avoid distance to scanner bias and limit occlusion bias. Target trees were semi-manually isolated and separated from the rest of the voxelized cloud. A module in the Computree software was used to perform these isolations [4]. The principle is to select cells belonging to the tree on a 2D XY grid (i.e. a raster) computed on vertical slices of the voxel cloud. The operator starts the selection at the base of the tree by choosing the size of the slice, then selects cells belonging to the tree and repeats this step for each slice along the Z axis until the top of the tree. For each slice, delineation of the crown of a target tree is based on the changes in density of the raster that appears at the border of the crown. The local environment structure was calculated using the part of the voxelized cloud which was included in a 60° angle reverse cone starting at the crown base of the target tree (Figure 1a and b). Two indices were thus used i) a competition pressure index (CPI) based on the density of points, their height and their distance to the target tree and ii) a competition homogeneity index (CHI) calculated using the Clark and Evans aggregation index on the XY projection plan. Then, 5 crown metrics were computed to estimate the crown structure using R software (Figure 1b and c) (1) crown ratio (2) material density inside the crown (3) crown openness, defined as the angle between the axis of symmetry and the vector between the crown base point and the most external point of the crown in the XY plane (4) sinuosity, calculated as the cumulative asymmetry of polygons fitted every 10 cm along the z axis (5) volume filled, quantified by the total crown volume. CHI, CPI and crown metrics were compared between pure and mixed stands using analyses of variance. The variability of the CHI, CPI and crown metrics according to stand type was investigated using mixed effect models.

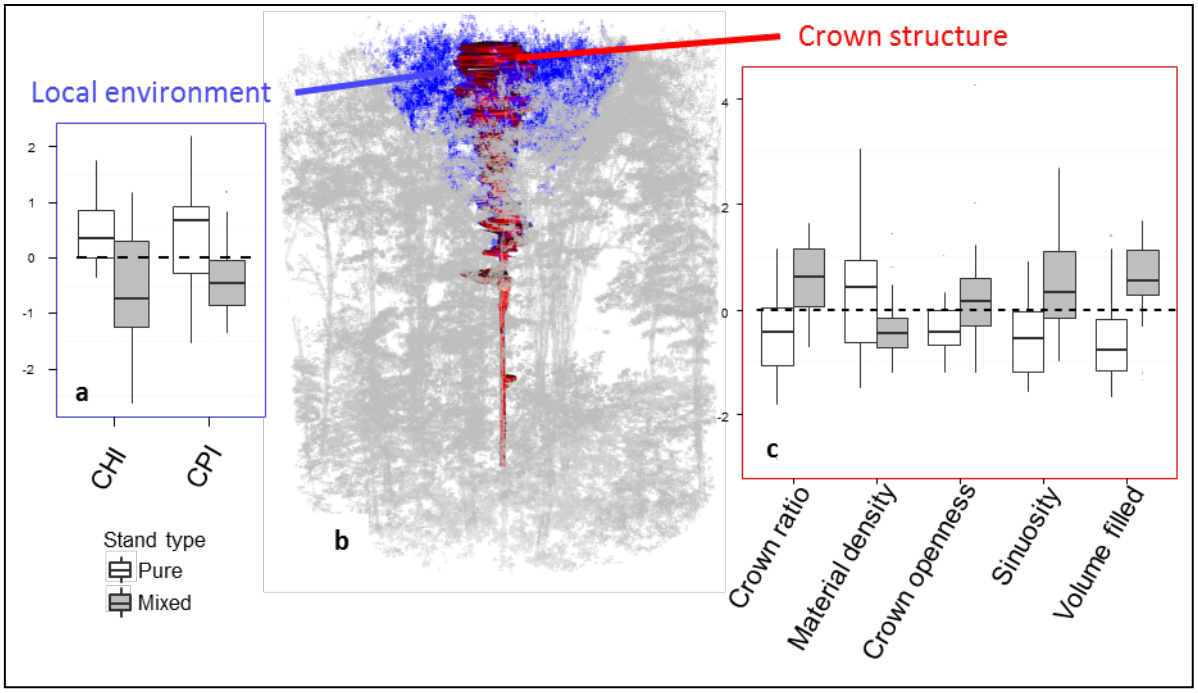


Figure 1: a) Local environment variability between pure and mixed stands b) TLS data of a target tree (red) and its local environment (blue) c) Crown metric variability between trees in pure and mixed stands

Results and discussion

Local environment variability

The CHI and CPI are lower in mixed stands (Figure 1a). The competition pressure is thus less important and the distribution of the material more clustered around trees in mixed stands. Sugar maple is a very shade tolerant species with a high light extinction coefficient [5], which could explain a higher CPI in pure stands. Similar results were observed for the shade tolerant species European beech [6]. The higher material densities inside the tree crowns observed in pure stands also supports this result. The lower CHI value in mixed stands can be explained by the high variability in crown shapes and strategies to acquire light, making the canopy more heterogeneous. Moreover, this variability and thus the potential crown trait complementarity for light acquisition could reduce the competition pressure in mixed stands when compared to pure stands.

Crown structure and implications on light interception

Statistically significant differences between stand composition (i.e. pure versus mixed) of the five studied crown metrics were observed (Figure 1c). In decreasing order of importance, trees have a higher crown ratio, filled more volume, have a higher sinuosity, a more open crown, a lower within-crown material density in mixed stands when compared to pure stands. These results should provide advantages to trees in mixed stands for increased light interception. Indeed, an increase in crown volume obviously increases the potential light intercepted by increasing the leaf area index. Moreover, higher crown openness probably increases the light penetrability in the lower part of the crown. Similar results were observed for European beech with a higher crown volume and flatter branch angles (which is similar to our crown openness) for trees in mixed stands when compared to pure stands [7]. The lower material density also suggest a more light penetrable crown in mixed stands, meaning that self-shading and self-pruning will be lower. The observed higher crown ratio in mixed stand supports this interpretation. Finally, the sinuosity is more important in mixed than in pure stands. This metric can be interpreted as the capacity of a tree to deviate from its symmetry axis in order to avoid neighbours and intercept more light.

Crown plasticity

All the crown metrics we quantified varied widely in response with stand type (Figure 1c), but also with the local environment (data not shown in the abstract). These results reveal a high sugar maple crown plasticity which could be a consequence of the feedback loop " Environment -> Growth -> Structure -> Environment" operating in forest ecosystems [7], i.e. the environment of a tree affects its growth which influences its crown structure which in return changes the canopy structure.

Conclusion

TLS data allows us to quantify local environments metrics reflecting the arrangements of material around a tree and to observe its impact on crown structure. Moreover, the three dimensional aspect of the crown structure metrics we used allow many functional interpretations, such as light interception. Validation of the tree isolation method is foreseen, as it is the basis of our results. Finally, our results could be used to further explain the potential benefits of mixed forests, at least in term of crown structure and light interception.

References

[1] E. R. Lines, M. A. Zavala, D. W. Purves, and D. A. Coomes, "Predictable changes in aboveground allometry of trees along gradients of temperature, aridity and competition," *Glob. Ecol. Biogeogr.*, vol. 21, no. 10, pp. 1017–1028, 2012.

[2] D. Barthelemy and Y. Caraglio, "Plant architecture: A dynamic, multilevel and comprehensive approach to plant form, structure and ontogeny," *Ann. Bot.*, vol. 99, no. 3, pp. 375–407, 2007.

[3] G. S. Biging and M. Dobbertin, "A Comparison of Distance-Dependent Competition Measures for Height and Basal Area Growth of Individual Conifer Trees," *For. Sci.*, vol. 38, no. 3, pp. 695–720, 1992.

[4] "Computree," *RD innovation ONF*. [Online]. Available: http://computree.onf.fr.

[5] C. D. Canham, A. C. Finzi, S. W. Pacala, and D. H. Burbank, "Causes and consequences of resource heterogeneity in forests: interspecific variation in light transmission by canopy trees," *Can. J. For. Res.*, vol. 24, no. 2, pp. 337–349, 1994.

[6] J. Metz, D. Seidel, P. Schall, D. Scheffer, E.-D. Schulze, and C. Ammer, "Crown modeling by terrestrial laser scanning as an approach to assess the effect of aboveground intra- and interspecific competition on tree growth," *For. Ecol. Manag.*, vol. 310, pp. 275–288, décembre 2013.

[7] D. Bayer, S. Seifert, and H. Pretzsch, "Structural crown properties of Norway spruce (Picea abies [L.] Karst.) and European beech (Fagus sylvatica [L.]) in mixed versus pure stands revealed by terrestrial laser scanning," *Trees*, vol. 27, no. 4, pp. 1035–1047, Aug. 2013.

Reducing uncertainties in above-ground biomass estimates using terrestrial laser scanning

Kim Calders[1,2], Andrew Burt[2], Glenn Newnham[3], Mathias Disney[2], Simon Murphy[4], Pasi Raumonen[5],Martin Herold[6], Darius Culvenor[7], John Armston[8,9], Valerio Avitabile[6], Mikko Kaasalainen[5]

[1]*National Physical Laboratory, UK; kim.calders@npl.co.uk*
[2]*University College London; and NERC National Centre for Earth Observation (NCEO), UK*
[3]*CSIRO Land and Water, Australia*
[4]*University of Melbourne, Australia*
[5]*Tampere University of Technology, Finland*
[6]*Wageningen University, The Netherlands*
[7]*Environmental Sensing Systems, Australia*
[8]*Queensland Department of Science, Information Technology, Innovation and the Arts, Australia*
[9]*University of Queensland, Australia*

Highlights: This work compares nondestructive estimates of AGB made using terrestrial laser scanning (TLS) with destructively harvested measurements. We show that estimates using our TLS approach agree to within 10% of the reference measurements. More traditional AGB estimates based on allometric equations show an underestimation of approximately 30%. Moreover, uncertainty in the TLS-derived estimates can be quantified accurately and explicitly, unlike that from allometric methods.

Key words: *Terrestrial LiDAR; biomass; destructive harvesting; tree reconstruction*

Introduction

Above-ground biomass (AGB) is a good indicator for forest productivity, carbon storage and sequestration of forests. The global distribution and magnitude of forest carbon sources and sinks remain uncertain and, as a result, carbon emissions based on these poorly-constrained stock estimates are even more uncertain, and variable across different scales. Remote sensing can be used to monitor AGB and carbon emissions at a large scale. Recent work of [1] and [2] used similar input data to generate pan-tropical satellite derived AGB maps. [3] reported substantial differences in mapped AGB between these two maps and spatial patterns were different when compared to field data distributed across the region. These disagreements demonstrate that accurate ground data are needed to improve the calibration and validation of global satellite derived AGB datasets.

Traditional field methods generally infer AGB based on indirect, empirical allometric (size-to-mass) relationships with tree parameters, such as tree height and diameter at breast height (DBH). [4] identified four types of uncertainty in AGB assessment using traditional forest inventory data: (i) errors in tree inventory, (ii) errors in the allometric equations, (iii) errors related to the size of the sampling plots, and (iv) errors related to the landscape-scale representation of these sampling plots. Terrestrial laser scanning (TLS) offers opportunities for a consistent and robust framework to support AGB estimates using 3D point clouds. TLS data has the potential to reduce errors (i) and (ii) by improving the traditional field methods to estimate AGB. It can also provide insight into errors (iii) and (iv), which are related to the spatial variance of the forest and its structure. We demonstrated in [5] that TLS inferred estimates of AGB agree better with 65 destructively sampled trees than traditional allometric equations. Our TLS approach uses quantitative structure models (QSMs) to estimate volume and showed a total AGB overestimation of 9.7% compared to an underestimation of 36.6 to 29.9% for the allometric equations. A key finding in [5] is that the errors for allometric estimates of AGB increase exponentially with increasing DBH, whereas the error for TLS estimates of AGB are independent of DBH.

The main limitation for efficiently using the method in [5] at plot level is the required visual evaluation of input parameters for the QSM tree modelling. A more objective QSM parameter setting is essential for applying this method in more challenging and complex biomes and larger plots. In this paper we compare a new automated approach for estimating the parameters required for QSM modelling and compare AGB results against the values obtained in [5].

Materials and method

Study area and data collection

Two plots in native Eucalypt Open Forest (dry sclerophyll Box-Ironbark forest) in Victoria, Australia were partially harvested in 2012-2013 to acquire accurate estimates of AGB of 65 trees. The plots had a 40 m radius and were established in Rushworth forest, with the main tree species in these plots being *Eucalyptus leucoxylon, E. microcarpa* and *E. tricarpa*. Measurements of fresh weight and tree structure were collected for each harvested tree. Dry weight to fresh weight conversion factors and basic density for each species were derived from a limited number of samples across a range of diameters. TLS data were collected with the RIEGL VZ-400 scanner, which records multiple return LiDAR data (up to four returns per emitted pulse). Full hemispherical scan data was collected at five scan locations per plot using a 0.06° angular sampling.

Deriving AGB from TLS data

Tree volume is directly inferred from 3D quantitative structure models (QSMs) of individual trees reconstructed from the TLS data; AGB is then inferred via basic density. Inferring tree volume consists of two steps: i) extracting single trees from the registered point cloud; and ii) 3D reconstruction of isolated trees using QSMs. The QSM method has two important input parameters: the cover patch size (*d)* and the minimum threshold of LiDAR points within a single patch for inclusion in the reconstruction (*nmin*). The reader is referred to [5] for full details of tree extraction method and the QSM modelling using visual evaluation of the input parameters. Volumetric estimation increases of 5000% are generally observed over the patch size range 0.02 to 0.09, so getting this value 'right' is critical. QSM results are far more stable with varying *nmin*, and more related to the point cloud density. As a result, we focus here on *d*. The following automated framework has been developed for autonomous optimisation of cover patch size *d*:

1. The TLS point cloud is reconstructed 10 times over the desired *d* range 0.02 to 0.09 at an increment of 0.05, with *nmin* set to 4.
2. For each of the 10 QSM models at each *d* value, 4 trunk cylinders at 7.5%, 10%, 12.5% and 15% of the trunk length are extracted from the QSM model. The coordinates of these 4 cylinders drive a pass-through filter to extract the TLS point cloud from which the QSM was formed.
3. To obtain trunk diameter estimates, a least squares circle fit is executed on the resultant clouds and the diameters compared with the cylinders as a percentage change to quantify model conformity to the cloud. A single result ($trunk_{match}$) is generated by averaging the 4 cylinders for each of the 10 models.
4. For each *d* value the mean and standard deviation are generated from the 10 models and then the coefficients of variation, *CV*, are calculated as standard deviation / mean.
5. The second highest *CV* (CV_{max2}) is identified across all *d* values and the lowest value that conforms to ($CV <= CV_{max2} * CV_{factor}$) and ($trunk_{match} >= trunk_{conformity}$) is selected as the optimised *d* where CV_{factor} = 0.5 and $trunk_{conformity}$ = 0.9. We use the second highest *CV* instead of the highest *CV* to avoid outliers.
6. If no optimised *d* is identified in step 4, the method falls back onto the *d* with the highest $trunk_{match}$.

We compare results derived using the above process with those derived from 'visual assessment' i.e. a manual inspection of the resulting QSM form for each value of *d*, and an heuristic decision to accept/reject, based on the analyst's experience and expectations [5].

Results and discussion

Individual tree estimates of AGB using the visual and automated optimisation assessment for QSM modelling are compared against destructively harvested reference measures of AGB in figure 1. The RMSE with respect to the 1:1 line is 171 kg for the visual QSM assessment and is 306 kg for the automated QSM assessment. The concordance correlation coefficient, CCC, computes the agreement on a continuous measure obtained by two methods and ranges between 1 (perfect concordance) and -1 (perfect discordance). Both the visual and automated approach show CCC ranging from 0.94 to 0.98.

The mean absolute deviation for the visual assessment is lower (110 kg) compared to the automated assessment (214 kg). This is not surprising since the visual assessment is better suited to address different issues with the input TLS point cloud. For example, a partially occluded lower stem (by leaves or other, smaller branches) can result in a higher *d* value via the automated optimisation (and an overestimation of volume) because $trunk_{match}$ will be relatively low. On the other hand, the volume can be underestimated when $trunk_{match}$ is high, but large branches are not modelled well. The maximum absolute AGB deviation values for the automated QSM assessment are still lower than those reported for the allometric equations (> 1500 kg) in [5]. The advantage of the automated approach is that it requires no user input and therefore can efficiently be applied to large forest stands.

[6] showed that tree-level uncertainty in AGB estimation from their allometric model was about of 50% the mean, but they showed that uncertainty dropped to 5 to 15% when considering whole-plot AGB. The total AGB

of the 65 trees in our study area is 75.70 t for the visual assessment and 73.73 t for the automated assessment. The destructive harvesting yielded 69.02 t, so the TLS QSM approach results in an overall overestimation of 6.8% (automated optimisation) to 9.7% (visual assessment).

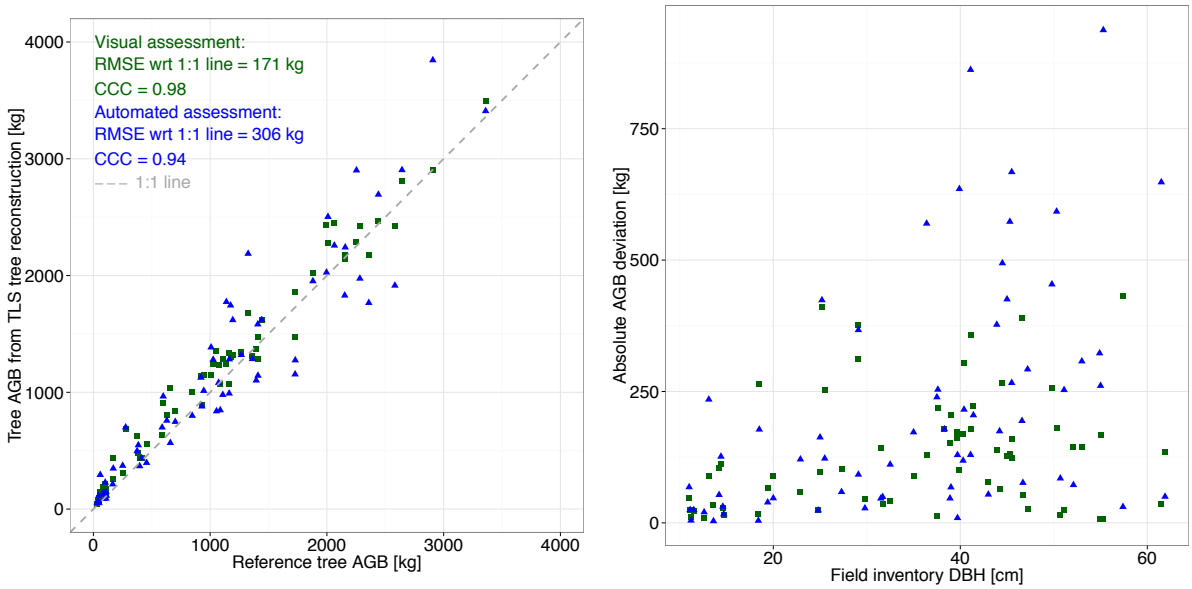


Figure 1: (Left) Comparison of destructively measured reference tree AGB with AGB inferred from TLS volume estimates through tree reconstruction and basic density information: visual QSM assessment vs. automated QSM assessment. (Right) Comparison of absolute tree AGB deviation.

Conclusion

In this work we contribute towards a nondestructive and robust method to infer AGB from TLS data. We demonstrate that AGB estimates from automated QSM parameter setting agreed well with those values derived from visual assessment. The uncertainties at tree level are higher for the automated QSM assessment, but the total AGB of the 65 sampled trees agrees to within 6.8% with the reference data. This is important for efficiently applying this TLS approach to larger areas or more challenging and complex biomes, such as tropical rainforests. Terrestrial laser scanning offers opportunities for a consistent and robust framework to support REDD+ monitoring capacities in developing countries. Limited research has been carried out in these countries to reduce the uncertainty of emission estimates of forest degradation activities, such as selective logging. Improving emission estimates relating to forest change requires better biomass estimates before and after change events at local levels [7]. The work presented here shows the possibility for applying TLS to larger areas more rapidly and consistently, as well as for combining with (and benchmarking) allometric estimates from traditional survey data, airborne lidar and other remotely sensed data.

References

[1] Baccini, A., Goetz, S.J., Walker, W.S., Laporte, N.T., Sun, M. et al. (2012). Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. Nature Climate Change, 2, 182–185.

[2] Saatchi, S.S., Harris, N.L., Brown, S., Lefsky, M., Mitchard, E.T.A. et al. (2011). Benchmark map of forest carbon stocks in tropical regions across three continents. PNAS, 108, 9899–9904.

[3] Mitchard, E.T.A., Feldpausch, T.R., Brienen, R.J.W., Lopez-Gonzalez, G., Monteagudo, A. et al. (2014). Markedly divergent estimates of amazon forest carbon density from ground plots and satellites. Global Ecology and Biogeography, 23, 935–946.

[4] Chave, J., Condit, R., Aguilar, S., Hernandez, A., Lao, S. et al. (2004). Error propagation and scaling for tropical forest biomass estimates. Philosophical Transactions of the Royal Society B: Biological Sciences, 359, 409–420.

[5] Calders, K., Newnham, G., Burt, A., Murphy, S., Raumonen, P. et al. (2015). Nondestructive estimates of above-ground biomass using terrestrial laser scanning. Methods Ecol Evol, 6, 198–208.

[6] Chave, J., Réjou-Méchain, M., Búrquez, A., Chidumayo, E., Colgan, M.S. et al. (2014). Improved allometric models to estimate the aboveground biomass of tropical trees. Glob Change Biol, 20, 3177–3190.

[7] Hill, T.C., Williams, M., Bloom, A.A., Mitchard, E.T.A. and Ryan, C.M. (2013). Are inventory based and remotely sensed above-ground biomass estimates consistent? PLoS ONE, 8, e74170.

International benchmarking on TLS methods for forestry applications: the project and the preliminary results

Xinlian Liang, Juha Hyyppä, Harri Kaartinen, Eetu Puttonen

Department of Remote Sensing and Photogrammetry, Finnish Geospatial Research Institute, FGI (former Finnish Geodetic Institute), 02431 Masala, Finland
(xinlian.liang, juha.hyyppa, harri.kaartinen, eetu.puttonen)@nls.fi

Highlights: The first international benchmarking project on TLS methods for forestry applications was launched in 2014. This paper presents the project overview and the preliminary statistical results.

Key words: *terrestrial laser scanning, TLS, single-scan, multi-scan, benchmarking, forestry.*

Terrestrial laser scanning (TLS) has been recently shown to be a promising technique for forest-related studies. Many tree attributes have been correlated with measurements from TLS data, and the accuracy of tree-attribute estimates is similar to, or even better than, national allometric models. TLS is anticipated to be a practical method for the future forest mapping and the data source for National Forest Inventories (NFIs).

EuroSDR launched in 2014 the benchmarking on TLS methods for forestry applications in order to understand the optimum data processing technique for future automated plot inventories. The emphasis in this benchmarking is on existing and modified algorithms, or manual measurements from TLS data. Meanwhile, this project is also open for all techniques in the research phase.

The objective of the TLS methods project is to evaluate the accuracy and feasibility of automatic, semi-automatic or manual tree extraction methods based on high-density TLS data. The sub-objectives include

- How much variation is caused in the tree parameter estimation by the processing methods to derive the tree measurements?
- How much variation is caused in the tree parameter estimation by the input TLS data, such as single- and multi-scan TLS data?
- What is the impact of forest conditions, e.g., density, species and ground vegetation?

The test forest is a southern Boreal Forest in Finland. 24 sample plots, 32-by-32m each, have been measured in the summer 2014. These forest plots have varying species, growth stages and management activities, and included both homogenous and less-managed forests. Five scans were made at the plot center and northeast, southeast, southwest and northwest directions, where the centre scan represents the single-scan approach and the merged five scans represent the multi-scan approach. Figure 1 and 2 show examples of the TLS data. The TLS data are exclusively used for this project before the publication of the final project report. After that, the point cloud data will be free of use for non-commercial research purposes.

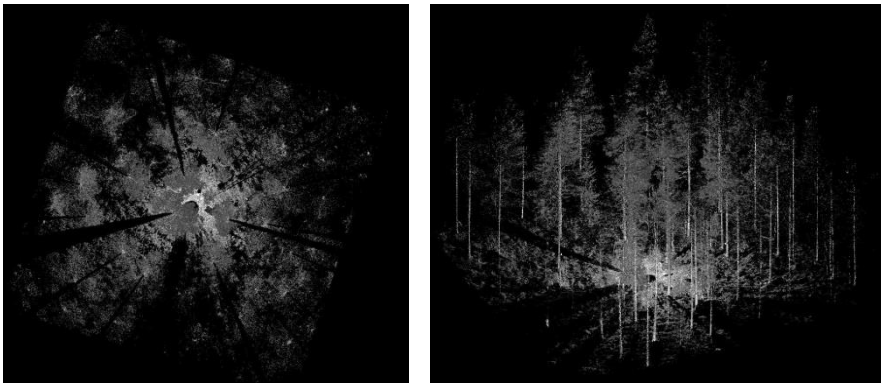


Figure 1: Single-scan TLS data of a test plot in 2D (left) and 3D (right).

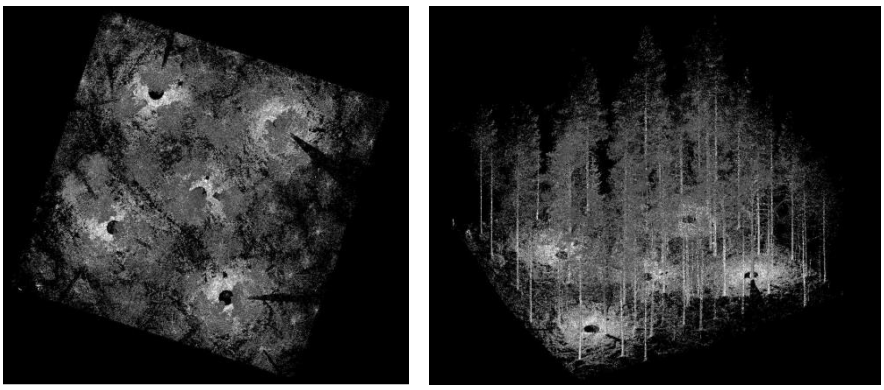


Figure 2: Multi-scan TLS data of a test plot in 2D (left) and 3D (right).

The plotwise parameters studied in this project include the tree position, height, DBH, stem curve and DTM. The tree measurements provided by the participants are being analysed. In the paper, an overview of the project and the preliminary statistical results are presented.

Seasonal change of leaf and woody area profiles in a temperate deciduous forest canopy from a dual-wavelength terrestrial lidar

Zhan Li[1], Alan H. Strahler[1], Crystal B. Schaaf[1, 2], David L. B. Jupp[3], Glenn Howe[4], Kuravi Hewawasam[4], Supriya Chakrabarti[4], Timothy A. Cook[4], Ian Paynter[2], Edward J. Saenz[2]

[1] *Department of Earth and Environment, Boston University, Boston MA, USA*
[2] *School for the Environment, University of Massachusetts Boston, Boston MA, USA*
[3] *CSIRO Marine and Atmospheric Research, Canberra, ACT, Australia*
[4] *Department of Physics and Applied Physics, University of Massachusetts Lowell, Lowell, MA, USA*

Highlights: Dual-wavelength point clouds of apparent reflectance from the Dual-Wavelength Echidna® Lidar (DWEL) offer direct, three-dimensional classification of leaves and woody materials. Two DWEL scans under a deciduous forest in spring (leaf-off) and summer (leaf-on) demonstrate seasonal change in height profiles of effective leaf and woody areas and area volume densities.

Key words: *Terrestrial lidar, TLS, DWEL, leaf area index, woody area index, deciduous forest*

Introduction

Forest structure interrelates with forest function and plays a critical role in regulating the exchange of energy, water, and carbon between the land surface and the atmosphere. Terrestrial laser scanners can characterize detailed forest structure rapidly and easily, providing ground truth for validation of airborne and spaceborne measurements of forest structure over large areas. This study uses scans from the Dual-Wavelength Echidna® Lidar (DWEL) to discriminate between leaves and woody materials directly in three-dimensional space using the spectral information from two simultaneously-pulsed infrared lasers. The direct discrimination of leaves, which constitute the carbon-fixing apparatus of the forest ecosystem, from other objects in forests, mainly woody materials in which carbon is stored, can improve the calibration of ecophysiological models of the carbon cycle. DWEL scans acquired at a temperate deciduous forest site in spring and late summer provide this discrimination and seasonal changes of effective leaf and woody areas and area volume density profiles with canopy height.

Instrument and Study Area

DWEL acquires full-waveform scans at both near-infrared (NIR, 1064 nm) and shortwave infrared (SWIR, 1548 nm) wavelengths with simultaneous laser pulses [1]. At the SWIR wavelength, laser power returned from leaves is much lower than from woody materials, such as trunks and branches, due to absorption by liquid water in leaves. In contrast, returned power from leaves and woody materials is similar at the NIR wavelength.

We scanned a deciduous forest site at Harvard Forest in central Massachusetts, USA, with the DWEL in both leaf-off (May 3rd 2014) and leaf-on (Sept 19th 2014) seasons. This 1-ha forest site is dominated by red maple (*Acer rubrum*), red oak (*Quercus rubra*) and white birch (*Betula papyrifera*), with an understory of these species accompanied by American beech (*Fagus grandifolia*), American chestnut (*Castanea dentata*) and others. We used two scans from the center plot of the site at the two seasonal dates to demonstrate the classification of leaves and woody materials in 3-D and seasonal change in proportions of leaf and woody area.

Methods

We first generated two calibrated point clouds of apparent reflectance at the two wavelengths (ρ_{nir} and ρ_{swir}) from each DWEL scan. This measure represents the reflectance of a diffuse scattering surface filling the laser beam that would return the observed power at the observed range. After merging the point clouds to obtain a single point cloud with dual-wavelength apparent reflectance values at each point, we classified the merged points into leaves and woody materials in two steps, firstly using spectral information, given by apparent reflectance, and secondly by spatial context information, given by 3-D distribution patterns of each point and its near neighbours. Our unsupervised classifier, the K-means clustering algorithm implemented in Scikit-learn package [2], first divided all points into 100 clusters using three variables: apparent reflectance values in the two wavelengths and a normalized difference index value ($NDI = (\rho_{nir} - \rho_{swir})/(\rho_{nir} + \rho_{swir})$). Where clusters included both leafy and woody hits, a supervised classification method using spatial context based on multiscale dimensionality criteria [3]labelled points into canopy hits (taken as leaf hits) and surface hits of trunks/coarse branches/ground materials.

We derived gap probabilities along canopy height z at multiple zenith angles θ, $P_{gap}(\theta, z)$ from apparent reflectance of points at both wavelengths respectively using the method of Jupp et al. [4]. However, in the case

of these Harvard Forest scans, laser power was insufficient for full coverage of the canopy; as a result, we excluded zenith rings at angles beyond 35°, where shots were not able to exit the top of the canopy (denoted H) and proper estimates of gap probability with height could not be obtained. We also excluded zenith rings from 0° – 10°, as these rings are biased by instrument placement away from large trunks and thick shrubs. As a result of these necessary omissions, we were unable to produce good estimates of the total leaf and woody area index and thus derived normalized leaf/woody area volume density (L/WAVD) profiles $\widetilde{f^L}(z)$ and $\widetilde{f^W}(z)$ using the ratio between effective woody area index (WAI) and leaf area index (LAI), λ, as

$$\lambda = \frac{\log P_{gap}^{W}(\bar{\theta}, H)}{\log P_{gap}^{L}(\bar{\theta}, H)}$$
$$\widetilde{f^L}(z) = \frac{1}{1+\lambda}\frac{\partial}{\partial z}\left(\frac{\log P_{gap}^{L}(\bar{\theta}, z)}{\log P_{gap}^{L}(\bar{\theta}, H)}\right), \widetilde{f^W}(z) = \frac{\lambda}{1+\lambda}\frac{\partial}{\partial z}\left(\frac{\log P_{gap}^{W}(\bar{\theta}, z)}{\log P_{gap}^{W}(\bar{\theta}, H)}\right) \quad (1)$$

Here, $\bar{\theta}$ indicates that λ, $\widetilde{f^L}(z)$ and $\widetilde{f^W}(z)$are average ratio values weighted by the solid angle of each zenith ring, rather than a mean angle [4]. It should also be noted that these are effective values because clumping or dispersion of leaves and branches may alter the observed gap probabilities.

Results and discussion

Classification of points

Comparison between equal-angle projection images of composite color and classification showed very good discrimination of leaves from woody materials. Spectral information classified most points well into leaves and woody materials in K-means clustering. Some mixed clusters occurred for two major reasons. First, some fine branches and small stems shared similar spectra with leaves, possibly because juvenile bark may have a higher water content than older woody materials. Second, a slight laser misalignment likely may have caused the simultaneous pulses of the two lasers to intercept different amounts of fine branches at far range, reducing the SWIR return. The points in these mixed clusters were then classified by spatial context information as noted above. Inspection of the images showed that many fine branches in the canopy were readily distinguished from leaves in both leaf-off and leaf-on scans, although leaf-off scans appeared to include more fine canopy branches misclassified into leaves. Fine branches of shrubs in the near range were largely classified correctly.

Normalized profiles of effective leaf and woody area

Figure 1 shows normalized effective leaf area and woody area indexes (L/WAI) and normalized effective leaf area and woody area volume density (L/WAVD) for the leaf-off and leaf-on scans; curves are derived from gap probabilities, which are independent of wavelength but dependent on proper calibration. Profiles are shown for both wavelengths; in most cases the NIR and SWIR curves are too close to distinguish clearly, indicating very good calibration of intensities to apparent reflectance values.

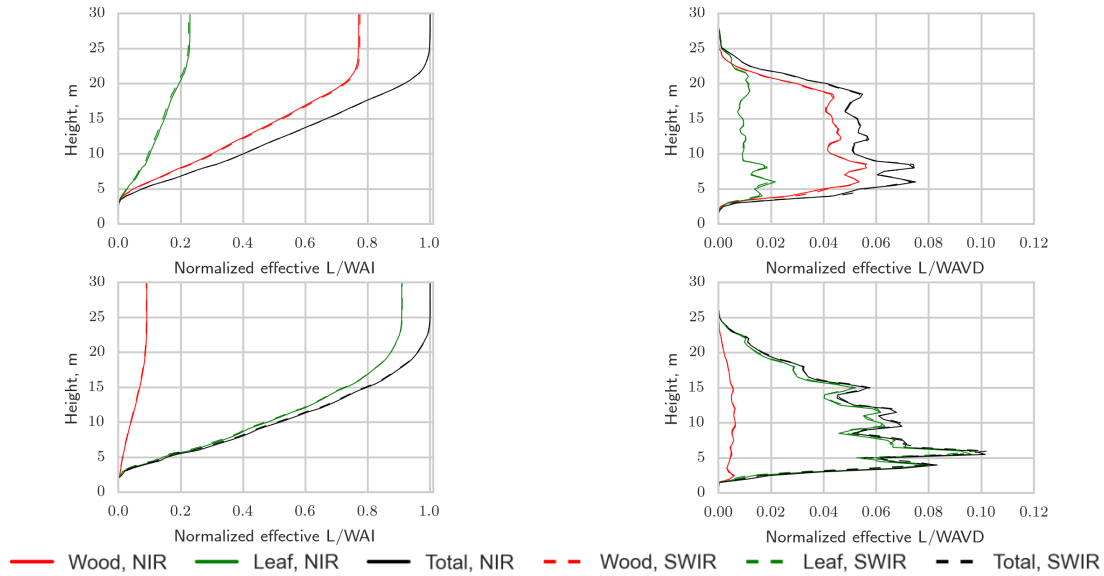


Figure 1 Normalized profile of leaf/woody area index (L/WAI, left column) and area volume density (L/WAVD, right column). Note x-axis values are proportions of leaves and woody materials, not absolute area index or area volume density. Upper row is the leaf-off scan and the lower row is the leaf-on scan. Red: woody materials; Green: leaves; Black: total. Solid lines are from NIR scans and dashed lines are from SWIR.

The LAI/WAI profiles (left column) are cumulative with height from the ground up. The leaf-off profile actually shows a significant proportion of leaf points (about 23 percent) for at least three reasons. First, there are several large white pines in the stand that provide a leaf signal while the deciduous trees are leafless. Secondly, the restricted range of zenith angles, coupled with the location of the scan centers as apart from nearby large trees, may bias the results against woody points of trunks and upper stems beyond the excluded 0–10° zenith rings. Lastly, the spatial similarity between fine branches and leaves in mixed spectrtal clusters leads to the classification of more points in the upper canopy as leaf hits than as branch hits. The leaf-on profile shows a much greater proportion of leaf area (about 91 percent) to woody area (about 9 percent), which is expected, but again, the leaf area may be overestimated.

The area volume density profiles show how the proportions of leaf and woody area change with height. The curves are scaled so that the integral of the total area volume density integrates to 1. In the leaf-off scan, the woody value is 2–3 times the leaf value; all three volume density profiles follow similar relative trends with height. After increasing as leaf and woody area is encountered directly above the instrument, volume densities peak at about 6 and 8 m, due to nearby understory crowns. Above this height, the three profiles remain relatively constant (with some variation), then decrease above about 18 m.

In the leaf-on scan, the woody area volume density is nearly constant with height, at about one-tenth of the total. The leaf and total curves follow each other closely, peaking in the lower canopy and then descending more or less uniformly from about 6–25 m. Some of this decrease may be due to weak or lacking leaf returns at higher zenith angles produced by insufficient laser power.

Conclusion

Spectral information, as calibrated apparent reflectance at NIR and SWIR wavelengths from DWEL scans, classifies leaves and woody materials including fine branches well in three-dimensional space. Some fine branches and small stems share similar spectra with leaves and can be further classified with shape information from point clouds, demonstrating the value of integrating bispectral information with 3-D spatial information. The discrimination of leaves and woody materials in space shows how leaf and woody area proportions change with height in the canopy, providing a clear picture of canopy structure.

Future work, planned to be completed and reported at SilviLaser 2015, will examine new scans made with higher laser power that will permit characterization of gap probability with height for zenith angles to 60° and beyond, allowing retrieval of leaf and woody plant areas, rather than proportions, as well as profiles that better sample the woody area of stand trunks and lower branches. This work was supported by the US National Science Foundation, under grant number MRI-0923389.

References

[1] Douglas, E. S., Martel, J., Li, Z., Howe, G., Hewawasam, K., Marshall, R. A., Schaaf, C. L., Cook, T. A., Newnham, G. J., Strahler, A., & Chakrabarti, S. (2015). Finding leaves in the forest: the Dual-Wavelength Echidna Lidar. Geoscience and Remote Sensing Letters, IEEE, 12(4), 776–780.

[2] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: machine learning in Python. Journal of Machine Learning Research, 12, 2825–2830.

[3] Brodu, N. & Lague, D. (2012). 3D terrestrial lidar data classification of complex natural scenes using a multi-scale dimensionality criterion: applications in geomorphology. ISPRS Journal of Photogrammetry and Remote Sensing, 68(0), 121–134.

[4] Jupp, D. L. B., Culvenor, D. S., Lovell, J. L., Newnham, G. J., Strahler, A. H., & Woodcock, C. E. (2009). Estimating forest LAI profiles and structural parameters using a ground-based laser called 'Echidna®.' Tree Physiology, 29(2), 171–181.

Estimation of tropical forest above-ground biomass using ALS and multispectral remote sensing data

Parvez Rana[1, 2, *], Basanta Gautam[3] and Timo Tokola[1]

[1]*School of Forest Sciences, University of Eastern Finland, P. O. Box - 111, FI – 80101, Joensuu, Finland*
[2]*Department of Ecosystem Science and Management, Texas A&M University, 1500 Research Parkway, Suite B-217, College Station, TX 77843, USA*
[3]*Arbonaut Ltd. Kaislakatu 2, FI – 80130, Joensuu, Finland*
**Corresponding author. Tel.: +1-979-676-9004; Email address: parvez_200207@yahoo.com (P. Rana)*

Highlights: This study evaluated the performance of airborne laser scanning (ALS), RapidEye and Landsat data in the estimation of above-ground biomass in a tropical forest of Nepal. The ALS model showed higher accuracy than RapidEye and Landsat for above-ground biomass (RMSE 56 vs. 93 vs. 96 tons/ha, respectively).

Key words: *Biomass, Tropical forest, LiDAR, REDD+.*

Introduction

Nepal experiences a wide range of climates, ranging from sub-tropical in the lowlands to arctic conditions in the high mountains. As one might expect, this wide variety of climatic conditions results in considerable variability in forest structure, architecture, and composition [1]. The latest National Forest Inventory (NFI) was conducted in the early nineties and is not sufficient to represent the present conditions, as Nepal has lost approximately 24.5% or 1,181,000 ha, of its forest cover between 1990 and 2010, corresponding to an average of 1.23% or 59,000 ha of forest per year [2]. Forest resource inventories are time-consuming and costly in Nepal because the large variety in forest conditions requires more plots to adequately capture this variability. In addition, access to the forest areas is another obstacle due to the steep mountainous terrain [3]. Without an accurate large-scale forest inventory method, future management plans could be negatively affected by inaccurate and imprecise growing stock estimates, which could lead to improper management decisions. This study investigated the estimation of above-ground biomass using airborne laser scanning (ALS), RapidEye and Landsat data in the southern part of Nepal. The accuracy of the predictions was evaluated by reference to the relative root mean square error (RMSE) and mean deviation.

Methodology

The Terai Arc Landscape region (Terai and Siwaliks) lies between 27.14 N and 29.08 N latitude and 80.15 E and 85.49 E longitude in southern Nepal. Terai is located in south and Siwaliks is located in north side of the study area. This region is well-known for productive services, ecological services, protective services and the secure livelihood it provides for the surrounding peoples. A network of 632 sample plots distributed in 20 ALS blocks over an area of 23,000 km^2 was surveyed between March 25th and May 20th, 2011. The ALS blocks were selected on the basis of their forest type, with the goal of maintaining uniform distribution of plots across the whole area. The field plots were then placed using systematic cluster sampling. Each ALS block (5 km × 10 km) contained six clusters, with eight plots in each cluster. Fixed circular plots (radius = 12.62 m), covering an area of 500 m^2, were utilized. The plots were divided into training (n=500) and validation (n=132) datasets; however both sets covered the variation in the forest characteristics over all the ALS blocks.

The ALS data were collected at daytime between March 14th and April 2nd, 2011 using a Leica ALS50-II laser scanning system operating at an altitude of 2200 m above ground level with a field view of 40°. The RapidEye images were acquired between 15th and 25th March, 2011. RapidEye imagery provides multispectral optical imagery on five bands (blue 440-510 nm, green 520-590 nm, red 630-685 nm, red-edge 690-730 nm and near infrared 760-850 nm) with a spatial resolution of 5m. Five Landsat 5 TM scenes were acquired. Each scene was a level 1T image (Standard Terrain Correction level) with a resolution of 30-m x 30-m. The images were acquired in January/February of 2010 and 2011.

The Landsat images were calibrated relative to each other utilizing the multiple linear regression approach presented by Tokola *et al.* [4]. The process mainly calibrates for differences caused by atmospheric conditions and sun-sensor angle effects present during an image acquisition. Local radiometric calibration was performed on the RapidEye images utilizing the method developed by Tuominen & Pekkarinen [5]. The RapidEye images were acquired on different dates and during different reflectance conditions, which affect the correlation with vegetation structure. The approach assures the RapidEye images will provide similar spectral responses over comparable forest conditions.

A combination of spectral bands, vegetation indices (normalized difference vegetation indices), and textural features by Haralick *et al*. [6] from the optical images were used in the above-ground biomass modeling. Each plot extent was utilized to extract a subset image of the spectral bands, vegetation indices, and textural features. The ALS-based canopy height (Height of the 10%, 20% … 100% percentile for first and last pulse points) and density metrics were employed in the above-ground biomass modeling. The area-based method was used to model the relationships between the field-measured above-ground biomass and the ALS canopy metrics [7].

In this step the sparse Bayesian (SB) was employed to construct the above-ground biomass model. The sparse Bayesian method chooses the best combination of independent variables to form a model by giving weights to the explanatory variables [3]. The RMSE and relative RMSE were used for validating the accuracy of the forest above-ground biomass estimates. Mean deviation was also utilized for determining the accuracy of the model. Mean deviation is the difference between the expected value of the model prediction and the true value. The statistical significance of the mean deviation was estimated utilizing a *t*-test. The mean deviation was considered to be significant if the absolute value of the *t* was greater than *t* corresponding with a probability of 0.05.

Results

The performance of the ALS, RapidEye and Landsat data is shown in Tables 1. The best model for the above-ground biomass consist of six ALS data predictors, while three and five predictors gave the best results with the RapidEye and Landsat data, respectively. The ALS model provided the best performance for all training plots, followed by the RapidEye and Landsat. The ALS data had the lowest relative RMSE of 31% followed by the RapidEye data of 52% RMSE and the Landsat data of 53%. The combination of ALS with RapidEye or Landsat were failed to improve the estimation accuracy of biomass.

Table 1: ALS, RapidEye and Landsat-based above-ground biomass estimation.

Above-ground biomass (ton/ha)	RMSE	$RMSE_{\%}$	Adjusted R^2	Mean deviation	p-value of *t*-test
ALS	56.0	31.3	76.8	–3.8	0.44
RapidEye	93.0	52.0	37.4	1.9	0.81
Landsat	96.1	53.7	33.1	2.0	0.81

The residuals of the sparse Bayesian models for all training plots are matched against the fitted values in Figure 1. Although there were several outliers, these were retained in the models in order to tolerate possible model error [8]. But statistical outliers were frequent in the RapidEye and Landsat models, which led to noise and non-constant variance in these cases. It can thus be said that the plot-specific standard error and standard deviation for the RapidEye and Landsat models did not fit with the data. The standard error of the means (thick lines) and the standard deviation (thin lines) were based on bins of data. The black thick lines indicate the mean value of these bins overlapped with the x-axis at 95% confidence intervals.

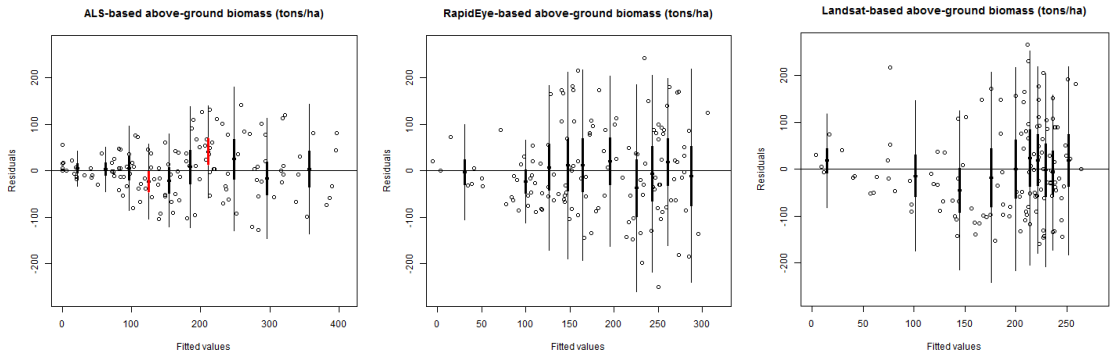


Figure 1: Residual plots of above-ground biomass (ton/ha). The thick lines show the standard error of the means and the thin lines the standard deviation.

Discussion and conclusion

This paper shows that the ALS models provided better accuracy than the RapidEye models for above-ground biomass in a mixed-species tropical forest of Nepal. It was promising that the ALS models explained 77% (R^2 value of 77%) of the variation in above-ground biomass. Such accuracy indicates that the ALS is probably the best available remote sensing method for the estimation of dense tropical forest attributes in REDD plus.

A mixture of studies have been conducted within the last 10 years on the possibilities for predicting forest variables by means of ALS data especially in Scandinavian forests, for instance, Maltamo *et al*. [10] and nowadays in tropical forests, for instance, Hou *et al*. [8]. Lower relative RMSE values were recorded from Scandinavian forests. However, the relative RMSE value was higher for tropical forests due to dense forests with

high understory vegetation. Moreover, Asner *et al*. [9] examined a universal ALS approach for tropical forest and a R^2 value of 80% recorded for above-ground carbon which comes close to our study, 77%.

ALS models slightly underestimated the attributes especially for the higher above-ground biomass values. The residual error in ALS model marginally increased for the greater stem volume and above-ground biomass values (Figure 1). Plots having high field measured value are likely to have dense canopy. It could be a factor that ALS pulses penetration rate to ground was low where the presence of highly dense under-stories vegetation and presence of complex mountainous forest in the study area. Therefore, some laser pulses are only reflected from the upper canopy and consequently the underestimation of tree height value. Similar trend was also observed in some other studies [8, 11].

On the contrary, the RapidEye models accuracy had higher error and underestimated the estimated field value for stem volume and above-ground biomass. The reason could be that *i*) the leaves start to drop as the trees during the first RapidEye images were taken and were already fallen when the last set of images were taken so that the lower NDVI values were recorded; *ii*) RapidEye sensor records the radiance coming mostly from the canopy surface. The fundamental theory is that the canopy allows penetration of the light but at the same time a percent of the incident energy is absorbed. If a dense forest consists of thick layer of needles/leaves just a portion of incident energy reaches the ground and bounces back towards the sensor. Thus most of energy recorded is that scattered by the upper part of the canopy and the variation is biomass beneath remains unrecorded [11].

With respect to the cost of data procurement, ALS was comparatively costly but more accurate than the RapidEye and Landsat, while ALS data is an alternative to achieve the accuracy requirements specified by Tier II and Tier III, as defined by the IPCC. Tout ensemble, the study outputs cast a promising light on the use of sparse Bayesian for forest parameter estimation with accuracy high enough and thereby to provide further support for the REDD methodology development. Finally, this high-tech forest resource mapping can help in the long-term to establish an automated forest monitoring system at the forest department and booster the sustainable forest management in Nepal.

Acknowledgements

We would like to thank the authorities responsible for the Forest Resource Assessment (FRA) project in Nepal for providing the data, and all the members of the project and field crew for their excellent data collection work in the field. We would like to thank Tuomo Kauranne, President of Arbonaut Ltd, Finland (through the LiDAR-Assisted-Multisource-Program Project). Arbonaut Ltd, Finland, has developed the LAMP method for large-scale remote sensing-based forest inventories. In addition, Jarno Hämäläinen, Jussi Peuhkurinen, Katja Gunia, Katri Tegel and Petri Latva-Käyrä (Arbonaut Ltd) also helped us during the data processing. We would also like to thank Lauri Korhonen, Jari Vauhkonen (University of Eastern Finland), and Md Latifur Rahman Sarker (Universiti Teknologi Malaysia) for their comments. This research work was funded by the Finnish Cultural Foundation (Regional Fund).

References

[1] Jha, B. N., & Paudel, G. (2010). REDD monitoring, reporting and verification systems in Nepal: gaps, issues and challenges. Journal of Forest and Livelihood, 9 (1), 21-32.

[2] FRA (2010). Global Forest Resources Assessment 2010. Country report. Rome, Italy.

[3] Rana, P., Korhonen, L., Gautam, B., & Tokola, T. (2014). Effect of field plot location on estimating tropical forest above-ground biomass in Nepal using airborne laser scanning data. ISPRS Journal of Photogrammetry and Remote Sensing, 94, 55–62.

[4]Tokola, T., Löfman, S., & Erkkilä, A. (1999). Relative calibration of multitemporal Landsat data for forest cover change detection. Remote Sensing of Environment, 68 (1), 1–11.

[5]Tuominen, S., & Pekkarinen, A. (2004). Local radiometric correction of digital aerial photographs for multisource forest inventory. Remote Sensing of Environment, 89 (1), 72–82.

[6]Haralick, M. R., Shanmugam, K., & Dinstein, J. (1973). Textural features for image classification. IEEE Transactions on Systems, Man and Cybernetics, 3 (6), 610–621.

[7] 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.

[8] Hou, Z., Xu, Q., & Tokola, T. (2011). Use of ALS, Airborne CIR and ALOS AVNIR - 2 data for estimating tropical forest attributes in Lao PDR. ISPRS Journal of Photogrammetry and Remote Sensing, 66, 751–942.

[9] Asner, P. G., Mascaro, J., Muller-Landau, H. C., Vieilledent, G., Vaudry, R., Rasamoelina, M., Hall, J. S., & Breugel, M. V. (2012). A universal airborne ALS approach for tropical forest carbon mapping. Oecologia, 168, 1147–1160.

[10] Maltamo, M., Packalén, P., Suvanto, A., Korhonen, K.T., Mehtätalo, L. & Hyvönen, P. (2009). Combining ALS and NFI training data for forest inventory: a case study in Kuortane, Western Finland. European Journal of Forest Research, 128, 305–317.

[11] Tegel, K. (2011). A comparison of Landsat-7 ETM+ and TerraSAR-X satellite imagery in estimating forest aboveground biomass in a two-stage sampling procedure. MSc dissertation, University of Helsinki. 70 p.

Mapping forest biomass in subtropical forests using small-footprint full-waveform LiDAR

Lin, LC, Cao[1, 2]; Nicholas, NC, Coops[2]; Guanghui, GS, She[1]; Honghua, HR, Ruan[1];

[1]*Co-Innovation Center for Sustainable Forestry in Southern China, Nanjing Forestry University, China*
[2] *Department of Forest Resources Management, University of British Columbia, Canada*

Highlights: This research is focused on sub-tropical forest which to date have had less focus than other forest biomes and examines the application of small-footprint full-waveform LiDAR. LiDAR metrics were extracted from composite waveforms for developing biomass estimation models in three broad forest types. The performances of these models are assessed and the importance of the metrics for predicting AGB investigated.

Key words: *Biomass; Full-waveform LiDAR; Composite waveforms; Subtropical forests*

Introduction

The subtropical forests biome accounts for approximately a quarter of the area of China and are particularly important for improving regional ecological environment and maintaining global carbon balance [1]. Despite their importance, there is still considerable uncertainty about carbon budgets within these subtropical forests. Remote sensing techniques can provide quantitative spatial explicit information and "wall-to-wall" observations for carbon stock mapping and monitoring. However, since subtropical forests are typically structurally-complex and carbon-dense ecosystems, optical sensors tend to saturate which inhibits reliable forest biomass estimates in these regions [2]. Light Detection and Ranging (LiDAR) is an active remote sensing laser technology capable of providing detailed, spatially explicit, three-dimensional information on vegetation structure. A large number of studies have demonstrated the potential of LiDAR to accurately estimate biophysical properties over a wide range of forest types [3].

Due to limitations in the electronics of many conventional airborne LiDAR systems, only surfaces that are sufficiently spaced apart can be distinguished into separate returns. Small-footprint full-waveform (FWF) LiDAR systems have become available commercially since 2004; these systems provide new opportunities for forestry studies. These FWF systems digitize and record the entire backscattered signal of each emitted pulse, and allow the recording of detailed geometric and physical properties of intercepted objects. In this study, the capability of ALS for estimating forest biomass was explored using FWF LiDAR data in a subtropical forest of Southeast China. Specific objectives are to (1) to assess a voxel-based composite waveform approach for processing FWF data; (2) to fit above ground biomass (AGB) estimation models with FWF LiDAR-derived metrics in three forest types (i.e., coniferous, broadleaved and mixed forests); and (3) to investigate the importance of these metrics in the models.

Materials and methods

The study area is located in Yushan Forest, southeast China (120°42′9.4″E, 31°40′4.1″N). It covers approximately 1103 ha, with an elevation range of approximately 20–261 m above sea level. The forest in Yushan belongs to the North-subtropical secondary forest with the dominated conifer species (e.g., Masson pine, Chinese fir and Slash pine etc.) and broad-leaved species (e.g., Sawtooth oak, Chinese sweet gum and Chinese holly etc.).

Plot data (30×30m) were collected during 2012-2013 in growing seasons under leaf-on conditions. At each plot, for trees with a DBH>5cm, diameter, height, height to crown base, and crown width in both cardinal directions were measured. Plot corners were located using dGPS in submeter accuracy. Species and crown classes of all measured trees were recorded. Above ground biomass (AGB) was calculated for individual trees based on species-specific allometric equations and field-measured DBH and height, and then summed to plot-level AGB. A total of 53 plots were used in this research.

The FWF LiDAR data were acquired on 17 August 2013 using a Riegl LMS-Q680i scanner on-board the LiDAR, CCD and Hyperspectral (LiCHy) Airborne Observation System. The system was operated at 900 m above ground level. Data was acquired at a 360 kHz pulse repetition frequency with a scanning angle of ±30° from nadir. The returned waveforms were recorded with a temporal sample spacing of 1 ns and beam footprint size of 0.45 m (nadir) in diameter. The average ground point distances of the dataset are 0.49 m in a single strip.

The returned waveforms were calibrated by W^e (the pulse width of the emitted pulse), I^e (the intensity of the emitted pulse) and S_i (the distances between the laser scanner and the reflecting object) [4]. To avoid the effects of off-nadir pointing on the waveform and to integrate non-vertical waveforms from different flight trajectories, a voxel-based composite waveform approach was applied to the full-waveform data [5]. This approach synthesized multiple raw waveforms into composite waveforms through the vertical space partitioning of forest canopies by voxels, and uses the maximum amplitude value to construct new composite waveforms (Figure 1). Total of 14 FWF metrics were derived based on the composite waveforms, i.e., height of median energy (HOME), waveform distance (WD), height/median ratio (HTMR), number of peaks (NP), roughness of outermost canopy (ROUGH), front slope angle (FS), return waveform energy (RWE). These metrics were all summarized as values of mean and standard deviation in each plot.

Results

Regression models were developed for estimating AGB of the three forest types. Stepwise selection was performed to select variables for the final models. The relative weights were calculated for each metric in the final models to find their importance [6]. Cross-validation was used to assess model accuracy. AGB were mapped using the forest type-specific regression models with a cell size of 30 × 30 m covering the entire study site.

Overall, AGB for all the forest types are generally well predicted (R^2=0.72-0.84, rRMSE=16.63-23.25%) (Table 1). The level of variation of AGB explained by forest type-specific models are 75-84%, higher than the levels for AGB in general models (72%). According to the results of relative weights analysis, *HOME*μ has the highest relative importance for each biomass predictive model.

Table 1: Biomass predictive models and statistics of accuracy assessment.

Models	Parameters and coefficients	R^2	rRMSE (%)
W_G	10.43 (*HOME*μ) + 2.28 (*FS*μ) - 3.39 (*HTMR*μ) −209.85	0.72	23.25
W_C	8.36 (*HOME*μ) −3.02 (*HTMR*μ) + 1.78(*FS*σ) −96.75	0.84	16.63
W_B	2.45 (*HOME*μ) + 2.21 (*HOME*σ) + 0.52 (*FS*σ) −8.43	0.78	18.16
W_M	-0.56 (*HOME*μ) −2.31 (*HTMR*μ) + 1.35(*NP*μ) + 85.32	0.75	20.72

Notes: W_G= AGB for all forests; Wc = AGB for coniferous forest; W_B = AGB for broadleaved forest; W_M = AGB for mixed forest; μ = mean; σ=standard deviation.

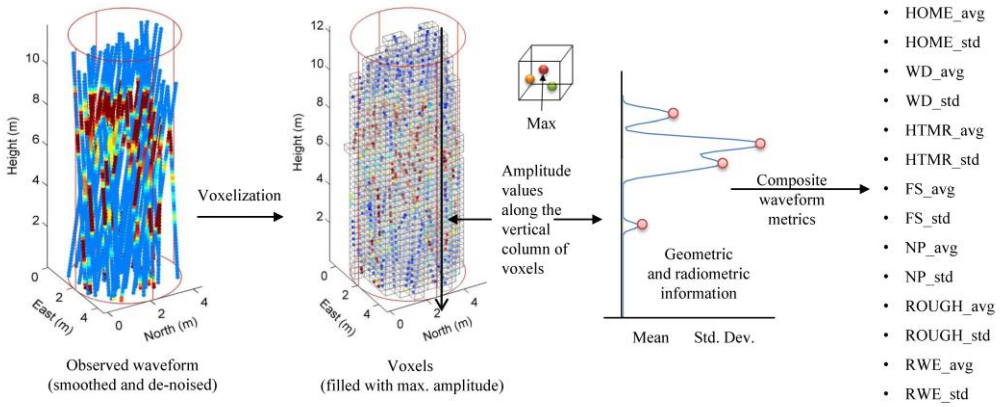


Figure 1: The process of composite waveforms formation and metrics extraction

According to the distribution map of AGB, relatively low AGB was found on steeper slopes, especially in the mountain ridges. Higher AGB values were predicted in most valleys, especially on the northern slopes where the dominant species prefer cooler summer temperatures with sufficient water supply. In some of the north–central areas, very high AGB were associated with old-growth broadleaved forests (e.g., Sawtooth oak). Coniferous-dominated forests contributed high AGB in areas with good site quality and relatively flat areas (e.g., stands dominated by Slash pine at the foot of southern part of the study area). There was a relatively large area of lower AGB in the northwest–central part of the study area, reflecting the presence of low-quality Chinese fir plantations on ridge tops with poor site quality with large canopy gaps and low stem density.

Discussion

In this research, AGB for all the forest types are generally well predicted. FWF LIDAR sensors record the detailed geometric and physical properties of forest canopy, which provides a good foundation for strong

relationships between the LIDAR-derived metrics and forest biomass. For FWF metrics analyzed in this study, the average value of HOME (i.e., waveform centroid to the ground) was the strongest predictors of AGB. Drake et al. (2002) demonstrated that HOME is sensitive to canopy openness and the vertical arrangement of canopy elements in the LVIS-footprint (0.05ha) and plot (0.25-0.5ha) level. In this research, we found similar principles at plot (0.09ha) level by using small-footprint (0.45m) full-waveform LiDAR data. It could be explained by: for forest canopy with densely packed woody materials and foliage, few LiDAR pulses can penetrate and reach the ground, thereby increasing HOME. Conversely, for canopy with more gaps, more LiDAR pulses can pass through, thus, reducing HOME. HOME is sensitive to these differences and as a result makes it a potentially good indicator for AGB.

In this study, the forest type-specific models had better performance than the general model (developed from all plots). This is consistent with previous studies which demonstrate that forests with highly variable species composition should be stratified to ensure accurate biomass predictions. Næsset (2004) reported that forest type did not have any significant impact on the estimated biomass models in boreal forest (in southeast Norway, dominated by Norway spruce and Scots pine) [7]. However, in our study, the complex forest conditions in subtropical forests contained greater species diversity, making the effects of tree-species composition (classified as forest types) significant. Overall, the models were more accurate for coniferous and broadleaved forests than in mixed forests across all models. This could be explained as: the dominating coniferous and broadleaved species in our research site have a relatively even-aged tree structure and higher homogeneous composition, i.e., less variation in tree height and density at the plot level. In contrast, the mixed forest has more variability and less accurate LiDAR-derived heights within plots.

Conclusions

This study shows that small-footprint full-waveform LiDAR data had the capability of mapping biomass in subtropical forests. We demonstrated that above ground biomass (AGB) estimations using forest type-specific models performed better than the general model. The metric of the average value of HOME was the strongest predictors of AGB, makes it a potentially good indicator (for AGB) in the subtropics.

Acknowledgements

We acknowledge grants from the National Natural Science Foundation of China (Grant No. 31400492). The authors gratefully acknowledge the foresters in Yushan National Forest Park for their assistance with data collection and sharing their rich knowledge and working experiences of the local forest ecosystems.

References

[1] Wang, X.H., Kent, M. & Fang, X.F. (2007). Evergreen broad-leaved forest in Eastern China: Its ecology and conservation and the importance of resprouting in forest restoration. For. Ecol. Manag, 245, 76-87.

[2] Lu, D., Chen, Q. & Wang, G.et al. (2012). Aboveground forest biomass estimation with landsat and lidar data and uncertainty analysis of the estimates. Int. J. For. Res, 2012, 1-16.

[3] Lindberg, E., Olofsson, K. & Holmgren, J. et al. (2012). Estimation of 3D vegetation structure from waveform and discrete return airborne laser scanning data. Remote Sens. Environ., 118, 151-161.

[4] Reitberger, J., Schnörr, C. & Krzystek, P. et al. (2009). 3D segmentation of single trees exploiting full waveform LIDAR data. ISPRS J. Photogramm. Remote Sens., 64, 561–574.

[5] Hermosilla, T., Coops, N.C. & Moskal, L.M. et al. (2014). Deriving pseudo-vertical waveforms from small-footprint full-waveform LiDAR data. Remote Sens. Lett. , 5, 332 – 341.

[6] Johnson, J. W. (2000). A heuristic method for estimating the relative weight of predictor variables in multiple regression. Multivariate Behavioral Research, 35, 1-19.

[7] Næsset, E. (2004). Estimation of Above- and Below-Ground Biomass in Boreal Forest Ecosystems. In Laser-Scanners for Forest and Landscape Assessment; ISPRS Archives. Freiburg, Germany, 145–148.

A nationwide forest attribute map of Sweden derived using airborne laser scanning data and field data from the national forest inventory

Mats Nilsson[1], Karin Nordkvist[1], Jonas Jonzén[1], Nils Lindgren[1], Peder Axensten[1], Jörgen Wallerman[1], Mikael Egberth[1], Svante Larsson[2], Liselott Nilsson[2], Johan Eriksson[2] and Håkan Olsson[1]

1. Department of Forest Resource Management, Swedish University of Agricultural Sciences, Umeå, Sweden, e-mail: name.name@slu.se.
2. Swedish Forest Agency, Jönköping, Sweden, e-mail: name.name@skogsstyrelsen.se

Highlights: A raster database with forest variables for all forest land in Sweden has been produced using existing airborne laser scanning and field data from the National Forest Inventory. The resulting stand level RMSEs for stem volume, basal area, mean tree height, and mean stem diameter were 17.2–23.3%, 15.0–18.3%, 6.2–9.7%, and 9.3–14.5%, respectively.

Key words: *nationwide forest database, national forest inventory, airborne laser scanning.*

Introduction

Sweden's land area covers 41 million ha, of which 22 million ha is productive forest land available for forestry. Approximately 40% of the growing stock in Sweden is Norway spruce *(Picea abies)*, 40% is Scots pine *(Pinus sylvestris)* and the remaining 20% are mostly broad-leaved species, in particular Birch (*Betula spp*.). The forest is divided into estates which are owned by 227 000 private forest owners and a few large companies.

The sample plot based National Forest Inventory (NFI) is the most important data source for collecting objective statistics regarding the forest resources on a national and regional level [1]. The NFI is carried out by the Swedish University of Agricultural Sciences (SLU).

The Swedish National Land Survey (Lantmäteriet) has carried out an airborne laser scanning (ALS) of almost all of Sweden, primarily with the purpose of creating a new national Digital Elevation Model (DEM). The data collected are available free of charge to authorities, and at a low cost to the private sector.

In order to stimulate activities in the forest sector, especially among private forest owners, increased availability of high quality and up to date spatial information about forest resources would be beneficial. The government recognised that data from the national laser scanning could provide such information and therefore provided funding for the production of a nationwide forest database based on these laser data. The project was led by the Swedish Forest Agency, which is the national authority for supervising the forest sector, and has been done in cooperation with SLU, which performed the method development and database production.

Objectives

One objective of the project presented here has been to produce a high resolution (12.5 m grid cell size) nationwide raster database with predictions of stem volume (m^3/ha), above ground tree biomass (ton/ha), basal area (m^2/ha), mean tree height (dm), and mean stem diameter (cm) by combining ALS data and field data from the Swedish NFI.

In addition a number of products that do not require training data, have been produced by the Forest Agency and by consultants contracted by them. These products include a tree height raster with a 2 m grid cell size, and maps of slope, hillshading and wet areas. That second part of the project is not included in this presentation

Laser scanning data

The Swedish National Land Survey's ALS campaign for a new national DEM started in 2009. By the end of 2014, 97 % of the productive forest land in Sweden had been scanned. The scanning is done with a posting density of 0.5–1 return/m^2, a maximum scanning angle of 20 degrees from nadir, and is organised in 387 blocks usually 25 by 50 km in size. The ambition is that each block should be scanned within a minimal time period using the same scanner. In total, 13 different scanner types were used in the national laser scanning. Most blocks have been scanned with Leica (74%) or Optech (25%) scanners and only 5 blocks were scanned with a Riegl scanner. The ambition has been to scan southern Sweden, which has more broadleaved forest, in the spring and autumn during leaf-off periods, whereas northern Sweden has mostly been scanned during the summer period.

The new national DEM (2 m * 2 m grid cell size) was used as ground reference when calculating vegetation heights for the ALS returns. Laser metrics such as height percentiles and vegetation ratio were calculated for 12.5 m * 12.5 m grid cells using Fusion software [2]. Differences in the laser point cloud density due to overlaps between adjacent scanning strips were removed before the laser metrics were calculated. This was done by

selecting data from the scanning strip that were collected with the smallest scanning angles when assigning ALS returns to each grid cell. Laser metrics were also calculated for the NFI plots.

NFI data

In the Swedish NFI approximately 9500 sample plots are field surveyed annually. Of these, 60 % are permanent plots with a plot radius of 10 m that are revisited every five years while the remaining 40 % are temporary plots with a 7 m radius. The NFI plots are positioned with GPS receivers giving a positional accuracy of about 5 m. During 2013 permanent plots located in productive forest and not on a stand boundary were re-visited and new coordinates were recorded using high precision GPS receivers giving a positional accuracy of about 1 m. Unfortunately a lack of funding restricted the collection of these more accurate plot coordinates to southern Sweden mainly. However, tests indicated that the need for highly accurate plot positions was less critical in northern Sweden due to the relatively low spatial variation in the forest landscape in the north.

Initial tests in northern, mid, and southern Sweden showed that the lowest RMSEs for stem volume and tree height in all parts of the country were obtained using only permanent plots, which had more accurate coordinate positioning and were larger than the temporary plots. In total, 11500 permanent plots from 2009 to 2013 located on productive forest land were used in the project. The NFI plot data were forecasted or backcasted to the same year as they were laser scanned. This was done using the Heureka forest planning system developed at SLU [3]. Plots that had been disturbed in the time period between scanning and field measurement were removed using information registered during the field inventory and a statistical outlier elimination procedure. In total, about 15% of the plots were removed.

Prediction of forest variables

Linear regression models were used to relate the selected forest variables, or transformations of the variables, to metrics derived from the ALS data. All estimates of models and prediction of grid cell values were done with an automated production line developed in the R statistical software running on Windows PCs. For each forest variable to be estimated were a set of one to six alternative models (depending of variable) used for all laser scanned blocks across Sweden. The models had only two or three explaining variables, the variables used were height percentiles, combination of height percentiles and vegetation ratio, and the standard vegetation of vegetation returns. The regression parameters were re-estimated for each block using field data from the approximately 350 closest permanent NFI plots. These plots were selected from the block to be estimated as well as nearby blocks scanned with the same scanner type (Leica, Optech, or Riegl) and with the same leaf condition (leaf on or leaf off). The choice between the one to six predefined models was done by maximizing the coefficient of determination. The regression models were then used to predict forest data for all grid cells within the block. Since the quality of the predicted values was relatively low in young forests, all variable values for grid cells with a predicted height less than 3 m was set to zero.

Validation data

Field data from 544 stands located in different parts of Sweden were used for validation. Each stand was field surveyed using 5 to 12 circular plots (5 m to 10 m radius) with regular spacing. The validation datasets comes from SLU's existing test sites and from two Swedish forest companies (Sveaskog and Holmen Skog), representing forests in southern, mid and northern Sweden.

Results and discussion

A nationwide raster database with predictions of stem volume, above ground tree biomass, basal area, mean tree height (basal area weighted), and mean stem diameter (basal area weighted) were produced. The validation results obtained for stem volume, basal area, mean tree height, and mean stem diameter are presented in Table 1.

The validation results are similar, or better, than those obtained with traditional subjective field surveys [4], but slightly worse than reported in many other Nordic studies based on laser scanner data [5]. The validation data set used for the stand level validation was however not ideal, in most cases were only 8 plots, most often with only an 8 m radius, used for estimation of rather large stands. Furthermore, there was up to a 3 year difference between the year of laser scanning and the year of field surveying, and the delineation of the validation stands were taken from existing forest maps. Thus the results presented in Table 1 would be better with higher quality validation data.

The most obvious problem with the estimates is that stem volume, basal area and biomass is overestimated for stands that is scanned during leaf on-season and contains much broadleaved forest. To remove outliers caused, for example, by cuttings and damages between the time of scanning and the time of inventory is an important processing step. The use of data from different scanner types and acquisition time points in the same regression model might be problematic from a theoretical point of view. The approach to only include data in each model that were scanned with the same scanner brand (Leica, Optech or Riegl), and were from the same season (leaf-on or leaf-off), and from a limited geographic area, worked well in practice. The use of geographically nearby reference plots was, however, not as important as we first anticipated. While separation of

scanner brands was found to improve the results, a further separation of scanner models, or even on scanner individuals was tested, but there was no clear indication that this would improve the results. One reason for this might be that the number of suitable field plots is reduced when more criteria are used for selection of the scanners to be used in a given model.

Table 1. Relative RMSE, number of stands, average number of plots per stand, and average stand size for the validation on stand level. The RMSEs have not been compensated for the stand level sampling error.

	Stem volume RMSE (%)	Tree height RMSE (%)	Stem diameter RMSE (%)	Basal area RMSE (%)	No. stands	Avg, no. plots per stand	Avg. stand size (ha)
Northern Sweden							
- Norrbotten	23.3	9.7	12.5	16.3	29	8	33.5
- Vindeln	17.2	7.7	9.3	-	30	10	11.5
- Umeå	20.2	6.2	14.5	15.2	11	-	4.1
- Örnsköldsvik	19.3	6.3	9.6	17.0	221	8	11.1
Mid Sweden							
- Iggesund	20.4	6.6	10.0	18.3	101	8	16.9
Southern Sweden							
- Götaland	19.0	9.1	11.3	15.0	152	8	8.5

An important part of the project has been to make the database available to forest owners, forest associations, forest companies, authorities, researchers, and the general public. Thus, all predictions can be downloaded as 12.5 m grid cell rasters free of charge at the Forest Agency's homepage (http://www.skogsstyrelsen.se/skogligagrunddata, at the moment only in Swedish). A tool that allows any stand to be manually delineated, and that prints a pdf file with all the estimated variables has also been developed and can be reached from this home page.

The cost of the project was in the order of 0.1 Euro/ha forest land. This low cost is due to the use of both existing laser data and existing field plots, as well as the fact that half the Swedish land area is productive forest. As a comparison, the cost of a similar data collection with manual methods would be in the order of 10 Euros/ha. Much of the time has been spent on development and testing different processing alternatives. Another large part of the project has been outreach activities including arranging meetings with end users all across Sweden. The actual processing of the data was highly automated and can thus be repeated in the future at an even lower cost.

The product has been very well received by both forestry professionals and private forest owners, most likely because the data are as good as, or better than, the data they are used to, but with a higher spatial detail, and often more recent. The two most frequent requests from the user are "Will there be a continuation with updated versions?" and "Can we get tree species as well?".

Conclusions

By using existing airborne laser scanner data and national forest inventory plots, a nationwide raster database with key forest variables has been produced for almost all forest land in Sweden. The accuracy of the estimated variables is sufficiently good for use in forest management planning. The product is being distributed free of charge and has been well received by the users. Advice to those that plan to make a similar product are to use as few scanners as possible, perform the scanning during leaf-off season, and make sure that training data and validation data are from as similar timepoints as possible.

References

[1] Axelsson, A-L., Ståhl, G., Söderberg, U., Petersson, H., Fridman, J. and Lundström, A. (2010). Sweden (Ch. 35). In: Tomppo, E. et al. (eds). National Forest Inventories. Pathways for common reporting. ISBN 978-90-581-3232-4.

[2] McGaughey, 2014. FUSION/LDV: Software for LIDAR Data analysis and Visualization. http://forsys.cfr.washington.edu/fusion/.

[3] Wikström, P, Edenius, L, Elfving, B, Eriksson, L. O., Lämås, T., Sonesson, J., Öhman, K., Wallerman, J., Waller, C., and Klintebäck, F. (2011). The Heureka forestry decision support system: An overview. Mathem Comp For & Natural-Resource Sci 3: 87-94.

[4] Ståhl, G. (1992). A study on the quality of compartmentwise forest data aquired by subjective inventory methods. Swedish Univ. of Agricultural Sciences. Dept. Of Biometry and Forest Management, Section of Forest Mensuration and Management, Report No. 24 (*In Swedish*).

[5] Næsset, E., Gobakken, T., Holmgren, J., Hyyppä, H., Hyyppä, J., Maltamo, M., Nilsson, M., Olsson, H., Persson, Å., and Söderman, U. (2004). Laser scanning of forest resources: the Nordic experience. *Scandinavian Journal of Forest Research.* 19, 482-499.

Updating the National Forest Attribute Map using stereo matching of aerial images, the national terrain model and data from the National Forest Inventory?

Jonas Bohlin, Johanna Blombäck, Jörgen Wallerman, Johan, E.S., Fransson

Department of Forest Resource Management , Swedish University of Agricultural Sciences , S-901 35 Umeå , Sweden.

Highlights: The possibility to produce nation-wide forest attribute maps has been explored using stereo matching of aerial images, the national terrain model and data from the National Forest Inventory. Results are similar to that of ALS based methods, proving image matching to be a future alternative method.

Key words: *aerial images, estimation, forest attributes, image matching, inventory, large area*

Introduction

Mapping of forest attributes using Airborne Laser Scanning (ALS) data has recently changed how forest companies acquire forest information in Sweden. For large forest holders, dedicated ALS campaigns together with simultaneous field inventories of training plots is feasible, but this is not cost-efficient for small forest holdings. In an effort to supply owners of small forest estates, and others, with ALS-based forest attribute maps, the Swedish Forest Agency (Skogsstyrelsen) is now mapping forest attributes for the whole of Sweden. This is performed using ALS data captured in the ongoing production of a new national digital elevation model and field sample plot data from the National Forest Inventory (NFI) [1]. The national collection of ALS data started in 2009 and has to the end of 2014 covered more the 95% of the productive forest land. However, this data are rapidly aging – every year 3.5% of the forest land is changed due to forest treatments like; clear-cutting, thinning or final felling of seedling trees. With this in mind, organisations providing information to the private forest owners (e.g., forest owner associations and the forestry agency) are now interested in possible future updates of the national forest attribute map.

Digital photogrammetry from aerial images has emerged as one possible technique because it generates 3D data similar to that generated by ALS. 3D data from digital photogrammetry describes the upper part of the canopy surface only, whereas ALS penetrates the canopy and also measures the vertical canopy structure and the ground level. However, studies [2,3] have shown that 3D data from aerial images can map forest attributes with accuracies similar to that of ALS based forest inventory. Since the Swedish National Land Survey annually acquires aerial images of about one third of the country, this is a highly interesting source of 3D data for large-scale mapping of forest attributes.

The aim of this study is to evaluate the accuracy of forest attribute mapping using stereo-matched standard aerial images from the national image acquisition program, in combination with sample plots from the NFI. Here, the attributes basal area weighted mean height (Hbw), basal area weighted mean diameter (Dbw), stem volume (Vol), and basal area (BA) are addressed using area-based estimation.

Material and Methods

The study sites are three image acquisition blocks, one in southern (E2) and two in the central part (N2 & Q2) of Sweden (Figure 1). Within each block, NFI sample plots were selected from the year of the aerial acquisition and the year before, in order to provide a sufficient sample of training data. For this study, both temporary (7 m radius) and permanent (10 m radius) plots were used.

The aerial images where captured using UltraCam Eagle, flying at 3700 m, generating images with a ground sampling distance of 0.25 m and with a forward overlap of 60% and a side overlap of 30%, ensuring that each point on the ground is represented in at least one stereo model. An image block is photographed in one season (i.e., the same year and in leaf-on or leaf-off condition) but the images acquisition dates may differ by many weeks due to poor weather conditions or technical problems. The images were block triangulated using bundle adjustment and radiometrically corrected by Lantmäteriet, as part of their operational aerial image production. Two different methods and software were used for photogrammetric matching of the aerial images to produce point cloud data: (*i*) SURE [4] which generates a height value for each pixel using cost-based matching, similar to the semi-global matching method; (*ii*) Match-T DSM [5] using sequential multi-matching where both least squares and feature-based matching were combined. Finally, the point cloud height values were transformed from height above mean sea level to height above ground level by subtracting the height of the ground provided by the national DTM from Lantmäteriet (2 m spatial resolution and 0.2 m RMS vertical accuracy).

For the modelling, metrics describing the point cloud data, such as height distribution and spatial density characteristics were calculated from the point clouds using the Fusion software package for every field reference plot.

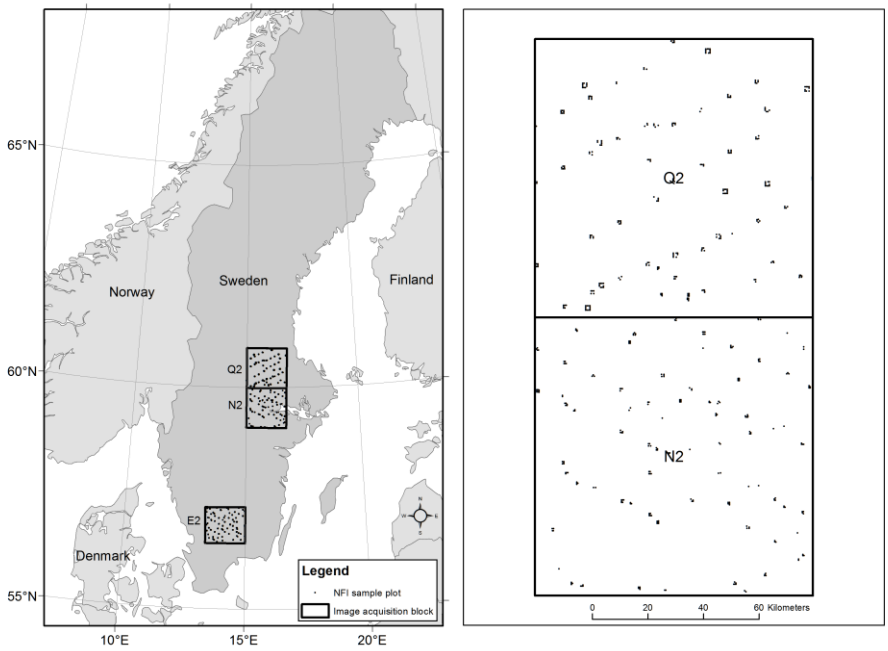


Figure 1, the three image acquisition blocks (E2, N2 & Q2) location on a background map of the Nordic countries to the left. The spatial distribution of the NFI sample plots for N2 and Q2, to the right.

The relationship between data from NFIs sample plots and metrics derived from the 3D point cloud was used to develop the regressions models. Validation of the estimation was done at plot level using leave one out cross validation.

Results and discussion

The results of the modelling are summarized in Table 1. The estimation accuracy in terms of RMSE varies little between the three study areas, this despite different seasonal data and different image matching algorithms.

Table 1: Describing the three study areas: E2; N2 and Q2, the image data and the image matching algorithms (IM) used and the results of the validation for each study area and forest attribute in terms of RMSE and bias at plot level using leave one out cross validation.

Study area	E2				N2				Q2			
Image data	Leaf-on, 2013				Leaf-on, 2014				Leaf-off, 2014			
IM	SURE (semi-global)				Inpho (feature based)				Inpho (feature based)			
n NFI plots	238				165				208			
	Hbw	Dbw	Vol	BA	Hbw	Dbw	Vol	BA	Hbw	Dbw	Vol	BA
RMSE	1.7	3.5	44.4	6.7	2.5	43.9	68.8	7.1	2.0	43.5	67.6	7.0
RMSE%	12.8	22.5	33.1	36.8	15.4	22.0	39.4	32.9	12.4	22.7	37.4	31.0
Bias	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.1	0.0
Bias%	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.1	0.0

The results presented in this study are similar to those acquired in ALS based forest attribute estimations. The National forest attribute map project has reported that the RMSE on plot level is about 9 – 11% for mean tree height; 14 – 16% for mean diameter; 17 – 24% for basal area and evaluations of stem volume prediction have given around 20% RMSE at stand level [1]. In a separate study [6], the study area E2 data set was validated at stand level using 28 independent stands, resulting in RMSEs 11.8% for Hbw, 18.9% for Dbw, 21.8% for Vol and 18.6% for BA. Similar RMSEs have been reported for ALS based forest in the Nordic countries, e.i., 2.5 – 13.6% for Hbw; 5.9 – 15.8% for Dbw; 9.3 – 16.6% for Vol and 8.6 – 13.2% for BA [7]. Thus, aerial images from the national image program together with sample plots from the NFI is clearly a feasible source of information to produce nation-wide forest attribute maps useful for forest management planning purposes.

Major advantages of this approach are the constant availability of new aerial images and the low cost of data. Furthermore, Lantmäteriet is planning to perform image matching, i.e., producing 3D data, for all images acquired 2015 and later as part of their annual image production. However, the weakness of the method is that the density of the forest is difficult to describe, which shows as a saturation of the estimation of high volumes as well as lower accuracies for volume and basal area. This is present in ALS based estimates where forest density is more accurately captured due to the canopy penetration of the active sensor. Further research must be aimed at investigating better describing density metrics as well as leaf-off season images impact on the forest attribute estimations. This because, when generating 3D data from leaf-off aerial images, the matched points tend to be generated on the ground in deciduous forests resulting in large differences between coniferous and deciduous areas for the height describing metrics used in modelling.

Conclusion

This study clearly shows that aerial images from the national image program together with sample plots from the National Forest Inventory can be used for large area forest attribute mapping.

References

1. Olsson, H.; Wallerman, J.; Nordkvist, K.; Jonzén, J.; Axensten, P.; Lindgren, N.; Nilsson, M.; Nilsson, L.-L.; Larsson, S. Production of a nationwide forest attribute map of Sweden using airborne laser scanning and national forest inventory plot data. In *ForestSAT2014: a bridge between forest sciences, remote sensing and geo-spatial applications*; Riva del Garda (TN), Italy, 2014.
2. Bohlin, J.; Wallerman, J.; Fransson, J. E. S. Forest variable estimation using photogrammetric matching of digital aerial images in combination with a high-resolution DEM. *Scand. J. For. Res.* **2012**, *27*, 692–699.
3. White, J. C.; Wulder, M. A.; Vastaranta, M.; Coops, N. C.; Pitt, D.; Woods, M. The Utility of Image-Based Point Clouds for Forest Inventory: A Comparison with Airborne Laser Scanning. *Forests* **2013**, *4*, 518–536.
4. Rothermel, M.; Wenzel, K.; Fritsch, D.; Haala, N. SURE: Photogrammetric Surface Reconstruction from Imagery. In; Berlin, Germany, 2012.
5. Anonymous *MATCH-T DSM 5.5 Reference Manual*; IMPHO GmbH: Stuttgart, Germany, 2013.
6. Blombäck, Johanna Update of forest management plans from Södra using stereo matching of aerial images, the national terrain model and data from the National Forest Inventory. Master thesis in Forest Management, Swedish University of Agricultural Sciences, in press.
7. Næsset, E.; Gobakken, T.; Holmgren, J.; Hyyppä, H.; Hyyppä, J.; Maltamo, M.; Nilsson, M.; Olsson, H.; Persson, Å.; Söderman, U. Laser scanning of forest resources: the nordic experience. *Scand. J. For. Res.* **2004**, *19*, 482–499.

Assessing a high-resolution woody vegetation map of New South Wales, Australia

Adrian Fisher [1, 2], Mike Day [3], Tony Gill [1, 3], Tim Danaher [1, 3], Adam Roff [3]

[1] *Joint Remote Sensing Research Program, University of Queensland*
[2] *Centre for Ecosystem Science, University of New South Wales (adrian.fisher@unsw.edu.au)*
[3] *Native Vegetation Information Science, Office of Environment and Heritage, New South Wales*

Highlights: A large-area, 5 m resolution map of woody vegetation extent was derived from satellite data and assessed against a large sample of airborne lidar data. The map had an overall accuracy of 90%, although errors varied greatly depending on vegetation type, cover, height, and fragmentation.

Key words: *forest mapping, accuracy assessment, operational applications*

Introduction

The Australian state of New South Wales (NSW) covers a large area (809,444 km^2), a wide range of climates (arid-temperate-subtropical) and a wide variety of vegetation types (shrublands, grasslands, woodlands and forests). The state government uses maps of woody vegetation (defined as trees and shrubs > 2 m in height) for a range of applications including vegetation community mapping, biodiversity studies and native vegetation regulation. Landsat TM/ETM+ data have previously been used for mapping woody vegetation extent across the states of Queensland [1], NSW [2], and Victoria [3], as well as across the globe [4]. The NSW government required a higher resolution map to improve woody vegetation detection, particularly in open woodlands, grasslands and highly modified areas. A method was developed to map woody vegetation extent from Satellite pour l'Observation de la Terre 5 (SPOT5) High Resolution Geometrical (HRG) data at 5 m resolution across NSW, and to validate this map against woody vegetation extent derived from airborne lidar data. By deriving spatially continuous information from a sample of lidar data covering the different vegetation formations found across all regions of NSW, the accuracy of the satellite-derived map with respect to vegetation type, cover, height, fragmentation and location was assessed.

Method

Satellite-derived woody vegetation extent

The SPOT5 HRG data consists of 4 multispectral bands (10 m pixels), and a panchromatic band (2.5 m pixels). A time series of one image per year for the period 2008-2011 was acquired, for a total of 1256 images. The multispectral imagery was corrected for atmospheric and bi-directional reflectance effects [5] and sharpened to 5 m pixels using the panchromatic imagery. The images were masked for cloud and cloud shadow, topographic shadow, and water. An estimate of foliage projective cover (FPC, defined as the fraction of the ground obscured by green leaf) was derived using a multiple linear regression model that related the multi-spectral reflectance to a reference dataset of FPC derived from the existing Landsat model. The probability of a pixel containing woody vegetation was determined using a binomial logistic regression model using the following parameters: mean FPC, mean red reflectance, variation in FPC over time, and vapour pressure deficit. The model was trained using 25930 observations of woody/non-woody vegetation interpreted from aerial imagery or SPOT5 HRG panchromatic images. Woody vegetation extent was mapped by applying thresholds to the probability images, with further editing by trained analysts. Each woody pixel was then attributed with the mean FPC value over time (Fig. 1).

Lidar-derived woody vegetation products

Many discrete return lidar surveys have been acquired over NSW for government use, covering 46,382 km^2 or 6% of the state. The surveys are not randomly distributed, as many lidar surveys were acquired over rivers, wetlands and urban areas, where high resolution elevation models were required for flood modelling. To ensure that the validation of the SPOT5 HRG product was not biased towards these areas, a subset was taken, where each of the 17 vegetation formation classes [6] within each of the 11 Local Land Service (LLS) regions was represented (Fig. 1). The lidar subset contained 119 tiles of data, from 34 different surveys, covering around 347 km^2, equivalent to more than 13.8 million pixels at 5 m resolution. The surveys were acquired throughout 2008-2013, by three lidar instruments. Flying heights ranged from 500-2250 m above ground, footprints ranged from 0.23-1.21 m, and mean pulse density ranged from 1.1-8.3 pulses/m^2. Maximum scan angle averaged 16°, with a maximum of 40° in one tile.

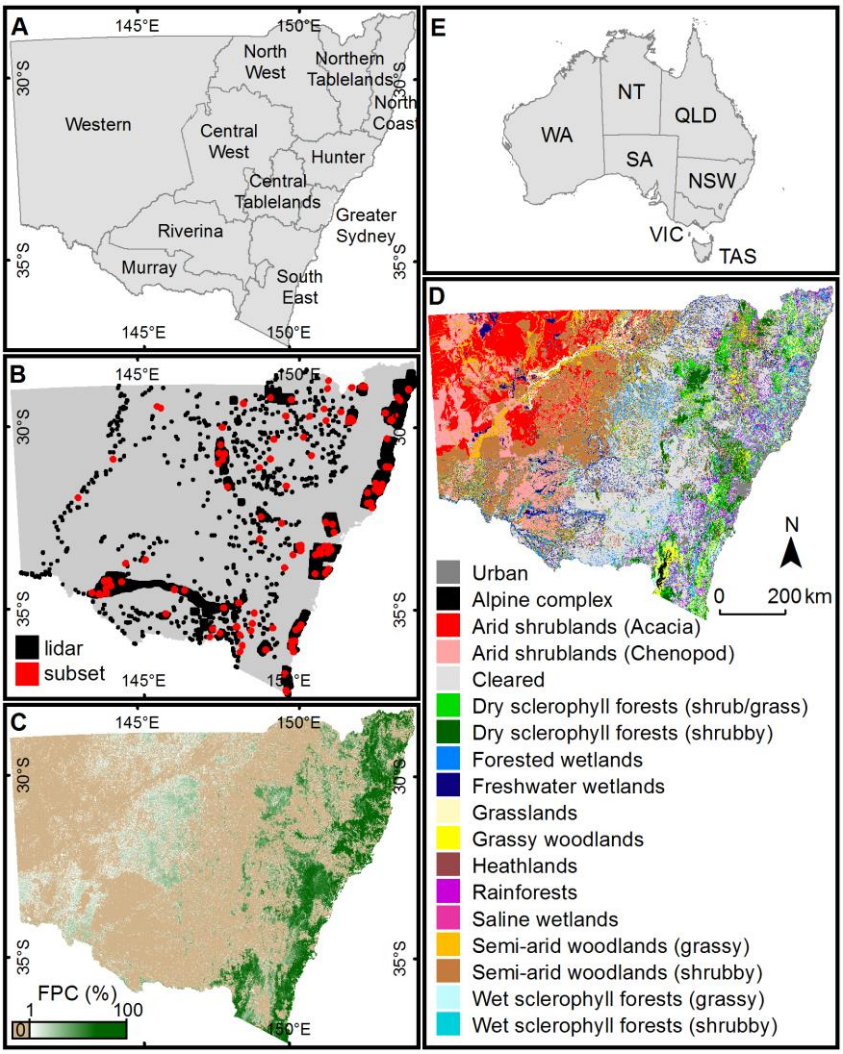


Figure 1: (A) Local Land Service regions of New South Wales (NSW). (B) Available lidar data for NSW from which the subset of reference data was sampled. (C) Satellite-derived foliage projective cover (FPC) for NSW. (D) Vegetation formations of NSW [6]. (E) Australian states.

Lidar data was sourced as LAS files with ground returns classified, which were converted to the Sorted Pulse Data (SPD) format [7] and spatially sorted into bins aligning with the 5 m SPOT5 HRG pixels. The height of each return above the ground was calculated using natural neighbour interpolation. Plant projective cover (PPC) was calculated for each bin, as the proportion of first returns from the canopy (> 0.5 m above the ground). Lidar woody vegetation extent was defined as pixels where PPC > 0 and at least one return was > 2 m above the ground. Maximum height of the vegetation was also derived as the 99th percentile of first return height. As each lidar tile created a spatially continuous map, the size of each contiguous woody/non-woody region, and the distance of each woody/non-woody pixel to the nearest pixel of the opposite class was also calculated.

Results

Using the lidar-derived woody vegetation extent as reference, the satellite derived woody vegetation extent had an overall accuracy, true positive rate (TPR) and users' accuracy > 90%, and a false positive rate (FPR) < 10%. The errors were not uniformly distributed across the state, with greatest omission in the semi-arid woodlands with grassy subformation (TPR of 41.6%), semi-arid woodlands with shrubby subformation (TPR of 53.1%), arid Chenopod shrublands (TPR of 54%) and cleared land (TPR of 55.6%), which are more common in the west. The least omission was observed in saline wetlands (TPR of 99.7%), dry sclerophyll forests with shrubby subformation (TPR of 98.7%), and wet sclerophyll forests with grassy subformation (TPR of 98.8%), which are more common in the east. In contrast, the greatest commission was found in the saline wetlands (FPR of 85.5%), heathlands (FPR of 66.8%), and dry sclerophyll forests with shrubby subformation (FPR of 58.8%) that are more common closer to the coast in the east, while the least commission was found in arid Chenopod shrublands (FPR of 0.5%) and grasslands (FPR of 1.5%) that are more common in the west.

Omission of woody vegetation pixels was more common where PPC was low and where vegetation was shorter. For example, there was 46% omission at 1-25% PPC and only 2% omission at 75-100% PPC, while there was 24% omission at 2-10 m height and 1% omission at > 30 m height. Omission of woody vegetation

pixels was also more common close to the edges of woody vegetation regions. For example, there was 39% omission of pixels that were < 10 m from a non-woody pixel and only 2% omission of pixels that were > 20 m from a non-woody pixel. The size of contiguous woody vegetation regions also influenced omission errors, with small regions having much greater omission than large regions. For example, there was 30% omission of pixels within woody regions that were < 1 ha and only 8% omission in regions > 3 ha.

Commission of non-woody vegetation pixels as woody was more common where maximum vegetation height approached 2 m. For example, there was 42% commission at 1.5-2 m height and only 6% commission at 0-0.5 m height. False positive woody vegetation pixels were also more common closer to the edges of woody regions, and within smaller non-woody regions. For example, there was 32% commission of pixels that were < 10 m from a woody pixel and only 2% commission of pixels that were > 30 m from a woody pixel, and there was 25% commission of pixels within non-woody regions that were < 1 ha and only 5% commission of pixels in regions > 3 ha.

Discussion

The assessment revealed that the type and magnitude of error in the satellite-derived woody vegetation extent was dependant on vegetation characteristics. Omission errors were greatest in arid and semi-arid regions where the vegetation is shorter, has thinner canopies and greater spacing between trees. Commission errors were greater in areas containing tall shrubs, such as heathlands, saline wetlands and dry sclerophyll forests with shrubby subformation. Both omission and commission errors were more common for pixels on the edges of woody regions, which may have partly been caused by differences in georeferencing between the satellite and lidar data. It was also observed that omission errors were common on the side of woody vegetation regions closer to the sun, and commission errors were common on the side in shadow, although this affect has not been quantified. The size of woody and non-woody regions containing errors was also found to have an effect, where omission and commission were more likely in smaller regions. This indicates that the SPOT woody vegetation extent has more errors in landscapes with fragmented, patchy woody vegetation, and where large continuous regions of woody vegetation contain small clearings. The accuracy of the SPOT woody extent map was also found to be similar to a previous Landsat based map, achieving slightly fewer omission errors, but with slightly more commission errors, despite the smaller spatial and spectral resolution of the sensor.

Conclusions

Using lidar-derived woody vegetation extent as reference data allowed the accuracy of the satellite-derived map to be assessed with respect to vegetation type, cover, height, and fragmentation. This informs users of the map, who can determine practical limitations based on their knowledge of local vegetation. Future work will assess the high resolution FPC values, once methods of deriving FPC from the lidar data are developed. This will also allow new satellite algorithms to be trained with the lidar data.

Acknowledgments

SPOT data were acquired from Airbus Defence and Space. Lidar data were supplied by NSW Land and Property Information and a number of commercial vendors funded from various sources. We owe a debt of gratitude to the numerous staff and volunteers who edited the maps.

References

[1] Armston, J. D., Denham, R. J., Danaher, T. J., Scarth, P. F., & Moffiet, T. N. (2009). Prediction and validation of foliage projective cover from Landsat-5 TM and Landsat-7 ETM+ imagery. Journal of Applied Remote Sensing, 3, 033540.

[2] NSW Office of Environment and Heritage (2008) Landsat woody extent and foliage projective cover (FPC) Version 2.1 (25 m). Available from http://mapdata.environment.nsw.gov.au/.

[3] Mellor, A., Haywood, A., Stone, C., & Jones, S. (2013) The Performance of Random Forests in an Operational Setting for Large Area Sclerophyll Forest Classification. Remote Sensing, 5, 2838-2856.

[4] Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O., & Townshend, J. R. (2013). High-resolution global maps of 21st-century forest cover change. Science, 342, 850-853.

[5] Flood, N., Danaher, T., Gill, T., & Gillingham, S.(2013). An operational scheme for deriving standardised surface reflectance from Landsat TM/ETM+ and SPOT HRG imagery for eastern Australia. Remote Sensing, 5, 83-109.

[6] Keith, D. A. (2004). Ocean Shores to Desert Dunes: The native vegetation of New South Wales and the ACT. Hurstville, NSW, Australia: Department of Environment and Conservation.

[7] Bunting, P., Armston, J., Lucas, R.M., & Clewley, D. (2013). Sorted pulse data (SPD) library. Part I: A generic file format for LiDAR data from pulsed laser systems in terrestrial environments. Computers & Geosciences, 56, 197-206.

Assessing the transferability of leaf area index estimation and the relative contribution to the dominant canopy within intensively managed Loblolly pine plantation forest from airborne LiDAR remote sensing within the south-eastern USA.

Matthew Sumnall[a], Alicia Peduzzi[b], Thomas R. Fox[a], Randolph H. Wynne[c], Valerie A. Thomas[a].

aVirginia Polytechnic Institute and State University, Department of Forest Resources and Environmental Conservation, 228 Cheatham Hall (Mail Code: 0324), Blacksburg, VA 24061, USA.
b United States Dept. of Agriculture – Forest Service, 507, 25th Street, Ogden, UT 84401, USA.
c Virginia Polytechnic Institute and State University, Department of Forest Resources and Environmental Conservation, 319 Cheatham Hall (Mail Code: 0324), Blacksburg, VA 24061, USA

Highlights:

- Demonstrates the transferability of leaf area index (LAI) prediction method over sites in the Southern-USA, and highlights the problems with the transferability of previous methods.
- Successfully demonstrates the vertical stratification of forest layers and the estimation of the contribution of the dominant canopy and understory to a sites total LAI.

Keywords: Loblolly pine; Forest management; airborne LiDAR; Leaf area index; vertical separation.

1. Introduction

The management of forested environments typically requires a great deal of information to ensure the objective of a site is achieved. For plantation forests this often means the production of biomass, or stemwood. The production potential for forest stands is correlated with the stands capacity to intercept light. This capacity, in turn, is a function of foliage production, lifespan and distribution. Of these, foliage production is believed to be the strongest link to productivity. Strong correlations between Leaf Area Index (LAI) and pine productivity have been reported. LAI is defined as half the total leaf surface area per ground surface area.

The vertical distribution of total leaf area within a forest stand will vary in forest environments which will impact the amount of light available to each tree. The density of vegetative components at higher levels will also affect the amount of light available for the growth within lower layers through shadowing. The sum total leaf area within a plot will include all dominant canopy and understory layer components, where the contribution of each layer can vary.

Both direct and indirect techniques exist for the assessment of LAI in situ, including destructive sampling and optical methods. Although these methods are considered accurate, they are time consuming, labor intensive and costly. Thus, they are impractical for measuring LAI continuously over large areas. The analysis of airborne discrete return laser scanner data (LiDAR) remote sensing can characterize both horizontal and vertical structures within forested environments. There exists a substantial amount LiDAR research studies which attempted to use the three-dimensional structural information from LiDAR to estimate LAI based on statistical methods. The estimation of vertical distribution of leaf area from airborne discrete return LiDAR has seen relatively little attention in terms of research, presumably do to the difficulty in the logistics of validating such vertical estimates across a study site.

A number of approaches have been applied to the estimation of vertical layer attributes, such as the heights of the dominant canopy and understory, using airborne discrete return LiDAR.

For example a number of studies utilize an approach where analysis is performed upon LiDAR returns which are summed at set heights, also known as volumetric pixels (voxels), in order to separate the vertical profile into different layers. Each voxel requires the horizontal and vertical extent to be defined by the user, for example 1×1×1m, allowing stacks of bins to be generated for each 1×1m of horizontal extent.

The aim of the current study was to develop a method to identify the vertical components of a given location and estimate the dominant canopy LAI component across multiple intensively managed Loblolly pine plantation forests subject to a number of different silvicultural management types using commercial airborne discrete return LiDAR technology at the field plot scale.

2. Materials and methods

2.1 Study sites and field data

A total of five study sites were visited which are located within North Carolina and Virginia, USA, in 2008 (88 plots) and 2013 (28 plots) in support of research studies investigating the role of intensive management in optimizing Loblolly pine (Pinus taeda L.) production. An additional three study sites were visited in Virginia, USA during 2014 (85 plots) and revisited in 2015. In each of the locations visited, field plots were established to sample various tree heights, planting densities and ages, in addition to sampling various understory densities, coverage and heights, and to include plots containing no understory. At each field location visited in between 2008 and 2014 tree height and LAI data was collected using the LiCor LAI-2000 Plant Canopy Analyzer on each plot during the summer. All field plots enumerated in 2014 were revisited during the winter (January) in 2015 to ensure leaf-off conditions for broadleaved understory vegetation.

2.2 LiDAR data

LiDAR data was produced from three separate acquisitions utilizing three different sensor designs and acquisition parameters: (i) 2008 – Optech ALTM 3100; (ii) 2013 – Riegl VQ-480; and (iii) 2014 – Leica ALS50, in parallel with field data collection. Plot level metrics relating to the distribution of return heights and intensity were generated. Plot-scale voxel stacks were created at the vertical separation of 0.2m, and a kernel smoothed regression line fit to the data, layer height parameters were extracted using local maxima and minima detection.

2.3 Statistical analysis

The LiDAR metrics generated from the 2008 and 2013 datasets were input into stepwise multiple regression analysis in order to generate a predictive model for LAI. This model was then validated against field metrics recorded in 2014. In order to remove the potential influence of understory LAI from the plot total, the plot locations visited in 2014-15 were used to generate a predictive model for estimating the difference in LAI between the summer and winter. This correction was then applied to winter LAI values of plot locations containing understory creating estimates of canopy only LAI.

2.4 Canopy only LAI

Layer height metrics from the voxel based analysis were utilized to re-compute those LiDAR metrics utilized in the LAI predictive models. Canopy only LAI estimates were then computed and compared to the field-based estimates of canopy only LAI.

3. Results and Discussion

Those LiDAR derived metrics (from 2008-2013) which calculated non-linear transformed indices using ratios of the sum of returns from different heights against one another, which used all returns, produced the strongest correlations (R^2 0.88; RMSE 0.36). The same model applied to the

data collected in 2014 also correlated strongly (R^2 0.86; RMSE 0.37). The results of the current study initially demonstrate the estimation of LAI across multiple disparate sensor or data acquisition configurations using a single model, and which offers the potential of a transferable model for similar study sites within the South-Eastern US, or at least those of managed Loblolly pine. Those indices derived from single, first-of-many and last returns or the sum of return intensities consistently produced lower correlations and exhibited a number of differences in terms of value distribution which indicates separate populations corresponding to the data from the two separate LiDAR acquisitions. This implies that indices developed using such data from disparate LiDAR sensor or data acquisition configurations are not transferable, or require acquisition specific calibration.

The method of estimating the dominant canopy LAI from LiDAR derived indices produced strong correlations between field estimates for (i) plots with understory R^2 0.72 (RMSE 0.35) and (ii) when this data was combined with plots containing no understory R^2 0.83 (RMSE 0.46). The application of this approach has could assist site management and augment existing estimates of stand LAI to assess the productivity of a dominant vegetation layer, and the effects of competing vegetation. Although there were some problems with the detection of vertical layers merged together, understory presence could be detected in the majority of cases. In terms of management objectives within the context of managed Loblolly pine forests, Fox et al. [1] suggests that the optimal level of pine peak-LAI (3.5 or greater) may require competition levels to be managed in order to reach this optimal level. An assessment of this type, from the plot to landscape scale, could now be potentially feasible using airborne LiDAR.

4. Conclusion

The application of discrete return airborne LiDAR hold potential for estimating total LAI with high accuracy horizontally when compared with multispectral image analysis, particularly in forests of high biomass, and in closed canopy conditions with understory components beneath. The results of the current study demonstrate the potential transferability of the LAI prediction model, the capability to detect vertical canopy layers and the prediction of LAI above height thresholds defined by those layers.

5. Acknowledgements

This research was possible thanks to the support from the Forest Productivity Cooperative. Funding for this work was provided in part by the Virginia Agricultural Experiment Station and the Program McIntire Stennis of the National Institute of Food and Agriculture, U.S. Department of Agriculture.

6. References

[1] Fox, T.R., Allen, H.L., Albaugh, T.J., Rubilar, R.A., and Carlson, C.A. 2007. Tree nutrition and forest fertilization of pine plantations of the southern United States. South J. Appl. For. 31(1): 5-11.

Mapping plant area index of tropical forest by Lidar: calibrating ALS with TLS

Grégoire Vincent[1], Cécile Antin[1], Jean Dauzat[2], Eloi Grau[3,4], Sylvie Durrieu[3]

[1]*IRD, UMR AMAP, Montpellier, 34000 France*
[2]*CIRAD, UMR AMAP, Montpellier, 34000 France*
[3]*CNES, UMR TETIS, Montpellier, 34000 France*
[4]*IRSTEA, UMR TETIS, Montpellier, 34000 France*

Highlights: We compare Plant Area Density (PAD) profiles derived from Terrestrial Laser Scanning (TLS) and contemporaneous Aerial Laser Scanning (ALS) in dense tropical forest. Poor sampling of the lower part of the canopy profile by ALS is mitigated by using a multiple resolution approach. Anisotropy of transmittance revealed by TLS allows further improvement of PAD estimates.

Key words: *TLS, ALS, fusion, Leaf Area Index, Tropical moist forest.*

Introduction

Leaf area Index (LAI, the one-sided green leaf area per unit ground surface area m^2/m^2) is a major driver of water and carbon exchanges and controls radiation extinction within the vegetation. It is therefore a key parameter in biogeochemical cycles and in vegetation dynamics. However, the measurement of Leaf Area Index in dense evergreen tropical moist forest is a considerable challenge. LAI estimates of moist tropical evergreen forests almost exclusively rest on indirect methods which all have serious limitations. Hemispherical photographs and optical devices such as the Leaf Area Meter LAI2200 are very sensitive to assumptions made on degree of clumping [1] and leaf angular distribution [2]. Besides, these methods do not allow distinguishing between photosynthetic and non-photosynthetic materials, leading to the estimation of PAI (Plant Area Index) rather than LAI. Remote sensing indices (NDVI, EVI) typically saturate at high LAI values [3]. The ability of Lidar to measure forest 3D structure, with an unequalled level of detail at both tree and plot level using Terrestrial Laser Scanning (TLS) systems and over large area with a coarser resolution using Airborne Laser Scanning (ALS) systems, opens up new possibilities to assess this key forest variable.

Material an methods

Multiple-return ALS provides spatially accurate information on laser beam progressive extinction during its downward path through the canopy. Therefore it has obvious potential for LAI estimation. A number of studies using different approaches have been conducted to evaluate the potential of ALS for this purpose [4-6]. In the present study we took a slightly different stance from previous works by 1) bringing in more complete information on the return signal provided by an up-to-date Full Wave Form ALS system (Riegl LMS Q560) combined with 2) precisely georeferenced multiple-return TLS (Riegl VZ400) acquisitions.

Using aerial laser scanning, backscattered energy is digitized and post processed to identify distinct echoes whose 3D position and amplitude are recorded [7]. From this information the projected area of objects intercepted by the laser beams are estimated, conditional to a number of regularity assumptions about vegetation optical properties.

A statistical analysis of ALS data showed that c. 50% of the variability of vegetation returns amplitude was explained by the echo rank and the total number of returns detected. Ray tracking of all laser shots within a voxel space and localization of their echoes allowed determining a local vegetation transmittance which was used to estimate local vegetation density (PAD m^2/m^3). However in our data set almost one third of the canopy volume (at 1 m^3 resolution) was not sampled by any laser pulse optical path. In such case or in the case transmittance was null, local transmittance was recomputed from the next coarser resolution and the process iterated to fill in the gaps.

Conversely all canopy voxels were sampled by at least 1000 shots using our dense multiple TLS acquisitions protocol. However because TLS data suffer from a fundamental indeterminacy (the generally unknown fraction of a given pulse which may not have been intercepted) a slightly different algorithm was used to compute the transmittance from TLS data. Transmittance in any voxel was computed as the weighted mean ratio of returns per shot (where the weight is the shot path length within a voxel) and single shot transmittance is either 0 or 1.
TLS multiple station acquisition and resulting omnidirectional sampling also provided a way of exploring transmittance dependence on shooting inclination angle. We analysed the transmittance directionality by sub-

setting TLS shooting elevations per 10-degree steps and analysed transmittance per direction along the entire canopy height profile.

Results and perspectives

Plant Area Density (PAD) was estimated from transmittance values assuming a spherical distribution function of foliage elements. Vegetation vertical profiles (mean PAD per vegetation layer) were established for six quadrats of 24mx24m in the scanned area (Figure 1). Three different profiles are shown per quadrat: ALS derived profiles obtained with or without using the multi-resolution approach as well as the TLS derived profiles. Multi-resolution processing efficiently corrected for most of the discrepancies between ALS and TLS profiles occurring in the lower canopy which was under-sampled by ALS. PAI estimates from TLS and multiresolution appeared to be within 5% of each other at the spatial scale considered.

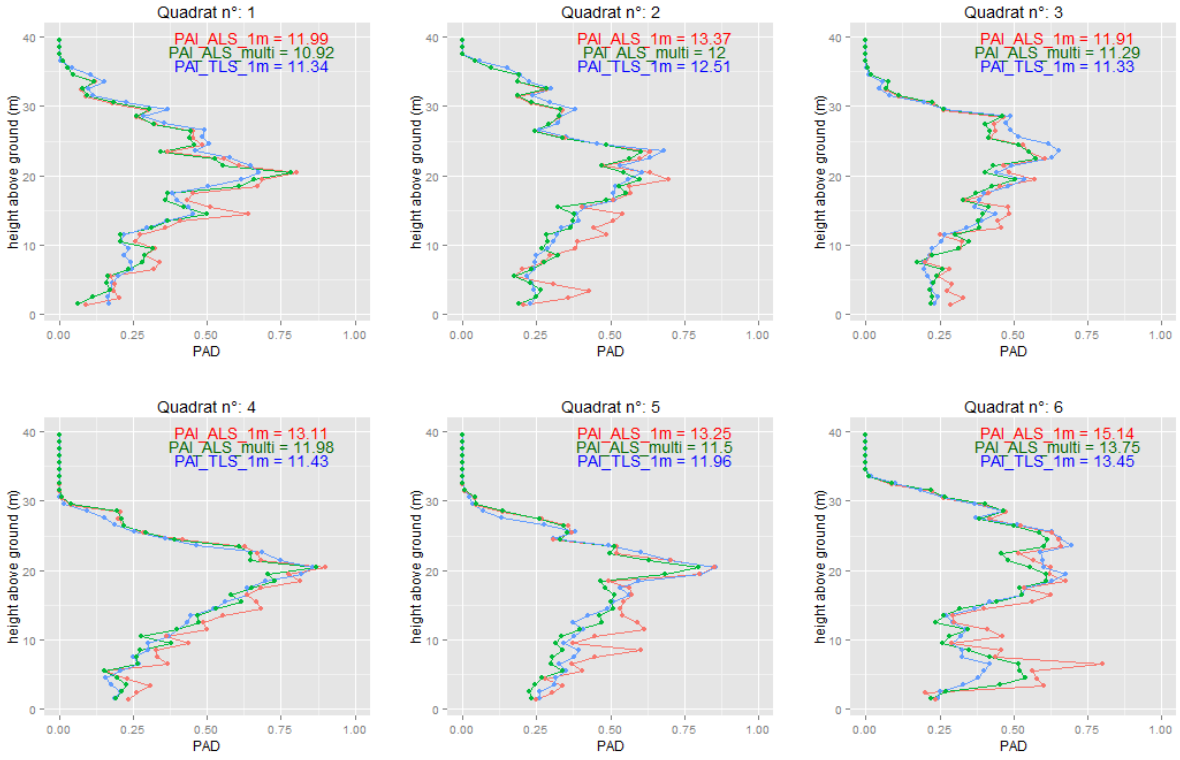


Figure 1: Vegetation density profiles (PAD) per 24mx24 m quadrats in dense tropical forest obtained either by Terrestrial laser scanner (TLS) or aerial laser scanner (ALS). Elementary volume (voxel) is 1 meter except for the ALS_multi case which uses mutiple resolution (here 1 to 4m) to fill in non sampled elementary voxels.

Variation in transmittance with inclination was found to be significant indicating that the actual leaf (or rather plant material) inclination distribution function was not spherical. Remarkably this anisotropic response was very similar from top to bottom of canopy. To evaluate the error made by assuming a spherical leaf inclination distribution function, we will adjust an ellipsoidal foliage element angle distribution and Plant Area Density to fit the observed directional transmittance values. The parameterized ellipsoidal distribution function will then be used to compute PAD from ALS data. The magnitude of the difference in PAD profiles resulting from the use of two leaf angle distribution functions will be estimated (underway).

Further improvement in ALS estimates may be obtained by taking into account the local variation in optical properties (reflectance) occurring between individual crowns. This should improve transmittance estimates from which PAD profiles are derived. Further developments will be needed to assess LAD profile instead of PAD profiles from both TLS and ALS data, e.g. by developing methods to separate foliage from large woody elements. Dense TLS sampling could then be used to evaluate the foliage clumping index to further improve LAI estimates.

Acknowledgements

The authors acknowledge funding from the French CNES/TOSCA program and from the Laboratoire d'Excellence CEBA (ANR-10-LABX-25)

References

[1] Ryu, Y., T. Nilson, H. Kobayashi, O. Sonnentag, B.E. Law & D.D. Baldocchi, (2010). On the correct estimation of effective leaf area index: Does it reveal information on clumping effects? Agricultural and Forest Meteorology, 150(3) , 463-472.

[2] Pisek, J., O. Sonnentag, A.D. Richardson & M. Mõttus. (2013). Is the spherical leaf inclination angle distribution a valid assumption for temperate and boreal broadleaf tree species? Agricultural and Forest Meteorology, 169(0), 186-194.

[3] Wang, Q., S. Adiku, J. Tenhunen & A. Granier. (2005) On the relationship of NDVI with leaf area index in a deciduous forest site. Remote Sensing of Environment, 94(2), 244-255.

[4] Morsdorf, F., B. Kötz, E. Meier, K.I. Itten & B. Allgöwer, (2006) Estimation of LAI and fractional cover from small footprint airborne laser scanning data based on gap fraction. Remote Sensing of Environment, 104, 50–61.

[5] Richardson, J.J., L.M. Moskal & S.-H. Kim. (2009) Modeling approaches to estimate effective leaf area index from aerial discrete-return LIDAR. Agricultural and Forest Meteorology, 149(6–7), 1152-1160.

[6] Detto, M., G.P. Asner, H.C. Muller-Landau & O. Sonnentag, Spatial Variability in Tropical Forest Leaf Area Density from Multireturn LiDAR and Modeling. Journal of Geophysical Research: Biogeosciences, 2015, 2014JG002774.

[7] Wagner, W., A. Ullrich, V. Ducic, T. Melzer & N. Studnicka (2006) Gaussian decomposition and calibration of a novel small-footprint full-waveform digitising airborne laser scanner. ISPRS Journal of Photogrammetry and Remote Sensing, 60(2), 100-112.

Retrieving forest canopy leaf area index by incorporating scan angle information from discrete aerial lidar data

Guang Zheng[1], Lixia Ma[1], Wei He[1], L. Monika Moskal[2]
[1]International Institute for Earth System Science, Nanjing University, China, 210023
[2]Remote Sensing and Geospatial Analysis Laboratory, Precision Forestry Cooperative, School of Environment and Forest Science, University of Washington Box 352100, Seattle, Washington, USA, 98115-2100

Abstract

The directional gap fraction (DGF) is an important factor for retrieving forest canopy biophysical structural parameters, bidirectional reflectance distribution function (BRDF), and radiation regime within and below a forest canopy. Laser beams actively emitted from light detection and ranging (lidar) system usually penetrate through a forest canopy with a specific off-nadir direction. In the present study, the range of scan angles of the aerial laser scanning (ALS) system was so wide (-29 ~ 29 degrees) that the incident angle of laser beams should not be neglected. We proposed a novel physically-based method to retrieve the DGF, extinction coefficient and effective leaf area index (LAIe) only based on the X, Y, Z coordinates of each point. Both single- and multi-flight-coverage forest plots were analyzed for computing the DGF both for each point stripe with same scan angle and the whole forest plot points. In addition, due to the uncompleted sample of the forest canopy surface and limited penetration ability of laser beams from ALS system, the ALS-based LAIe could predict 71.92 % variation of digital hemispherical photography (DHP) based LAIe estimation results with the linear regression statistical model as $Y = 1.042X + 0.899 (N = 18,\ p < 0.001)$, however, the ALS-based LAIe tended to underestimate the results from DHP-based method, especially for the forest plots with high density and wide scan angle.

Key words: Directional gap fraction, effective leaf area index, aerial laser scanning, scan angle

1. Introduction

Leaf area index (LAI) is one of the most key parameters for different ecological process-based and bidirectional reflectance distribution function (BRDF) models [1-5]. Beer-law is the theoretical foundation for LAI retrieving from either ground or remotely sensed data based algorithms [6-11], which quantitatively characterizes the light attenuation during penetrating through a forest canopy with the non-randomly distributed foliage elements. It could be defined as:

$$\ln(P(\theta)) = -G(\theta)L_e / \cos(\theta) \quad (1)$$

where θ is the zenith angle of incoming solar beams; *Le* is the effective leaf area index (LAIe) [12] which could be converted into true LAI by multiplying clumping index (Ω) [13] associated with tree species; P(θ)and G(θ)/cos(θ) are the gap fraction and extinction coefficient of the foliage elements for the incoming solar beams with θ zenith angle and full-range azimuthal angles (0 ~ 360 degrees), respectively.

In order to differentiate the FOV- and directional- based measurements and provide generic terms for different applications, we defined the “angular gap fraction (AGF)” and “directional gap fraction (DGF)”. The AGF is capable to describe the gap fraction of a forest canopy region within the same zenith angle and whole or partial azimuthal angles from ground-based perspective such as the whole or partial annulus rings of a digital hemispherical photo. However, the DGF is more appropriate to characterize the spatial variation of the view angle dependent gap fraction for a forest canopy observed from the sky. Lidar has been widely used to retrieve forest canopy structural parameters due to the build-in 3D mapping ability by measuring the transmittance distances of laser beams between the scanner and target objects at the speed of light [14-16].

The current work attempted to develop an approach to retrieve the extinction coefficient based on point cloud data generated using ALS with considering scan angle information. By calculating the DGF and extinction coefficient with considering the scan angle information for each point stripe of a forest plot point cloud data, the LAIe results could be obtained for both each point stripe and the whole forest plot. The objectives of this paper are as follows:
Developing the physical based methods to estimate the DGF, extinction coefficient, and LAIe of a forest point cloud data for parallel incident solar beams or laser beams with same scan angle emitted from ALS system; and

(1) Exploring the variation of the DGF of a forest canopy by changing the view angles from sky by using the multi-coverage forest plot point cloud data generated from ALS system.

2. Materials and methods

2.1 Study area

Washington Park Arboretum (WPA) located in south of University of Washington, Seattle, USA, which is an urban heterogeneous forest with over 20,000 trees, shrubs and vines (Figure 3), the topographic elevation variation of WPA falls in the range from 10 m to 48.47 m, and the slope variation is less than 15%. 27 circular plots with 30 m radius were selected in this study with the LAIe ranging from 0.19 to 6.48, and a range of 4,000 different tree species.

2.2 DGF estimation

The non-ground forest point cloud restored into original spatial distribution pattern was then divided into different stripes according to scan angles. The bounding box of each point cloud stripe with same scan angle was first

reconstructed with the minimum and maximum value of X, Y, Z coordinates as the corner point coordinates, and the geometric center of the bounding box was set as the new origin of the Cartesian coordinate system. By assuming that the incoming solar beams always come from nadir direction, by rotating the point cloud data, we can simulate the different directions of incoming solar radiation. The direction of incoming parallel laser beams would be the nadir direction, however, the relative position between the incident laser beams and points and the spatial distribution pattern of the points was kept the same as their original situation. Then, the voxelization process was implemented for the rotated point stripes in order to estimate the DGF, a number of voxels would be produced in 3-D space to explore the forest point distribution pattern.

2.3 Extinction coefficient estimation

For the incident parallel laser beams with fixed direction, the extinction coefficient of a forest stand could be obtained by finding the orientation distribution of the "effective foliage elements". Although only partial surface was sampled due to the occlusion effect for some single flight coverage forest plots, the sampled points represented the effective foliage elements in terms of the penetration of parallel incident beams with specific direction. Thus, extinction coefficient could be obtained through the ellipsoidal model [17] after characterizing the foliage elements orientation distribution. In the present work, the method developed by Zheng and Moskal [18] was employed to retrieve the foliage element orientation distribution by computing the normal vectors for every point within a neighbor region of a forest plot point cloud data. The normal vectors of each point within a neighbor region was reconstructed for a forest plot point cloud data. The leaf angle distribution could be characterized by computing the normal vectors for given point cloud data.

2.4 LAIe estimation

Dividing forest plot point cloud into different stripes according to different scan angles of incident solar or laser beams, the LAIe for the whole forest stand could be obtained by summing up the LAIe for each stripe. The LAIe for each stripe was retrieved based on the equation 1 through computing the DGF and directional extinction coefficient.

$$-\ln P(\theta,\alpha) = G(\theta,\alpha)\times L_e / \cos(\theta) \tag{2}$$

where θ is the zenith angle of incoming solar or laser beams, α is the azimuthal angle of foliage element, Le is the LAIe. Therefore, the LAIe for the whole plot could be obtained through computing the following equation:

$$\begin{aligned} Le &= \omega_1 Le_1(\theta_1,\alpha) + \omega_2 Le_2(\theta_2,\alpha) + \cdots + \omega_n Le_n(\theta_n,\alpha) \\ &= -\left(\omega_1 \cdot \cos\theta_1 \cdot \frac{\ln P(\theta_1,\alpha)}{G(\theta_1,\alpha)} + \omega_2 \cdot \cos\theta_2 \cdot \frac{\ln P(\theta_2,\alpha)}{G(\theta_2,\alpha)} + \cdots + \omega_n \cdot \cos\theta_n \cdot \frac{\ln P(\theta_n,\alpha)}{G(\theta_n,\alpha)} \right) \end{aligned} \tag{3}$$

where $\omega_1, \omega_2, \ldots, \omega_n$, were proportion of the i-th Le of the point stripe with same scan angle to the total Le of the whole forest plot, it could be computed as the ratio of the number of points in each stripe over the total point number of each forest plot.

3. Results

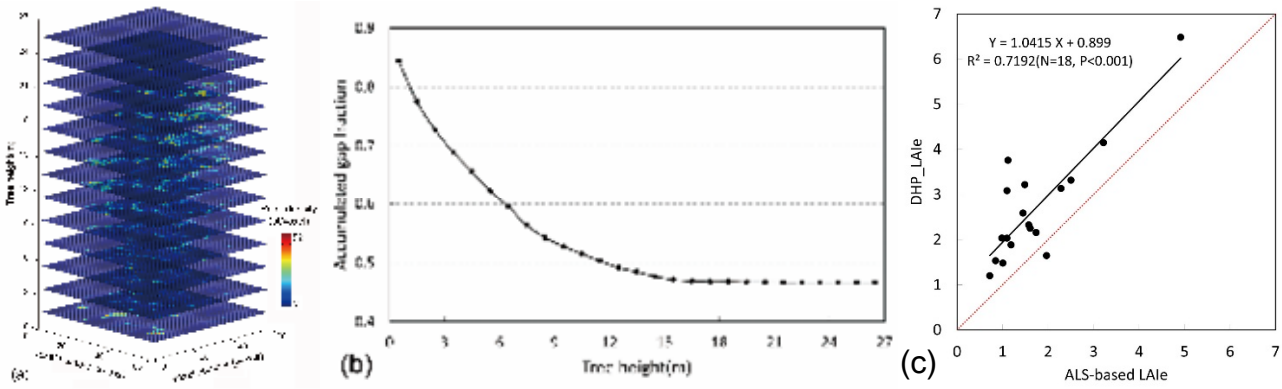


Figure 1. DGF variation along different tree height. Inset (a) is the point density distribution map along different elevations; (b) is the accumulated directional gap fraction from the bottom to top of plot 29 with multi-scan coverage; (c) is the comparison between DHP-based LAIe and the ALS-based directional LAIe by incorporating scan angle information.

4. Discussions

4.1 Anisotropy of view angle dependent DGF

The anisotropy of DGF indicated the importance of the view direction in terms of DGF computation, especially for the lidar-based BRDF study. The actively emitted laser beams usually penetrated through a forest canopy with specific

direction and sampled partial forest canopy surface, thus, the DGF is an important proxy to characterize the light penetration, radiation regime, and the geometric optical characteristics of a forest canopy.

4.2 Sensitivity analysis of voxel size on DGF

Changing of voxel size will alter the number of points in each voxel, and further affect the DGF in each slice plane. Thus, it is necessary to explore the optimum voxel size in terms of the DGF estimation based on the point cloud data generated from ALS. By changing the side length of cubic voxel from 0.1 to 5.0 m, the vertical DGFs were calculated accordingly we found that there is a relative stable range existing ranging from 0.5 m to 1.5m.

5. Conclusions

By retrieving the LAIe with considering the scan angle information through computing the DGF and extinction coefficient for different forest plots, we found that DGF was a view angle dependent variable which suggested that the scan angle of laser beams should not be ignored in terms of the LAIe estimation based on the point cloud data generated from ALS system. The ALS-based DGF will be a key factor to investigate the directional reflectance characteristic of a forest canopy based on the physical-based BRDF model driven by ALS produced point cloud data.

References:

[1] Y. Shao, J. Pan, and B. Sun, "Study on characteristics of soil organic carbon decompositions and carbon pool under different vegetations," *Journal of Soil and Water Conservation*, vol. 19, pp. 24-28, 2005.

[2] J. M. Chen and S. G. Leblanc, "A four-scale bidirectional reflectance model based on canopy architecture," *Ieee Transactions on Geoscience and Remote Sensing*, vol. 35, pp. 1316-1337, Sep 1997.

[3] C. Hopkinson, J. Lovell, L. Chasmer, D. Jupp, N. Kljun, and E. van Gorsel, "Integrating terrestrial and airborne lidar to calibrate a 3D canopy model of effective leaf area index," *Remote Sensing of Environment*, vol. 136, pp. 301-314, Sep 2013.

[4] W. Qin and S. A. Gerstl, "3-D scene modeling of semidesert vegetation cover and its radiation regime," *Remote Sensing of Environment*, vol. 74, pp. 145-162, 2000.

[5] J. Liu, J. M. Chen, J. Cihlar, and W. M. Park, "A process-based boreal ecosystem productivity simulator using remote sensing inputs," *Remote Sensing of Environment*, vol. 62, pp. 158-175, Nov 1997.

[6] T. Hirose, "Development of the Monsi-Saeki theory on canopy structure and function," *Annals of Botany*, vol. 95, pp. 483-494, Feb 2005.

[7] G. Zheng and L. M. Moskal, "Retrieving Leaf Area Index (LAI) Using Remote Sensing: Theories, Methods and Sensors," *Sensors*, vol. 9, pp. 2719-2745, Apr 2009.

[8] H. Tang, M. Brolly, F. Zhao, A. H. Strahler, C. L. Schaaf, S. Ganguly, G. Zhang, and R. Dubayah, "Deriving and validating Leaf Area Index (LAI) at multiple spatial scales through lidar remote sensing: A case study in Sierra National Forest, CA," *Remote Sensing of Environment*, vol. 143, pp. 131-141, 2014.

[9] G. Zheng, L. M. Moskal, and S. H. Kim, "Retrieval of Effective Leaf Area Index in Heterogeneous Forests With Terrestrial Laser Scanning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, pp. 777-786, February 2013.

[10] K. G. Zhao and S. Popescu, "Lidar-based mapping of leaf area index and its use for validating GLOBCARBON satellite LAI product in a temperate forest of the southern USA," *Remote Sensing of Environment*, vol. 113, pp. 1628-1645, August 2009.

[11] J. M. Chen, P. M. Rich, S. T. Gower, J. M. Norman, and S. Plummer, "Leaf area index of boreal forests: Theory, techniques, and measurements," *Journal of Geophysical Research-Atmospheres*, vol. 102, pp. 29429-29443, Dec 1997.

[12] J. M. Chen and T. A. Black, "Measuring leaf area index of plant canopies with branch architecture," *Agricultural and Forest Meteorology*, vol. 57, pp. 1-12, Dec 1991.

[13] J. M. Chen and J. Cihlar, "Plant canpy gap-size analysis theory for improving optical measurement of leaf-area index," *Applied Optics*, vol. 34, pp. 6211-6222, Sep 1995.

[14] M. A. Lefsky, W. B. Cohen, S. A. Acker, G. G. Parker, T. A. Spies, and D. Harding, "Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests," *Remote Sensing of Environment*, vol. 70, pp. 339-361, Dec 1999.

[15] J. E. Means, S. A. Acker, B. J. Fitt, M. Renslow, L. Emerson, and C. J. Hendrix, "Predicting forest stand characteristics with airborne scanning lidar," *Photogrammetric Engineering and Remote Sensing*, vol. 66, pp. 1367-1371, Nov 2000.

[16] J. B. Drake, R. O. Dubayah, D. B. Clark, R. G. Knox, J. B. Blair, M. A. Hofton, R. L. Chazdon, J. F. Weishampel, and S. D. Prince, "Estimation of tropical forest structural characteristics using large-footprint lidar," *Remote Sensing of Environment*, vol. 79, pp. 305-319, Feb 2002.

[17] G. S. Campbell, "Extinction coefficients for radiation in plant canopies calculated using an ellipsoidal inclination distribution," *Agricultural and Forest Meteorology*, vol. 36, pp. 317-321, Apr 1986.

[18] G. Zheng and L. M. Moskal, "Leaf Orientation Retrieval From Terrestrial Laser Scanning (TLS) Data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, pp. 3970-3979, April, 2012.

Validation of continental scale vertical plant profile mapping using waveform lidar airborne laser scanning

John Armston[1,2], Peter Scarth[2], Richard Lucas[3,2]
Philip Lewis[4,5], Mathias Disney[4,5], Stuart Phinn[2]

[1]*Remote Sensing Centre, Queensland Department of Science, Information Technology and Innovation, Australia*
[2]*Joint Remote Sensing Research Program, School of Geography, Planning and Environmental Management, University of Queensland, Australia*
[3]*School of Biological, Earth and Environmental Sciences, University of New South Wales, Australia*
[4]*Department of Geography, University College London*
[5]*NERC National Centre for Earth Observation (NCEO)*

Highlights: A simple radiative transfer model was adapted to map directional gap probability at high spatial resolution (1m) using waveform lidar datasets from Australia's Terrestrial Ecosystem Research Network (TERN)'s remote sensing program (Auscover). These products validated maps of vertical plant profiles for the Australian continent derived through integration of ICESat GLAS waveforms with ALOS PALSAR and Landsat data products.

Key words: *Airborne laser scanning, waveform lidar, large area mapping, uncertainty*

Introduction

Vertical plant profiles are important descriptors of canopy structure and current research is exploring ways of combining these with radar and optical remote sensing data products to predict above-ground biomass (AGB) over large areas. In Australia, an approach has been developed to produce large area maps of vertical plant profiles by extrapolating waveform lidar estimates of vertical plant profiles from ICESat/GLAS using large area segmentation of ALOS PALSAR and Landsat satellite image products [1]. A preliminary national product has been developed under the auspices of the Japanese Space Agency (JAXA) Kyoto & Carbon (K&C) Initiative (http://www.eorc.jaxa.jp/ALOS/en/kyoto/kyoto_index.htm). However independent validation is required to assess large area segmentations, validate retrieved height and cover metrics, and identify limitations to be targeted in future development.

Airborne laser scanning data may be used to estimate vertically resolved directional gap probability $P_{gap}(\theta, z)$ and derive vertical plant profiles equivalent to the ICESat/GLAS products produced by [1]. We previously developed and validated a method for the direct retrieval of $P_{gap}(\theta, z)$ from waveform lidar using data acquired from airborne laser scanning platforms [2]. In contrast to discrete return data, the $P_{gap}(\theta, z)$ retrievals were insensitive to differences in sensor and survey configuration but estimates were biased when local spatial variation in ground reflectance was highest at distances greater than the footprint size but less than the plot size. Here we extend the method to account for local spatial variance in ground reflectance and map over large areas. The approach has been applied to Terrestrial Ecosystem Research Network (TERN) Auscover data for independent validation of the Australian vertical plant profile mapping.

Australian Vertical Plant Profile Mapping

ALOS PALSAR L-band Fine Beam Dual (HH and HV) polarisation data mosaics at 25 m spatial resolution, provided through the JAXA K&C Initiative, were combined with an Australian mosaic of persistent green vegetation cover derived from Landsat sensor time-series data [1]. These data were segmented to map clumps of contiguous vegetation structure and reduced to a maximum of 4000 unique classes using *k*-means based cluster analysis. These classes were then intersected with ICESat/GLAS L2 (Release 33) Global Land Surface Altimetry Data (GLA14) data acquired between 2003 and 2009 and obtained from the National Snow and Ice Data Centre (NSIDC). All GLAS laser shots were corrected for differences in footprint size, laser output power, vegetation/ground reflectance, and terrain slope. These laser shots were aggregated for each cluster and processed to mean vertical profiles of $P_{gap}(\theta, z)$. Vertical plant profiles were then derived and height of peak foliage density (H_{peak}) and height percentiles (95th, 75th, 50th, and 25th) extracted. To produce a national map, each clump in the original segmentation was then assigned the mean vertical profile corresponding to the nearest cluster centroid.

Methods

Airborne laser scanning data collection

The 11 TERN Auscover sites are located in significant Australian biomes across a wide range of environmental conditions from wet-tropics rainforest to low open woodlands and shrublands. Airborne lidar surveys have been undertaken at these sites since February 2012 and are ongoing. Each survey used a waveform recording RIEGL LMS-Q560 operating at a nominal altitude of 300m and with >50% overlap between flight paths, resulting in > 10 pulses per square metre. Each site nominally covers a 5 by 5 km area, and so often provides complete sampling of vertical plant profile map segments. These data are open access and may be obtained from http://www.auscover.org.au/data/product-list.

Waveform lidar product development

Waveform lidar processing was implemented for large datasets using an open source processing framework developed in Python and based on the sorted pulse data model [3]. Decomposition of waveforms, calibration and estimation of above-ground heights for individual returns are described by [2] and references therein. High spatial resolution (1 m) DEM, canopy height model, and canopy mask image products were also derived for interpretation of errors in the Australian vertical plant profile mapping.

Vertical profiles of the directional canopy gap probability $P_{gap}(\theta, z)$ were derived for each laser pulse using a model developed by [2]:

$$P_{gap}(\theta, z) = 1 - \frac{\sum_{z=z_i}^{z=\max(z)} R_{v,i}}{R_v} \frac{1}{1+\frac{R_g}{J_0\rho_g - R_g}} \quad (1)$$

where R_v is the integral of canopy return intensity, R_g the integral of ground return intensity, and $J_0\rho_g$ the ground volumetric backscattering coefficient accounting for transmission effects. $J_0\rho_g$ may be estimated from the integral of unimodal ground returns and averaged for a local area [2]. We assumed θ was equal to zero for the present analysis. To avoid the assumption of constant ρ_g in a local area, we used a robust 2D smoothing spline based on the discrete cosine transform [4] to derive grids of ρ_g at the native spatial resolution of the ALS datasets (1 m).

Using these estimates of ρ_g, $P_{gap}(\theta, z)$ were derived for each grid cell. Estimates of ρ_v were also calculated by: $\rho_v = -R_v\,\rho_g/(R_g - \rho_g)$. The $P_{gap}(z)$ estimates were then aggregated to the Australian vertical plant profile map image segments and equivalent profiles calculated. Estimates of height of peak foliage density (H_{peak}) and height percentiles (95th, 75th, 50th, and 25th) were then extracted from these vertical plant profiles for comparison.

Results and Discussion

Waveform lidar data products

Maps of ρ_g, ρ_v, and $P_{gap}(\theta, z)$ at 1 m spatial resolution are shown in Figure 1(a-c) for the Karawatha Forest Park TERN Auscover site in South East Queensland. The ρ_g and ρ_v maps show spatially variable ground reflectance and relatively constant canopy reflectance for the Eucalyptus spp. dominated canopy. Validation of $P_{gap}(\theta, z)$ estimates using field measurements across all TERN Auscover and state government agency sites also highlight the need for modelling spatially variable ground reflectance over assuming a constant in the estimation of $P_{gap}(\theta, z)$. Estimates of the height of peak foliage density derived from the $P_{gap}(\theta, z)$ profiles at 1m spatial resolution are shown in Figure 1(d).

Validation of ICESat/GLAS mapping

Initial results indicate a closer correspondence between relative apparent vertical plant profiles than absolute vertical plant profiles. An example is provided in Figure 1(f), where height percentiles from the relative $P_{gap}(z)$ profiles are similar. However the vertical plant profile from the K&C map show a higher level of cover compared to the reference (RIEGL) profile. This may be at least partially explained by the date difference between the airborne and spaceborne data captures.

Initial validation show r^2 values between 0.4 and 0.7 for the different height metrics at the Karawatha Forest Park TERN Auscover site. However the comparison needs to be expanded to all sites to capture a greater range of vegetation height and cover classes. As well as expand the comparison across all TERN Auscover and State

Government monitoring sites, current work is testing if within-cluster vertical plant profiles, metrics and distributions are consistent between map segments.

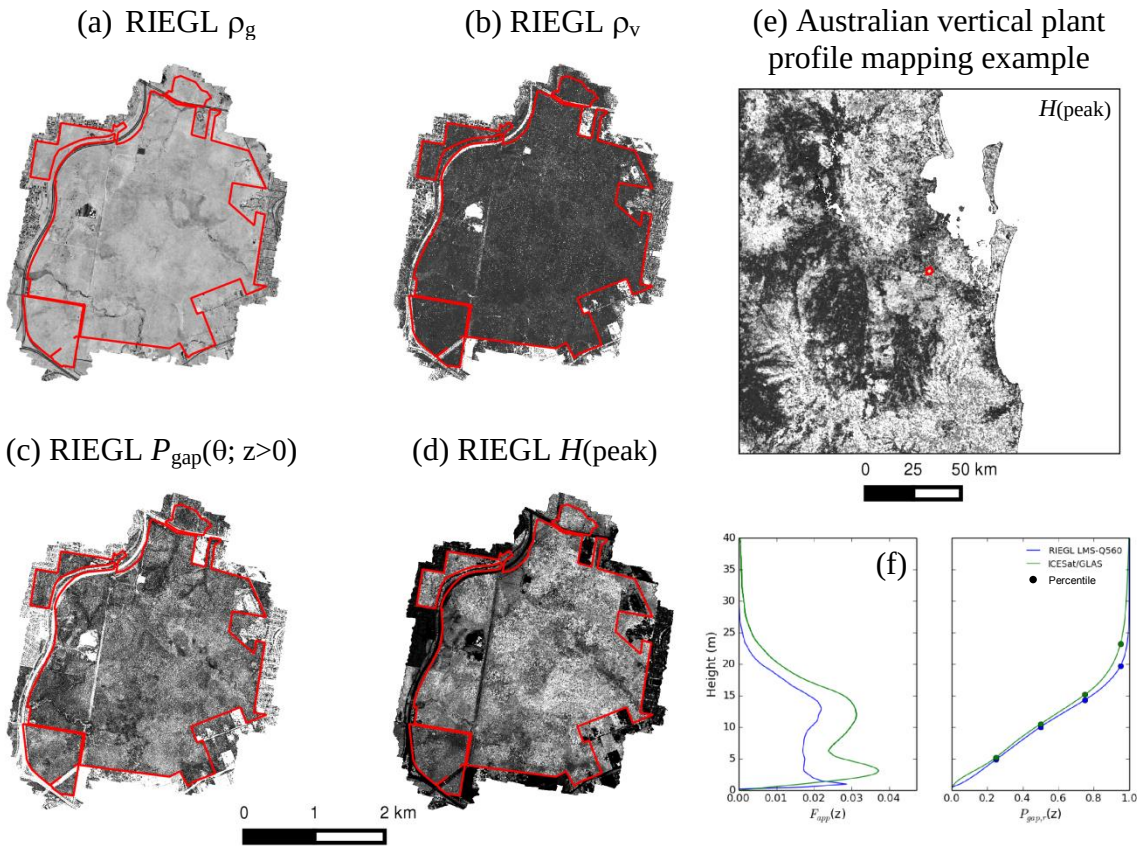


Figure 1: Example of airborne waveform lidar products derived for the Karawatha Forest Park TERN Auscover site in South East Queensland, Australia (a-d) for validation of K&C Australian vertical plant profile mapping. A regional example (e) of the height of peak foliage density from the ICESat/GLAS mapping, and an example (f) comparison to the airborne (RIEGL) mapping for a single image segment, are also shown. The extent of the Karawatha Forest Park site is shown in red.

Conclusion

Outcomes of this work include an approach to account for local spatial variance in ground reflectance in the direct estimation of $P_{gap}(\theta, z)$ using the simple radiative transfer model of [2], and application of this approach to large waveform lidar datasets generated from airborne laser scanning. This permits direct estimation of vertical $P_{gap}(\theta, z)$ profiles, derived apparent vertical plant profiles, and related parameters at finer spatial resolution and over larger areas than previous studies. These are being applied to validate the K&C vertical plant profile mapping for Australia and the results at all TERN Auscover sites will be presented at SilviLaser 2015. Ongoing research is evaluating existing airborne laser scanning archives held by Australian state and federal government agencies for directly calibrating vertical plant profile and AGB mapping for the Australian continent.

References

[1] Scarth, P., Armston, J., and Lucas, R. (2014). Continental scale forest and woodland structure mapping using Landsat, ALOS PALSAR and GLAS ICESat. ForestSAT 2014: A bridge between forest sciences, remote sensing and geospatial applications, Riva del Garda, Italy.

[2] Armston, J., Disney, M., Lewis, P., Scarth, P., Phinn, S., Lucas, R., Bunting, P., and Goodwin, N. (2013). Direct retrieval of canopy gap probability using airborne waveform lidar. Remote Sensing of Environment, 134: 24-38.

[3] Bunting, P., Armston, J., Clewley, D., and Lucas, R. (2013). Sorted pulse data (SPD) library - Part II: A processing framework for LiDAR data from pulsed laser systems in terrestrial environments. Computers & Geosciences, 56: 207-215.

[4] Garcia, D. (2010). Robust smoothing of gridded data in one and higher dimensions with missing values. Computational Statistics and Data Analysis, 54: 1167-1178.

Voxel based occlusion mapping and plant area index estimation from airborne laser scanning data

Daniel Kükenbrink, Reik Leiterer, Fabian D. Schneider, Michael E. Schaepman, Felix Morsdorf

Remote Sensing Laboratories, University of Zurich, Winterthurerstrasse 190, CH-8057, Zurich, Switzerland

Highlights: We introduce a ray-tracing based approach for mapping occluded areas in airborne laser scanning (ALS) data for a temperate mixed forest in Switzerland. Furthermore, the approach showed promising results towards a three-dimensional retrieval of plant area index (PAI) from ALS data.

Key words: *ALS, forestry, PAI, ray-tracing, voxel traversal*

Introduction

Accurate three-dimensional information of canopy structure contributes to better understanding of radiation fluxes within the canopy and the physiological processes associated with it. Small-footprint airborne laser scanning (ALS) data proved valuable for characterizing the three-dimensional structure of forest canopies [1]. Nevertheless, very few studies actually analysed what the ALS system was able to actively observe and which parts were occluded from the system due to dense vegetation or scanning patterns [2]. The information about occluded areas could be of importance and could serve as a quality layer for characterizing the three-dimensional structure of the forest canopy. We therefore introduce a ray-tracing based approach for characterizing the ALS observation pattern inside a three-dimensional voxel grid. The study was performed on the Laegeren test site (47º 28´ 43.0 N, 8º 21´ 53.2 E, Switzerland), a temperate mixed forest characterized by steep slopes and high species diversity. Both, leaf-on and leaf-off ALS acquisitions were available for analysing seasonal differences in occluded areas.

As a key canopy structural characteristic, plant area index (PAI) serves as an important input parameter for radiative transfer modelling. PAI is defined as half the total area of leaves and woody materials per unit ground area [3]. ALS proved valuable in deriving PAI using simple approaches based on the Beer-Lambert law of light extinction and the estimation of penetration rate of the laser pulses through the canopy (e.g. in [4]). By relying on physically based approaches, requirements for field calibration is minimized which is of special interest for large area sampling. In this study, we analysed the possibility of deriving voxel-based PAI estimation from ALS measurements using the proposed ray-tracing approach.

Materials and Methods

ALS data were acquired over the Laegeren test site under leaf-off (April 10th 2010) and leaf-on (August 1st 2010) conditions. Both campaigns were flown with a nominal height of 500 m above ground and a beam divergence of approximately 0.5 mrad resulting in a footprint size of 0.25 m. The wavelength of the laser was 1550 nm. A flight strip overlap of approximately 50% finally leads to a mean point density of 20 m^{-2} in the leaf-off and 40 m^{-2} in the leaf-on dataset. The positional accuracy of the ALS data was <0.15 m in vertical and <0.5 m in horizontal direction. Plot-level PAI estimates from digital hemispherical photographs (DHPs) were available for validation purposes.

For the estimation of the voxel-based PAI, the scene was first divided into a three-dimensional voxel grid with a predefined voxel dimension. Afterwards, each laser pulse was traced through the voxel grid using a simple voxel traversal algorithm first introduced in [5] and adapted for the use with terrestrial laser scanner (TLS) in [6]. The algorithm was adapted for the use with ALS, considering multiple returns per pulse. For each voxel, the number of pulses able to traverse the voxel with (N_{hit}) and without (N_{miss}) a laser return was recorded. Additionally, the number of pulses theoretically traversing the voxel that were obstructed from the surrounding canopy was registered as the number of occluded pulses (N_{occ}). By considering N_{hit}, N_{miss}, and N_{occ}, one is able to extract a voxel grid with classifications, providing the observable and occluded regions for that particular ALS survey. The classification scheme follows the approach described in [6] and is further described in Table 1.

Considering N_{hit} and N_{miss}, a penetration rate through the voxel can be estimated. Combined with Beer-Lambert law of light extinction, a PAI can be estimated adapting the approach described in [4] to a voxel based approach:

$$PAI = -\frac{1}{k} * \ln\left(\frac{N_{miss}}{N_{miss} + N_{hit}}\right) \quad (1)$$

Table 1: Classification of voxel cells (following [6])

	Number of		
	Returns (N_{hit})	Penetrations (N_{miss})	Occlusions (N_{occ})
Observed	>0	>=0	>=0
Empty	=0	>0	>=0
Hidden	=0	=0	>0
Unobserved	=0	=0	=0

where k is the extinction coefficient defined by:

$$k = \frac{G(\theta, \theta_L)}{\cos\theta} \tag{2}$$

where G is the projection of foliage area that depends on the leaf normal angle (θ_L) and the incidence angle of the laser pulse (θ) [7]. For the PAI estimation, N_{miss} and N_{hit} were further weighted according to the number of returns registered by the pulse, following the assumption, that a pulse with n returns can be divided into n equal parts (e.g. for a pulse with 7 returns, each return counts as $1/7^{th}$ N_{hit}). N_{miss} and N_{hit} were further weighted according to their path length through the considered voxel. For the case, where N_{miss} is zero, the above model saturates. Therefore, N_{miss} was replaced by the smallest possible value greater than zero (e.g. 1/7 as the maximum number of returns per laser pulse was 7).

Results and Discussion

In Figure 1 the voxel classification output of the proposed voxel traversal algorithm is shown for a 300 m transect with one voxel depth (voxel dimension = 1x1x1 m). Only voxels observed with a laser return (N_{hit}) and occluded voxels (N_{occ}) are shown. Occluded voxels are shown in yellow to red. Observed voxels with a registered laser return in them are shown in green to blue. The redder the occluded voxels are, the more pulses would have theoretically traversed the voxels, if they were not obstructed by the canopy. The bluer the voxels are, the more laser returns were registered by the ALS system inside the corresponding voxel. The results for both leaf-on and leaf-off conditions are shown. The occluded areas are much more dominant under leaf-on than under leaf-off conditions, as the laser pulses are much more obstructed by the denser canopy compared to the

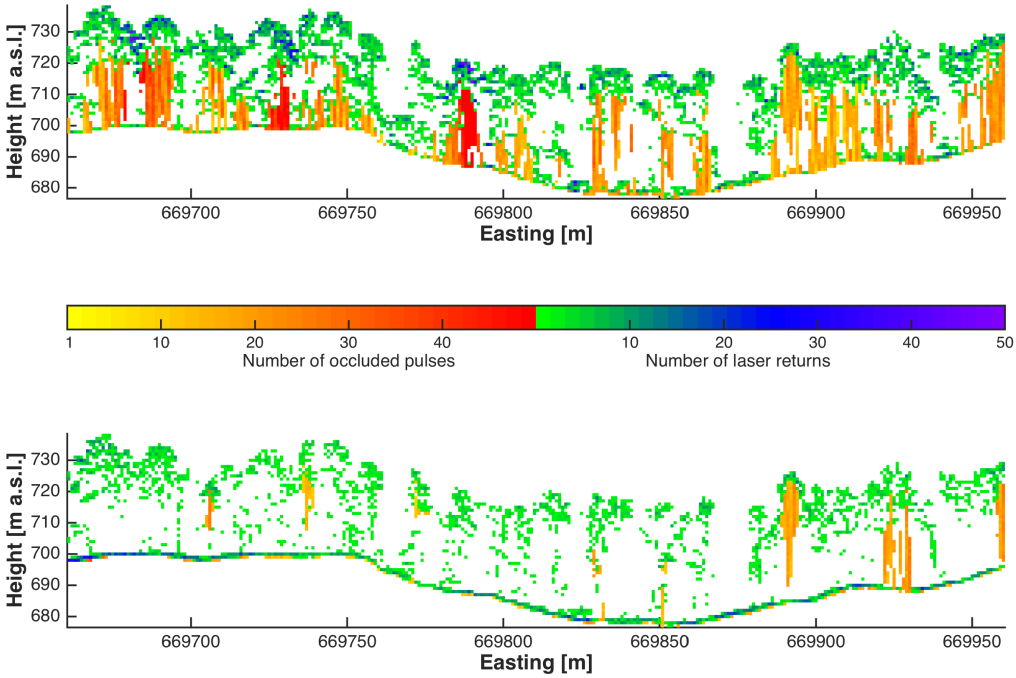


Figure 1: Voxel classification result for a 300 m transect with one voxel depth (1 m voxel dimension). Yellow to red denotes occluded voxels not visible for the ALS system and green to blue show the number of laser returns inside voxels, that were observed from the ALS. Top: Voxel classification for leaf-on condition. Bottom: Voxel classification for the same transect in leaf-off condition. Coordinates are in the Swiss national grid CH1903 / LV03

leaf-off case. The remaining occluded areas in the leaf-off case are mostly due to coniferous trees still obstructing laser pulses from penetrating into the canopy. In both conditions, scanning patterns (scan-angle, incidence angle, local incidence angle, pulse density) are believed to be a further cause for occlusion in the ALS dataset, especially in rugged terrain as found in our study site. Furthermore, the chosen voxel dimension affects the detected occluded area, as the possibility for a pulse traversing a voxel increases with increasing voxel size.

As the voxel classification results for leaf-on and leaf-off conditions are to some extent complementary, occluded areas in the leaf-on case could be filled using the information from the leaf-off condition. This is especially of importance when trying to characterize understory vegetation from ALS measurements [2].

The proposed PAI estimation approach showed promising results, with good correlation to the available plot-level PAI estimates from DHP measurements, when aggregating the vertically resolved PAI values to a two-dimensional map. Nevertheless, some voxels showed unrealistically high PAI values when a large amount of laser returns were registered for the corresponding voxel, but no – or just very few – laser pulses were able to penetrate the voxel. This is mostly the case for voxels which are totally filled e.g. by large branches or a tree trunk. Approaches to overcome these issues are currently investigated. For instance, the incorporation of the available full-waveform features (e.g. intensities, full width at half maximum FWHM) and the estimation and compensation of a clumping index considering the distribution of the laser returns inside the voxel could be used in this context.

The positional accuracy of the laser point-cloud is assumed to be a source of uncertainty in this ray-tracing approach, especially when considering small voxel sizes of less than 1 m. We therefore restricted the minimum voxel size to 1 m. Increasing voxel dimensions were found to increase the robustness of the PAI retrieval.

Conclusion

We proposed a simple voxel-traversal algorithm for mapping occluded areas from small-footprint ALS inside three-dimensional space and tested the algorithm for the use in retrieving vertically resolved PAI estimates. The proposed approach showed that the amount of occluded voxels is highly dependent on the season of ALS acquisition. Furthermore, laser acquisition patterns were found to have an influence on the amount of occluded areas inside the forest canopy. The proposed PAI estimation approach showed promising results when aggregated and compared to 2D estimation of PAI using hemispherical photographs. Yet, for very dense vegetation and voxels totally filled with wood materials, the proposed approach highly overestimates PAI, leading to unrealistically high values for the analysed study site. The next steps will include the analysis of the clumping inside each voxel and trying to compensate for the PAI overestimation. Also further analysis on the influence of pulse density, scan and incidence angle on the occluded areas and the penetration depth of the laser pulses will be conducted. We further plan to incorporate the retrieved PAI voxel grid in a radiative transfer model to analyse the complex three-dimensional radiative budget inside the forest canopy.

Acknowledgement

This study has been supported by the European Facility for Airborne Research in Environmental and Geo-science (EUFAR), grant agreement number 312609, Joint Research Activities (JRA1) HYLIGHT project.

References

[1] Schneider, F. D., Leiterer, R., Morsdorf, F., Gastellu-Etchegorry, J.-P., Lauret, N., Pfeifer, N., & Schaepman, M. E. (2014). Simulating imaging spectrometer data: 3D forest modeling based on LiDAR and in situ data. Remote Sensing of Environment, 152, 235-250.

[2] Korpela, I., Hovi, A., & Morsdorf, F. (2012). Understory trees in airbone LiDAR data – Selective mapping due to transmission losses and echo-triggering mechanisms. Remote Sensing of Environment, 119, 92-104.

[3] Chen, J. M., Black, T. A., & Adams, R. S. (1991). Evaluation of hemispherical photography for determining plant area index and geometry of a forest stand. Agricultural and Forest Meteorology, 56, 129-143.

[4] Solberg, S., Brunner, A., Hanssen, K. H., Lange, H., Næsset, E., Rautiainen, M., & Stenberg, P. (2009). Mapping LAI in a Norway spruce forest using airborne laser scanning. Remote Sensing of Environment, 113 (11), 2317-2327.

[5] Amanatides, J., & Woo, A. (1987). A fast voxel traversal algorithm for ray tracing. Eurographics, 87 (3), 1-6.

[6] Bienert, A., Queck, R., & Schmidt, A. (2010). Voxel Space Analysis of Terrestrial Laser Scans in Forests for Wind Field Modeling. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 38 (Part 5), 92-97.

[7] Ross, J. (1981). The radiation regime and architecture of plants stands. Dr. W. Junk Publishers, The Hague.

Detecting leaf water content using intensity data from terrestrial laser scanner

Xi Zhu, Andrew K. Skidmore, Roshanak Darvishzadeh, Tiejun Wang

Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands. E-mail: x.zhu@utwente.nl

Highlights: We investigated the possibility of estimating leaf water content (LWC) using radiometric corrected intensity data from terrestrial laser scanner (TLS). Firstly, a reflection model was used to remove specular intensity caused by leaf surface at perpendicular angle; and then reference targets were applied to correct the incidence angle effect. Results show the accuracy of the estimation of leaf water content was significantly improved after the radiometric correction (R^2=0.757).

Key words: *leaf water content, laser intensity, leaf surface, specular intensity, incidence angle*

1. Introduction

Vegetation water status plays an important functional role in forest fire susceptibility and propagation modelling as well as the plant physiological status assessment[1]. Passive remote sensing techniques provide an efficient and non-destructive way for estimating vegetation water content at both leaf and canopy levels, but it can only be used when the naturally occurring energy (i.e., sunlight) is available and could be negatively affected by solar illumination, atmospheric conditions, soil and understory vegetation reflectance[2]. The terrestrial laser scanner (TLS) intensity is insensitive to light and weather conditions to some extent and has the potential to retrieve leaf water content (LWC) underneath the forest canopy and its exact spatial distribution.

However, using TLS intensity to estimate LWC is complicated by leaf angle distribution, because the incidence angle between laser beam and the leaf normal strongly affects TLS intensity[3]. The incidence angle effect is dominated by target reflectance, while the variations in surface topography that are smaller than the laser footprint on the target do not have a significant influence on the incidence angle effect, suggesting that the incidence angle effect can be corrected using functions that depend on the object brightness[4]. In other words, radiometric correction of TLS should allow object brightness to be accurately normalized to reflectance.

Before correcting the angle effect, the relationship between leaf water content and intensity at an angle perpendicular to the TLS-object nadir is necessary. However, a challenge to attain such a relationship between TLS intensity at perpendicular angle and LWC is to remove the specular intensity component, as this part of intensity is caused by the leaf surface. Though the contribution of specular intensity from the leaf surface may have little effect on total intensity from different incidence angle, it has a noticeable effect at perpendicular angle[5]. At perpendicular angle, there is a considerable fraction of specular component which varies due to the difference of surface features. Thus, for LWC estimation, the intensity at perpendicular angle may be affected by leaf surface which needs to be corrected.

The overall objective of this study is to assess the potential of TLS intensity data for estimating LWC. Specifically we aim to 1) examine the correlation between LWC and TLS intensity at perpendicular angle; 2) investigate the surface feature influence on intensity and assess the application of reflection models on leaf intensity for building the correlation between LWC and TLS intensity at perpendicular angle; 3) investigate the leaf angle effect on intensity variation and evaluate the prediction accuracy of correction methods involving brightness feature.

2. Materials and methods

2.1 Materials and measurements

Four Spectralon panels manufactured by Labsphere, Inc. with nominal reflectance values of 12%, 25%, 50%, 99% were observed at incidence angle between 0° to 80°, with steps of 10 degrees. Leaves were detached from eight plant species and oven dried for modelling. In total, we obtained 463 samples with different levels of water content. They were flattened and stuck to a board for intensity measurement at a perpendicular angle to obtain the correlation between water content and intensity. Four other plant species were used for validation.

The LWC was calculated using the following formula[6];

$$\frac{MW - MD}{A}\left(g / cm^2\right)$$

where MW is the mass of the wet leaf and MD is the mass of completely dried leaf, A is the leaf area.

The laser scanner used in this paper was the time-of-flight scanner RIEGL VZ-400 that employs a SWIR (1550nm) laser which is sensitive to water content [7]. The system has a beam divergence of 0.35 mrad, a range accuracy of 5 mm, an effective measurement rate of 122,000 meas./sec and a maximum range of 160m at 20% reflectance in a high speed mode.

2.2 Reflection model for TLS intensity simulation

A Linear combination of the Lambertian model and Beckmann law was used here to simulate the reflection [8]. It provides a comprehensive theory that can be applied to a wide range of surface conditions ranging from smooth to very rough. The Lambertian reflection model defines the diffuse reflection for the dull, matte surfaces, while Beckmann law models specular reflection properties.

$$I = f\left(k_d \cos\alpha + \frac{(1-k_d)}{\cos^5\alpha} e^{-\frac{\tan^2\alpha}{m^2}}\right)$$

where I is the intensity, f represents intensity at normal incidence angle, k_d is the fraction of diffuse reflection, α is the incidence angle, m describes the surface roughness, m typically takes values between 0 (smooth surface) and 0.6 (rough surface).

2.3 Reference targets based model for incidence angle correction

The intensity was normalized here to normal angle of 0°. In order to achieve this, observed intensity at different incidence angle for the four panels with different intensity at perpendicular angle (I(0°)) are stored in a lookup table (LUT). The LUT is interpolated to provide I(0°) for any specific incidence angle and observed intensity, so I(0°) can be linearly interpolated from the model for a given intensity recorded by the instrument and incidence angle.

3. Results

3.1 Correlation between LWC and TLS intensity at perpendicular angle before removal of specular intensity

TLS intensity is correlated with LWC for 7 plant species individually, while for *Zamioculcas zamiifolia*, there is no significant correlation which could be attributed to saturation and its surface structure (Figure 1: left). Combining 8 plant species together, TLS intensity has no correlation with LWC before removal of specular intensity. Even when they have the same LWC, the intensity is quite different, which could be caused by different leaf surface texture. Smooth and shiny leaf surfaces would exhibit higher TLS intensity than rough and matte leaf surfaces, as they are more affected by specular intensity at perpendicular angle at the same LWC level.

3.2 Effect of removal of specular intensity on the correlation between LWC and intensity at perpendicular angle

TLS intensity has been simulated to exclude the specular part in order to attain a more general model for LWC prediction using the reflection model. *Zamioculcas zamiifolia* shows the lowest diffuse fraction (a = 0.3358), which partially causes the poor correlation between its LWC and TLS intensity at perpendicular angle (Figure 1: left).

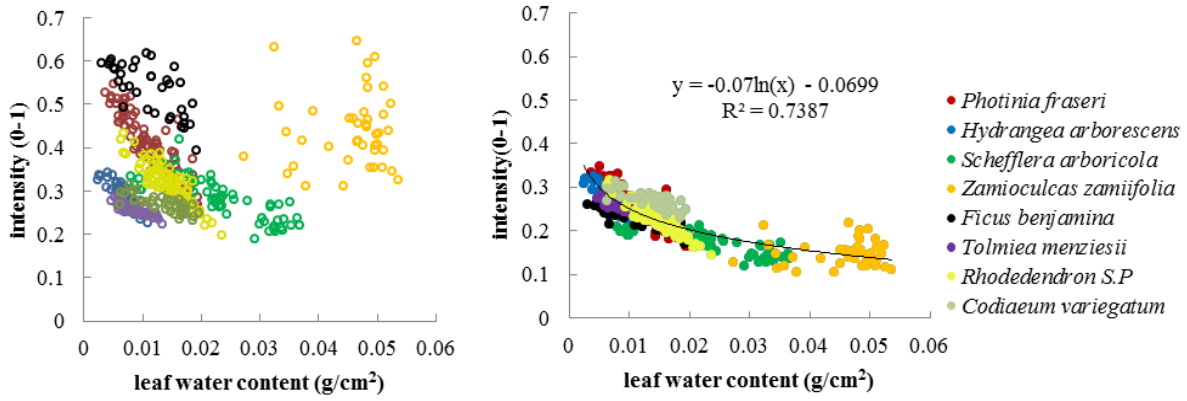


Figure 1: Correlation between LWC and TLS intensity at perpendicular angle (left: before removal of specular intensity, right: after removal of specular intensity)

TLS intensity at perpendicular angle was corrected by subtracting the specular component. As expected, species that have higher intensity at the same LWC level such as *Zamioculcas zamiifolia* have higher specular intensity, after correction the intensity drops to the same level as other species, while the intensity of species which are less shiny did not drop much after correction due to lower specular fraction. Therefore, after removal of specular intensity, when the samples from the 8 plant species were pooled together, more accurate R^2 value of 0.7387 was obtained (Figure 1). TLS intensity decreases as LWC increases, while it becomes less sensitive when LWC is high as the trend line is becoming flat (logarithmic relationship).

3.3 Correction of incidence angle effect on TLS intensity for LWC prediction

The prediction of LWC was done on 4 other plant species based the relationship obtained in 3.2 (Figure 1). The result of LWC estimation was overestimated by using uncorrected TLS intensity with an R^2 value of 0.3903. After correction of angle effect using reference targets, the accuracy was significantly improved (Table 1).

Table 1: LWC prediction using TLS intensity before and after correction of leaf angle effect

LWC prediction	R^2	RMES
Before correction	0.390	0.0080
After correction	0.757	0.0031

4. Discussion

The results of this study show that the radiometrically corrected TLS intensity data can be used to estimate LWC with acceptable accuracy (R^2 0.757; RMES 0.0031).

Before removal of the specular component, there is no significant relationship for 8 species pooled together. After correcting for the specular effect, a significant correlation emerged for the combined species. This is mainly due to the difference of surface feature. Some species are more shiny so that at perpendicular angle they exhibit stronger specular intensity which is not related to LWC. Another possible reason is that the leaf internal structure is also affecting TLS intensity which also needs to be adjusted. However, this is beyond the scope of this study.

The correlation between LWC and TLS intensity at perpendicular angle follows a logarithmic function, indicating that when the LWC is high, TLS intensity is less sensitive to LWC, suggesting there is a saturation problem for LWC estimation using TLS intensity.

Correction of incidence angle effect using reference targets improved the accuracy of LWC estimation, indicating the correction is effective. In this study, brightness is involved into the angle effect correction, since it is the dominant factor. Surface roughness also affects intensity decay with incidence angle, waveform data might be a good choice to get surface roughness information.

References

[1] Scriber, J. M. (1977). Limiting effects of low leaf-water content on the nitrogen utilization, energy budget, and larval growth ofHyalophora cecropia (Lepidoptera: Saturniidae). Oecologia, 28, 269-287.

[2] Wessman, C. (1994). Estimating Canopy Biochemistry through Imaging Spectrometry. In Imaging Spectrometry — a Tool for Environmental Observations (pp. 57-69). Netherlands: Springer Netherlands.

[3] Eitel, J. U., Vierling, L. A., & Long, D. S. (2010). Simultaneous measurements of plant structure and chlorophyll content in broadleaf saplings with a terrestrial laser scanner. Remote sensing of environment, 114, 2229-2237.

[4] Krooks, A., Kaasalainen, S., Hakala, T., & Nevalainen, O. (2013). Correction of intensity incidence angle effect in terrestrial laser scanning. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, 1, 145-150.

[5] Sinclair, T., Schreiber, M., & Hoffer, R. (1973). Diffuse reflectance hypothesis for the pathway of solar radiation through leaves. Agronomy Journal, 65, 276-283.

[6] Danson, F. M., Steven, M. D., Malthus, T. J., & Clark, J. A. (1992). High-spectral resolution data for determining leaf water content. International Journal of Remote Sensing, 13, 461-470.

[7] Tucker, C. J. (1980). Remote sensing of leaf water content in the near infrared. Remote Sensing of Environment, 10, 23-32.

[8] Beckmann, P., & Spizzichino, A. (1987). The scattering of electromagnetic waves from rough surfaces. New York: Artech Print on Demand.

Using LiDAR area-based approach and spatial optimization to delineate harvest areas

Adrián Pascual[1], Timo Pukkala[1], Francisco Rodríguez[2, 3], Sergio de Miguel[1, 4]

[1] University of Eastern Finland, School of Forest Sciences, PO Box 111, 80101 Joensuu, Finland

[2] föra forest technologies. C/ Oreste Camarca, 4. 42004 Soria, SPAIN.

[3] Dpt. Producción Vegetal y Recursos Forestales. Universidad de Valladolid. SPAIN

[4] Departament de Producció Vegetal i Ciència Forestal, Universitat de Lleida-Agrotecnio Center (UdL-Agrotecnio), Av. Rovira Roure, 191, E-25198 Lleida, Spain

Highlights: The study showed how LiDAR data can be used in forest planning to solve spatial problems such as harvest block formation. These blocks were generated using LiDAR to achieve detailed spatial resolution for estimating forest characteristics and generate small management units from which homogeneous harvest blocks were formed.

Key words: *forest planning, spatial optimization, ALS, precision forestry.*

Introduction

One change of paradigm in forest planning is related on how treatments are applied to forest areas. In this regard, the size of forest management units may be reduced and, in consequence, forests may be considered as the sum of homogeneous small-sized forest units (e.g., pixels).

Using LiDAR exponentially increases the amount of information for improving the accuracy of forest inventories. Even with low-density pulse, LiDAR-based methods succeed in estimating forest characteristics [1]. This enhancement in the predictability of forest attributes is the starting point to form management units as the aggregation of small homogeneous units [2].

The combination of GIS, computational analysis and spatial optimization enlarges forest management planning applications. The worldwide use of LiDAR in forest inventory methods broadens the optimal resolution of forest planning problems that copes with xyz coordinates such as define homogeneous harvest blocks within a forest area. Simulated Annealing (SA), among the heuristics optimization methods, has been previously used to optimise the aggregation of cutting areas in combination with other forest planning objectives [3].

The aim of this study was to create harvest blocks using LiDAR to define management units and discuss the use of alternative implementations of the SA algorithm for this purpose.

Materials and methods

The study area is located in in the Iberian System Mountains of Soria. Scots pine (*Pinus sylvestris*) and maritime pine (*Pinus pinaster*) are the main species. Field plots and LiDAR data were collected at the same time aiming to predict forest stand attributes by using LiDAR metrics at plot level as predictors. This method is known as the area-based approach [4]. A set of models was used to estimate forest characteristics along a 30-year planning horizon.

Simulations based on the models developed were used to generate treatment alternatives for a three-10-year-period management plan for which multiple *n* objectives such as a periodical cutting target or maximising the aggregation of cutting areas were defined. The optimization was conducted by maximizing a utility function using SA algorithm with both 1-neighbor and 2-neighbor alternatives. The difference is related on how the algorithm starts checking solutions (i.e., 2-movement alternative checks if one solution could improve the utility function value by making two simultaneous moves and substitutes the previously accepted solutions).

Results and discussion

The spatial optimization results indicated that the utility function value increases as freezing temperature decreases for both SA alternatives. Although both SA alternatives provided similar results, their behaviour differs considerably. While the 1-neighbour implementation rapidly increased the utility until a value relatively close to the maximum, the 2-neighbor alternative improved the utility more slowly but the value was higher than the first alternative.

Light, normal and heavy thinnings were aggregated to display visually the results aiming to create large enough harvest areas by taking into account practical and operational restrictions. Thinning areas are presented over the growing stock volume layer generated from the models based on LiDAR-metrics (Figure 1). These harvest blocks were highly influenced by the change of forest cover within the forest which can led to obtain non-aggregated results, as shown in Figure 1. To avoid this problem, more weight was given to minimise the total length of cut and non-cut calculation unit borders.

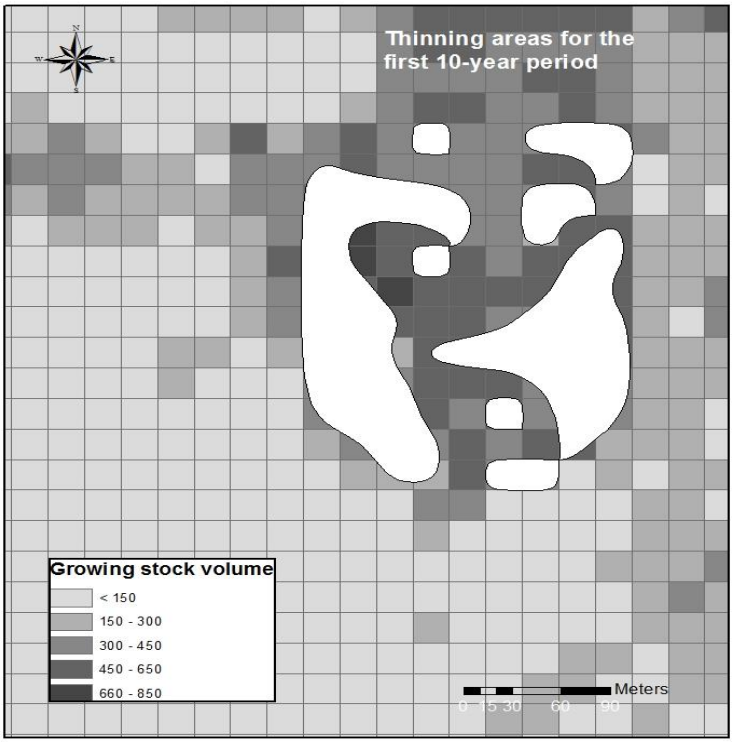


Figure 1: Thinning areas for the first period

The presented methodology can boost the use of LiDAR in current forest planning methods since LiDAR-based inventory is increasingly used in forest inventory nowadays. This research work showed how the combined use of LiDAR and growth and yield models can assist precision forestry by providing high spatial accuracy from the xyz point cloud information to be used in the delineation of treatment areas within a forest.

References

[1] Jensen, J. L. R., Humes, K. S., Conner, T., Williams, C. J., & Degroot, J. (2006). Estimation of biophysical characteristics for highly variable mixed-conifer stands using small-footprint lidar, 1138, 1129–1138.

[2] Heinonen, T., Pukkala, T., Vauhkonen, J., & Maltamo, M. (2011). Dynamic Treatment Units in Eucalyptus Plantation. Forest Science 57(5): 416–426.

[3] Pukkala, T., & Kurttila, M. (2005). Examining the performance of six heuristic optimisation techniques in different forest planning problems. Silva Fennica, 39(1), 67–80.

[4] 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.

Forest aboveground biomass estimation based on LiDAR and scaling model in China

Yuan Zeng, Dan Zhao, Yujin Zhao, Bingfang Wu

Key Lab of Digital Earth Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing, China

Highlights: We select 5 typical study sites for collecting airborne LiDAR data and estimate the forest aboveground biomass, then scale to the 30m scale by Landsat TM data and 250m scale based on MODIS, canopy height, vegetation cover and forest classification data. The output is China's forest aboveground biomass map 2010.

Key words: *forest aboveground biomass, airborne LiDAR, scaling model, canopy height*

The objective of this study is to estimate the China's forest aboveground biomass (AGB) using airborne LiDAR and scaling models. Based on the type and location of forest in China, total 8 districts were separated for building the AGB estimated models, and 5 typical forest study sites (100x100km) were selected for collecting the airborne LiDAR data (100-200km^2) , also 26 general forest study sites (50x50km) have detailed field measurements, see Figure 1.

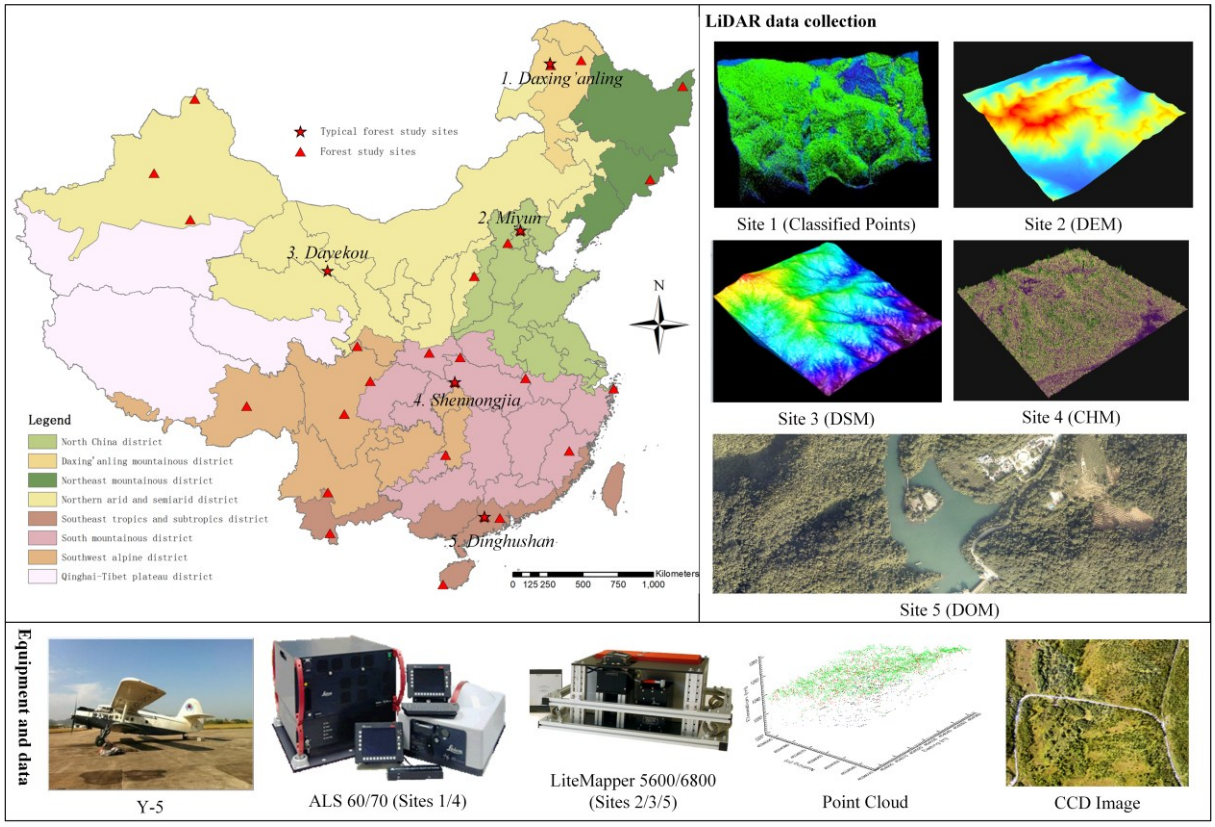


Figure 1: 8 districts and 5 typical forest study sites with airborne LiDAR data collection in China

The forest AGB was estimated firstly in the flight area. Tree height and density were extracted from LiDAR point clouds. A stepwise regression model was built combining field-measured AGB, canopy height percentile and crown density. Then, for each typical study site, according to the forest classification and AGB estimation in the LiDAR flight area, a scaling model was used for retrieving the AGB at 30m scale by Landsat TM data. For estimating AGB at the district scale, a pixel based canopy height data are needed, which could be calculated based on the extracted height from the space-borne GLAS footprints and MODIS BRDF data. Finally, a scaling model combining 30m AGB, 250m MODIS time series data, continuous canopy height map, fraction of vegetation cover and forest classification data was built up for mapping the AGB in each district. Figure 2 shows the output, 2010 forest AGB map of China.

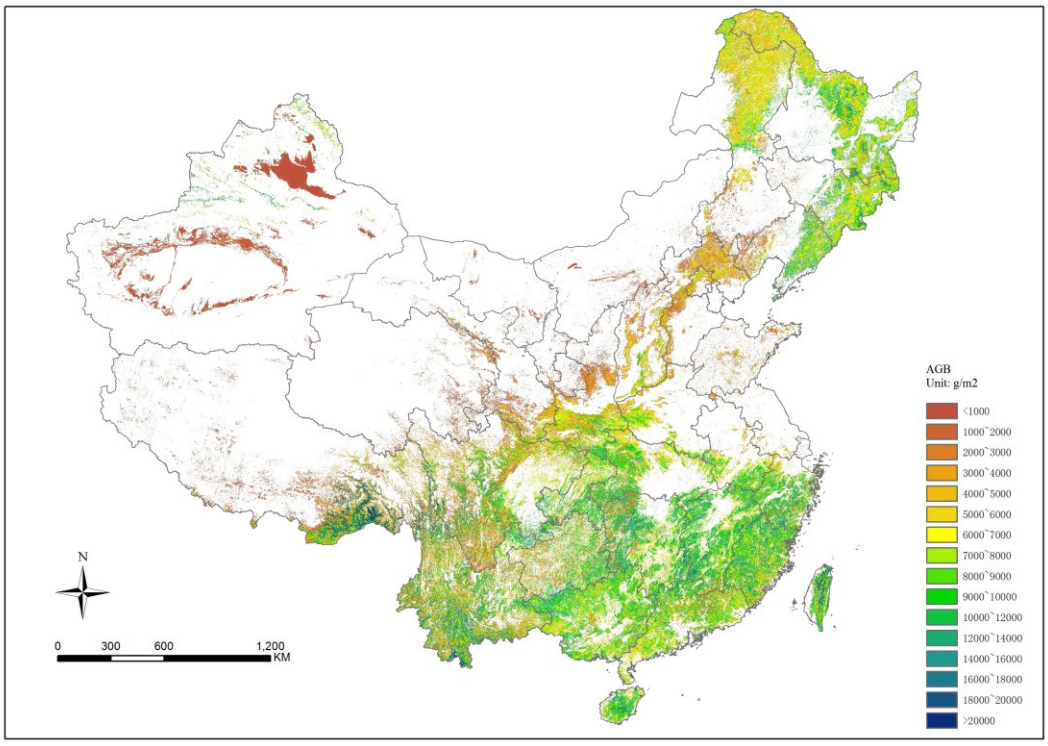


Figure 2: China's forest aboveground biomass map 2010 with 250m spatial resolution

We use the independent field data to do the validation in each district. Currently, the total amount of 2010 China's forest AGB is 14.7 billion ton based on our study. This result still needs further systematic analysis and calibration.

Extending the scale of applicability of models predicting forest structure and composition from airborne LiDAR measurements

Jean-Romain Roussel[1], John Caspersen[2], Martin Béland[1], Alexis Achim[1]

1. *Faculté de Foresterie, Géographie et de Géomatique, Université Laval, Québec, 2405 rue de la terrasse, Québec, Qc, Canada G1V 0A6*
2. *Faculty of Forestry, University of Toronto, 33 Willcocks Street, Toronto, ON, Canada M5S 3B3*

Highlights: Several issues restrain the possibilities to make large scale predictions of forest structure and composition from LiDAR measurements. In an attempt to produce a generic model applicable at a large scale this study will focus on the general effects of LiDAR settings and of cluster size on model predictions.

Key words: *LiDAR, modeling, large scale application, LiDAR settings, forest structure, forestry management, predictive model.*

Introduction

In line with current needs to obtain precise information about the structure, composition and health of the forest resource, LiDAR technology is increasingly being used in forest surveys. Although it can generate measurements impossible to obtain by ground technicians, both by their quantity and their technical nature, the use of LiDAR requires the development of robust mathematical models or algorithms to interpret the data. Current models can make accurate predictions in well-defined contexts, but several issues restrain the possibilities to make large scale predictions. Reasons for this are that: (i) calibrations are generally based on ground inventories made over a limited area, (ii) models are highly dependent on the LiDAR settings (i.e. flight altitude, pulse frequency, point density, emitted energy, sensors capabilities, etc.), which often vary between datasets and (iii) models are calibrated for various plot sizes, and the forest characteristics of interest (e.g. structure and composition) can be expected to vary according to the scale of observation. In this context, this project will focus on developing new approaches towards building predictive models that will ultimately become independent of LiDAR settings and remain valid for different scales of observation.

Content

Changes in the behaviour of the LiDAR signal when settings vary were partially described by different authors (e.g. [1]–[6]). These studies often focus on the vertical distribution of returns or on the effect on the digital elevation model (DEM). There is existing knowledge on the effect of a decrease of the density of points on the accuracy of the DEM, but there is little information on the effect of flight altitude, point density, pulse frequency, etc. on other derived variables such as the horizontal distribution of returns or the return intensities. Based on detailed inventories made across Canada, we will describe how and why derived variables change with point density, flight altitude and ground inventories. Our intent is to identify some generalities valid at a large scale, which could be used to develop an algorithm applicable to different LiDAR datasets acquired with different settings. Through this we hope to make predictive models of forest structure and composition applicable at the scale of an entire forest biome.

Another challenge is to collate datasets that used various plot sizes in ground inventories because variables such as structure or composition cannot be assumed to be independent of the scale of observation. This also has implications on the choice of the rasterization size to describe the forest. Based on an extensive spatially-explicit ground inventory, we will describe (i) the effect of the rasterization on forest structure and (ii) the effect of rasterization on predictive models. Preliminary results show that forest structure (i.e. composition, basal area and tree size distribution) is unstable for small areas around 400 m^2, a common area used for space rasterization in forestry (e.g. [2], [7]–[11]) . This means that for a cluster of 400 m^2, a small difference in size (+25%) or a small difference in positioning (moving the cluster by 25% of its dimension) the assessment of forest structure can vary widely. On the contrary, the assessment of forest structure in bigger observational units of 1000 or 2000 m^2 is less sensitive to such variation. This preliminary result suggests that predictions could be highly dependent of small differences in the raster grid used for mapping the predictions.

Previous studies have also attempted to develop generic models that remain applicable outside of their range of calibration. A new approach developed by [12] appears to be compatible with this idea. [12] presents a predictive model that does not predict the forest characteristics from LiDAR data, but instead attempts to predict LiDAR data from forest characteristics. The idea of the current study is to reverse this model. If it is possible to make a mechanistic prediction of what would be the LiDAR signal for a given forest, it should also be possible to use this mechanistic model to predict the forest characteristics from existing LiDAR data. This approach seems promising for producing a general model valid over a wide scale. However, if it is possible to predict the LiDAR data from a description of the forest, it remains that two different forests might generate similar LiDAR signals. So, from a given dataset, it might not be possible to identify the 'original' forest. Mathematically, this means that the model could be a surjective function. Our work will be focused on this question. Using numerical simulations, we will quantify if and how this surjectivity problem may prevent the reversing of the model. If the model appears to be reversible, the work will focus this question. In the other case the work will be focused on the improvement of the model presented by [12] until this 'one-to-many' problem is avoided.

References

[1] X. Yu, J. Hyyppä, H. Hyyppä, and M. Maltamo, "Effects of flight altitude on tree height estimation using airborne laser scanning," in *International Archives Of Photogrammetry Remote Sensing And Spatial Information Sciences*, 2004, vol. XXXVI, pp. 96–101.

[2] E. Næsset, "Effects of different flying altitudes on biophysical stand properties estimated from canopy height and density measured with a small-footprint airborne scanning laser," *Remote Sens. Environ.*, vol. 91, pp. 243–255, 2004.

[3] A. Kukko, S. Kaasalainen, and P. Litkey, "Effect of incidence angle on laser scanner intensity and surface data.," *Appl. Opt.*, vol. 47, pp. 986–992, 2008.

[4] E. Næsset, "Assessing sensor effects and effects of leaf-off and leaf-on canopy conditions on biophysical stand properties derived from small-footprint airborne laser data," *Remote Sens. Environ.*, vol. 98, pp. 356–370, 2005.

[5] E. S. Anderson, J. A. Thompson, D. A. Crouse, and R. E. Austin, "Horizontal resolution and data density effects on remotely sensed LIDAR-based DEM," *Geoderma*, vol. 132, no. 3–4, pp. 406–415, Jun. 2006.

[6] K. K. Singh, G. Chen, J. B. McCarter, and R. K. Meentemeyer, "Effects of LiDAR point density and landscape context on estimates of urban forest biomass," *ISPRS J. Photogramm. Remote Sens.*, vol. 101, pp. 310–322, 2015.

[7] K. Lim, C. Hopkinson, and P. Treitz, "Examining the effects of sampling point densities," *For. Chron.*, vol. 84, no. 6, pp. 876–885, 2008.

[8] E. B. Racine, N. C. Coops, B. St-Onge, and J. Bégin, "Estimating Forest Stand Age from LiDAR-Derived Predictors and Nearest Neighbor Imputation," *For. Sci.*, vol. 60, no. 1, pp. 128–136, Feb. 2014.

[9] J. Boudreau, R. F. Nelson, H. A. Margolis, A. Beaudoin, L. Guindon, and D. S. Kimes, "Regional aboveground forest biomass using airborne and spaceborne LiDAR in Québec," *Remote Sens. Environ.*, vol. 112, no. 10, pp. 3876–3890, Oct. 2008.

[10] 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. For. Res.*, vol. 38, no. 5, pp. 1095–1109, May 2008.

[11] V. Thomas, P. Treitz, J. H. McCaughey, and I. Morrison, "Mapping stand-level forest biophysical variables for a mixedwood boreal forest using lidar: an examination of scanning density," *Can. J. For. Res.*, vol. 36, no. 1, pp. 34–47, Jan. 2006.

[12] R. A. Spriggs, M. C. Vanderwel, T. A. Jones, J. P. Caspersen, and A. David, "Towards inferring stand structure from remote sensing data: a simple canopy model to predict lidar return distributions," pp. 1–47 in press.

Automatic Tree Detection from Laser Scanning Point Clouds

Beril Sirmacek

Department of Geoscience and Remote Sensing, Delft University of Technology, Stevinweg 1, 2628CN Delft The Netherlands
B.Sirmacek@tudelft.nl

Highlights: A novel method is introduced to detect trees automatically from laser scanning point clouds. The method consists of these steps; local ground level identification, point cloud classification, individual tree detection. The algorithm performed well on mobile laser scanning (MLS), airborne laser scanning (ALS) and terrestrial laser scanning (TLS) point clouds.

Key words: *Laser Scanning, Point Clouds, Tree Detection, Classification*

Introduction

Development and increasing availability of laser scanning technology has been providing new opportunities to researchers to develop new computer vision and 3D modelling applications in order to reconstruct objects of interest in computer environment. In the outdoor environment, trees are one of the most important objects of interest to monitor, model and measure. Trees do vital job for well-being of all living species. They provide beauty, food, protected space for animals, but which is more important, they absorb carbon dioxide and provide oxygen which contributes to environmental circulations. Therefore, for scientists, for municipalities and other government agencies, it is important to track numbers of trees and to measure their physical attributes for scientific analysis such as carbon dioxide absorbing volume. Thanks to laser scanning technology which provides opportunity to do all these physical measurements in the computer environment without disturbing the nature and also by spending less man effort. In the last decade, scientists have offered algorithms to detect trees from laser scanning point clouds and to generate 3D models of them in the computer environment. McDaniel et al. [1] have developed an automatic individual tree detection method for terrestrial laser scanning point clouds. The proposed approach is interesting since its methodology has similarities with the method which is introduced in this paper. First, the method segments the ground plane using the input point cloud. After fitting a ground surface model on the lower points, tree trunks are detected. Finally, k-means classification method is used for classifying rest of the points to assign them to the closest tree trunk. The researchers have tested their approach on point clouds of different environments. They have observed that the algorithm performed best, when the ground is more or less flat without steep slopes. Unfortunately, dealing with pole, street light, sign like objects have not addressed in their article. Lalonde et al. [2] have extracted 3D features from point clouds to classify the points as trees and non-trees. Afterwards, they have applied a segmentation algorithm to identify individuals. Finally, they have used 3D primitive geometric shapes such as cylinders in order to generate 3D tree models. They have used the models to estimate the breast height diameter on terrestrial and airborne point clouds. Gorte and Pfeifer [3] have generated a range image from point clouds in order to segment trees. After segmentation of individual trees, they mapped the points into a voxel space, where it is segmented into branches using mathematical morphology. The voxel size selection has played a very crucial role in well performance of the algorithm. Raumonen et al. [4] proposed a method for automatically extracting approximate tree branch measures from point clouds. The method worked on the assumption that a tree can be locally approximated with cylinders without using voxel spaces. These examples demonstrate that most of the previously introduced methods depend on voxel space creation or fitting pre-defined geometrical models to the point clouds. The voxel space creation based methods are generally very sensitive to the voxel size selection. Furthermore, the methods which are using geometrical models also need specific parameter selection and they have a risk of missing trees which have different trunk shapes than the pre-defined models. Besides their complexity and high sensitivity to the parameter selection, these earlier approaches generally need relatively higher computation time. In addition to that, most of the time they are tested

in regions which include only trees as objects. This literature survey indicates that there is still a need for fast and reliable tree detection algorithm development for processing large data sets which include trees but also many other types of objects.

New mathematical approaches are developed to achieve high quality tree detection results with rather less implementation difficulties. This paper introduces a method to detect individual trees from laser scanning point clouds. As the first step, points of the input point cloud are classified into two classes; as “tree” and “non-tree” points. This classification result is used to generate a 2D voting matrix which is an imaginary gridded surface that is positioned approximately on the ground layer. The vote value of each grid is calculated by calculating the point density in the 3D space which is bounded by that grid. The high value grids generally correspond to the tree trunks which have very high point densities. Therefore, detecting local maxima of the voting matrix help us to detect tree trunks. Afterwards, a classification method is applied to assign the rest of the points to tree trunks. The classification method simply classifies the lowest points to the ground by assigning a label “0”. If a point is in a certain proximity of a tree trunk then a tree ID number is assigned to the point, otherwise again the label “0” is assigned which indicates that the point is not belong to any tree. In this way, the resulting point cloud is generated by holding the x, y, z coordinates of the points and a label in for each point in an extra column. When false coloring is applied to this point cloud, each one of the detected individuals can be visualized in a different color which makes the visualization more informative. In Figure 1, we provide an example output to show how the resulting point clouds look like. We test our method on TUDelft campus point clouds which are acquired by different sensors as; mobile, airborne and terrestrial laser scanners. We also provide discussion on parameter selection, down-sampling methods and their effects on the results, dependency on the sensor and scale, and big data processing possibilities. False detection of the objects which are very close to the tree trunks remains as a challenge which needs to be improved in future studies. Nevertheless, the experimental results indicate robustness of the proposed algorithm and its possible usage when fast and big data process is necessary.

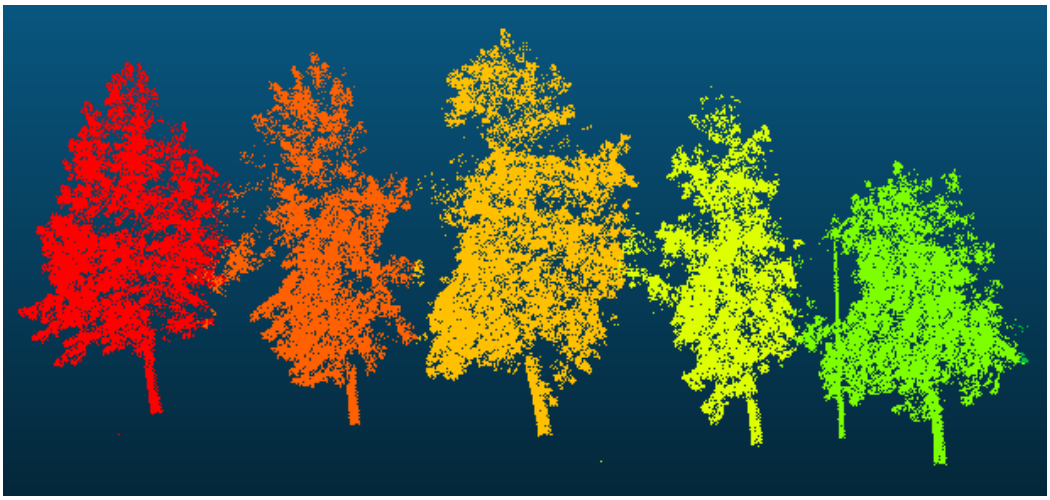


Figure 1: A tree detection example (detected individuals are indicated with a different color codes)

ACKNOWLEDGEMENTS

This research is funded by the FP7 project IQmulus (FP7-ICT-2011-318787) a high volume fusion and analysis platform for geospatial point clouds, coverages and volumetric data set. We also give thanks to FUGRO company in the Netherlands for providing us mobile laser scanning point clouds acquired from a part of the TUDelft campus area.

References

[1] McDaniel, M.W., Nishihata, T., Brooks, C.A., Salesses, P., & Iagnemma, K. (2012). Terrain classification and identification of tree stems using ground-based LiDAR. Journal of Field Robotics, Vol/ 29 (6), pp. 891-910.

[2] Lalonde, J., Vandapel, N., & Hebert, M. (2006). Automatic three-dimensional point cloud processing for forest inventory. Technical Report CMU-RI-TR-06-21, Robotics Institute, Carnegie Mellon University.

[3] Gorte, B.G.H., & Pfeifer, N. (2004). Structuring laser scanned trees using 3D mathematical morphology. ISPRS International Archieves of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 35, pp. 929-933.
[4] Raumonen, P., Kaasalainen, S., Kaasalainen, M., & Kaartinen, H. (2011). Approximation of volume and branch size distrubition of trees from laser scanner data. ISPRS International Achieves of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 3812, pp. 79-84.

Development of an algorithm to generate pit-free Digital Surface Models from LiDAR

Anahita Khosravipour[1], Andrew K. Skidmore[1], Martin Isenburg[2] and Tiejun Wang[1]

[1]*Department of Natural Resources, Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, P.O. Box 217, 7500 AE, Enschede, The Netherlands (a.khosravipour@utwente.nl, a.k.skidmore@utwente.nl, t.wang@utwente.nl).*
[2]*rapidlasso GmbH, Casparigasse 16, 97286 Sommerhausen, Germany (martin@rapidlasso.com).*

Highlights: We present a novel algorithm for generating pit-free Digital Surface Models directly from raw point clouds. It can generate both raster and TIN outputs. It considers all points (not only first returns) and avoids pit creation by "freezing" higher-up triangles as it constructs a pit-free TIN using incremental constrained Delaunay.

Key words: *LiDAR, Point cloud, Forestry, Digital Surface Model, TIN, Raster, Data pits.*

Introduction

Single tree-level forest information is required for a variety of applications (e.g., ecological, hydrological, and meteorological) as well as forest management activities (e.g., species identification, gap analysis, pest management, volume and carbon estimation) at the local and regional scales [1]. In recent years, collecting 3D point data with small-footprint airborne LiDAR (Light Detection And Ranging) systems has become a common remote sensing technique for extracting individual tree parameters, because it allows to capture the three-dimensional structure of tree elements such as branches and foliage with great detail and positional accuracy [2]. From dense discrete-return LiDAR points, we can generate both a Digital Terrain Model (DTM) that corresponds to the ground elevation, as well as a Digital Surface Model (DSM) that captures the elevation of the vegetation and of men-made objects [2]. In combination, they allow extracting the height, the shape, and the area of above-ground features that make it possible to derive, for example, tree canopy characteristics.

The standard way to generate a LiDAR derived-DSM for applications in vegetation analysis is to construct a Delaunay TIN using *only* the first returns that usually correspond to the echoes reflected from the uppermost vegetation layer and then rasterizing the TIN onto a grid with user-defined pixel size [3]. However, a number of problems have been identified that limit DSMs in their capacity to accurately retrieve individual tree crown attributes by using first returns. The problems of using all first returns for constructing the Delaunay TIN is that whenever there are points with similar *x* and *y* coordinates but very different *z* elevations they will form needle-shaped triangles that appear as spikes in the TIN [4]. These spikes create extreme irregularities in canopy surface elevation that geometrically distort the structure of the raster DSM [5]. That first returns can be close in *x* and *y* but distant in *z* comes from (a) merging of different LiDAR flight lines, (b) laser beams penetrating deep into the canopy before producing a first return and (c) off-nadir scan-angles that give the LiDAR system a clear view of the top of the tree as well as the ground beneath. A number of researchers report that the presence of first returns penetrating the canopy creates valleys within tree crowns, making trees less distinguishable [6]. These downward spikes – also called "pits" – that are present in a first-return TIN reduce the accuracy of subsequent extraction of forest inventory information such as treetop detection and crown segmentation [3].

In order to address this issue, researchers have recommended several algorithms that attempt to minimize the influence of data pits [3-5]. However, these typically are raster-based techniques that are applied to the final (normalized) DSM after generating it from first returns only. In the following we present a novel and efficient algorithm that uses all LiDAR returns (instead of only the first returns) to generate a pit-free Digital Surface Model by constructing a Triangulated Irregular Network using an incremental constrained Delaunay method.

Material and method

Our algorithm uses the raw point cloud (i.e., un-classified and un-normalized) as input and generates a pit-free raster-based DSM and (optionally) its corresponding pit-free TIN-based DSM as output. The algorithm incrementally constructs the pit-free TIN by avoiding the creation of downward spikes in the first place. This is achieved by "freezing" higher triangles whose three edges are shorter in the *x*-*y* plane than user-specified threshold, called the *freeze constraint*. The core idea is to give preference to connect LiDAR returns into TIN triangles that are nearby not only in *x* and *y* values, but also in *z*. For this, our algorithm iteratively inserts LiDAR points in batches of elevation intervals, for example, 1, 3, or 5 meters, into a constrained Delaunay triangulation starting with the highest elevations (Figure 1). After each batch insertion, the algorithm "freezes" all those small

triangles whose three edges are shorter (in the *x*/*y* plane) than the freeze constraint (Figure 1). When inserting the next lower batch, all LiDAR points falling into "frozen triangles" are simply ignored.

The freeze constraint essentially covers-up the pits by freezing smaller triangles that are higher up in the canopy. If the value is too small, few triangles are frozen and "pits" reappear. If the value is too large, many frozen triangles will also cover the morphological structure of the crown or even connect nearby crowns. We use pulse spacing statistics to pick a suitable freeze constraint. The pulse spacing determines the size of the smallest feature that can be captured by the LiDAR survey and the expected *x*-*y* length of edges in the final TIN.

The results is a "pit-free" TIN that can either be used directly or be rasterized into a "pit-free" raster DSM. A prototype of this algorithm has been implemented and will soon be available as an integral part of the LiDAR processing modules of LAStools [7]. The proposed algorithm has capability to create a pit-free DSM (either TIN or raster) from any kind of acquired point clouds datasets (e.g., LiDAR or dense-matching photogrammetry).

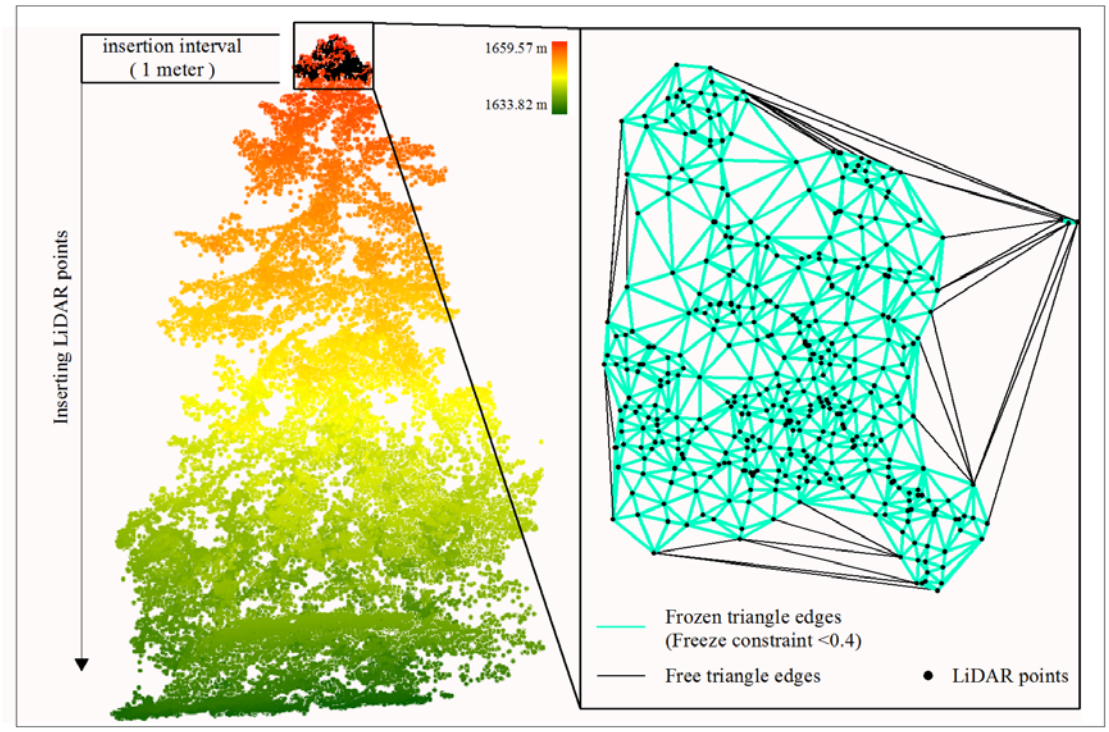


Figure 1: Constructing a pit-free TIN for an individual tree. The algorithm iteratively inserts LiDAR points in batches of elevation intervals (1 meter). The freeze constraint for freezing triangles in the triangulation that is being generated is set to 0.4 meters.

We applied our algorithm to LiDAR data collected in the Bois Noir (Black Wood) forest in the Barcelonnette basin located in the southern French Alps, which is mainly covered by coniferous plantation forests. The LiDAR data were acquired using a helicopter during the leaf-on season with a small footprint full-waveform LiDAR system (RIEGL VQ-480i), which was developed specifically for mapping mountainous forested area. The system was using a laser pulse repetition rate of 300-kHz and a scan width of 60° and performed on-line full-waveform analysis to extract up to five discrete returns for each pulse. The mean point density was 160 point/m^2. In order to demonstrate the robustness of our algorithm, we artificially created a lower-density version (mean pulse density of 7 point/m^2) of the LiDAR data using simple thinning algorithm [3].

Results and discussion

We ran our algorithm using a freeze constraint of 0.4 m (and 1.5 m) and an insertion interval of 1.0 m on our high-density (and low-density) data set to generate DSM rasters with a grid size of 0.15 m (and 0.5 m) that were sufficient for recognizing a minimum crown diameter of 0.50 m. The result demonstrates that our algorithm is able to efficiently remove all irregularities in the canopy surface from high- and low-density LiDAR data sets. Figure 2 shows the DSMs that are generated from our pit-free algorithm compared with standard DSMs. As you can see, the algorithm successfully removed all pits – small dark squares – by constraining the TIN construction process to assure that only LiDAR returns that are nearby when hitting the canopy at various levels are connected in the final DSM. However, the pits are clearly visible in the standard DSM that is generated by simply constructing the TIN using LiDAR first returns. Our result indicate that our algorithm is especially successful for low-density LiDAR data full of pits such as our artificial example data which corresponds to a small footprint laser beam sampling the canopy with low density. In our experiments, the algorithm effectively removed the pits as well as preserved the edges of canopy gaps and crowns.

The proposed algorithm has several advantages over our previous algorithm for constructing pit-free CHMs or height-normalized DSMs [3]: (a) it uses the information of all LiDAR returns (instead of only the first returns) to generate a DSM, (b) it can be utilized with any type of input data be it LiDAR or point clouds derived via photogrammetry, (c) it can operate on raw point clouds (unclassified and un-normalized), (d) it does not generate large amounts of temporary partial rasters and, therefore, it is much faster – especially on steep terrains, and (e) it cannot only produce pit-free DSM rasters but (optionally) also output the corresponding triangulated version – pit-free TINs which can be directly used in TIN-based approaches [8].

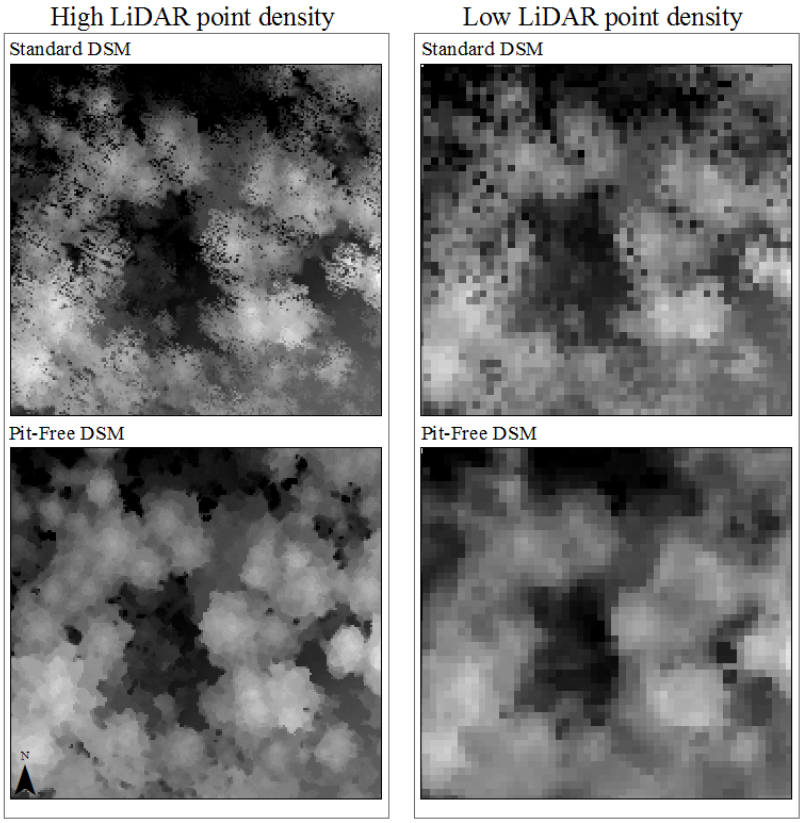


Figure 2: The pit-free DSMs versus standard DSMs on high- and low-density LiDAR data.

References

[1] Yao, W., Krzystek, P., & Heurich M. (2012). Tree species classification and estimation of stem volume and DBH based on single. Remote Sensing of Enviroment, 123, 368–80.

[2] Hyyppä, J., Hyyppä, H., Leckie, D., Gougeon, F., Yu, X., & 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, 1339–66.

[3] Khosravipour, A., Skidmore, A.K., Isenburg, M., Wang, T., & Hussin, Y.A. (2014). Generating Pit-free Canopy Height Models from Airborne Lidar. Photogrammetric Engineering & Remote Sensing, 80, 863–872.

[4] Axelsson, P. (1999). Processing of laser scanner data—algorithms and applications. ISPRS Journal of Photogrammetry and Remote Sensing, 54, 138–47.

[5] Ben-Arie, J.R., Hay, G.J., Powers, R.P., Castilla, G., & St-Onge, B. (2009). Development of a pit filling algorithm for LiDAR canopy height models. Computers and Geosciences, 35, 1940–49.

[6] Pitkänen, J., Maltamo, M., Hyyppä, J., & Yu, X. (2004). Adaptive methods for individual tree detection on airborne laser based canopy height model. International Archives of Photogrammetry, Remote Sensing & Spatial Information Sciences, 36,187–91.

[7] rapidlasso GmbH. (2015). LAStools – fast tools for LiDAR processing, http://rapidlasso.com.

[8] Vauhkonen, J., Holopainen, M., Kankare, V., Vastaranta, M., & Viitala R. (in press). Geometrically explicit description of forest canopy based on 3D triangulations of airborne laser scanning data. Remote Sensing of Environment.

Dual-Wavelength Echidna Lidar: First deployments at TERN sites around Australia

Michael T. Schaefer[1,2], David L. Jupp[2], John Armston[3], Alan Strahler[4], Crystal Schaaf[5], Zhan Li[4], Glenn Howe[6], Kuravi Hewawasam[6], Jason Martel[6], Glenn J. Newnham[7], Jenny L. Lovell[8], Ewan S. Douglas[9],

[1]*Precision Agriculture Research Group. School of Science and Technology, University of New England, Armidale, NSW 2351, Australia*
[2]*CSIRO Land and Water GPO Box 1666, Canberra, ACT 2601, Australia*
[3] *Remote Sensing Centre, Department of Science, Information Technology, Innovation, Ecosciences Precinct, 41 Boggo Road, Dutton Park, QLD 4102, Australia.*
[4] *Department of Earth & Environment, Boston University, Boston, MA, USA*
[5] *School for the Environment, University of Massachusetts Boston, Boston, MA, USA*
[6] *Department of Physics and Applied Physics, University of Massachusetts Lowell, Lowell, MA, USA*
[7] *CSIRO Land and Water, Clayton South, VIC, Australia*
[8] *CSIRO Marine and Atmospheric Research, Hobart, TAS, Australia*
[9] *Astronomy Department, Boston University, Boston, MA, USA*

Highlights: A dual-wavelength echidna lidar (DWEL) has been deployed at three terrestrial ecosystem research network (TERN) supersites across Australia, enabling researchers to successfully distinguish trunks/branches from leaves. For each of the scan locations, the 3D point clouds have been used to generate foliage profiles as well as LAI and PAVD information.

Key words: *DWEL, TLS, foliage profiles, LAI.*

Introduction

A new terrestrial forest scanning lidar has recently been deployed in Australia to three separate terrestrial ecosystem research network (TERN) calibration validation sites to evaluate forest structure. The dual-wavelength echidna lidar (DWEL) is one of two specialised instruments in the world that has been developed specifically to characterise forest structure at fine spatial scales while distinguishing between leafy and woody vegetation which is not as easily achieved using a single wavelength laser scanner. This is achieved by the coaxial alignment of two laser beams of separate wavelengths that display different reflectance properties when interrogating leafy and woody vegetation. DWEL is also capable of capturing partial hits from soft targets (leaves) in the forest canopy due to the retention of the laser pulse shape from these hits and analysis of the fully detected and digitised waveform.

This paper evaluates the first deployments of DWEL around Australia at selected TERN field sites and discusses some of the initial results from these three deployments.

Materials and methods

Instrumentation

The DWEL is a stationary terrestrial laser scanner that uses a rotating mirror to scan in a vertical circle as the whole instrument head rotates in the azimuth direction on its tripod base. This provides full angular coverage of the upper hemisphere as well as much of the lower hemisphere around the instrument. Concurrent, coaxial laser pulses produced by the two onboard lasers (1064 nm and 1556 nm) are long (approximately 5.0 ± 0.1 ns at the full-width half-maximum position of the pulse) and sharply peaked which allows full sampling of the return pulse waveform shape as well as easy interpolation for accurate distance retrievals to scanned objects. Additionally, the lasers have been individually calibrated to provide apparent reflectance measurements. The beam divergence of the lasers and scanning resolution are adjustable, with options of 1.25, 2.5 and 5 mrad available for the lasers and 1, 2, and 4 mrad for the scanning resolution. The resulting data set can be used to examine structural forest parameters, such as the mean tree diameter at breast height (DBH), stem density, basal area, canopy height, leaf area index, foliage profiles and above-ground biomass using approaches originally developed and used with the Echidna validation instrument [1]–[4].

Data Collection

The DWEL instrument has been deployed at three TERN AusCover supersites around Australia (with each site consisting of a 5 km by 5 km area). Within each of these sites are usually a series of long term monitoring plots that are of interest to a variety of different research groups.

Processing steps

The full waveform data produced by DWEL must go through an initial suite of processing steps before it is usable for scientific analysis. The steps are as follows: 1) convert the raw DWEL binary waveform data to hdf5 format; 2) derive laser direction data from the DWEL rotary encoders and assign to waveform data; 3) the data is then filtered and corrected for any saturation using IDL; 4) 3D point clouds are produced of the individual scattering events with x, y, z coordinates and associated apparent reflectance attributes; 5) point clouds are converted into SPD format; and 7) foliage profiles and LAI estimates are produced using the SPDtools python bindings [4],[5].

Results

Figure 1 displays colour composite images of the mean return values returned in a 4 mrad angular grid at the two wavelengths in selected scans from the 3 locations. Trunks and branches have a similar return power at both wavelengths, while leaves appear much darker at 1556 nm due to absorption by water in the leaves. Thus, trunks appear to be much brighter than the leaves except for the cases where the bark on the trees are either wet or appear to contain moisture. Gaps in the vegetation canopy are characterised by blue colouring.

(a)

(b)

(c)

Figure 1: Colour composite images of the mean return values in the dual-wavelength scans. Red, 1556 nm; green, 1064 nm: blue, canopy gaps. The horizontal dimension is azimuth angle and the vertical dimension is the zenith angle. (a) Site SS at Tumbarumba in Bago State Forest, New South Wales, (b) site Gold0101 in D'Aguilar National Park near Brisbane, Queensland, (c) Log decay site in the Warra Long-Term Ecological Research site, Tasmania.

Concurrent TLS scans were made with a Riegl VZ-400 laser scanner at the QLD and TAS sites allowing comparison of the 3D point clouds themselves as well as the foliage profile and LAI measurements from each scanner, a sample set of these profiles are displayed in Figure 2.

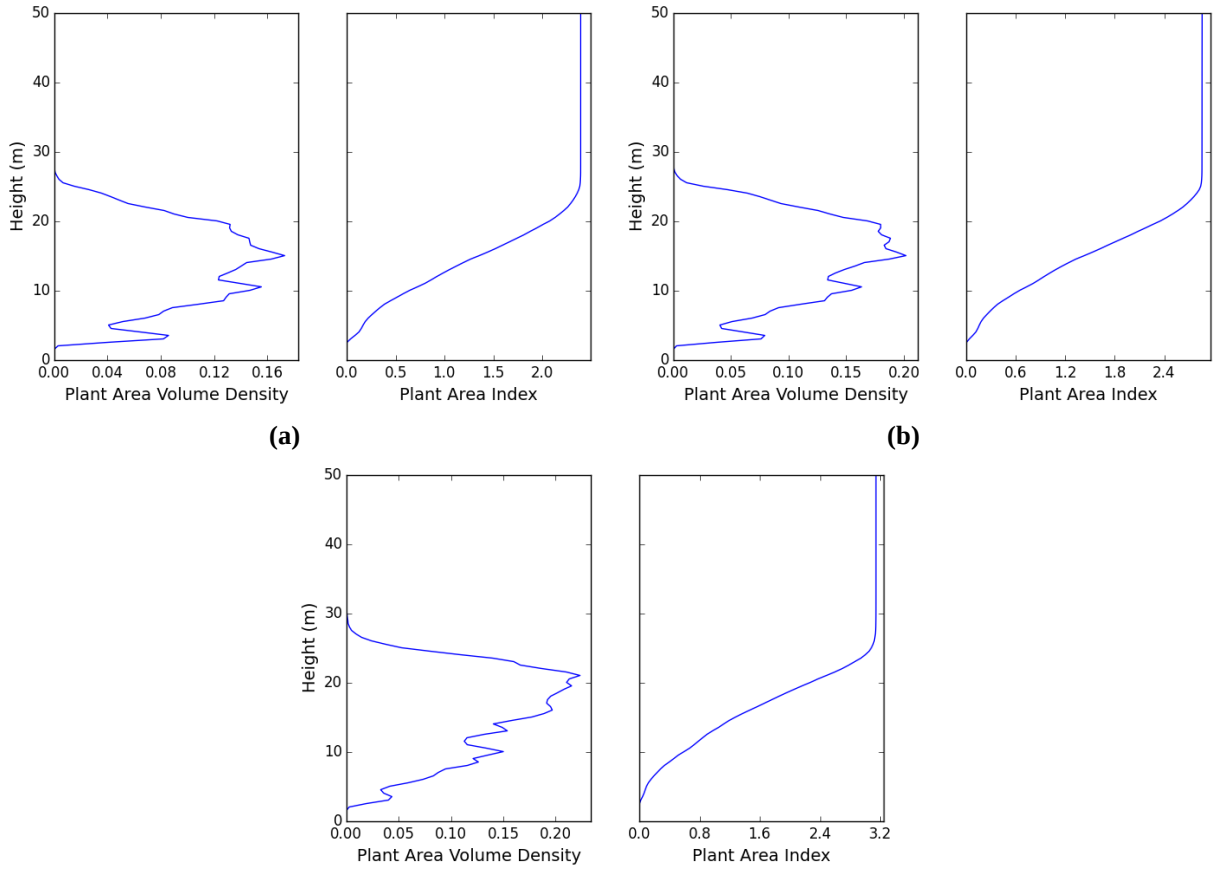


Figure 2: An example of foliage profiles from (a) DWEL 1064nm, (b) DWEL 1556 nm and (c) Riegl VZ-400 (1556 nm) at the QLD Gold0101 field site created using the python binding of the SPDtools software suite.

Foliage profiles of the DWEL and Riegl VZ-400 TLS instruments appear to be of a similar shape, particularly in the 1556 nm waveband. The extended range of the Riegl sensor results in an improved delineation of the canopy elements at range. The 1064 nm waveband of the DWEL instrument depicts a slightly different shape than that of the Riegl and DWEL 1556 nm band. The reason being that the 1064 nm is sensitive to both leaves and woody vegetation while the 1556 nm waveband is most sensitive the woody vegetation.

Conclusions

The dual-wavelength terrestrial lidar scans that are displayed in this research show very good contrast between leafy and woody vegetation when scanned using the DWEL instrument. Consequently, this allows foliage profiles and the plant area index to be calculated that are sensitive to changes in the amount of green or woody vegetation in a scan.

References

[1] Douglas, E. S., Martel, J., Li, Z., Howe, G., Hewawasam, K., Marshall, R. A., Schaaf, C. L., Cook, T. A., Newnham, G. J., Strahler, A., & Chakrabarti, S. (2015). Finding Leaves in the Forest: The Dual-Wavelength Echidna Lidar. IEEE Geoscience and Remote Sensing Letters, 12(4), 776-780.

[2] Lovell, J. L., Jupp, D. L. B., Culvenor, D. S., & Coops, N. C. (2003). Using airborne and ground-based ranging lidar to measure canopy structure in Australian forests. Canadian Journal Remote Sensing, 29, 607-622.

[3] Strahler, A. H., Jupp, D. L., Woodcock, C. E., Schaaf, C. B., Yao, T., Zhao, F., Yang, X., Lovell, J., Culvenor, D., Newnham, G., Ni-Miester, W., & Boykin-Morris, W. (2008). Retrieval of forest structural parameters using a ground-based lidar instrument (Echidna®). Canadian Journal of Remote Sensing, 34(sup2), 426-440.

[4] Zhao, F., Yang, X., Schull, M. A., Román-Colón, M. O., Yao, T., Wang, Z., Zhang, Q., Jupp, L. B., Lovell, J. L., Culvenor, D. S., Newnham, G. J., Richardson, A. D., Ni-Meister, W., Schaaf, C. L., Woodcock, C. E., & Strahler, A. H. (2011). Measuring effective leaf area index, foliage profile, and stand height in New England

forest stands using a full-waveform ground-based lidar. Remote Sensing of Environment, 115(11), 2954-2964.

[5] Calders, K., Armston, J., Nenham, G., Herold, M. & Goodwin, N. (2014). Implications of sensor configuration and topography on vertical plant profiles derived from terrestrial lidar. Agricultural and Forest Meteorology, 194, 104-117.

Integrating terrestrial laser scanner data into forest management decision processes

Robert Schneider[1], Richard A. Fournier[2], Guillaume Giroud[3], Alexa Bérubé-Deschênes[1], Hugues Power[4], Eduardo Bittencourt[5], and Joan E. Luther[6]

[1] *Chaire de recherche sur la forêt habitée, Département de biologie, chimie et géographie, Université du Québec à Rimouski (UQAR), 300 allée des Ursulines, Rimouski,QC, G5L 3A1, CANADA*
[2] *Centre d'Applications et de Recherches en Télédétection (CARTEL), Département de géomatique appliquée, Université de Sherbrooke, 2500, boul. de l'Université, Sherbrooke, QC, J1K 2R1, CANADA*
[3]*Direction des inventaires forestiers, Ministère des Forêts, de la Faune et des Parcs, Gouvernement du Québec, 5700, 4e Avenue Ouest, local A.108, Quebec City,QC, G1H 6R1, CANADA*
[4]*Direction de la recherche forestière, Ministère des Forêts, de la Faune et des Parcs, Gouvernement du Québec, 2700, rue Einstein, Quebec City,QC, G1P 3W8, CANADA*
[5]*Consortium en foresterie Gaspésie-Les-Îles, 37, rue Chrétien, C. P. 5, Gaspé, Québec, G4X 1E1, CANADA*
[6]*Natural Resources Canada, Canadian Forest Services – Atlantic Forestry Centre, 26 University Drive, Corner Brook, NL, A2H 5G4, CANADA*

Highlights: Forest management decisions are becoming more complex as several criteria must be considered. Practitioners thus require more detailed information. This project will develop a decision support system for silviculturists based on terrestrial laser scanner data. The system will have two modules: a resource evaluation system and a residual stand growth simulator.

Key words: *terrestrial laser scanner, resource evaluation, stand growth, silviculture*.

Introduction

Forest managers need more detailed information on the forest resource in order to optimize stumpage allocation to the mills. The 3D information obtained from terrestrial laser scanner (TLS) offers promising initial results in order to evaluate standing tree dimensions [1] and stand wood properties [2]. Accordingly, the objective of this project is to develop a decision support system based on TLS data. More specifically, the project will:

- Develop algorithms to estimate wood property attributes from tree metrics associated with structural variables such as crown attributes, branchiness, or inter-tree competition,
- Link these external tree metrics as well as non-TLS metrics to wood fibre attributes,
- Predict tree growth with both TLS and non-TLS metrics,
- Synthesize the procedures within a decision support system which can work with both TLS and usual inventory data to:
    - Evaluate standing timber external and internal wood properties, and
    - Simulate stand development on a 15 to 20 year horizon.

Materials

The project will use data from 2 different sources: (**DB1**) white spruce (*Picea glauca*) plantations which are ready for a first commercial thinning from eastern Quebec and (**DB2**) mature black spruce (*Picea mariana*) stands from Newfoundland published in [2].

In **DB1**, a total of 86 sample plots were scanned with a Faro Focus 3D scanner. All living trees were cored (i.e. 1810 trees), in order to obtain the past 5 and 10 year increments. An additional 10 plots were also scanned with the Faro scanner, where a total 80 balsam fir (*Abies balsamea*) and 80 white spruce trees were felled, and product recovery simulated using Optitek© mill simulator developed by FPInnovations.

A subset of the data used for a published study on the estimation of wood fiber attributes [2] were retained for **DB2.** It contains information on 76 black spruce trees from 16 plots. The plots were scanned with a Zoller and Fröhlich Imager 5006i. A wood core was acquired for each sample tree at breast height (1.3 m) and processed by FPInnovations for wood property measurements. Of all the possible wood attributes, wood density, microfibril angle, module of elasticity and fibre length were retained.

Methods

Wood quality attributes

The importance of wood properties varies with tree age. For trees harvested during commercial thinning, such as those available in **DB1**, the juvenile core is important. Wood product quality is thus not determined by the fibre attributes, but through lumber visual grading. The most important factor which downgrades lumber is knot size. Hence efforts will focus on establishing crown and branch attributes from the TLS data. The TLS information will thus be used to develop branch characteristic attributes, such as branch density (e.g. nb. of branches/m) and branch size estimation. Comparisons between product recovery using sawmill scanners, raw field TLS data and TLS taper derived models will be compared, in order to estimate the gain in precision of taper algorithms. Other tree characteristics such as stem taper and sweep will also be obtained from the TLS data. These metrics, as well as the product recovery, will be related to standard inventory data and TLS-based metrics.

Wood fibre attributes, as those available in **DB2**, are more important to assess for mature trees. Pith-to-bark wood fibre attributes (e.g. density, module of elasticity, microfibril angle, fiber length) will be modelled for each tree. The parameters of each model will then be related to inventory and TLS-based metrics to establish a general predictive model which will take into account site variability, tree characteristics and plot-level information. The following metrics will be extracted from the TLS data and included in the model: micro-site information such as local slope and stem density, canopy information like the Rumple Index, and local crown competition indices obtained with voxel density-distant based method.

Tree growth model

Tree growth information, available from the core measurement of **DB1** and **DB2**, will be used to calibrate a growth model based on potential growth [3]:

$$\text{increment} = \text{potential growth} \cdot \text{red}_{\text{comp}} \cdot \text{red}_{\text{size}} \quad (1)$$

where the increment of an individual tree is determined by its potential growth for a given site, which is reduced by competition (red_{comp}, a function defined between [0,1]) and its size (red_{size}, a function defined between [0,1]). Tree competition will be quantified through spatially and non-spatially explicit competition indices. Mortality models are already available (managed stands: [4]; natural stands: [5]), and will be adapted to the project needs. A first step will be to use the TLS cloud points to obtain the tree map. However, TLS-based competition indices developed in the wood quality section will be explored in predicting tree growth, once the permanent sample-plots will have grown 5 years after the first measurement (e.g. in 2018).

Decision support system

The algorithms and models developed in each section will form the basis of the decision support system (Fig. 1). As TLS is not widely available for forest managers, allometric algorithms will be developed in order to enable the input of conventional inventory data. These algorithms will be expanded to allow use of spatially-explicit data like forest inventory maps or airborne lidar data. TLS algorithms and processing will be carried out in the CompuTree platform (http://computree.onf.fr), whereas the growth and wood quality modules will be programmed in the Capsis platform [6].

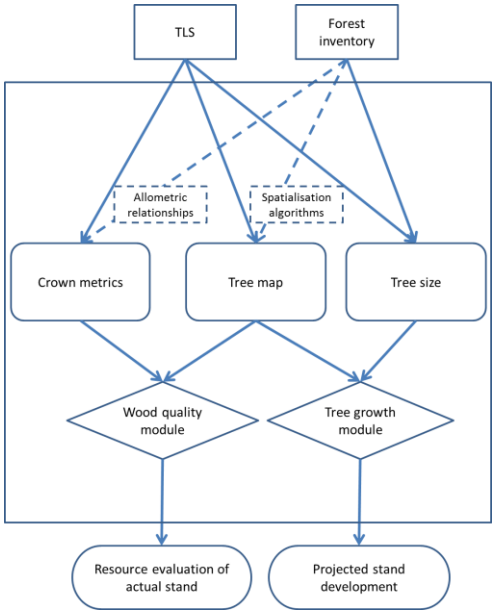


Figure 1. Schematic representation of the decision support system components.

An important outcome of the project is to provide forest managers with tools to evaluate the tree properties related to wood product quality through the wood quality module. These tools will be able to estimate volume and quality of the harvested trees for different silvicultural entries. With the tree growth module, it will be possible to project the development of the residual trees, 15 to 20 years forward.

Acknowledgements

The authors would like to thank the Fonds de recherche du Québec - Nature et Technologies for funding. They would also like to underline the contributions of the Newfoundland Fibre Project for TLS data and the Forest Industry Competitive Advantage Project led by Corner Brook Pulp and Paper Limited for the fibre attribute data.

References

1. Othmani A, Piboule A, Krebs M, et al. (2011) Towards automated and operational forest inventories with T-Lidar. Hobart , Australia, pp 1–9
2. Blanchette D, Fournier RA, Luther JE, Côté J-F (2015) Predicting wood fiber attributes using local-scale metrics from terrestrial LiDAR data: A case study of Newfoundland conifer species. For Ecol Manag 347:116–129.
3. Biging GS, Dobbertin M (1992) A comparison of distance-dependent competition measures for height and basal area growth of individual conifer trees. For Sci 38:695–720.
4. Penner M, Swift DE, Gagnon R, Brissette J (2006) A stand density management diagram for balsam fir in New Brunswick. For Chron 82:700–711.
5. Fortin M, Langevin R (2010) ARTEMIS-2009 : un modèle de croissance basé sur une approche par tiges individuelles pour les forêts du Québec. Direction de la recherche forestière
6. Dufour-Kowalski S, Courbaud B, Dreyfus P, et al. (2012) Capsis: an open software framework and community for forest growth modelling. Ann For Sci 69:221–233.

Analysis of roundwood surface for detection of inner defects.

V-T Nguyen [1,2], T. Constant [1,2], D.Elhareth, [1,2] A. Piboule[3], F. Colin [1,2]

[1] *AgroParisTech, UMR 1092 LERFOB, 54000 Nancy, France*
[2] *INRA, UMR 1092 LERFOB, 54280 Champenoux, France*
[3] *ONF, RDI, 54000 Nancy, France*

Highlights: Different methods for analysing roundwood surfaces described by T-LiDAR were proposed. The aim is to link external defects affecting the surface roughness to the presence of inner defects in order to grade logs or standing trees. The relevance of the methods were tested against 3D models of the inner structure obtained from X-Ray CT scanning of logs

Key words: Terrestrial Lidar, X-Ray Scanning, Roundwood , Roughness, Grading

Introduction

In forestry, T-Lidar technology provides high detailed information about the geometrical structure of a forest plot or an individual tree. This geometrical information can be used for different purposes going from biomass estimation until the feeding of functional plant structure models.[1].

In this work, the aim was the extraction of information contained in 3D data, concerning the grading of roundwood or standing trees. The literature about wood quality features extracted from Lidar data is scarce. The characteristics of branches have naturally been included into many works looking at the reconstruction of single trees from the first works [2], to the last ones [3]. Nevertheless, some works were more focused on the wood quality: Schütt [4] developed the concept of extracting quality criteria from Lidar data, by using neural networks. To our knowledge subsequent works have not been published. Other papers focused on branchiness through the characterization of existing branches [4] or by using the morphology of the logs [5] of branch scars for assessing the clearwood radius [6].

Here, the focus was given to the tree part, i.e. the main bole, providing relevant matter for timber industry, slicing or rotary cutting processes, or stave mills requiring information about wood quality impacting commercial values. Different levels of information are useful for grading logs. Of course, the most important level is the volume itself since the volume of products is directly linked. Nevertheless, the accurate estimation of the volume is still a open question even if methods based on the fitting of geometrical primitives are already available [3]. But characteristics describing the overall shape of the volume are also important because the sawing pattern can be optimized according to features such as cross-sectional, or longitudinal shapes, taking into account ovality, curvature and flexuosity, respectively. Furthermore for standing trees information about the inclination of the trunk, as well as curvature, is an indicator for biomechanical processes affecting the inner wood quality. (See. Noyer et al. In Silvilaser 2015).

In this beginning project the aim was to look after additional information corresponding to local defects which have to be taken into account for grading roundwood. as specified by the different norms (EN-1316 EN-1927). Among the defects relevant for roundwood grading, the geometrical information is not always sufficient, Reflectance is an additional information provided with no additional cost in the acquisition process, at the opposite coloured 3D clouds imply additional costs, mainly caused by the duration of the acquisition. This colour information can be useful for detecting some defects especially on the cross-section of logs, it could be also relevant for some outer defects as debarking, but in this first step of this project, only 3D information without colour, was concerned. But, the possible use of this additional information was kept in mind for improving the processing in further steps.

A large part of the defects we are looking for belongs to the rameal system of a tree. They go from from a living sequential branch until the branch scar at the bark surface, passing through all the intermediate stages. But epicormic formation such as burls or shoots are also targeted even if their size can be a limiting characteristic for detecting them against the resolution of the data.

The main objective of the work presented here concerns the development and an original evaluation of the performance of an algorithm dedicated to the detection of areas prone to correspond to a defect from the analysis of the surface roughness of the trunk or log surfaces.

Material and Methods :

Sixteen one meter long logs belonging to different species, and coming from a larger database were used for testing the performance of the algorithm. The particularity of this database is to gather T-LiDAR and X-Ray scans of the same logs. In previous works [7], the X-Ray scans were analysed and provide a geometrical model of the inner structure of the rameal system of each log, identifying the shape, the size, and the type of each element.

The work-flow for identifying clusters of 3D points susceptible to be a defect at the surface of a log described by T-Lidar was as follows : Firstly, the definition of the principal axes of the 3D cloud were calculated for allowing to work in a cylindrical coordinate system, and for easily defining angular sectors along the log. Then the general trend of the radial variation vs the longitudinal position was fit by a second order polynomial for each angular sector. This trend defines a quadratic surface which is used as a reference to compute differences in terms of radius and normals with the 3D points. The statistical distribution of these differences is used to define automatic thresholds taking different features of the distribution into account for detecting the suspected areas. Among different methods, we can notice the extraction of extreme values based on the assumption of a statistical parametric distribution of the differences, and another ones using the asymmetry of the distribution for detecting abnormal differences.

The location of the suspected areas is compared to the reference data coming from X-Ray scanning of the log, giving information about the location, the depth and the dimensions of the defects, and concerning branches as well as buds. This comparison is achieved thanks to a matching algorithm of both types of scanned data.

All the developments required to achieve the comparison with T-Lidar data were implemented into the Computree platform (http://www.computree.onf.fr) within the Lerfob plugin but this part is not yet distributed

Results :

We present here first results concerning one algorithm detecting suspected areas. The latter is based on the statistical distribution of the distance of 3D points to the reference surface. These tests were carried out on eight logs of different species. Different parameter values proportional to the standard deviation σ were tried to segment potential defect areas. The Figure 1a shows the best results, indicating that the best multiplicative coefficient was around 1.6 σ . The defects present in the eight logs were branches often limited to a knot occurrence, then a little number of epicormic shoots and finally numerous buds, with a very small size and difficult to detect.

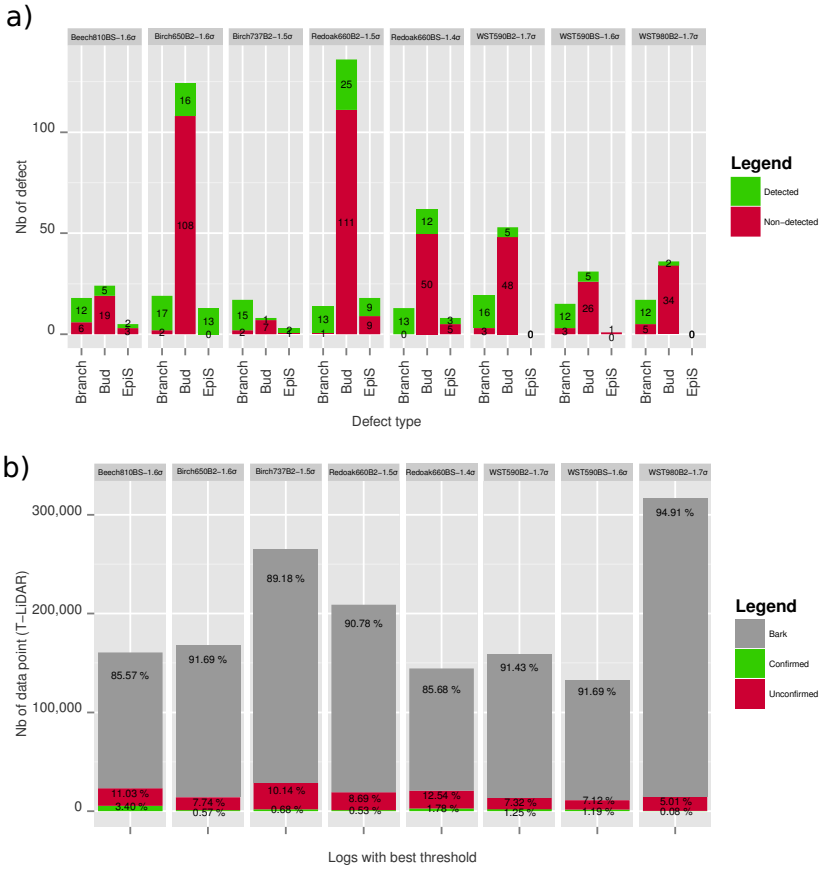


Figure 1 :First results on eight logs for the algorithm detecting clusters of points suspected to be a defect and based on the distance to reference surfaces built by angular sectors..

The Figure 1b presents the total number of 3D points describing each of the eight logs, in grey the points which were not detected as a defect, in green the points identified and confirmed as belonging to a potential defect, an in red the points detected as potential defects but not confirmed.

This work is on progress, and different segmentation methods are currently tested.

References

[1] Dassot, M., Constant T., and Fournier M.. "The Use of Terrestrial LiDAR Technology in Forest Science: Application Fields, Benefits and Challenges." *Annals of Forest Science* (2011): 1–16.

[2] Pfeifer, N., Gorte, B., Winterhalder, and D. "Automatic Reconstruction of Single Trees from Terrestrial Laser Scanner Data." *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences* 35, no. 2 (2004): 114–19.

[3] Raumonen, P., Kaasalainen M., Akerblom M. , Kaasalainen Sanna, Kaartinen H., Vastaranta M., Holopainen M., Disney M., and Lewis P.. "Fast Automatic Precision Tree Models from Terrestrial Laser Scanner Data." *Remote Sensing* 5, no. 2 (February 2013): 491–520. doi:10.3390/rs5020491.

[4] Schütt, C., Aschoff T. , Winterhalder D., Thies M., Kretschmer U., and Spiecker.H. "Approaches for Recognition of Wood Quality of Standing Trees Based on Terrestrial Laser Scanner Data." In *ISPRS*, 36:W2, 2004. http://www.isprs.org/proceedings/xxxvi/8-w2/schuett.pdf.

[5] Thomas, L, Thomas R.E. , "A Graphical Automated Detection System to Locate Hardwood Log Surface Defects Using High-Resolution Three-Dimensional Laser Scan Data." In *Proceedings, 17th Central Hardwood Forest Conference, Newtown Square, PA, USA. US Department of Agriculture, Forest Service, Northern Research Station, General Technical Report NRS-P-78*, 92–101, 2011. http://www.nrs.fs.fed.us/pubs/gtr/gtr-p-78papers/11thomasp78.pdf.

[6] Stängle, S; M., Brüchert F, Kretschmer U., Spiecker, H and Sauter U; H.. "Clear Wood Content in Standing Trees Predicted from Branch Scar Measurements with Terrestrial LiDAR and Verified with X-ray CT." *Canadian Journal of Forest Research* (October 7, 2013): 131007143804004. doi:10.1139/cjfr-2013-0170

[5] Klemmt, H-J, Seifert T., Seifert S., Kunneke A., Wessels B., and Pretzsch H.. "Assessment of Branchiness in a Pinus Pinaster Plantation by Terrrestrial Laser Scanner Data as a Link between Exterior and Interior Wood Properties." In *Proceedings of Silvilaser 2010,*. Freiburg, n.d.

[7] Morisset, J-B, Mothe F., and Colin F.. "Observation of Quercus Petraea Epicormics with X-ray CT Reveals Strong Pith-to-bark Correlations: Silvicultural and Ecological Implications." *Forest Ecology and Management* 278 (August 15, 2012): 127–137.

Optimized process chain for mountainous forest and non-forest classification using airborne lidar data

Qingwang Liu[1], Zengyuan Li[1], Yong Pang[1], Kailong Hu[1, 2]

[1] *State Laboratory for Forest Remote Sensing and Information Techniques, Research Institute of Forest Resource Information Techniques, Chinese Academy of Forestry, Beijing, 100091, P. R. China;*
[2] *College of Geoscience and Surveying Engineering, China University of Mining & Technology (Beijing), Beijing, 100083, P. R. China*

Abstract: The study area was dominated by subtropical evergreen and deciduous forest. The point cloud data was acquired with experimental lidar system (LRS200A) on December, 2014. This paper developed an optimized process chain to classify different returns from forest and non-forest, and empirical parameters could produce remarkable classification results.

Key words: *forest, classification, lidar, point cloud, process chain*

1. Introduction

Mountainous forests with complex canopy structure characteristics are difficulty to measure with ground-based methods. Airborne lidar can precisely obtain terrain features and canopy three dimensional structure features by transmitted and received laser energy. The forest canopy returns of laser pulses are different from non-forest returns, such as ground, building, high voltage line & tower, etc. Most works [1-4] focus on the terrain extraction, individual tree segmentation, stand parameters invasion, etc. It's specially supposed use finer classification of forest and non-forest on mountainous area. According to the data process chain [5], the forest and non-forest returns can be well classified under certain conditions. But it's very difficult to distinguish forest and high voltage line & tower from return points. This paper developed an optimized process chain to check and refine classification result of forest and non-forest.

2. Material

The study area locates at Dongbao Mountain, Jingmen, Hubei province, P. R. China (see Figure 1, the red slant rectangle indicates flight area in Figure 1(b)), lies within the subtropical and humid climate zone, characterized by mountain evergreen and deciduous forest. The geolocation extent are 112:01:08E-112:13:46E, 30:58:18N-31:07:42N, and terrain varies from 71-340 m.

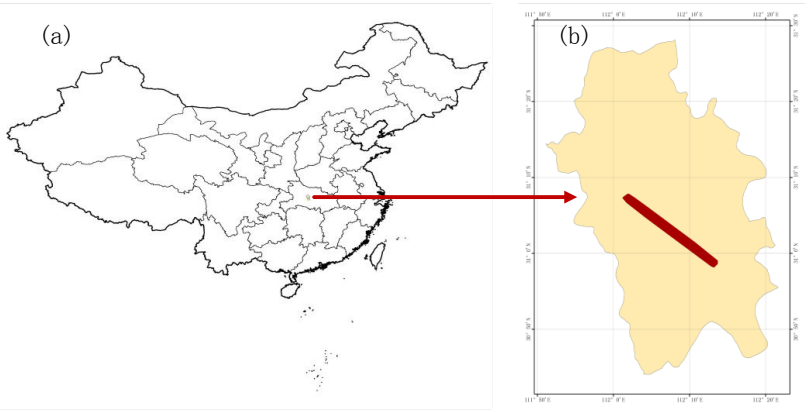


Figure 1 Study area site

Lidar point cloud data were acquired using experimental lidar system (LRS200A) on December, 2014. Laser wavelength is 1064 nm, pulse repeat frequency is 200 kHz, see Table 1 for more detailed descriptions. The mean of sampling density is 2.16 points per square meter. Point cloud density distribution fluctuated along the flight lines (see Figure 2).

Table 1 Laser scanning system specification

Laser scanning system			
Model	LRS200A	**PRF**	200 kHz

Wavelength	1064 nm	**Beam divergence**	0.3 mrad
Scan mode	oscillating mirror sinusoidal modulation	**Max scan angle**	45 deg
Returns	4 (3cm)	**Max scan frequency**	100 Hz
Working range	50～1000 m	**Height accuracy**	<15cm (1000m)

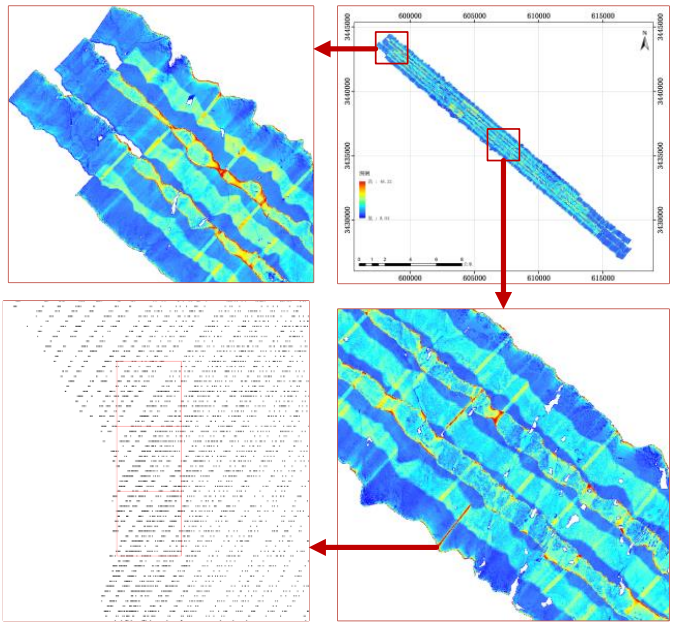


Figure 2 Point cloud density distribution

The point cloud data includes echo returns from forest, terrain, building, high voltage line & tower, road, etc. There are almost no return points in the water area due to the mirror reflection or absorption from reservoirs and rivers. Laser point coordinates were solved with IMU and DGPS information.

3. Methods

Point cloud classification mainly depend on the 3D spatial structure characteristics. Laser returns from terrain normally under returns from other objects above ground surface, such as vegetation, buildings, except for some pits or holes. It is difficult to distinguish vegetation and other mixed objects with similar heights. Ordinary process methods had some disadvantages to accurately classify forest and non-forest returns dealing with complex circumstances. An optimized process chain was developed by repeatedly experimental method (see Figure 3).

The optimized process chain's inputs were point cloud data tiles. Ground point algorithm was used to extract ground returns from default returns. The initial ground returns were refined by low point algorithm with different parameter settings. The non-ground returns were classified as near-ground, understory, medium canopy, high canopy, and emergent returns using above ground algorithm with a serial parameter settings. Return check method was used to distinguish non-forest returns from the initial forest returns.

I Prepare

Point cloud data should be well geometry corrected and tiled according to specified tile size. Data tiles take advantage of reduced file size, overlapped returns of adjacent flight lines, convenient data check, etc. Considering the memory limits of different computers, large flight data files were tiled into small files. Each tile contains returns from any flight lines if the returns fall into the tile extent.

II Ground point algorithm

Ground points are assumed as some local low points that are confident hits on the ground within specified window size. Ground point algorithm classifies ground points by iteratively building a triangulated surface model [6]. The search window size will influence detected ground points. Small window size tends to keep little variation of steeple slope, while large window size cut sharp ups and downs area and get smoother surface.

III Low point algorithm

Low points lies under ground surface, which may be some returns from small pits or holes. Low points will produce false ground returns, and cause pits effect on the triangulated surface. The common case of low points is bits of low returns or tens of low returns underground surface. Tens of low returns detections should use large window size to avoid misclassify non-low returns. Bits of low returns detection can get better results with small windows size.

IV Above ground algorithm

Above ground points are classified as different types according to the return heights above the ground surface. The near-ground points are regarded as returns from rocks, grasses, etc. The understory points represent returns from shrubs or little trees. Forest canopy returns are classified as medium canopy and high canopy points. Emergent points may be returns from extraordinary trees, high voltage line & tower, etc. Return check can be used to distinguish non-forest returns.

V Return check method

The automatic classification result of ground, forest, non-forest should be carefully checked to reduce misclassified points. Return check is conducted by visually examine the horizontal and vertical distribution of different returns with eyes. Ground point check include the large gaps of horizontal distribution, and the sink & salient areas on TIN surface. Low building points mainly confounded with understory returns. Building roof points have flat features. High voltage line & towers mixed with forest canopy, which are very difficult to distinguish from canopy returns with sparse distribution. Emergent points are classified to solve this problem, which are far less than medium canopy and high canopy points. Sparse returns from high voltage line & towers could be fast checked to discriminant from extraordinary trees returns.

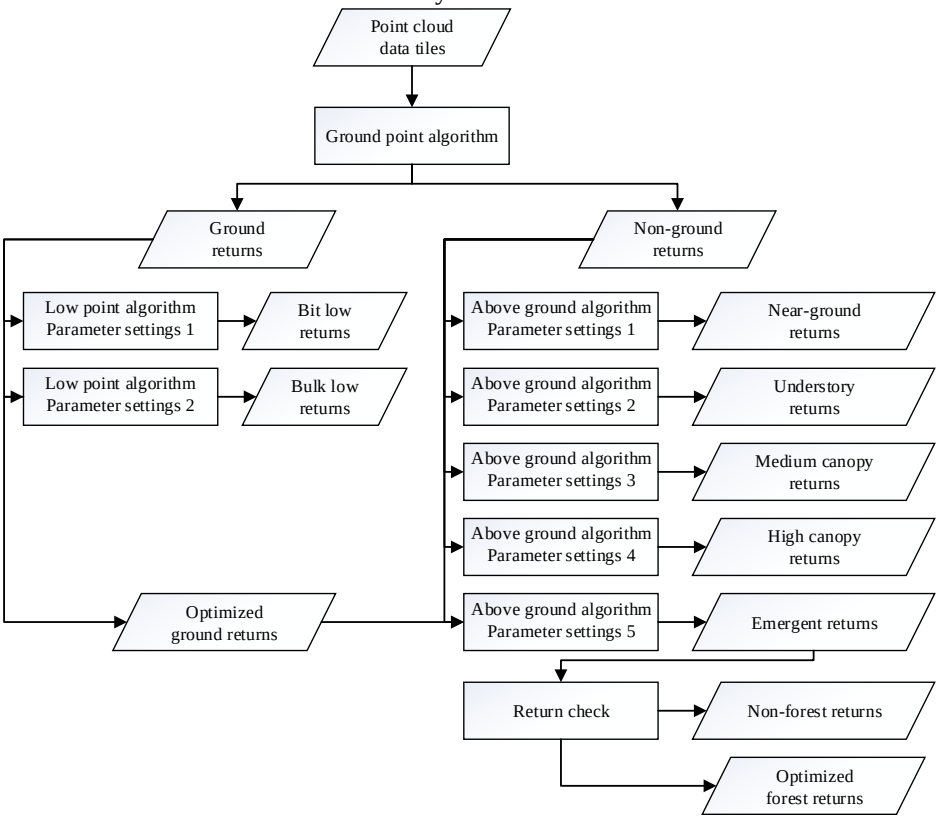


Figure 3 Optimized process chain for forest and non-forest classification

4. Results

The representative point data tiles were tested to obtain remarkable result using different parameter settings. The ground point algorithm's window size selects 30m to keep moderate variations of terrain and sweep small building on the slope. The low point algorithm use 15m window size and 50 points threshold to classify tens of low points, 5m windows size and 2 points threshold for bits of low points. The above ground algorithm use 0.5m, 2m, 6m and 26m to separate different vegetation type. Recommended parameter settings for optimized process chain on study areas see Table 2.

Table 2 Recommended parameter settings for optimized process chain on study areas

Objects	Parameters	Recommended values
Data tiles	Tile size	1km * 1km
	Edge buffer size	30m
Mountain ground points	Max. window size	30m
Tens of low points	Max. window size	15m
	Max. point count	50
	Distance threshold	0.5m
Bits of low points	Max. window size	5m
	Max. point count	2

	Distance threshold	0.5m
Near-ground points	Min. height	0m
	Max. height	0.5m
Understory points	Min. height	0.5m
	Max. height	2m
Medium canopy points	Min. height	2m
	Max. height	6m
High canopy points	Min. height	6m
	Max. height	26m
Emergent points	Min. height	26m
	Max. height	99m

Laser return points were classified using optimized parameters and refined by visual analysis. The initial ground surface were created using ground classification algorithm and eliminated some pseudo ground returns from pits or holes by low point detection algorithm. The ground tiles were checked wall to wall by visual analysis. Near-ground, understory, medium canopy, high canopy and emergent points were classified by height above ground. Low buildings were usually less than 2 m and contained in understory canopy class. There haven't found efficient method to deal with such mixed buildings up to now. It's advised to use only the initial understory points and turn off other class display while visually extracting such buildings. High voltage lines & towers often mixed in the emergent class, and visually profile analysis method should be adapted to correctly distinguish such objects. Figure 4 shows the sample area of laser point classification. Raw point cloud were rendered by elevation, different classes selected easily distinguished colour scheme to facilitate visual analysis.

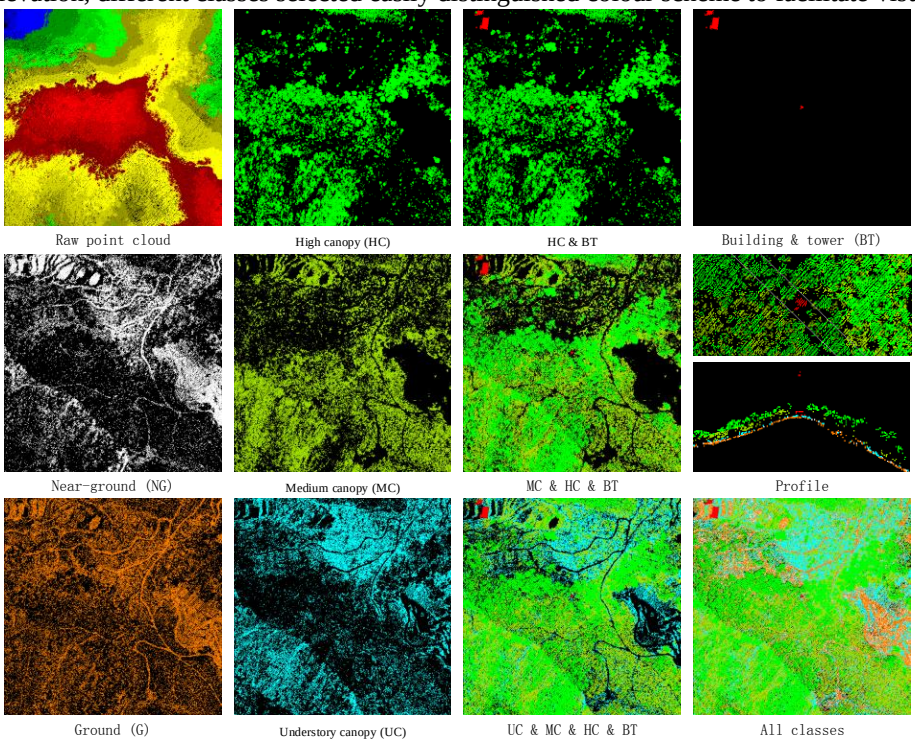


Figure 4 Sample area of laser point classification

5. Discussion

Mountainous forest return contains terrain information. The ground point algorithm's parameter settings should consider the terrain variations and artificial objects, such as buildings. The count of low point underground surface varied with the width and depth of pits or holes. The whole process chain should use representative area to train suitable parameter settings applying to other point cloud data tiles.

Near-ground points are returns from uneven ground or small objects. The overlap area between different flights also contribute to near-ground points due to the measuring errors. Near-ground points were easily influenced by flight conditions, laser scanning system and terrain etc. There were more uncertainness involved in such returns, which should be exclude from understory canopy. Near-ground were less usable while eliminating low buildings from understory by visual analysis. According to experience, near-ground points were classified by height threshold between 0 -0.5 m above ground surface on mountainous region without too cliffy slopes.

Emergent canopy points are returns from objects that are above normal canopies. Most dominate trees and sub-dominate trees composed normal canopy. Emergent canopy represent some higher dominate trees. High

voltage line & tower usually higher than forest canopies to ensure the safety of power transmission. Laser returns from emergent canopy and other higher objects have similar spatial features, which are difficulty to separate using recognition algorithms. Sometimes, low vegetation grow on break ridges or steep cliffs may be misclassified as emergent canopy points due to the outstanding part being cut off by ground extraction algorithms. It was very useful to distinguish laser returns from emergent canopy and high canopy in case of correctly recognizing non-forest objects with visual analysis method.

Ground and vegetation points could be used to generate canopy height model (CHM), which represent the spatial variation of canopy surface. The misclassified building were easily recognized on CHM with regular features by visual checking, while the leaky high voltage tower were difficulty to determined due to fewer cells existed. Statistics analysis will find possible errors within CHM. Vecterize of suspicious cells will enhance the interpretation of lanky objects.

6. Conclusion

Laser returns from ground surface should refined using the combination of ground point and low points algorithm. It's recommended to classify vegetation returns above ground into near-ground, understory, medium canopy, high canopy and emergent points. Near-ground class will facilitate distinguishing returns from low buildings and understory, while emergent class increase the ability of recognition between higher dominate trees and other lanky objects. The recommended parameter setting of optimized process chain should be adjusted according to representative cloud point data at different situations. Visual analysis should be used to correct misclassification and improve the process accuracy.

The optimized process chain for mountainous forest and non-forest classification obtain remarkable results using local parameter settings. The airborne lidar point cloud data were tiled and processed tile by tile. Stratified canopy assisted to check returns from non-forest objects, increased the distinguishability of different returns and accuracy of classification.

Automatic classification algorithms could seldom correctly distinguish returns from forest canopy and non-forest objects, such as small buildings, high voltage line & tower, according to the spatial relationship of points. More robust algorithm should be developed to solve this problem.

Acknowledgements

We are grateful to persons participating in data acquisition and processing for the airborne experimental campaign. This work was supported by the National Natural Science Foundation of China (Grant No. 41201334), the National High Technology Research and Development Program of China (Grant No. 2013AA12A302), the National Key Technology Research and Development Program of China (Grant No. 2012BAH34B02).

References

[1] Wulder M. A., White J. C., Nelson R. F., Næsset E., Ørka H. O., Coops N. C., Hilker T., Bater C. W., Gobakken T. (2012). Lidar sampling for large-area forest characterization: A review. Remote Sensing of Environment, 121, 196–209.

[2] Véga, C., Durrieu, S., Morel, J., Allouis, T. (2012). A sequential iterative dual-filter for Lidar terrain modeling optimized for complex forested environments. Computers & Geosciences, 44, 31–41.

[3] Véga, C., Durrieu, S. (2011). Multi-level filtering segmentation to measure individual tree parameters based on Lidar data: Application to a mountainous forest with heterogeneous stands. International Journal of Applied Earth Observation and Geoinformation, 13(4), 646–656.

[4] Côté, J.-F., Fournier, R. A., Frazer, G. W., Niemann, K. O. (2012). A fine-scale architectural model of trees to enhance LiDAR-derived measurements of forest canopy structure. Agricultural and Forest Meteorology, 166–167, 72–85.

[5] Liu, Q., Li, Z., Chen, E., Pang, Y. (2010). Study on the Forest Mapping Method Using Airborne LIDAR. First Chinese Symposium on LIDAR Remote Sensing Applications, 204 – 208.

[6] TerraSolid. (2014). TerraScan User's Guide.

Comparison between volume of harvested logs and volume estimated by Airborne LiDAR analysis

Katsumasa OONO[1], Akihito SATO[2]

[1]*ASIA AIR SURVEY CO.,LTD. E-mail:kat.oono@ajiko.co.jp*
[2]*Forest Agency HOKKAIDO Regional Forest Office E-mail: Akihito_sato@rinya.maff.go.jp*

Highlights: Aim of this study is to investigate if significant relation exists between the volume of harvested log and the volume estimated by LiDAR data analysis. The result shows that such a relation exists and therefore LiDAR analysis is considered to have potential as a method to estimate log volume.

Key words: *Airborne LiDAR, Individual tree detection, DBH estimation, volume estimation*

Introduction

LiDAR survey has potential to replace field sampling survey to estimate harvested log volume. However, although stem volumes derived from LiDAR analysis were compared with calculated volume by field plot survey in Japan[1-3], no research was conducted to compare the LiDAR derived log volume and the actual volume of harvested log measured in logging yard.

In this study, log volume calculated by three methods was compared to verify the reliability of LiDAR method. The three methods are, LiDAR survey, sampling survey in the field and actual measurement of harvested log. Sampling survey in the field is a common practice in Japan to estimate log volume before harvest. Study area is planted Japanese larch in Hokkaido, northern Japan.

Target area and Data used

The target area of our study is located in eastern part of HOKKAIDO. The size of target area is 1,500ha. The dominant species is planted Japanese larch. LiDAR data for our analysis was acquired with 4 points/m^2 density on June 28th, 2012. 30 plots were prepared for calibration and making correlation to the equation of DBH estimation. Size of a calibration plot is 0.04ha. Further, sample field survey was carried out in more than 2% of area of every compartment, which is the base unit for forest management in Japan. This type of sample field survey is common in Japan to estimate total volume of harvested log. On the other hand, thinning harvest in around 180ha has been carried out for two years. Total log volume of compartment was also measured at a place where harvested logs were gathered and this log volume was used for validation.

Method

To estimate exact volume of Japanese larch, the target area was classified into two categories - larch forest and others - by using aerial photos and Laser Forest Type Image[4]. Laser Forest Type Image is generated using LiDAR intensity data, tree height data and crown texture data. The image shows different forest types by different colours.

Individual tree was detected by LiDAR analysis[5] in Japanese larch forest. The feature of this individual tree detection method is to use crown surface degree to detect tree tops instead of DCHM. Height of detected individual tree was measured from the LiDAR data. Crown polygons were drawn by Water-Shed algorithm. There is significant relation between crown size and breast height diameter (DBH) and correlation equation of estimating DBH is made by field measured DBH and crown size derived from LiDAR analysis. Three correlation equations are prepared using three kinds of crown size (1.Crown area, 2.Crown surface area, 3.Crown volume). The determination coefficients (R^2) of three equations are compared, and one equation with the highest R^2 is chosen for DBH estimation. Stem volume is calculated by volume equation which needs tree height and DBH as input data. Stem volumes derived from LiDAR analysis and log volume area compared. An Expected log volume are calculated by be multiplied with thinning ratio and merchantable ratio.

Results and discussion

First, the validation was carried out by comparing data obtained at 30plots. Number of stems, tree height, DBH and stem volume were calculated by both LiDAR survey and field plot survey. If the four variables measured by field plot survey are set as 100%, the results of LiDAR survey were as follows. Number of tree was

in the range of 80% to 120%, the average was 97% and RMSE was 1.1. The correlation coefficient was 0.92 and this result indicates that tree height measurement by LiDAR has the same level of accuracy with field plot survey.

Tree height was in the range of 90% to 110%, the average was 97% and RMSE was 1.1. The correlation coefficient was 0.92 and this result indicates that tree height measurement by LiDAR has the same level of accuracy with field survey.

Three types of correlation equation were prepared to compare DBH measured in the field with three kinds of crown sizes. Figure 1 shows relation between DBH measured in the field and crown size. Crown surface area (square maker) has highest R^2. In this study, therefore, crown surface area and following equation are used for DBH estimation.

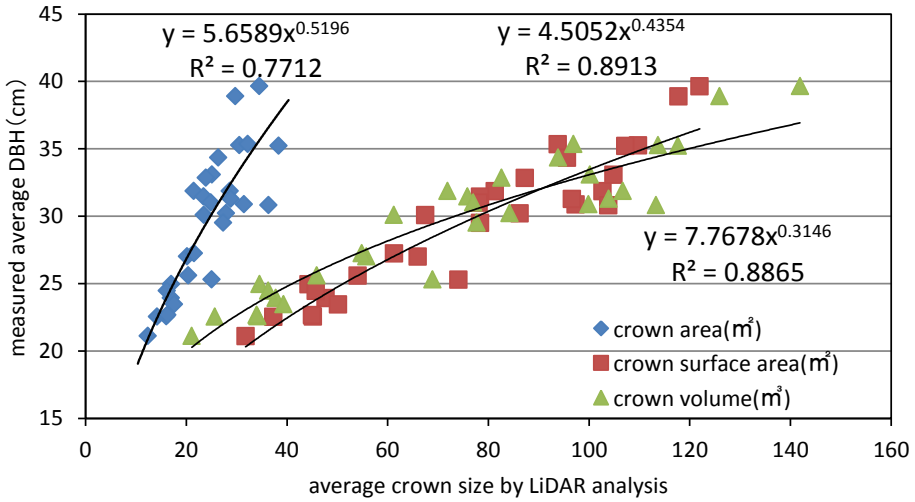


Figure 1: Relation between average crown size and average DBH measured in the field

$$Y = 4.5052X^{0.4354} \quad \text{..... 1}$$

Here, Y is estimated DBH (cm) and X is crown surface area (m^2).

The results of DBH estimation were in the range of 92% to 110%, the average was 100% and RMSE was 1.7. Considering that the average DBH of the plots is around 30 cm, DBH estimation by crown surface has the same level of accuracy with field survey.

Stand stem volume was in the range of 80% to 136%, the average was 101% and RMSE was 1.3. From RMSE result, the error was roughly 12%, because average volume of 30 plots is 10.6m^3.

Secondly, total log volume was compared with two types of volume estimation. One is sample field survey and another is wall-to-wall LiDAR analysis. The result is summarized on Table1.

Table 1: Comparison of log volume and expected volume

ID	name of compartment	compartment area (ha)	actual log volume (m^3)	sample field survey			LiDAR analysis		
				stand volume (m3)	expected log volume (m^3)	(%)	stand volume (m3)	expected log volume (m^3)	(%)
1	202TO	1.38	41.68	81.43	60.74	146	34.53	25.76	62
2	203HO	7.93	313.40	447.07	332.99	106	492.06	367.19	117
3	203HO1	3.16	51.23	100.88	75.42	147	41.43	30.98	60
4	203HO3	2.56	96.28	139.21	103.93	108	92.15	68.86	72
5	210I	36.15	902.80	1,235.98	925.87	103	1,264.78	945.46	105
6	210RO	3.48	108.38	159.78	119.36	110	128.07	95.72	88
7	210HA	43.24	1,360.08	1,400.76	1,042.32	77	1,976.80	1,477.73	109
8	210HA2	4.38	145.04	171.59	128.39	89	173.19	129.61	89
9	213RO	20.05	970.08	972.14	618.44	64	1,314.35	835.75	86
10	213RO1,2,3,4	159.36	6,770.52	10,297.44	6,558.39	97	10,256.51	6,526.85	96
Total		179.41	10,759.48	15,006.28	9,965.85	93	15,773.87	10,503.90	98

In forest of small compartment, the accuracy tends to be low, because it is difficult to match actual harvested area and calculation area of LiDAR volume. In forest of large compartment, the accuracy of volume estimation was in the range of 86% to 109%. The accuracy of volume estimation from typical sample survey sometimes becomes low (ex. ID 7), because the sample area might not be the area which has average volume of compartment.

Conclusion

In this study, it was confirmed that the analysis of individual tree detection has significant relation between LiDAR and field survey. Further, wall-to-wall LiDAR volume estimation method shows some potential to estimate total log volume as an alternative method to field sampling survey. However, in this study, only a few compartments were used for the comparison and target area was not large. Therefore, the authors plan to carry out additional case studies in various types of forests and number of harvested area.

References

[1] Tmoaki, T., Kazukio, Y., Yoshimichi, S., & Masashi, T., (2005). Predicting individual stem volumes of sugi (Cryptomeria japonica D. Don) plantations in mountainous areas using small-footprint airborne LiDAR, Journal of Forest Research, 305-312

[2] Yohei, N., Humiki, H., & Kenji, O., (2010). Estimation of coniferous standing tree volume using airborne LiDAR and passive optical remote sensing, Journal of Agricultural Meteorology, 111-116

[3] Keiko, I., Junichi, I., Takeshi, S., Yukihiro, M., & Katsunori, K., (2010) Estimating stand volume in broad-leaved forest using discrete-return LiDAR: plot-based approach, Landscape and Ecological Engineering, 29-36

[4] Katsumasa, O., Fumihiko, I., & Satoshi, T., (2012). Invention of a method to use airborne LiDAR intensity and DCHM data for forest type classification, Silvilaser2012, 158

[5] Katsumasa, O., Yoichi, N., & Atsushi, H., (2008) An improved method of individual tree detection using airborne LiDAR, Silvilaser2008, 508-516

Automatic Tree Breast Height Diameter Estimation from Laser Mobile Mapping Data in an Urban Context

Mónica Herrero-Huerta[1] and Roderik Lindenbergh[2]

[1]*Department of Cartographic and Land Engineering, University of Salamanca, Higher Polytechnic School of Avila, Hornos Caleros 50, 05003, Avila, Spain. Phone number: 0034-920353500-3820, monicaherrero@usal.es*
[2] *Department of Geoscience and Remote Sensing, Delft University of Technology, 2628 CN Delft, The Netherlands.*

Highlights: This study focuses on developing an innovative methodology to automatically estimate the diameter at breast height (DBH) of urban trees sampled by a Laser Mobile Mapping System (LMMS). The high-quality results confirm the feasibility of the proposed methodology, providing scalability to a comprehensive analysis of urban trees.

Keywords*: breast height diameter, LiDAR, urban trees, laser mobile mapping system, estimation.*

Introduction

Accurate measures of structural parameters of urban trees and monitoring their changes over time are essential to inventory of urban trees and growth models, modelling of carbon cycle and management systems of urban trees [1]. The estimation of these variables has traditionally been done with field inventories via pilot samplings. The advent of recent remote sensing technologies has opened up a new field of possibilities for carrying out such work through non-destructive methods, providing advantages regarding the economic costs involved, the time invested and estimation errors.

Among the various urban tree measurements, DBH (stem diameter at 1.3 m height) is an important tree inventory attribute because it serves as a fundamental parameter in tree allometry and estimation of basal area, thus providing valuable information about individual trees and tree stand structure. Many countries store the DBH of urban trees in cadaster databases for monitoring purposes .

The application of LMMS (Laser Mobile Mapping System) enables to fast and accurately capture 3D data of individual urban trees along the road. However, no research has been done on retrieving DBH using this technology. LMMS operates at the suitable work scale between manual and airborne LiDAR measurements. The latter fails to capture the complete vertical distribution of the canopy [2]. Terrestrial Laser Scanning has been used for estimating tree parameters[3], but being time-consuming compared with LMMS in an urban context. However, the method proposed has to face additional difficulties due to the specific acquisition geometry of the technology mentioned, such as the partial 3D data (only one side view), or having to deal with occluding vegetation, leading to underestimations compared to manually collected field data [4].

The aim of the present work is to develop an efficient and precise methodology in a novel way, to obtain DBH of urban trees destined for inventories based on point clouds coming from LMMS. Additionally, the influence of the density of points on the accuracy of deriving DBH estimations were investigated.

Materials and Methods

Laser Mobile Mapping System

For this work, data acquired by Fugro Drivemap is considered. This LMMS is composed of two high performance laser scanners type Riegl VQ250, a four-wheeled all-terrain vehicle and a navigation unit. The Global Positioning System has 10 Hz of positioning rate and the laser pulse rate is 1.333·106 pulses/s, being the maximum density 115000 points/m^2, the field of view 360^{o}x 26.8^o, the ranging accuracy less than 2cm and the maximum range 30 m.

Ground Inventory Data

The ground truth of the DBH was established using a measurement tape by a qualified operator at 5 centimeter precision. It should be remarked that the studied trees are less than 10 meters from de road.

DBH tree extraction Methodology

Before estimating the DBH, the extraction of individual trees of the point cloud is necessary through an automated workflow by voxel analysis (not the aim of this study). For the DBH tree extraction methodology the following steps are performed for each tree (Figure 1):

1). Outliers removal.

A statistical analysis on each point's neighbourhood is performed by assuming a Gaussian distribution. All points whose mean distance is outside an interval defined by the global mean distance and standard deviation are considered as outliers and removed from the dataset.

2). Extraction of trunk data.

The limit between the trunk and canopy is calculated by the curvature change of the histogram that represents the number of filled voxels of 0.25 cm^3 size as a function of height.

If the histogram shows there are ramifications in the trunk, k-means is used to assign a cluster (k clusters) to each data point ($\mathbf{x}_i$,i=1...n). K-means is a clustering method that aims to find the positions μ_i,i=1...k of the clusters that minimize the Euclidean distance from the data points to the cluster (Equation 1):

$$\arg min_c \sum_{i=1}^{k} \sum_{x \in c_i} d(x, \mu_i) = \arg min_c \sum_{i=1}^{k} \sum_{x \in c_i} ||x - \mu_i||_2^2 \quad (1)$$

where $\mathbf{c}_i$ is the set of points that belong to cluster i.

3). Determining the orientation of each trunk by Principal Component Analysis (PCA) and projecting the trunk points into a plane orthogonal to the axis corresponding to the principal direction.

A classical approach consists in performing a PCA of the 3D coordinates of the point cloud of the trunk. This statistical analysis uses the first and second moments of the point cloud, and results in three orthogonal vectors centered on the centroid of the point cloud. The PCA synthesizes the distribution of points along the three dimensions, and thus models the principal directions and magnitudes of variation of the point distribution around the center of gravity [6].

4). Fitting a circle to the projected trunk points by using RANSAC (RANdom SAmple Consensus) to mitigate the influence of outliers [7].

DBH was retrieved from laser datasets by circle fitting at different height bins. This can be done because of the assumption that the diameter is not significantly varying along a short length of the stem for mature trees.

Points that deviated most from a fitted circle were considered noise and removed for DBH estimations through the RANSAC method.

5). Quality control by obtaining the RMSE of the DBH estimated and measured at different height bin sizes.

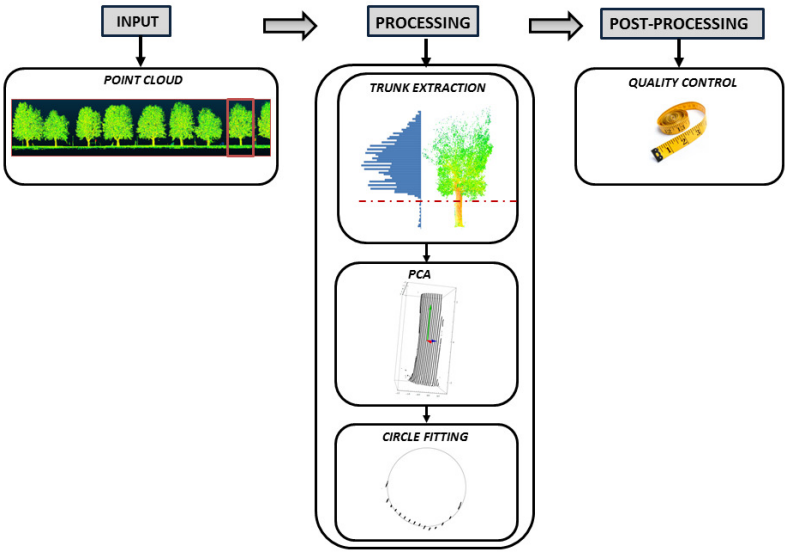


Figure 1. Workflow of the DBH Methodology

Experimental Results

The study area is located at the Delft University of Technology campus (lat. 52º00´ N; long. 4º22´ E; 1 m asl, The Netherlands) and tree species include Plantanus and Quercus. Data collection was performed on the 4th of November 2014 covering an area of approximately 750 x 1200 m on a day without wind. The dataset consists of 53,958,666 points, with the x, y and z coordinates and intensity. A pre-processing step of this dataset was carried

out in order to obtain the individual trees of the zone of interest. After that, the point cloud has 1,860,156 points with no intensity. There are 29 trees in total, 14 Platanus and 15 Quercus. The diameters of the first one are bigger than the others.

The accuracy of DBH estimation is influenced by several factors. Different height bins for determining the number of points to fit the circle, were tested to derive reliably DBH estimations. In addition, DBH estimation errors were confronted with the number of points to fit the circle but a relation was not seen.

Table 1 collects the robust statistical parameters of the linear regression stablished between the actual and the estimated DBH at different height bin sizes. The probability of success to achieve a right DBH defined by this parameters, is also included. From this analysis, it is possible to conclude that 20 cm height bin provides better results, from 1.20 m to 1.40 m height of each individual tree.

Table 1. Correlation coefficients and probability of success at different bin sizes

Circle Fitting Height Bin (m)	R^2	RMSE (cm)	Probability of Success (%)
1,25-1,35	0,792	5,67	62,1
1,20-1,40	0,837	4,35	79,3
1,10-1,50	0,688	6,45	75,9
1,00-1,60	0,600	12,12	86,2

Conclusions

This research presents a non-invasive method with application to urban tree inventory that consists of DBH estimation. The method is fast, reliable, robust and objective, and could serve to take decisions in advance regarding actions to be taken in urban trees and planning the management and maintenance optimally. This study also discussed the influence of the number of points to fit the circle and circle fitting height bin size on the influence of DBH estimation accuracy, validated against field measurements. Regarding these variables, the results show that there is no significant relation between the number of points and the DBH estimation errors. A 20 cm height bin provides better results, reaching a R^2 value of 0.84 and a RMSE of 4.35 cm with a probability of success of 79.3%. In addition, the results show that working only on one side of the tree, the visible side from the road, is still feasible for approximating the DBH with an accurate and precise fit.

This methodology can be extrapolated to a comprehensive study of urban trees.

It should be mentioned that the proposed methodology is part of a robust and efficient workflow which considers the automated, large scale extraction of tree sizes and locations sampled by a laser mobile mapping system. It includes tree location, level height and crown width throw a sensitivity analysis.

Acknowledgements

Authors would like to thank FUGRO GeoServices B.V. for providing the dataset. The result is partly funded by the IQmulus project (FP7-ICT-2011-318787).

References

[1] Næsset, Erik. "Practical large-scale forest stand inventory using a small-footprint airborne scanning laser." Scandinavian Journal of Forest Research 19.2 (2004): 164-179.

[2] Lim, K., Treitz, P., Wulder, M., St-Onge, B., & Flood, M. (2003). LiDAR remote sensing of forest structure. Progress in physical geography, 27(1), 88-106.

[3] Srinivasan, Shruthi, et al. "Terrestrial Laser Scanning as an Effective Tool to Retrieve Tree Level Height, Crown Width, and Stem Diameter." Remote Sensing 7.2 (2015): 1877-1896.

[4] Moskal, L. M., & Zheng, G. (2011). Retrieving forest inventory variables with terrestrial laser scanning (TLS) in urban heterogeneous forest. Remote Sensing, 4(1), 1-20.

[5] Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A k-means clustering algorithm. Applied statistics, 100-108.

[6] Van Goor, B., Lindenbergh, R. C., & Soudarissanane, S. S. (2011, August). Identifying corresponding segments from repeated scan data. In ISPRS Workshop Laser Scanning 2011, Calgary, Canada, 29-31 August 2011; IAPRS, XXXVIII (5/W12), 2011. International Society for Photogrammetry and Remote Sensing (ISPRS).

[7] Schnabel, Ruwen, Roland Wahl, and Reinhard Klein. "Efficient RANSAC for Point Cloud Shape Detection." Computer graphics forum. Vol. 26. No. 2. Blackwell Publishing Ltd, 2007.

Multi-wavelength Airborne Laser Scanning for Characterization of Tree Species

Eva Lindberg [i], Christian Briese [i, ii], Michael Doneus [iii, iv],
Markus Hollaus [i], Anke Schroiff [v], Norbert Pfeifer [i]

[i] Vienna University of Technology, Department of Geodesy and Geoinformation, Gußhausstraße 27 –29, A-1040 Vienna, Austria (Eva.Lindberg, Christian.Briese, Markus.Hollaus, Norbert.Pfeifer@geo.tuwien.ac.at)
ii EODC Earth Observation Data Centre for Water Resources Monitoring GmbH, c/o Vienna University of Technology, Gusshausstrasse 27-29, A-1040 Wien, Austria
iii Department of Prehistoric and Historical Archaeology, University of Vienna, Franz-Kleingasse 1, A-1190 Wien (michael.doneus@univie.ac.at)
iv LBI for Archaeological Prospection and Virtual Archaeology, Hohe Warte 38, A-1190 Wien
v YGGDRASIL Diemer, Dudenstraße 38, D-10965 Berlin, Germany (ankeschroiff@yahoo.de)

Highlights: This study analyses ALS data with three different wavelengths (1064 nm, 1550 nm and 532 nm) from a forest area to investigate the potential for tree species classification. Including both spectral and geometric information from ALS data may characterize tree species better than geometric information from one wavelength only.

Key words: *LiDAR, reflectance, vegetation, forestry, NIR, NDVI.*

Introduction

Data from airborne laser scanning (ALS) are being used for classification of tree species in forests [1,2]. From interpretation of aerial images, it is well known that different tree species reflect light at different wavelengths [3]. A multi-wavelength laser scanner would provide spectral as well as geometric information about the vegetation, enabling detailed analysis of the backscattering object surface. Hence, classification of tree species is one potential application of multi-wavelength laser scanning. A combination of green and NIR/SWIR wavelengths is likely to be useful for vegetation analysis since the reflectance of plants is highest around 550 nm and in the NIR/SWIR spectrum due to their chlorophyll content. The Normalized Difference Vegetation Index (NDVI) derived from the visible red and NIR spectral bands is known to yield high values for vegetated areas [3]. Vegetation indexes could be calculated in a similar way from laser scanning data with multiple wavelengths.

An operational multi-wavelength ALS system is now available from Optech with near-infrared laser light (NIR; 1064 nm), short-wavelength infrared laser light (SWIR; 1550 nm) and a visible green wavelength (532 nm) [4]. The new system will most likely be important for land cover mapping, including vegetation mapping. However, to perform optimal classification, the laser data should be radiometrically calibrated to eliminate the effect of sensor, flight and individual echo dependent acquisition geometry and present atmospheric conditions. The process of radiometric calibration can be applied to discrete as well as full waveform ALS data if the output power is known or can be assumed to be constant. Full waveform ALS data has the additional advantage that the echo width of the received signal can be derived.

This study describes characterization of tree species from radiometrically calibrated ALS data with three wavelengths: 1064 nm, 1550 nm and 532 nm [5]. The ALS data are collected with three separate scanners and two separate flights but the aim is to get close to multi-wavelength laser scanning.

Materials

Study area

The study area is part of the LBI ArchPro case study area Carnuntum [6] and is located in deciduous riparian forest along the Danube in the northeast of Austria (Lat. 48 ° 07’ N, Long. 16 ° 52’ E). The main tree species are European ash (*Fraxinus excelsior*) and hybrid poplar (*Populus x canadensis*). Other tree species include box elder (*Acer negundo*), bird cherry (*Prunus padus*), and English oak (*Quercus robur*).

ALS data

ALS data were collected with separate instruments with different wavelength as described in Briese et al. [5]. On May 24th, 2013, data were collected with the ALS instruments RIEGL VQ-820-G (laser wavelength: 532 nm) and RIEGL VQ-580 (1064 nm). On May 28th, 2013, data were collected with the ALS instruments RIEGL VQ-820-G (laser wavelength: 532 nm) and RIEGL VQ-480i (1550 nm). The point density (last echo) of all data sets was more than 20 points per m². The ALS data covered an area of approximately 800 ha. All initial signal

processing (i.e., full-waveform decomposition, georeferencing, strip adjustment, and radiometric calibration) is described in Briese et al. [5].

Field data

An initial field visit was done on September 26th, 2013. Seven homogenous forest stands with one dominant tree species were identified and their positions were measured with a hand-held GPS. A part of the laser-scanned area was selected for a more detailed field inventory based on the presence of homogenous forest stands with at least 50-100 m extent. The area selected for detailed inventory was approximately 28 ha and located on flood-plain land. Three transects were placed in the selected area to cover a variety of tree species, tree height, and stem density. On November 7th, 2013, 22 field plots with 10 m radius were placed with 100 m distance along the transects. For each field plot, the number of trees of each species was recorded. The proportion of each tree species was calculated in relation to the total number of trees. The positions of the center of the field plots were measured using a Differential Global Positioning System (DGPS).

Features of ALS data relative to tree species

The following raster layers were derived from the ALS data (Figure 1).

- Normalized digital surface model (nDSM) with 1×1 m raster cells from RIEGL VQ-580 and RIEGL VQ-480i
- Maximum reflectance of first return (r) with 0.5×0.5 m raster cells from all sensors. The values derived from the RIEGL VQ-580 on May 24th were called r_{580} and the values derived from the RIEGL VQ-480i and RIEGL VQ-820-G on May 28th were called $_{r480i}$ and $r_{820\text{-}G}$ respectively.

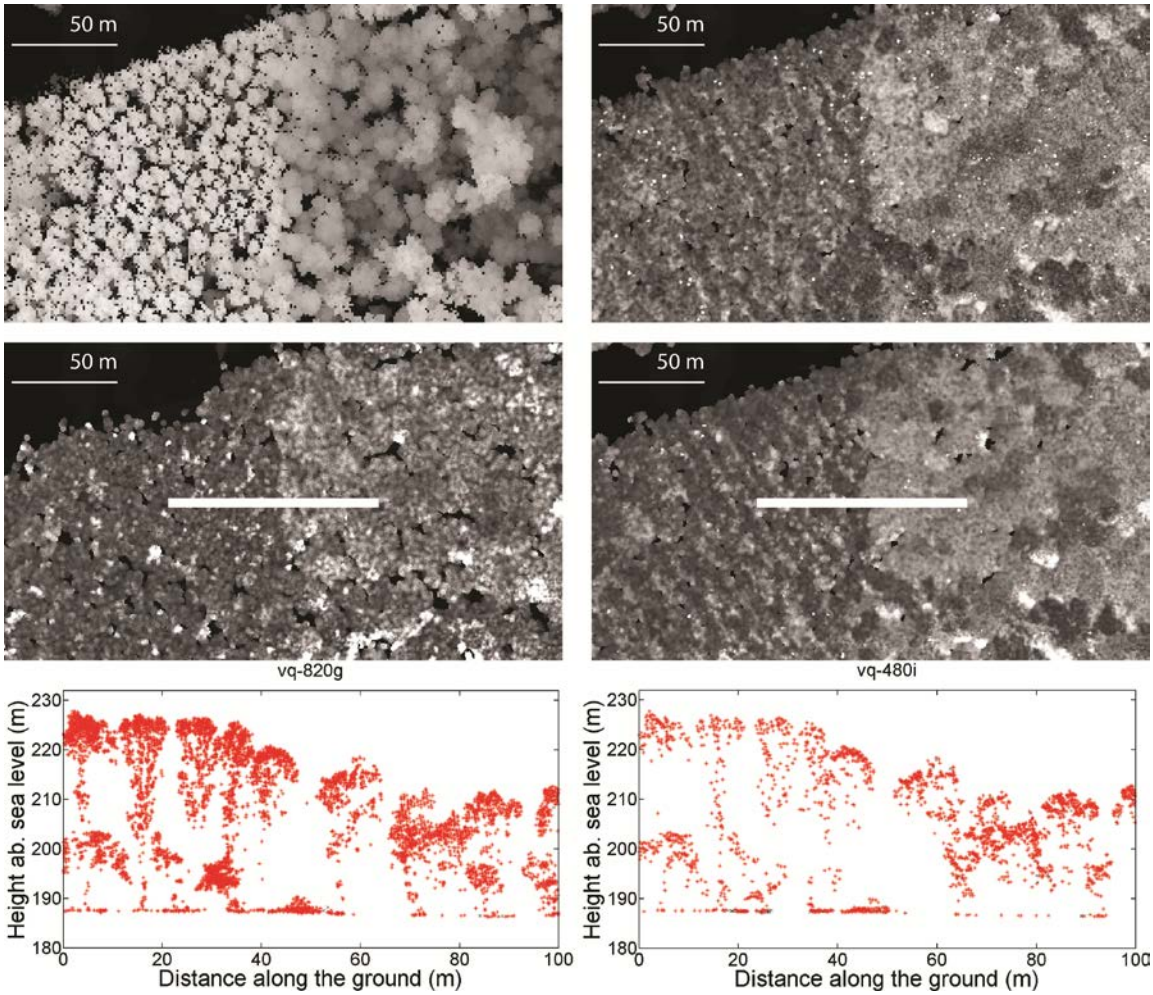


Figure 1: A selected area with a stand of poplar to the left and a stand of European ash to the right. The subfigures are the nDSM derived from the RIEGL VQ-580 data (top left; values ranging from 0 to 45 m), the reflectance from the RIEGL VQ-580 data (top right; values ranging from 0 to 0.5), the reflectance from the RIEGL VQ-820-G data on April 28th (middle left; values ranging from 0 to 0.04), the reflectance from the RIEGL VQ-480i data (middle right; values ranging from 0 to 0.25), a transect from the RIEGL VQ-820-G data (bottom left) and a transect from the RIEGL VQ- 480i data (bottom right). Each transect is 5 m wide and 100 m long and is shown as a white line in the corresponding raster.

Characterization of tree species

The main tree species in the area can be distinguished from the reflectance of all three wavelengths. The point density was slightly higher for the RIEGL VQ-820-G data than for the other two wavelengths, but the proportion of returns at different height above the ground appears to be similar for all wavelengths. Also the level of the detail of the tree crowns appears similar, and delineation of tree crowns or groups of trees could be feasible from all wavelengths.

The RIEGL VQ-820-G scanner has higher sensitivity than the two other scanners since it is primarily intended for bathymetric mapping. It also has a twice as big angular divergence and hence footprint. This means that one pulse can intersect a greater number of surfaces and that the sensor will record many of these surfaces as echoes, resulting in a greater number of echoes for each emitted pulse. Although some of the surfaces are not in the middle of a footprint, the return signal will still be stored at the position corresponding to the middle, resulting in less accurate positioning and more noise. It is worth noting, that the small openings in the canopy visible in the reflectance maps do not appear equal for all datasets. Not only point density, but also the footprint (diameter and energy distribution) may be the cause of that. Also the new Optech multi-wavelength ALS system has a bigger angular divergence for the green wavelength, so similar results can be expected for that.

Based on the field data, we will derive the reflectance and additional variables for tree crowns of the different species. By summarizing the values over the different tree species, a signature can be calculated for each species. We will compare the signatures of the different species to derive features that can be used for tree species classification.

Acknowledgements

This study was funded partly by the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (FORMAS), partly by the European Community's Seventh Framework Programme (FP7/2007–2013) under grant agreement No. 606971, the Advanced_SAR project, and partly by the EU-Programme Marie Curie Industry-Academia Partnerships and Pathways (EU-FP7-IAPP) under grant agreement No. 251234, the ChangeHabitats2 project. We would like to thank the Ludwig Boltzmann Institute for Archaeological Prospection and Virtual Archaeology (archpro.lbg.ac.at) for the support in data provision, especially Dr. Geert Verhoeven and Martin Wieser. Furthermore, we would like to thank Riegl Laser Measurement Systems GmbH for the acquisition of the laser data.

References

[1] Lindberg, E., Eysn, L., Hollaus, M., Holmgren, J., & Pfeifer, N. (2014). Delineation of tree crowns and tree species classification from full-waveform airborne laser scanning data using 3-d ellipsoidal clustering. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 7(7), 3174-3181.

[2] Hollaus, M., Mücke, W., Höfle, B., Dorigo, W., Pfeifer, N., Wagner, W., Bauerhansl, C., & Regner, B. Tree species classification based on full-waveform airborne laser scanning data. In Proceedings of SilviLaser 2009, College Station, Texas, USA, 14–16 October, 2009; pp 54-62.

[3] Lillesand, T., Kiefer, R.W., & Chipman, J. (2007). Remote sensing and image interpretation. 6th edition ed.; New York: John Wiley & Sons, Inc.

[4] Optech Titan. http://www.optech.com/index.php/product/titan/

[5] Briese, C., Pfennigbauer, M., Ullrich, A., & Doneus, M. Multi-wavelength airborne laser scanning for archaeological prospection. In XXIV International CIPA Symposium, Strasbourg, France, 2-6 September, 2013.

[6] Trinks, I., Neubauer, W., & Doneus, M. (2012). Prospecting archaeological landscapes. In Progress in cultural heritage preservation, Ioannides, M.; Fritsch, D.; Leissner, J.; Davies, R.; Remondino, F.; Caffo, R., Eds. Springer Berlin Heidelberg: Vol. 7616, pp 21-29.

Challenges in measuring forest attributes from Araucaria trees with ALS data

João P. Pereira[1], Marcos B. Schimalski[2], Veraldo Liesenberg[2]

[1]*FeLis, Albert-Ludwigs University Freiburg, Germany, joao.pereira@felis.uni-freiburg.de*
[2]*FARS, State University of Santa Catarina, Lages, Brazil*

Highlights: Forest attributes were retrieved from *Araucaria angustifolia* trees using ALS LiDAR (Airborne Laser Scanning – Light Detection and Ranging) data. Thiessen Polygon method was used as an alternative to region growing methodology in tree detection. Diameter at breast height (DBH), crown diameter and height were this procedure's results.

Key words: *laser scanner, Thiessen polygon, natural forest, araucaria, region growing*

Introduction

In natural forests it is a challenge to measure irregular crown trees even in open field. From *Araucariaceae*, *Araucaria angustifolia* is one example of irregular crown trees. *A. angustifolia* is classified as an endangered species. Therefore, there is the necessity for methodologies to accurately measure these trees. These data can be an important resource to collaborate with a management project for araucaria's conservation. This paper aimed to measure variables from *Araucaria angustifolia* trees using LiDAR data employing a new methodology called Thiessen Polygon method. Developed by [1], this methodology came to serve as an alternative to region growing method. As a hypothesis, this methodology allied to ALS LiDAR data can be a possible alternative for measuring trees easily.

Material and methods

The study area is located in the city of Painel, Santa Catarina state, Brazil. Inserted in the Araucaria Forest, this area contains a large occurrence of *Araucaria angustifolia* (araucaria). Due to the species' morphology even in open field is hard to measure these trees accurately. Measurements were taken from isolated trees of *A. angustifolia* and compared to the ones obtained using ALS data. To avoid field measurement errors only isolated trees were included in the experiment. In total 4 subareas were selected containing only isolated *araucarias*.

Thiessen polygon method is based on the same principle as region growing for tree delineation. First calculates the local maxima (seeds) using ArcGIS. With the seeds is possible to obtain the Thiessen polygons. Empty spaces between the trees can be removed using a binary gap mask. At the end of this process is possible to obtain total height and crown diameter. DBH is calculated using crown diameter in linear regression techniques.

Tree delineation was assessed through hit rate, omission and commission indexes. Field and ALS measurements were compared as an assessment form as well.

Results and discussion

Table 1 shows similarity between field data and ALS measurements and ranges from 75 to 92%. Except for the third area, the hit rate value was low due to a higher number of trees per hectare. The tree crown form might also be a contributing factor for the lower hit rate, especially in more dense forest areas.

Table 1: Assessment of tree measurements and delineation.

Assessment		area 1	area 2	area 3	area 4
Trees delineation	Omission rate (%)	9,7	8,7	24,8	8,3
	Commission rate (%)	4,8	7,2	5,7	8,3
	Hit rate (%)	90,3	91,3	75,2	91,7
Field measurements	DBH (cm)	$32{,}63_a$	$31{,}12_a$	$34{,}40_a$	$32{,}54_a$
	Total height (m)	$10{,}56_b$	$10{,}25_b$	$11{,}15_b$	$10{,}48_b$
	Crown diameter (m)	$8{,}50_c$	$8{,}20_c$	$9{,}40_c$	$8{,}75_c$
ALS measurements	DBH (cm)	$32{,}41_a$	$31{,}51_a$	$33{,}10_a$	$32{,}14_a$
	Total height (m)	$10{,}12_b$	$10{,}05_b$	$10{,}92_b$	$10{,}41_b$
	Crown diameter (m)	$8{,}02_c$	$8{,}10_c$	$9{,}05_c$	$8{,}82_c$

Equal letters in the same column do not differ statistically in Dunnet's test at 0,10 of reliability.

DBH difference between field and ALS measurements was from 0,22cm to 1,3cm. Total height difference varied from 7cm to 23cm. Crown difference varied from 0,10m to 0,48m. It is possible to say the difference between measurements presented a good behaviour, showing the potential of the ALS technology applied to forestry.

In general, the magnitude of hit rate accuracy corresponds well with previous experiments conducted by other authors. The same methodology was used in Barcelonnete, France and reached a hit rate of 96% [1]. With other methodologies, other authors reached a hit rate of 40% [2], 61% [3] and 86,7% [4].

Figure 1 shows the delineation results. Good results are presented in Figure 1a, 1b and 1d. However, Figure 1c, the blue arrow shows one misidentification. In this particular case, there is a group of close trees overlapping their crowns causing therefore trees misidentification. The green arrow shows a dense corridor of trees. This occurrence can also be a contributing factor to the hit rate decrease.

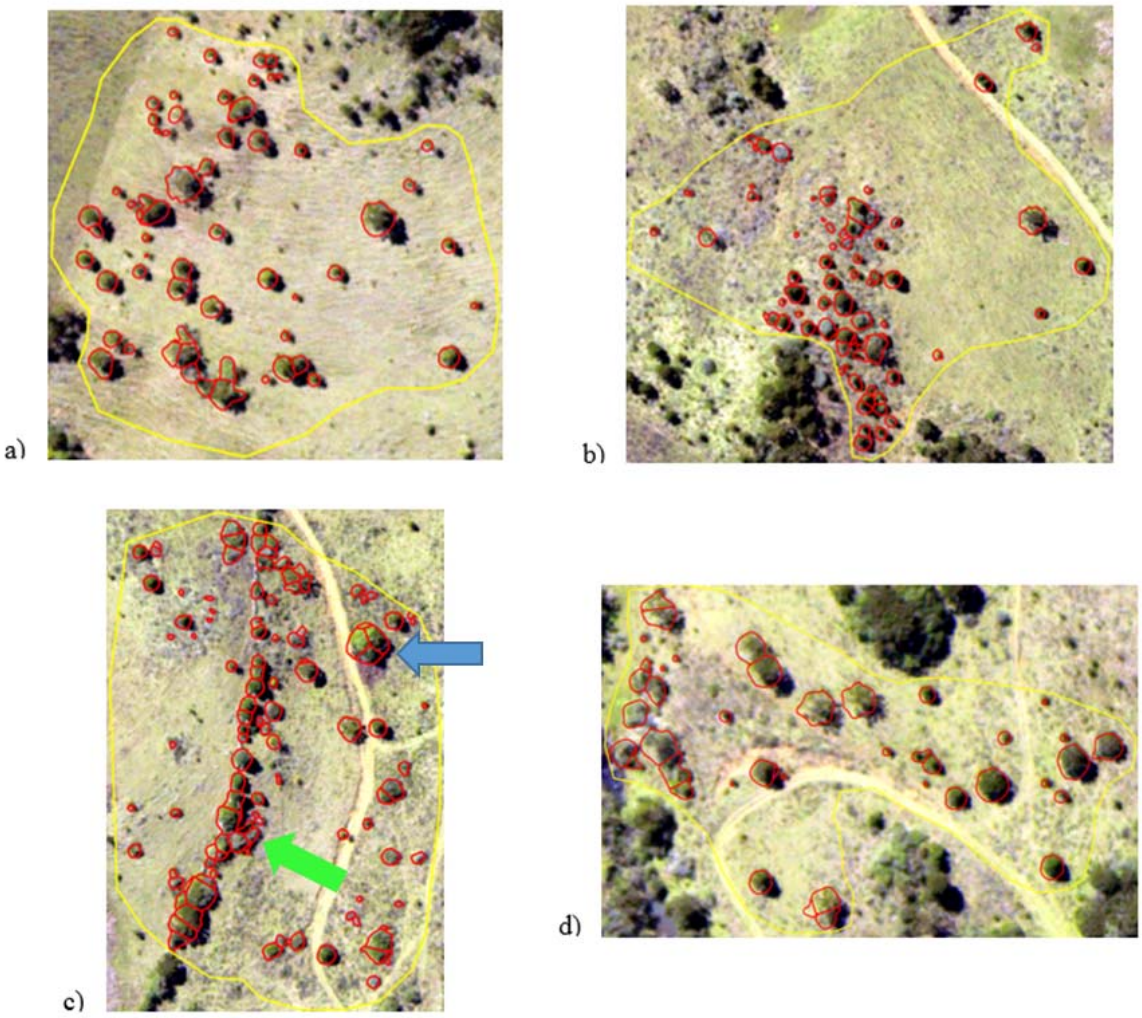


Figure 1: Isolated trees delineation. a): area 1; b): area 2; c): area 3; d): area 4.

Conclusion

DBH and total height measurements are possible to be obtained with ALS for Araucaria angustifolia trees. Thiessen Polygon methodology showed good results when compared to other authors. These results show potential to use both ALS and Thiessen Polygon methodology for further forest management practices in Mixed Ombrophilous forest in Southern Brazil.

Acknowledgements

We are grateful to the Aeroimagem S/A company from Curitiba, Brazil for the ALS LiDAR dataset. We are also grateful to CAPES (Coordination for the Improvement of Higher Level Personnel) and CNPq (National Council for Scientific and Technological Development) for providing research support.

References

[1] Kumar, V. (2012). Forest inventory parameters and carbon mapping from airborne LiDAR. University of Twente, 1-104.

[2] Kwak, D. A., Lee, W. K., Lee, J. H., Biging, G. S., Gong, P. (2007). Detection of individual trees and estimation of tree height using LiDAR data. Journal of Forest, 12, 425-434.

[3] Pitkänen, J., Maltamo, M., Hyypä, J., Yu, X. (2004). Adaptive method for individual tree detection on airborne laser based canopy height model. Remote Sensing and Spatial Information Sciences**,** 36, 187-191.

[4] Reitberger, J., Heurich, M., Krzystek, P., Stilla, U. (2007). Single tree detection in forest areas with high-density LiDAR data. Remote Sensing and Spatial Information System**,** 36, 139-144.

Scikit-image for trees local maxima detection

João P. Pereira[1], Holger Weinacker[1], Barbara Koch[1], Marcos B. Schimalski[2]

[1]*FeLis, Albert-Ludwigs University Freiburg, Germany, joao.pereira@felis.uni-freiburg.de*
[2]*FARS, State University of Santa Catarina, Lages, Brazil*

Highlights: Tree top detection using laser scanner data is very important when tree modeling and measurement are concerned. However, software used for this purpose are expensive and scarce. This project aimed to assess Scikit-image for tree local maxima detection. Scikit-image is a Python based package using Numpy and Scipy.

Key words: *local maxima, LiDAR, scikit-image, python.*

Introduction

Due to the diffusion of LiDAR (Light Detection and Ranging) technology, new methodologies are necessary to support this advance. There are software available with the capacity to detect local maxima in forest areas. Although, some are expensive and others limited in their options. Scikit-image ([1]) is a Python based algorithms package for image processing created from the association between Numpy and Scipy packages. With it, the user is not limited to what a software can offer and it is possible to create algorithms based on each necessity. This paper aimed to assess the applicability of Scikit-image for trees local maxima detection. As a hypothesis, we believe with the right configuration Scikit-image can be a valuable tool to be used in forestry allied with LiDAR data. However, this is the first time Scikit-image is been tested for forest applications in local maxima detection context. There is the necessity for more tests to assess Scikit-image's performance and robustness.

Material and Methods

The study area is located in the city of Painel, Santa Catarina state, Brazil. It is inserted in the Araucaria forest with the presence of high density forest fragments and open field. Tree height was the variable measured on the field to compare with the local maxima automatically calculated. LiDAR data had an average point density of 7 point/m². Local maxima detection was calculated in two different forms, through TreesVis and Scikit-image local maxima detection algorithms. Both methods used the digital surface model (DSM) for treetop detection. DSM was calculated in TreesVis with a geometric resolution of 0,5m. In Scikit-image, the local maxima detection used several python packages including Numpy, Scipy, GDAL and Matplotlib. Total height measured in the field was compared with the treetops calculated in TreesVis and Scikit-image to assess the method used. The reason to use TreesVis in this project is to compare the results of Scikit-image with an already established algorithm for local maxima detection.

Results and discussion

Table 1 shows there are enough statistical evidences to state that field tree height, TreesVis and Scikit-image local maximas do not differ. Both TreesVis and Scikit-image showed a lower tree height average. That can be explained due to the filtering necessary to smooth the DSM before treetop detection. Nevertheless, difference stood between 0,09 and 0,17m for TreesVis and Scikit-image, respectively.

Table 1: Average local maxima for each method

	Total Height (m)	Number of trees detected	Commission	Omission	Hit rate
Field (F)	11,70a	66	-	-	100%
TressVis (T)	11,61a	60	0,0%	9,1%	90,9%
Scikit-Image (S)	11,53a	66	7,6%	7,6%	92,4%
F - T	0,09	6	-	-	-
F - S	0,17	0	-	-	-

Average height values followed by the same letter in the line do not differ statistically according to Dunnet's test at 90% of reliability

Evaluating the hit rate, both TreesVis and Scikit-image showed good results. TreesVis had no commission error and 9,1% of omission. Scikit-image was unable to detect 6 trees, resulting in an omission of 7,6%.

Omission in both algorithms was result of low density LiDAR data, which is a critical factor in DSM calculation and posterior treetop detection. Commission for Scikit-image also resulted in 7,6% due to the presence of small bushes and the configuration used in Scikit-image's algorithm. Further studies may reduce this value when a better configuration is found. Hit detection reached 90,9 and 92,4% for TreesVis and Scikit-image, respectively. Increasing the density of LiDAR data could improve this results, achieving a hit rate close do 99%.

Consulting other references, [2] achieved a hit rate of 96% using Thiessen Polygon method, [3] reached a 61% hit rate, [4] detected 70% of the trees correctly and [5] located 87,3% of the trees in the analysed area. Thus, is possible to say Scikit-image and TreesVis reached a satisfactory and coherent result compared to the literature.

Figure 1 shows the performance of both algorithms tested. Figure 1a shows the result of Scikit-image processing. Blue arrows indicate errors of commission, caused by the algorithms configuration and very low vegetation occurrence. Yellow arrow shows one of the omission cases caused by the close proximity from small trees to big trees. Besides that, Scikit-image algorithm for local maxima detection proved efficient. All high trees were correctly detected, showing the potential of this package. Further tests should be made to assess the behaviour of this package with other forest formations. Scikit-image, due to its flexibility, can also be implemented in different software solutions. TreesVis, for example, could implement this algorithm along with the already existing one.

Green arrows in Figure 1b emphasize the omission errors with TreesVis. Although, the omission's reason with TreesVis is important to quote. All misdetections happened with low trees and/or trees close to each other. This behaviour is common in low density LiDAR data, causing the omission due to lack of information from tree crowns in the DSM model.

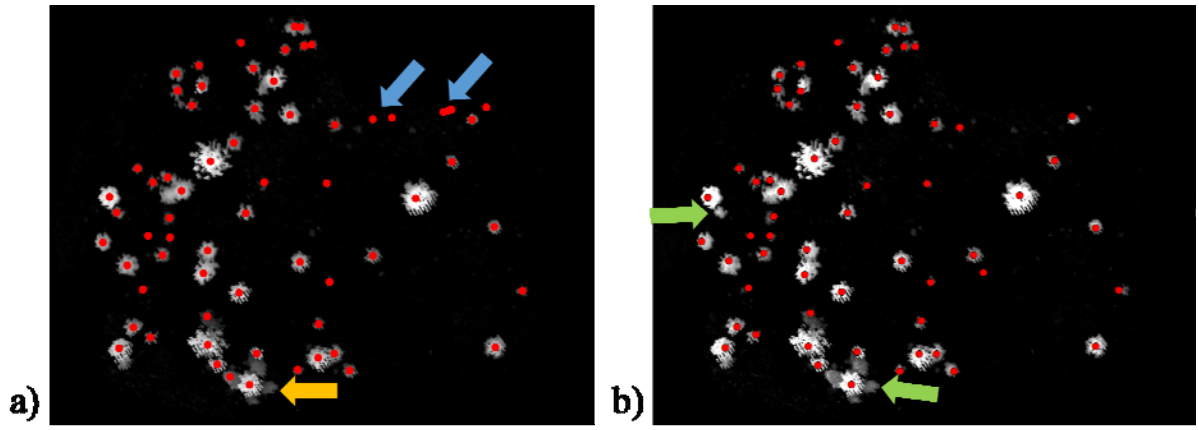


Figure 1: Local maxima detection. a) local maxima calculated with Scikit-image; b) local maxima calculated with TreesVis.

Conclusion

Scikit-image is a promising tool for local maxima detection using LiDAR data. It presented good behaviour and has shown good results under the appropriated configuration. Although, further studies are necessary to improve the results and to adapt the algorithm to different kinds of forests (i.e. coniferous and dense coniferous forests; broad leafs dense forests).

Acknowledgements

We are grateful to the Aeroimagem S/A company from Curitiba, Brazil for sharing the ALS LiDAR data, which without it, to perform this work would not be possible. We are also grateful to the National Council for Scientific and Technological Development (CNPq) and the Science without Borders program for providing the necessary resources to complete this project.

References

[1] Van der Walt , S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, F. D., Yager, N., Gouillart, E., Yu, T. (2014). Scikit-image: image processing in Python. PeerJ, 1-18.

[2] Kumar, V. (2012). Forest inventory parameters and carbon mapping from airborne LiDAR. University of Twente, 1-104.

[3] Reitberger, J., Heurich, M., Krzystek, P., Stilla, U. (2007). Single tree detection in forest areas with high-density LiDAR data. Remote Sensing and Spatial Information System**,** 36, 139-144.

[4] Rahman, M. Z. A., Gorte, B. G.H. (2009). Tree crown delineation from high resolution airborne LiDAR based on densities of high points. ISPRS Workshop Laser Scanning, 28, 123-128.

[5] Koch, B., Heyder, U., Weinacker, H. (2006). Detection of individual tree crowns in airborne LiDAR data. Photogrammetric Engineering and Remote Sensing, 74, 357-363.

Forest change detection using airborne waveform lidar data

Yong Pang, Zengyuan Li, Luxia Liu, Hao Lu, Chen Bowei, Qingwang Liu

Institute of Forest Resource Information Technique, Chinese Academy of Forestry, Beijing 100091, China; Tel: 86-10-62888847 Fax: 86-10-62888315, E-mail: pangy@ifrit.ac.cn

Highlights: We used multi-temporal ALS data to monitor different categories forest changes, which including forest growth and logging changes, understory loss during forest tending, and forest thinning intensity.

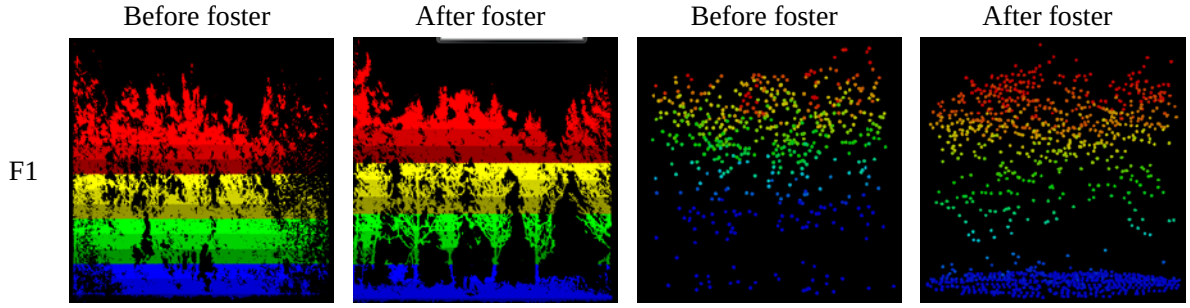
Key words: *airborne waveform lidar, forest change, growth, logging, tending, thinning*

Introduction

As a dynamic ecosystem, forests are always in changing status. These happens during forest ecosystem succession and forest management activities. Some changes are not apparent through direct observation while some are observable. Lidar provides an efficient way to characterize structure related parameters which are usually related to forest change phenomena.

Content

We designed experiments to detect forest changes using multi-temporal waveform lidar datasets. In Yuxi of Yunnan province, we detected forest growth and logging changes through airborne lidar data from 2007 and 2014. In Puer of Yunnan province, the understory loss during forest tending was quantified and mapped through combination of TLS and ALS data collected before and after tending activity. In Xiuning of Anhui province, forest thinning intensity was estimated using lidar data collected before and after thinning management event. Lidar metrics variables, foliage vertical profile, and estimated volume density were used to characterize these changes quantatively. The performance of these test sites shows good potentials for change detection especially caused by forest management activities.



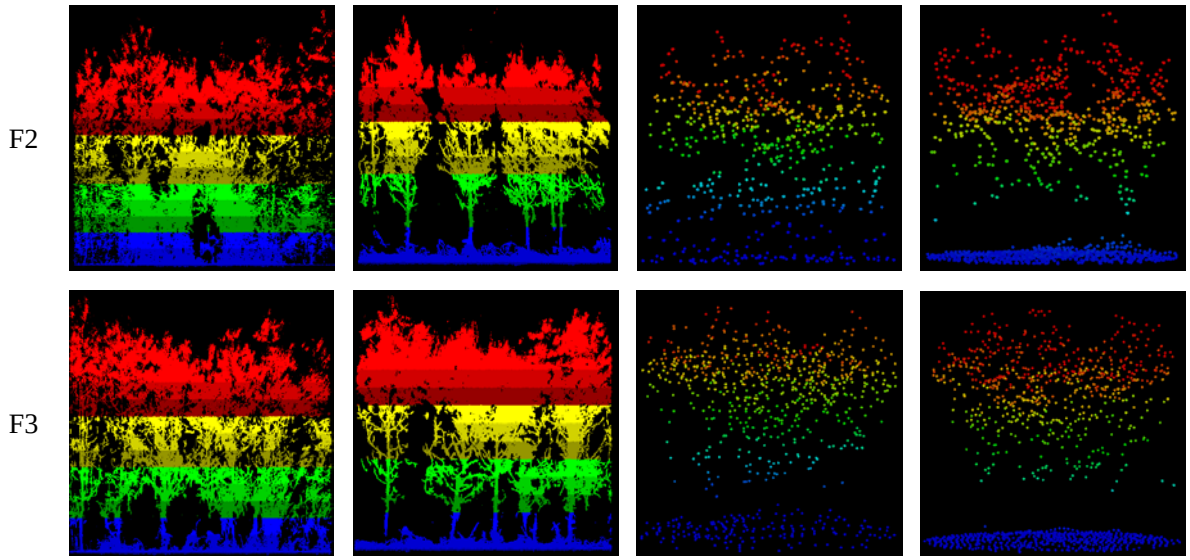


Fig.1 The comparison of TLS and ALS point clouds before and after forest tending

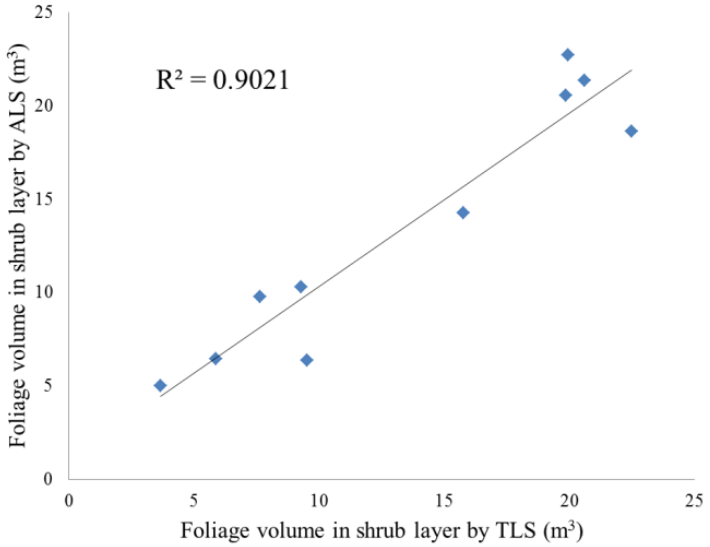


Fig.2 The regression result between ALS and TLS estimated volume of shrub layer

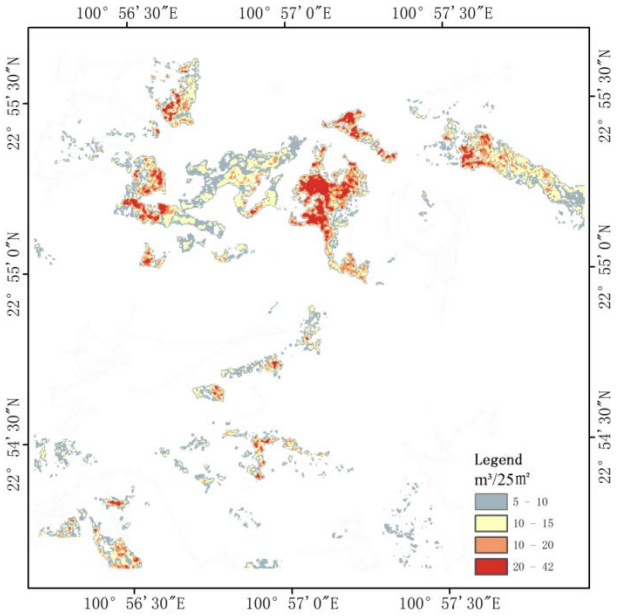


Fig.3 The change picture of the foliage elements volume in shrub layer

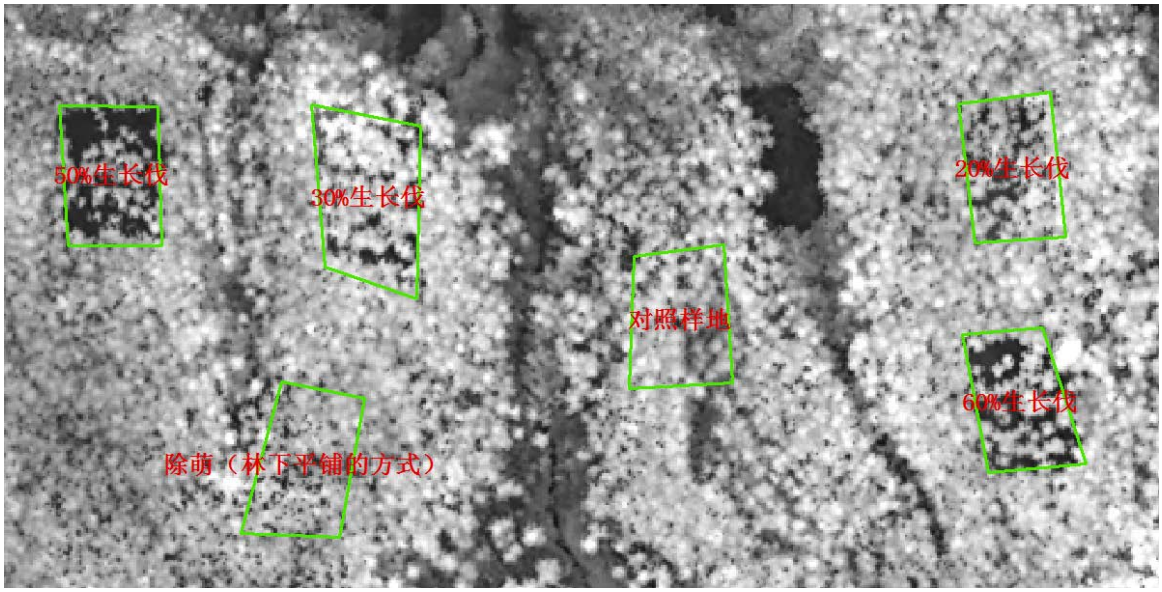


Fig.4 CHM comparison of different forest thinning intensity plots

Assessment of models validity domain for predictions of dendrometrical parameters from lidar data.

Bock J.[1], Piboule A.[2], Munoz A.[1]

1 : ONF, Pôle RDI de Chambéry, 42 quai Charles Roissard, F-73023 Chambéry.
2 : ONF, Pôle RDI de Nancy, 11 rue Ile de Corse, F-54000 Nancy

Highlights: How could we know if maps predictions have been obtained from models operating in their calibration range or in an extrapolation mode? The study proposes a method defining the validity domain of a model using boundaries of the calibration metrics and distances of the pixels to be predicted from the metrics used in the model calibration stage.

Key words: *validity domain, Lidar, model selection, uncertainty maping.*

Introduction

Mapping forest parameters from LIDAR data, involves a key role played by the model selection process. Many studies [1] have based this selection on an overall low mean square error (RMSE) of the model. AIC and BIC criteria are also frequently used since they further put a negative weight on complex models. Once selected, models are applied to predict forest parameters for every pixels of the area covered by the LIDAR data.

Since forest managers need reliable maps and considering that model accuracies are more reliable within their calibration ranges, we propose a method depicting the frontier of the calibration range.

The use of a convex hull has been proposed by [2]. In this study, we evaluate methods for characterizing the calibration range of a model used to predict basal area (BA) from LIDAR data. The approach compares convex hulls, alpha shapes, and distances calculations (including prediction residuals and Mahalanobis distance) which highlight the distance between the LIDAR metrics of the calibration and the prediction samples. This multi-criteria approach is used to assess model robustness, and estimate limitations in map usage when models operate outside of their calibration range or when prediction pixels are too far from calibration plots.

Material et methods

The study site is located in Vercors (France – Dept 38) where two LIDAR survey have been produced in September 2010 and August 2011, by the same operator, using the same acquisition parameters: mean pulse emission of 12 pts/m², obtained with a RIEGL LMS-Q560 at a flight altitude of 600 m. The area is covered mainly with uneven stands composed of Fir (*Abies alba*), Spruce (*Picea abies*) and Beech (*Fagus sylvatica*). Some coppice of Beech are still present, but cover less than 10% of the area. Trees were measured between 2011 and 2013, *i.e.* 1 or 3 growing seasons after the LIDAR survey. The diameter at breast height (DBH) has been measured for all trees starting at 17.5 cm diameter. Plots that were established only in the public forests, were georeferenced using GNSS receiver with differential correction. The dependent variable BA (m²/ha) was calculated by summing all tree section areas, divided by the plot area.

Explanatory variables were LIDAR metrics calculated from 3D points cloud normalized with a TIN (Triangulated Irregular Network) of ground points. Nine metrics families have been considered: height percentiles, distribution statistics, height differences to mean height, cumulative point density within fixed or variable height slices, means of local height maxima, canopy volume and penetration rate within slices. Some filters have been apply to select all, first or lasts pulses, ground pulses, or pulses above 2 or 1 m height.

The most accurate models have been selected after testing all combinations of 4 explanatory variables within additive models. Three models were selected according to multi-criteria optimization: precision (*i.e.* model with the lowest RMSE obtained by cross-validation), with low correlation between predictors (VIF <5) and low effect of dendrometrical, topographical and acquisition parameters. At least, 2 methods to define models validity domain have been tested.

The first method is based upon the definition of the outer boundaries of the metrics used for model calibration. As the retained models contained 4 predictors, they were first summarized by the first axes of a Principal Component Analysis (PCA). Three methods of boundaries adjustment have been tested: a convex hull [2,3], and two alpha-shapes (with alpha values of 0.5 and 1). Once these boundaries adjusted, it was then possible to detect pixels either inside or outside these calibration boundaries.

The goal of the second method is to calculate, in the predictor space (*i.e.* 4 dimensions) of each model, the Mahalanobis distance between each point of the target sample and the reference one (*i.e.* the calibration plots).

The Mahalanobis distance was preferred to the Euclidian one, since it takes into account the correlation between LIDAR metrics. The Mahalanobis distance was obtained using the yaImpute package (R-Core Team).

Thanks to the geographical coordinates carried by each predicted plots, uncertainty maps were produced. From these maps, spatial analysis were performed to evaluate validity domains and Mahalanobis distances, and to investigate their relation with forest types recorded by the National Forest Inventory (NFI) as well as stand properties (public or private forests).

Results

Three models were retained among all tested models, having low correlation between explanatory variables (VIF < 5). Error rates of these 3 models were low: coefficient of variation ranging between 17 and 25% of relative RMSE. Model ① showed the lowest RMSE: 5.1 m²/ha, but was sensitive to all external parameters examined (pulse density, topographical conditions and stand density). Model ② had a slightly larger relative RMSE (19%), but the influence of external parameters were limited to stand density. Model ③ had an even larger coefficient of variation (22%), and was shown to be insensitive to stand density, but remained sensitive to other dendrometrical, topographical and LIDAR acquisition parameters.

The first 2 PCA axes performed on each set of 4 dependent variables explain 80%, 83% and 82% of the variability, for models ①, ② and ③ respectively. Calibration plots can be projected in the space of these principal components and external envelops were adjusted to define the calibration domain of each model. According to type of envelope used (*i.e.* convex hull or alpha shape), the rate of prediction plots within the validity domain varies from 59 to 93% for convex hull (table 1), from 52 to 86% for alpha-shape (alpha =1) and between 52 to 80 % when an alpha of 0.5 is used. Spatial analysis show that in private forest, prediction plots fall more frequently outside the calibration range as compared to public forests (table 1).

Analysis of Mahalanobis distance to the nearest calibration plot illustrate that prediction plots of model ① are further from the calibration range, compare to predictions from model ② or ③. Irrespective of the model considered, spatial overlay with forest types illustrates a greater Mahalanobis distances in private forests than in public forests (figure 1). Spatial overlay with NFI stand types shows that models were relatively stable among forest compositions for model ② or ③. A high dispersion of Mahalanobis distances can be observed with model①, particularly for decidous stands.

Discussion / Conclusion

Model selection only based on RMSE criteria could lead to choose a biased model in some part of the study area especially when models are used in an extrapolation mode. An illustrative example was presented with model ① in private forests or in decidous stands (*i.e.* these conditions were not part of the calibration sample). We have shown that model robustness can be assess using: the boundaries of the model calibration range, dispersion of prediction plots and Mahalanobis distances to the nearest calibration plots. Using the geographical coordinates of the predicted plots, uncertainty maps can be performed. These maps allow to identify areas where models are used in an extrapolation mode. These areas could be removed for further analyses or the extent of the calibration domain be completed in these areas in order to improve maps precision.

Even if a method defining the boundaries of the calibration domain have already been used to illustrate the extent to which the field measurements encompass landscape structural conditions[2], the method presented here allows to classify prediction plots as within or outside the validity domain. Uncertainty maps can be produce using these new indices. Nevertheless, the type of boundary envelops influence the resulting maps. Furthermore, prediction plots can be outside the calibration domain but near a calibration plot. Conversely, a prediction plot can be inside the calibration range but far from a calibration plot depending how compact the calibration sample is in terms of LIDAR predictors.

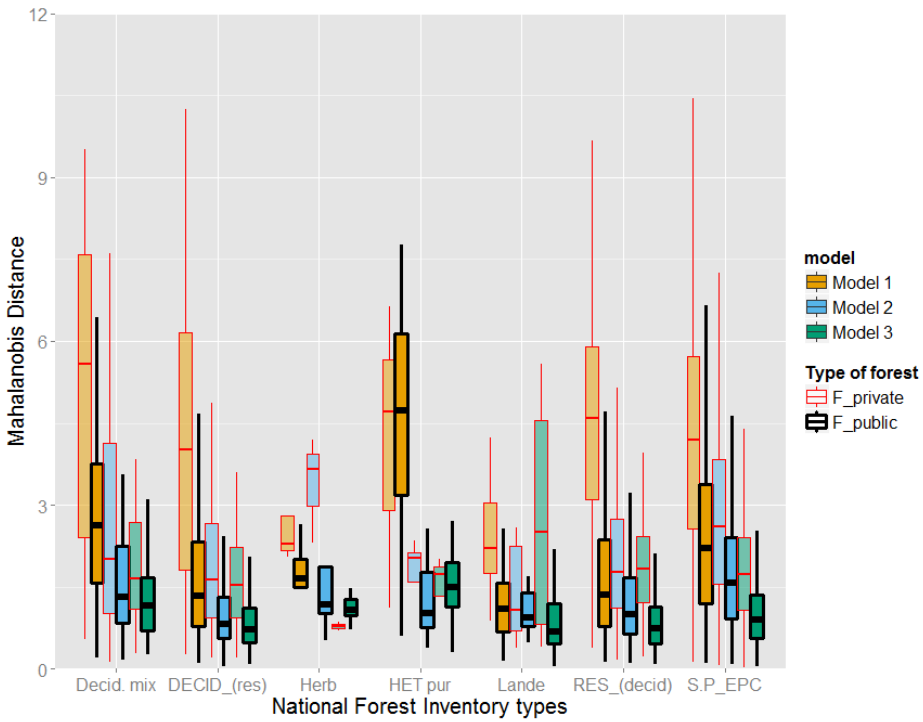
The approach using Mahalanobis distance allows to work in the calibration space of the LIDAR predictors. This further complements the information obtained by the establishment of the boundaries envelops obtained by either convex hull, or alpha-shape technics.

		F_private	F_public
Model 1 = f(dHmean_sup2m ; Nrel.6m12m ; dHmean_lst ; TxPen1m0.5H99_0.8H99_lst.tot)	CHULL	93%	91%
	alpha=0.5	57%	63%
	alpha=1	83%	86%
Model 2 = f(d9 ; d3_lst ; Hmv3_Bd ; TxPen1m0.5H99_0.8H99_lst.upp)	CHULL	59%	89%
	alpha=0.5	47%	80%
	alpha=1	52%	86%
Model 3 = f(dcum03 ; d0 ; Hcv_lst ; TxPen1m0.5H99_0.8H99_only.upp)	CHULL	64%	88%
	alpha=0.5	52%	76%
	alpha=1	55%	80%
Number of prediction plots		20694	39413

Table 1 : Proportion of prediction pixels found within the calibration range for the 3 models retained, according to stand properties (public or private forests) and the method of calibration envelope establishment (Chull = Convex Hull , Alpha Shape with an alpha = 0.5 or 1).

Figure 1 : Distribution of Mahalanobis distances to the nearest calibration plot for three prediction models of BA, depending on the forest types (NFI code) and stand property (Public or Private), with :

- Model ① TxPen1m0.5H99_0.8H99_lst.tot.Nrel.6m12m.dHmean_lst.dHmean_sup2m
- Model ②, TxPen1m0.5H99_0.8H99_lst.upp.d9.d3_lst.Hmv3_Bd
- Model ③ TxPen1m0.5H99_0.8H99_only.upp.dcum03.d0.Hcv_lst



References

1. Naesset, E. Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. *Scand.* J. *For. Res.* **19,** 164–179 (2004).
2. White, J. C. *et al. A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach*. (2013). at <https://cfs.nrcan.gc.ca/publications?id=34887>
3. Fekety, P. A., Falkowski, M. J. & Hudak, A. T. Temporal transferability of LiDAR-based imputation of forest inventory attributes. *Can. J. For. Res.* **45,** 422–435 (2015).

Exploring G-LiHT data to predict individual tree LAI from tree height

Huaguo Huang, Kan Huang

Key Laboratory for Silviculture and Conservation, Ministry of Education, College of Forestry, Beijing Forestry University, Beijing100083, People's Republic of China

Highlights: Goddard's LiDAR, Hyperspectral and Thermal (G-LiHT) data were explored to extract height and effective leaf area index (LAI) of single trees and investigate the relationship between them. In three forested areas, high correlation (R^2=0.68) was found between single tree mean or maximum LAI and its tree height.

Key words: *G-LiHT, LAI, tree height, site index.*

Introduction

G-LiHT is a portable, airborne imaging system that simultaneously maps the composition, structure, and function of terrestrial ecosystems using lidar, image spectroscopy and thermal cameras [1]. With the full name as NASA Goddard's LiDAR, Hyperspectral and Thermal Imager instrument, G-LiHT data is a freely available online(http://gliht.gsfc.nasa.gov), which is a unique data source to enable data fusion at fine-scale (< 1 m). Despite a few applications[2], G-LiHT data has not been fully discovered.

With the increasing demands for precise LAI and the aid of high resolution satellite images available, the LAI of individual trees (LAI_{tree}) is possible to be produced for forest management [3]. LAI_{tree} can be used to reconstruct three-dimensional (3D) scenes for radiative transfer modeling [4]. However, it is difficult to validate satellite-derived LAI_{tree} if using traditional LAI measurement methods (hemi-photos) . Therefore, most studies used LiDAR data as surrogate measurements [5]. Terrestrial LiDAR system (TLS) has very high density, it has been successfully used to derive LAI_{tree} at plot scales . If with similar high density, airborne LiDAR is also capable of extracting LAI_{tree} for larger area than TLS. Based on its high density of point cloud, the objective of this paper is to exploit G-LiHT capability to extract single tree LAI..

Sites and Data

G-LiHT data collected in three sites in USA were used, which are the Howland Forest site (N45°13'58.98", W68°44'44.64") in Maine, the Prospect Hill site of Harvard Forest (N42°32'35.31", W72°10'41.73") in Massachusetts, and the Smithsonian Environmental Research Center site (N38°53'34.42", W76°33'20.91") in Maryland respectively. LiDAR used a 300 kHz multi-stop scanning LiDAR with a 60° field of view and 10 cm diameter footprint, and the site was flown with 50% overlap in north–south and east–west directions to achieve a mean return density of up to 50 laser pulses/m^2. Leaf-on data in June were downloaded and used. The hyperspectral images with 402 bands (417-1000 nm) were acquired simultaneously with LIDAR.

G-LiHT provided three levels (L1, L2, and L3) of products. We used the L3 product containing returns ("point clouds"), DTM (Digital Terrain Model), CHM (Canopy Height Model, see Fig. 2), LiDAR metrics, and common vegetation index. Among them, point cloud data were stored in LAS file format. Public released point clouds were re-sampled and processed as smaller data volumes with around 1 km tiles for mapping. The point density (points/m^2) of first returns is from 0 to 23 with a mean value around 5. Scan angle is between -30 to 30 degree.

Methods

Fours major steps were used to calculate LAI_{tree} (Fig. 1).

Step 1 -- Four high resolution (0.4 m) image products were generated, including CHM, total number of first returns denoted as N_{tot}, number of ground first returns denoted as N_{grd}, 95th percentile of first returns denoted as H_{95} from point clouds using BCAL LiDAR Tools.

Step 2 -- Based on CHM, individual tree crowns were segmented and tree heights were extracted using TreeVaW software [6, 7] . A list of trees, noted as LISTREE, with parameters of position (x, y), height H_{max}, and radius R was created.

Step 3 -- For a tree i in LISTREE, its vertical crown projection mask was created. Within this mask, the number of N_{tot} and N_{grd} were counted as $N_{tot,\,i}$ and $N_{grd,\,i}$ and the maximum value of H_{95} was calculated. A near nadir canopy gap fraction of p_i was estimated as:

$$p_i = \frac{N_{grd,i}}{N_{tot,i}} \quad (1)$$

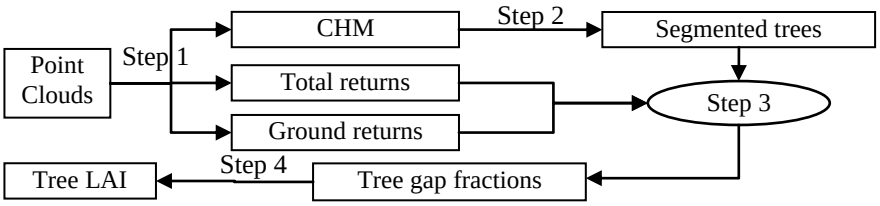


Figure 1: Four steps to retrieve LAI_{tree} from point clouds

Step 4 -- Assuming a spherical leaf angle distribution of tree *i* and an incidence zenith angle θ, the gap fraction was theoretically related to LAI_{tree} by Beer-Lambert law similar to previous studies [8]:

$$p_i = \exp\left(-\frac{0.5 \times LAI_{tree}}{\cos\theta}\right) \quad (2)$$

The LAI was then inversed from Eq. (2):

$$LAI_{tree} = 2.0 \times \cos\theta \times \ln p_i \quad (3)$$

For simplicity, nadir incident ($\theta=0$) is used. H_{95} was used to remove those "bad" trees with errors larger than 5 m between H_{95} and CHM-extracted H_{max}.

To explain the relationship between each tree's height and LAI, three basic assumptions were defined:

A1: Three basic types of tree crown shells, cylinder, ellipsoid or cone, were assumed to construct any tree crown with a maximum radius *R* and a crown length *L*.

A2: The crown length *L* is linear with tree height H_{max}: $L = a \times H_{max}$

A3: The leaf area of a tree (LA_{tree}) is linearly correlated to its crown volume (V_{tree}) for trees with the same specie type: $LA_{tree} = g \times V_{tree}$.

Based on assumptions (A1) and (A2), the volume of a tree can be calculated as:

$$\mathbf{V_{tree} = f \times \pi \times R^2 \times H_{max}} \quad (4)$$

Where *f* is *a* for cylinder, *a*/2 for ellipsoid, *a*/3 for cone, or [*a*/3, *a*] for combinations of these three crown types. Then, assumption (A3) is used to compute the leaf area (or LAI) of a tree:

$$\mathbf{LA_{tree} = g \times V_{tree} = g \times f \times \pi \times R^2 \times H_{max}} \quad (5)$$

$$\mathbf{LAI_{tree} = LA_{tree} / (\pi \times R^2) = g \times f \times H_{max}} \quad (6)$$

Note that *f* is a specie-dependant constant, while *g* is assumed specie-independent but varying with phenology, which needs to be inversed or predefined here. As a result, $\boldsymbol{LAI_{tree}}$ is linearly with H_{max}.

Results

Figure 2 shows scattering plots between H_{max} and LAI_{tree} for the three sites using G-LiHT derived information, where significant linear correlations are generally found (R^2>0.35). Note that H_{max} and LAI_{tree} are symbolized in sub-figures with H and LAI for convenience. The slopes of the linear equations are from 0.12 to 0.311. The best result (R^2=0.68) is in the second tile of the PH site, which shows evenly sampled tree heights. On SE site, the worst result (R^2=0.36) appears with uneven and fewer tree samples located in the second tile.

Significant errors of predicted LAI_{tree} are visible, which somehow reduces the accuracies of predictions. However, it is interesting that the upper boundaries of predicted LAI_{tree} are quite stable for all tiles. Therefore, we put all nine groups' results into one figure (Fig. 3a), which shows a clear upper boundary. Besides, an improved R^2 (0.68) is also found by integrating the results of nine tiles. The upper boundary can be fitted using a log equation or second order polynomial curve (Eq. 7 and 8). Fig. 3b shows that the LiDAR derived LAI values are positively related to vegetation index.

$$\mathbf{LAItree = 3.344 \times ln(Hmax) \quad (R2=0.98;\ RMSE = 0.40)} \quad (7)$$

$$\mathbf{AItree = -0.004521 \times (Hmax)2 + 0.4205 \times Hmax + 3.11 \quad (R2=0.99;\ RMSE = 0.31)} \quad (8)$$

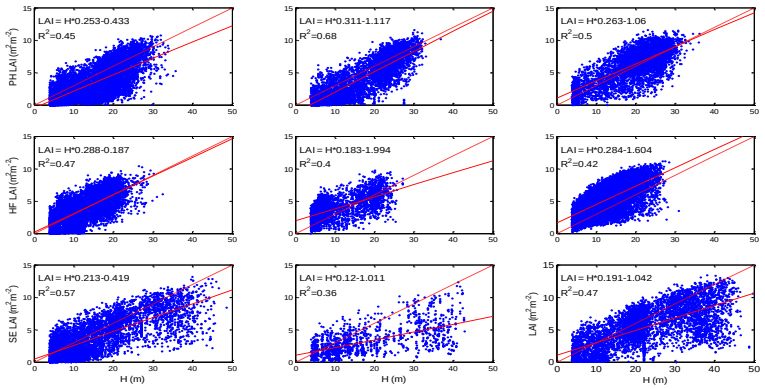


Figure 2: Scattering plots between H_{max} (Horizontal axis, H) and LAI_{tree} (vertical axis, LAI) at nine tiles in three sites; dash lines representing a line with slope 0.3 and zero interception; solid lines are fitted lines.

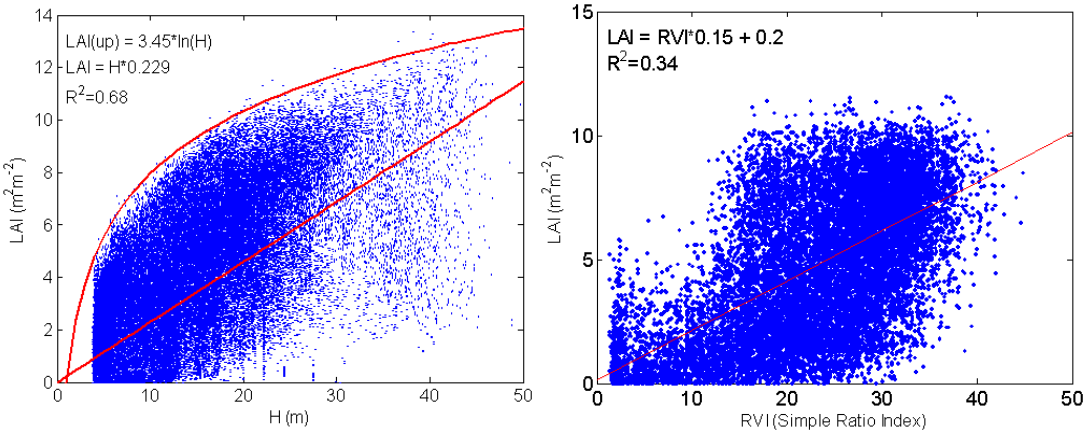


Figure 3: (a) Scattering plot of tree height (H) and LAItree for all nine tiles; (b) Scattering plot of LAI and RVI (simple ratio vegetation index) from hyperspectral images.

Conclusion and discussions

Based on G-LiHT data, it has been demonstrated that mean effective leaf area index (LAI) of single tree was highly correlated to single tree height (R^2=0.68). Besides, the potential maximum LAI values are more correlated, which may be explained by site quality. This conclusion is limited on the test in eastern forests of USA, but it provides a direction to allocate leaf area in forests from coarser to finer resolution. Next studies will evaluate the findings over wider eco-zones. The specie-induced variations will also be explored.

The upper boundary of predicted LAI is commonly visible for all tested tiles. It should have some explanations. The maximum LAI value for a tree height can be seen as the ideal growth limit based on the site quality and best competition status for light and soil fertilization. This explanation is similar with the concept of site index.

References

[1] B. D. Cook, L. A. Corp, R. F. Nelson, E. M. Middleton, D. C. Morton, and J. T. McCorkel.(2013). NASA Goddard's LiDAR, Hyperspectral and Thermal (G-LiHT) Airborne Imager. Remote Sensing. 5, 4045-4066.

[2] L. I. Duncanson, B. D. Cook, G. C. Hurtt, and R. O. Dubayah.(2014). An efficient, multi-layered crown delineation algorithm for mapping individual tree structure across multiple ecosystems. Remote Sensing of Environment. 154, 378-386.

[3] S. D. Roberts, T. J. Dean, D. L. Evans, J. W. McCombs, R. L. Harrington, and P. A. Glass.(2005). Estimating individual tree leaf area in loblolly pine plantations using LiDAR-derived measurements of height and crown dimensions. Forest Ecology and Management. 213, 54-70.

[4] H. Kobayashi, R. Suzuki and S. Kobayashi.(2007). Reflectance seasonality and its relation to the canopy leaf area index in an eastern Siberian larch forest: Multi-satellite data and radiative transfer analyses. Remote Sensing Of Environment.,

[5] G. Zheng and L. M. Moskal.(2012). Computational-Geometry-Based Retrieval of Effective Leaf Area Index Using Terrestrial Laser Scanning. Ieee Transactions on Geoscience and Remote Sensing. 50, 3958-3969.

[6] H. Kaartinen, J. Hyyppa, X. Yu, M. Vastaranta, H. Hyyppa, and A. Kukko.(2012). An International Comparison of Individual Tree Detection and Extraction Using Airborne Laser Scanning. Remote Sensing. 4, 950-974.

[7] S. C. Popescu and R. H. Wynne.(2004). Seeing the trees in the forest: Using lidar and multispectral data fusion with local filtering and variable window size for estimating tree height. Photogrammetric Engineering And Remote Sensing. 70, 589-604.

[8] S. Solberg, A. Brunner, K. H. Hanssen, H. Lange, E. Naesset, and M. Rautiainen.(2009). Mapping LAI in a Norway spruce forest using airborne laser scanning. Remote Sensing of Environment. 113, 2317-2327.

Obtaining forest description for small-scale forests using an integrated remote sensing technique

Cong (Vega) Xu, Bruce Manley, Justin Morgenroth

New Zealand School of Forestry, University of Canterbury, New Zealand

Highlights: In New Zealand, over 70% of plantation forests are managed by large-scale forest owners with the means to employ best management practices. In contrast, the 30% of plantation forests managed by small-scale forest owners are unlikely to afford regular area and yield assessment, which becomes the main hurdle for obtaining detailed forest description. This research explores a cost-effective remote sensing approach (fusion of airborne LiDAR, aerial photography and RapidEye satellite imagery) to obtain detailed area and yield information for small-scale forests in the Wairarapa region, New Zealand.

Key words: *Airborne LiDAR, RapidEye, Aerial photo, satellite imagery, sensor-fusion*

Introduction

Forestry is the third largest export earner and a significant industry in New Zealand. A majority of plantation forests (70%) are owned by large-scale forest owners with more than 1000 ha of forest area (MPI, 2012). Large-scale forests are managed professionally with regular assessment of forest areas and yields through aerial photography and inventories in order to support management decisions. On the other hand, small-scale forest owners (less than 1000 ha) are reported to own 30% of the total plantation area (MPI, 2012). The forest description of small-scale forests is believed to be less reliable due to inconsistent area definition and management practices. The areas reported may well be gross areas rather than net stocked areas, and the yield is assumed based on large-scale forests from the same region. This is likely to result in an overestimation of yield for New Zealand's small-scale forests. Small-scale forest owners are unlikely to afford regular area and yield assessment, which becomes the main hurdle for obtaining detailed forest description for those small-scale plantations.

Geospatial data including both vector and raster layers in New Zealand are increasingly becoming accessible for public use. In addition, a number of regional councils have acquired low to medium density LiDAR and high resolution aerial photographs, which could potentially be used to enhance current forest description for small-scale forests. This research intends to evaluate the possibility of deriving more detailed forest description of small-scale forests utilising free or low-cost remote sensing datasets (LiDAR, aerial photo and RapidEye satellite imagery), to accurately delineate net stocked area and predict yield production for small-scale forest owners without incurring substantial costs for the owners.

Materials and Methods

The study area is the Wairarapa area within the Wellington Region, which is located in the south-east of the North Island of New Zealand (Figure 1). The region was targeted for this research as half of the plantations in the region are owned by small-scale forest owners. The Wellington Regional Council acquired wall-to-wall discrete airborne LiDAR and high-resolution (0.3 m) aerial photos. Low-cost RapidEye scenes are also used in the study to provide an extra image source. Due to the extensive coverage of the study area, the study area is divided into equal sized (3.6 x 2.4 km) grids (n=???) and a stratified sampling approach was applied to select samples for plantation mapping. The forest plantation defined under the Land Use Carbon Analysis System (LUCAS) was superimposed on sampling grids so that each grid can be classified into one of the plantation cover classes: <10 ha, 10-100 ha and > 100 ha. Twenty samples were then randomly selected within each forest class, which resulted in 60 grids selected in total for conducting area assessment. Additionally, three randomly selected grids from each forest class (i.e. 9 grids in total) were used as the training datasets for developing forest mapping workflow. These are independent from the 60 sample datasets (See Figure 1).

RapidEye-derived reflectance, textural information and vegetation indices and LiDAR extracted surfaces (Canopy Height Model, Digital Elevation Model, intensity, slope, elevation standard deviation) were used as

inputs to segment and classify forest plantation areas in the 60 sample grids, using an Object-Based Image Analysis (OBIA) decision tree classification approach. Digitised plantations with aerial photos were used as the 'ground truthing" data to determine the classification and mapping accuracy of individual sensors and combined sensors.

In New Zealand, forest owners who enter the Emissions Trading Scheme (ETS) are required to provide inventory data for plots preselected by the Ministry of Primary Industries (MPI) for carbon reporting purpose. This inventory data may be the only plot data available for small-scale forests. Therefore, the carbon plots were selected as the plot data for use in this study. The authors re-measured 124 plots covering a wide range of forest age classes and management regimes. The measurement involved diameter and height for all trees in the plot, and we relocated the plot centre with a Trimble GeoXH in order to minimise the location offset. Forest volume is modelled based on the correlation developed between LiDAR/RapidEye metrics and the ground data using both parametric and non-parametric approaches. This study also evaluates the effect of using a consumer grade GPS and high grade GPS on the accuracy of forest height and volume estimation.

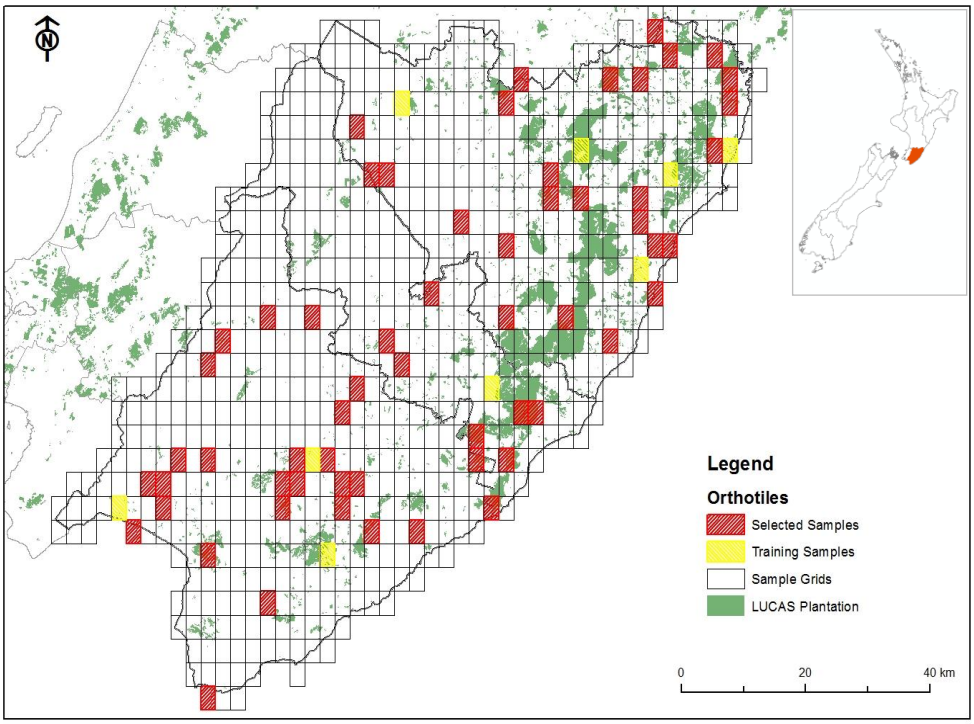


Figure 1: Study area.

Results

Both LiDAR and RapidEye derived surfaces have proved useful in delineating net stocked plantation areas, combining surfaces from both LiDAR and RapidEye have improved the classification accuracy further, mainly because the canopy height information more accurately differentiated between plantation forests and other land cover types.

Acknowledgements

Thanks to The New Zealand School of Forestry for providing research funding. Thanks to the Wellington Regional Council and Landcare Research for providing LiDAR and aerial photos, thanks for AAM Group and Indufor Asia Pacific for delivering RapidEye images. Thanks to Dr Michael Watt for valuable suggestions on the project and Humin Lin for her support in field data collection.

References

MPI. (2012). National Exotic Forest Description. Wellington: Ministry of Primary Industries.

Comparison of a low-cost photogrammetric method (Photo-Panoramas) to Terrestrial Laser Scanning for measurement of vegetation structure

Jasmine Muir[1,2], John Armston[1], Ben Ward[3,4], Ben Sparrow[3,4], Stuart Phinn[2], Peter Scarth[1,2]

1Queensland Department of Science, Information Technology, Innovation and the Arts, Australia
2University of Queensland, Australia
3University of Adelaide, Australia
4Terrestrial Ecosystem Research Network, Australia

Highlights: This paper compares a photogrammetry method (photo-panoramas) and Terrestrial Laser Scanning (TLS) for measuring stem count, diameter at breast height (DBH), and basal area. These parameters were underestimated by both methods. TLS achieved higher accuracies (47% stem detection) than photo panoramas (23%) at higher basal area (30 m^2/ha).

Key words: *terrestrial laser scanning, forestry, vegetation structure, photogrammetry*

Introduction

Measuring vegetation structural attributes is important for forestry, carbon monitoring, biodiversity conservation, fuel load management and ecosystem function [1-4]. Traditionally these metrics are collected using manual field measurements. In recent years however, Lidar (light detection and ranging) measurements collected by Terrestrial Laser Scanning (TLS) have been used successfully to characterize vegetation structure metrics [5-9].

Here we explore whether a low cost photogrammetry method (known as photo-panoramas), can measure stand structure metrics to the same level of accuracy as a high cost survey grade TLS (Riegl VZ400). The structural metrics measured are stem count, basal area and diameter at breast height (DBH). Direct manual measurements were collected at two Eucalypt woodland sites, near Brisbane, Queensland, Australia, and compared to structural metrics derived from TLS and the photo-panoramas method. The photo-panorama method is attractive from a cost perspective and results from our analysis provide guidance as to whether the method is applicable in Eucalypt woodland sites.

Method

Study Site and Data

Two Eucalypt woodland sites, near Brisbane, Queensland, Australia were measured during a field campaign carried out by the Terrestrial Laser Scanning International Interest Group (TLSIIG) in July-August 2013 (http://tlsiig.bu.edu). Two plots were chosen with co-incident field data, TLS and photo-panorama analysis results (Kara001 Centre and Gold0101North).

Stem location (bearing and distance to plot centre) and DBH (at 1.3m) was manually measured within a 25m radius of the plot centre. A basal area wedge with factor 1.0 was used to calculate field basal area.

TLS data was acquired at each of the two sites using a Riegl VZ-400 instrument with a beam spacing of 0.06degrees and pulse repetition frequency of 300 Khz. The Riegl VZ-400 instrument is a time-of-flight sensor, and uses a laser operating at 1550nm (near infra-red). It has on board waveform processing, which automatically detects up to four returns. A single horizontal scan was collected (30^0 – 130^0 zenith) at each plot.

The photo-panorama method involves taking photos to form a 360 degree panorama at three evenly spaced locations, 1.45m from the plot centre. The method has been developed as part of the Terrestrial Ecosystem Research Network (TERN) Ausplots Rangelands survey protocol [10]. A Canon D50 camera attached to a Gigapan © rotating stage atop a camera tripod, was used to take 21 photos to form each of the panorama images.

Data Analysis

A Digital Terrain Model (DTM) was produced from the TLS data to allow extraction of points at 1.3m (breast height) in the point cloud. Ground returns were classified by creating a minimum z surface at 0.5m resolution using all returns, and then passing this through an iterative morphological grey opening filter. The morphological structure element was progressively increased from an initial window of 0.5 to 10m. Adjacent pixels with a slope greater than 45degrees were masked from analysis. Original minimum points within 0.5m of

the resulting surface were classified as ground points and natural neighbour interpolation was used to interpolate areas of no data [11-12].

Point data was extracted at 1.3m above the DTM. To remove non-trunk hits only first and last returns with relative reflectance greater than -10dB were used. A 5cmbuffer around each point was dissolved to produce point clusters. Clusters were considered if they contained more than 10 points. Circles were fitted to each cluster [13] and considered to be stems if the difference between the minimum x and minimum y value for the point cluster, and the expected closest point to the plot centre on the fitted circle was less than a specified threshold (i.e. this takes advantage of the geometry of single scans where all points are on the side of the circle towards the scanner), and if the radius of the fitted circle was within a specified percentage of the distance for all points in the cluster to the fitted circle centre. All fitted circles meeting these assumptions were presumed to be stems for further analysis.

The photo-panorama method identifies the positions of corresponding features in the three panorama photos taken around the plot centre. This enables reconstruction of camera positions and 3D feature location [14]. Tree trunks are identified using an automated process and cylinders fitted to estimate DBH, and basal area.

Results and Discussion

Results of the analysis for each site are shown in Table 1. TLS achieved a stem classification accuracy of ~ 73% and ~47% at the Kara001 Centre and Gold0101 North plots respectively, while the photo-panorama method achieved accuracies of ~60% and ~ 23% at the same sites. Basal area was underestimated by both TLS and phot-panoramas at both sites, but the estimate from TLS was much closer to the field measured values within 4.4 m^2 and 5.4m^2, as opposed to differences of 9.5 m^2 and 19 m^2 for photo-panoramas at Kara001 Centre and Gold0101 North. Mean DBH for the plot was estimated within ± 1cm at Kara001 Centre and within ± 2cmat Gold0101 North for both TLS and photo-panoramas.

Table 1: Comparison of field measurements, TLS and Photo-Panorama stem count, basal area and mean DBH for the Kara001 Centre and Gold0101 North plots.

		Stem Count	Basal Area (m2/ha)	Mean DBH (m)
Kara001 Centre	Field	128	21.00	0.15
	Riegl VZ400	94	16.59	0.16
	Photo Panoramas	78	11.50	0.14
Gold0101 North	Field	194	30.00	0.18
	Riegl VZ400	91	24.61	0.20
	Photo Panoramas	44	11.03	0.16

Stem detection accuracy for both the TLS and photo-panoramas determines the accuracy of the structure metrics in relation to field data. The smaller number of stems compared to field measured, by both these methods is influenced by occlusion, with stems behind those identified not able to be viewed. TLS performed consistently across the sites, however results from the photo-panorama method were considerably different to field stem counts at the higher basal area site Gold0101 North. This suggests that the photo-panorama method to calculate basal area may not be applicable at sites of high basal area, or may need further refinement to work in these areas. In addition improved accuracies from TLS have been shown possible when using multiple co-registered scans as opposed to a single scan [15 -16]. However using multiple scans increases the scanning field time due to placement of reflector targets, and scanner setup overheads. The accuracy of using single scans for stem detection has also been shown to improve when an occlusion factor is applied [17]. .

Further Analysis and Conclusion

Vegetation structure metrics such as stem count, DBH and basal area can be measured using TLS and photogrammetry techniques. Our analysis suggests that photo-panoramas may not be suited to use in plots with high basal area. Further analysis across sites with a range of basal areas will enable determination of appropriate basal areas for which this method can be used. In addition the method to account for occlusion in the photo-

panorama results is still in development, and the use of an occlusion factor with the TLS measurements will be investigated. It is expected that use of an occlusion factor will improve accuracy results for both photo-panoramas and TLS structure metrics.

The cost-advantage of using the photo-panorama method over the Riegl VZ400 TLS will only be of benefit if the method can be proven to work across a range of vegetation communities. The results of our analysis suggest that the photo-panorama method may not be applicable in Eucalypt woodlands of high basal area. In these environments field methods (such as the basal area wedge) may be quicker and more reliable. However TLS offers the advantage of detailed 3D reconstruction from which a variety of additional structural metrics can be extracted. These can be used in comparison of structure for different vegetation communities, and as inputs for calibration and validation of remote sensing products derived from airborne and satellite data.

References

[1] Harmon, M. E., Franklin, J. F., Swanson, F. J., Sollins, P., Gregory, S. V., Lattin, J. D., Cummins, K. W. (1986). Ecology of Coarse Woody Debris in Temperate Ecosystems. Advances in Ecological Research, 15, 133–302.

[2] Lindenmayer, D. B., Margules, C. R., & Botkin, D. B. (2000). Indicators of biodiversity for ecologically sustainable forest management. Conservation Biology, 14(4), 941–950.

[3] Parker, G. G. (1995). Structure and microclimate of forest canopies. In M. D. Lowman & N. M. Nadkarni (Eds.), Forest canopies (pp. 73–106). Boston, MA: Academic Press.

[4] Spies, T. A. (1998). Forest structure: A key to the ecosystem.pdf. Northwest Science, 72(Special Issue No. 2), 34–39.

[5] Lovell, J. L., Jupp, D. L. B., Newnham, G. J., & Culvenor, D. S. (2011). Measuring tree stem diameters using intensity profiles from ground-based scanning lidar from a fixed viewpoint. ISPRS Journal of Photogrammetry and Remote Sensing, 66(1), 46–55.

[6] Raumonen, P., Kaasalainen, M., Åkerblom, M., Kaasalainen, S., Kaartinen, H., Vastaranta, M., Lewis, P. (2013). Fast Automatic Precision Tree Models from Terrestrial Laser Scanner Data. Remote Sensing, 5(2), 491–520.

[7] Tansey, K., Selmes, N., Anstee, A., Tate, N. J., & Denniss, A. (2009). Estimating tree and stand variables in a Corsican Pine woodland from terrestrial laser scanner data. International Journal of Remote Sensing, 30(19), 5195–5209.

[8] Yang, X., Strahler, A. H., Schaaf, C. B., Jupp, D. L. B., Yao, T., Zhao, F., … Ni-Meister, W. (2013). Three-dimensional forest reconstruction and structural parameter retrievals using a terrestrial full-waveform lidar instrument (Echidna®). Remote Sensing of Environment, 135, 36–51.

[9] Yao, T., Yang, X., Zhao, F., Wang, Z., Zhang, Q., Jupp, D., Strahler, A. (2011). Measuring forest structure and biomass in New England forest stands using Echidna ground-based lidar. Remote Sensing of Environment, 115(11), 2965–2974.

[10] White, A., Sparrow, B., Leitch, E., Foulkes, J., Flitton, R., Lowe, A. J., & Caddy-Retalic, S. (2012). AusPlots Rangelands survey protocols manual. Retrieved from http://tern.org.au/AusPlots-Rangelands-Survey-Protocols-Manual-pg23944.html

[11] Zhang, K., Chen, S.-C., Whitman, D., Shyu, M.-L., Yan, J., & Zhang, C. (2003). A progressive morphological filter for removing non ground measurements from airborne LIDAR data. IEEE Transactions on Geoscience and Remote Sensing, 41(4), 872–882.

[12] Pirotti, F. (2013). State of the Art of Ground and Aerial Laser Scanning Technologies for High-Resolution Topography of the Earth Surface. European Journal of Remote Sensing, 66–78.

[13] Taubin, G. (1991). Estimation Of Planar Curves, Surfaces And Nonplanar Space Curves Defined By Implicit Equations, With Applications To Edge And Range Image Segmentation, IEEE Trans. PAMI, 13, 1115-1138.

[14] Sparrow, B., Ward, B., Schaefer, M., Armston, J., Muir, J., Caddy-Retalic, S., Thurgate, N., Lowe, A., Phinn, S. (2013). Basal area from photos: is it possible? American Geophysical Union, Fall Meeting 2013, 12/2013.

[15] Trochta, J., Král, K., Janík, D., & Adam, D. (2013). Arrangement of terrestrial laser scanner positions for area-wide stem mapping of natural forests. Canadian Journal of Forest Research, 43(4), 355–363.

[16] Liang, X., Litkey, P., Hyyppa, J., Kaartinen, H., Vastaranta, M., & Holopainen, M. (2012). Automatic Stem Mapping Using Single-Scan Terrestrial Laser Scanning. IEEE Transactions on Geoscience and Remote Sensing, 50(2), 661–670.

[17] Strahler, A. H., Jupp, D. L. B., Woodcock, C. E., Schaaf, C. B., Yao, T., Zhao, F., … Boykin-Morris, W. (2008). Retrieval of forest structural parameters using a ground-based lidar instrument (Echidna(R)). Canadian Journal of Remote Sensing, 37(Supplement 2), S426–S440.

Comparing individual tree crown delineation and species identification derived from photogrammetric and airborne lidar point clouds

Benoît St-Onge[1], Félix-Antoine Audet[1], Jean Bégin[2]

1. *Department of Geography, University of Quebec at Montreal, Canada*
2. *Department of Wood and Forest Sciences, Laval University, Canada*

Highlights: We compared the performance of individual tree delineation, as well as species identification (softwood vs. hardwood) based on 3D data alone, obtained from canopy height models derived respectively from photogrammetric and lidar point clouds. The two sets of results were found to be similar, but were influenced by viewing geometry.

Key words: *Photogrammetry, image matching, lidar, delineation, species.*

Introduction

Photogrammetric point clouds (PPC) are becoming denser and more accurate as airborne acquisition technology and image matching methods evolve. Large forward overlap (e.g. 80%) allows for multiray matching and reduces the visual difference between consecutive images. New, more sophisticated image matching techniques reduce errors in the image correspondence solution. What is more, PPCs not only store the XYZ coordinates but also the multispectral information extracted from the image. These recent progresses are particularly significant for forest mapping because the matching of forested scenes is among the most difficult image correspondence problem. It is therefore a domain where the biggest improvements can be expected.

Introduced in 2001 [1], the combination of photogrammetric digital surface models and lidar digital terrain models (DTM) for creating hybrid canopy height models (CHMs) is becoming more common [2]. In this approach, the PPCs are normalized using lidar DTMs to produce *X,Y,Height* point clouds which are used as is, or in interpolated raster form. Some authors have used such hybrid data in classical area-based approach (ABA) with success to characterize forest structure in Norway [3] or Canada [4]. Applying individual tree crown delineation (ITC) on photogrammetric CHMs is just starting (e.g. [5], [6]) and little is known about the potential of this approach. Individualized crowns would allow using the multispectral information of the PPCs (in addition to the 3D data) to identify species and health state on a single tree basis.

The objective of the present study is to assess the potential of applying ITC methods to PPC by comparing the results derived from ALS and PPC. We first compared the overall 3D statistics of both datasets, and then confronted the respective lidar-based, and photogrammetric-lidar ITC results in terms of delineation performance, and utility for distinguishing softwood and hardwood trees based on 3D data alone.

Methods

Study region and data

The study area is the Montmorency Forest, a mixed boreal research forest located 80 km north of Quebec City, Canada (41°12'N, 71°11'W). Softwood trees (mainly *Abies balsamea*) dominate the landscape, but hardwood trees are also present (e.g. *Betula papyrifera*). Most trees do not exceed 20 m in height. The topography is rather hilly, with locally strong slopes. The airborne lidar data was acquired using an Optech ALTM3100 discrete return sensor in August 2011 with an incident pulse density of 7.13/m^2. The airborne images were captured in June of 2012 using an UltraCAM XP camera from Microsoft (formerly Vexcel), with a forward overlap of approximately 80%. The resulting ground pixel size was 10 cm. Very few changes occurred between the two acquisition years and test sites were chosen only in areas without changes.

Preprocessing

Various image matching software were evaluated in preliminary tests. We chose Pix4DMapper (Pix4D SA, Switzerland) for the high quality of the resulting PPC (high 3D definition, very few artefacts). The absolute orientation of the images was calculated from ground control points extracted from the lidar data. Once oriented, the images were matched twice using two strategies, one with at least two intersecting rays, and the other by forcing the use of a minimum of three rays. The resulting point clouds, as well as the lidar first returns were then converted to rasters by keeping the highest return in each 25 cm pixel. Pixels without any point were left null. These uninterpolated digital surface models were converted to CHMs by subtracting the lidar DTM. The resulting uninterpolated CHMs were kept for analysis, but an interpolated version of the PPC-based and lidar

CHMs were created using inverse distance weighting in which the non-null pixels used in the interpolation were selected based on a natural neighbour schemes.

Analysis

Our first goal was to assess the relationship between the ALS and PPC datasets in terms of correlation, completeness, and accuracy prior to tree delineation. The correlation and RMSE between the two uninterpolated CHMs was calculated for every pixel that was non-null in both CHMs. The height differences between these two CHMs were also computed. This analysis was done over 5 different sites, having different forest structures and areas varying from 12 to 54 ha. The two versions of the image matching (minimum of 2 and 3 intersecting rays respectively) were used in these comparisons.

ITC was performed on the interpolated CHMs, both on the lidar and PPC (created using a minimum of 2 rays). We used for this an in house algorithm that relies on height-adaptative gaussian filtering, local maxima detection, edge detection, and region growing around the local maxima, controlled by stopping conditions based on edges, tree size and shape. The quality of the ITC on the lidar was checked by visual interpretation. The lidar-ITC is used "as is" in this study solely for the purpose of evaluating the quality of the PPC-based results. The latter was judged through visual comparisons to gain an understanding of the qualitative differences between the two versions of the delineated crowns, and quantitatively by comparing the individual tree counts and frequencies per height bin of 2 m. The average slope of the profile of each delineated crown was calculated respectively from the lidar and PPC-based CHMs. The slope value was used to identify species class (softwood vs. hardwood) using an empirically derived threshold. For ground truth, sample crowns were identified as softwood or hardwood by photo-interpretation, resulting in 1256 and 1560 interpreted crowns respectively for PPC and lidar segmentation, half of sample trees were used for identifying the optimal threshold, and half for assessing the performance of species class identification.

Results

An example of the resulting PPC is presented in figure 1 a, and the lidar and PPC-based CHMs appear respectively in figure 1 b and c. The first results (not presented here) indicated that forcing a minimum of three intersecting rays during matching left many small patches without any points, so only results based on a minimum of two rays are presented. We see from figure 1 b & c that the lidar and PPC-based CHMs are very similar.

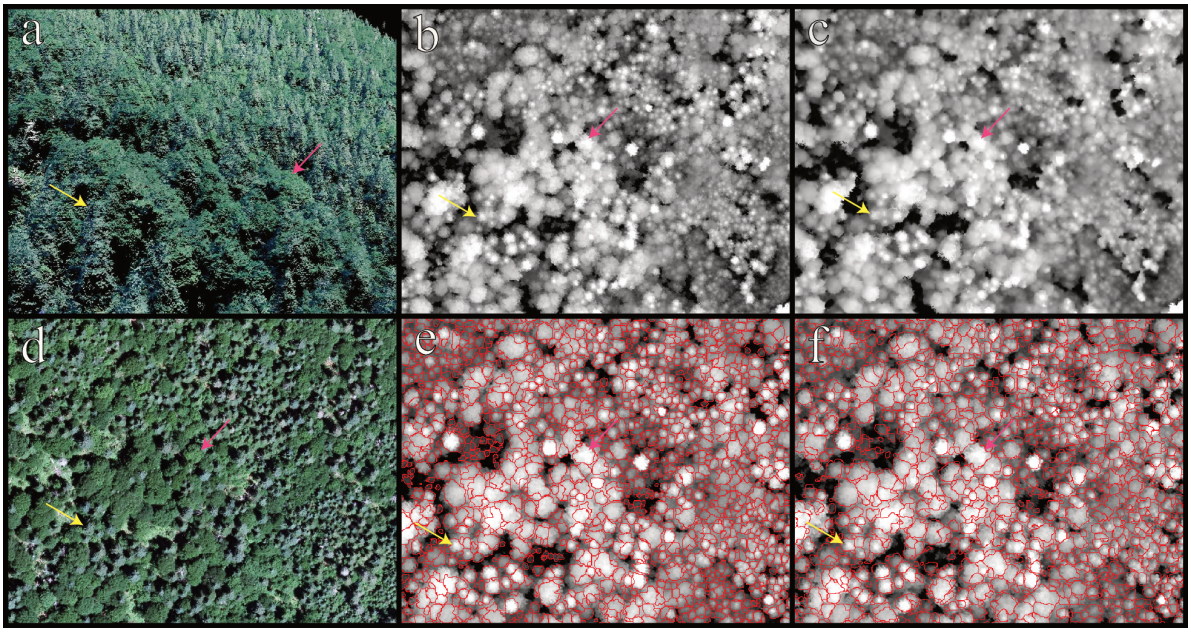


Figure 1 : Canopy height models (CHMSs) and crowns delineation results for one site of the study region: a) oblique view of a sample PPC, b) lidar CHM, c) PPC-based CHM, d) orthoimage corresponding to the CHMs, e) lidar crowns over the lidar CHM, f) PPC-based crowns over the lidar CHM. The red and yellow arrows designate the same two trees across panels for comparison purposes.

In comparing the uninterpolated CHMs on a pixel-wise basis, the following results were obtained over the 5 sites: a R^2 of 0.55-0.68, and RMSE of 2.68-2.06 m between the pixel heights of the respective CHMs, and height differences ranging from -30.23 to 35.98, with a mean of 1.52-3.32 m and a median of 0.11-2.28 m. Table 1 shows the pixel counts and frequencies per height class respectively for the uninterpolated lidar- and PPC- based CHMs. We see that the absolute pixel counts differ mostly for the shortest and highest classes (<4 m and >20m), but that in terms of proportion of pixels, a large difference is seen only for the shortest class.

Respective tree delineations results are very similar (figure 1, e&f). Some small trees that were distinguished in the lidar-based delineation were merged in the PPC-based results. A smaller number of trees were found in the PPC-based CHMs (Table 1). The difference in percentage per height class doesn't exceed 2.70%.

Table 1: Statistics of the PPC and lidar CHMs and individual tree crowns.

Pixel-wise	< 4 m	4-8 m	8-12 m	12-16 m	16-20	> 20 m
Lidar count	117237	155081	185606	166668	92318	17133
PPC count	30371	138356	179260	165110	75613	7553
Lidar %	15.79	21.13	25.29	22.70	12.58	2.33
PPC %	5.09	23.20	30.06	27.69	12.68	1.27
Delta %	-10.88	2.07	4.77	4.99	0.10	-1.06
Crown-wise						
Lidar count	39	1265	2972	2212	1616	708
PPC count	59	1022	2325	1939	1365	378
Lidar %	0.44	14.36	33.73	25.10	18.34	8.03
PPC %	0.83	14.42	32.8	27.36	19.26	5.33
Delta %	0.39	0.06	-0.93	2.26	0.92	-2.70

A confusion matrix between softwood and hardwood trees (not presented here) yielded an overall classification accuracy of 67.78%, with a kappa coefficient of 0.30 using lidar, and of 73.85%, with a kappa coefficient of 0.44, using the PPC-based CHMs.

Discussion and conclusion

Due to progress in image matching techniques, hybrid CHMs in which the surface was generated using PPCs are now very similar to lidar CHMs. Differences appear mostly in the areas of the stereo-models that are viewed more obliquely where smaller trees surrounded by taller ones, and the far side of taller elongated crowns, are missed. This results in number of photogrammetric points being generated at low heights between trees much lower than in the lidar case. This may explains the large difference in pixel-wise counts between PPCs and lidar. It however does not in general preclude applying ITC to PPC-based CHMs. We could demonstrate that although the resulting absolute tree counts somewhat differed, the percentage of tree crowns per height category was very similar. Furthermore, species class (softwood vs hardwood) could be well identified on an individual crown basis using only the 3D data. We hypothesize that further refinements in species identification will be possible by also using the multispectral information contained in the PPCs. This opens the way for individual tree forest inventory in which structural and species information is extracted based on photogrammetric-lidar CHMs.

Acknowledgements

We thank Dr. Chris Hopkinson (Lethbridge U., Canada) for his scientific and financial collaboration in acquiring the lidar data. Our gratitude also goes to Pix4D for offering Pix4DMapper at a discounted price.

References

[1] St-Onge, B. A., & Achaichia, N. (2001). Measuring forest canopy height using a combination of lidar and aerial photography data. *International Archives Of Photogrammetry Remote Sensing And Spatial Information Sciences*, *34*(3/W4), 131-138.

[2] White, J. C., Wulder, M. A., Vastaranta, M., Coops, N. C., Pitt, D., & Woods, M. (2013). The utility of image-based point clouds for forest inventory: A comparison with airborne laser scanning. *Forests*, *4*(3), 518-536.

[3] Gobakken, T., Bollandsås, O. M., & Næsset, E. (2014). Comparing biophysical forest characteristics estimated from photogrammetric matching of aerial images and airborne laser scanning data. *Scandinavian Journal of Forest Research*, 1-14

[4] Pitt, D. G., Woods, M., & Penner, M. (2014). A comparison of point clouds derived from stereo imagery and airborne laser scanning for the area-based estimation of forest inventory attributes in boreal Ontario. *Canadian Journal of Remote Sensing*, *40*(3), 214-232.

[5] Tompalski, P., Wezyk, P., & Weidenbach, M. (2014). A Comparison of LiDAR and Image-derived Canopy Height Models for Individual Tree Crown Segmentation with Object Based Image Analysis. In *5th Geobia Conference Thessaloniki. Publ. in South-Eastern European Journal of Earth Observation and Geomatics* (Vol. 3, No. 2S).

[6] Waser, L. T., Ginzler, C., Kuechler, M., Baltsavias, E., & Hurni, L. (2011). Semi-automatic classification of tree species in different forest ecosystems by spectral and geometric variables derived from Airborne Digital Sensor (ADS40) and RC30 data. *Remote Sensing of Environment*, *115*(1), 76-85.

CARTOMOB: an integrative GIS tool for forest management and logging operations based on LiDAR data

Thomas Carrette, Alain Thivolle-Cazat

thomas.carrette@fcba.fr
Institut Technologique FCBA, Domaine universitaire, 38044 Grenoble, France

Highlights: CARTOMOB is a new integrative GIS tool created for forest-based practitioners and which takes as input LiDAR-based forest information and other relevant information sources. This information is then enhanced into maps customized for the user and indicators relevant for forest managers and logging companies.

Key words: *Analytic tool, Forest management, Decision support- system*

Introduction

Recent evolution of remote sensing technologies such as LiDAR and development of new geomatic tools open up new scope of applications in the field of forest management and logging.

The main advantages brought by such technologies and computational model are precise topographical information, qualification of the resource, continuous forest mapping and forest road networks detection. Nerveless such information is not easily exploited to its full potential. In the FORESEE research project, a GIS based solution called CARTOMOB has been developed. It integrates such input, handles its processing in an analytical tool and delivers results in a decision support system.

Methods

Information provided by LiDAR is combined with other geographic layers. This integrative tool matches LiDAR information with public databases or information layers regarding species distribution or cadastre... Hence, a technical and economic study of the forest resource is carried out. The objectives are to characterize and spatialize wood availability, as well as to analyse the physical and economical feasibility of its logging.

To achieve this, three modules based on mathematical models are run alternately to estimate i) volumes available by product type (silvicultural regimes), ii) technical accessibility and iii) economic aspects such as financial balance of the operation. The silvicultural module is based on wood rate scaling available in the area. This analyse can be drilled down to as many species as desired, but in this first version of the software only four main species were detailed. Regarding technical accessibility, the SYLV'ACCESS software developed by IRSTEA has been directly encapsulated in CARTOMOB. This module enables the creation of "smart buffer" to identify the area accessible by a skidder and the distance the machine will have to travel.

At the end of the process, it is possible to easily visualize and estimate available wood volume, inaccessible plots, and economical interest for owners, contractors and public body. This service can answer many needs voiced-out by forest managers, logging contractors, timber industries, forest owners and natural conservators. CARTOMOB can be used in many practices by professional or institutions.

- For professional companies: commercial exploration (location of the volume by species and type of wood: sawn timber or pulp wood…), support decision for logging operation (cost, aggregation of site operation…) and detection of under-exploited areas (inactive owners).

- For institutions: support decision when preparing the creation of new forest roads (Area with high unreachable volume or high cost of mobilization…), sub-plot analysis of the resource, inventory and statistics for an area selected or a type of proprieties (see Table 1)...

Table : Type of results available for forest management purpose.

Area and nature of the plot	Volume Sawn Timber	Volume Pulpwood	Global Volume Available	Volume unreachable wood
<= 4 ha				
Private owner	68 076	46 243	114 305	23 982
Public owner	65 727	43 023	108 730	20 712
>= 20 ha				
Private owner	21 491	58 960	80 450	14 127
Public owner	769 341	575 627	1 344 914	156 598
4 < s < 20 ha				
Private owner	26 882	27 093	53 972	23 737
Public owner	811 822	500 230	1 312 069	190 019
Total	**1 763 339**	**1 251 175**	**3 014 440**	**429 175**

All volume are in cubic meters

Implementation

CARTOMOB has been tested in over 900 square kilometres of forests in the northeast of France on the Vosges area. Feedback was collected from private and public forest managers. According to them, analyses enabled by the tool clearly show its relevance. CARTOMOB is very evolutional and simple, on purpose, with a calculation time of 45 minutes for all this area. The maps provided by the tool were simple to use in the field and the information layers were easily integrated in practitioners' own GIS software. (See figure 1 for an example of output map).

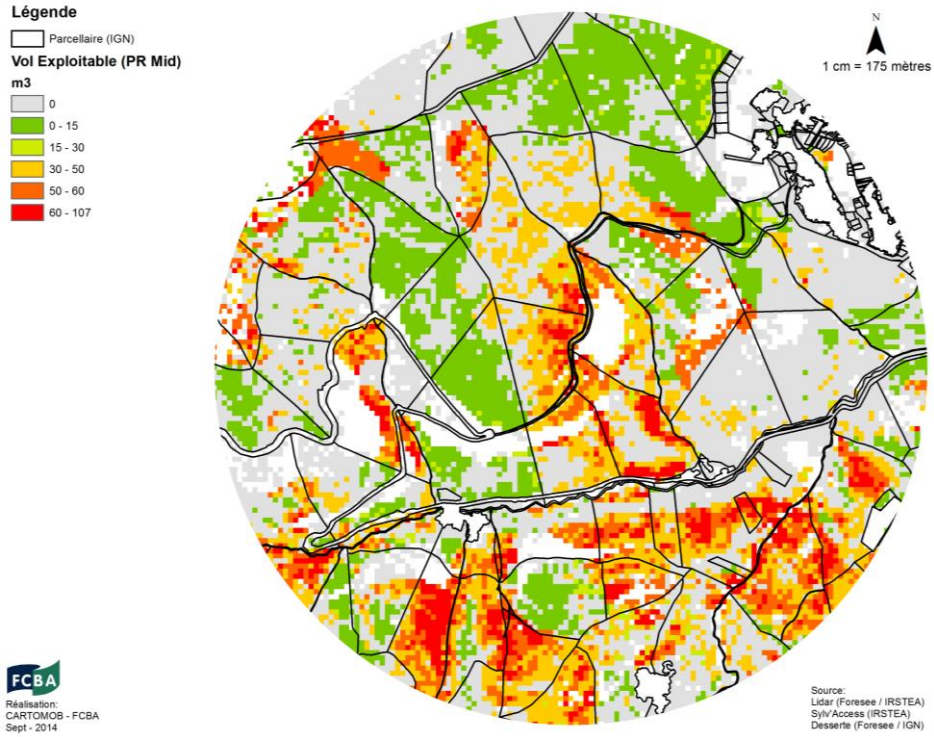


Figure 1: Output map on volume available in a sub-area of the north-east test site

This tool has been developed in the ARCGIS environment. Based on a python script, the structure of the tools has been conceived in a way to allows implementation and evolution on each part of the three modules. The major advantage of this tool is still its ability to integrate. In the near future more information will be integrated such as hydrographic network or fertility. Hence, all developments in connection with the LiDAR mapping and forest mapping can be provided to forest-based practitioners who are not in the habit of handling such data.

Analysis of the motion induced by tree growth from T-LiDAR measurements: a case study for Beech poles (*Fagus sylvatica* L.)

E. Noyer [1, 2] and T. Constant [1, 2]

[1] *INRA, UMR 1092 Laboratoire d'Etude des ressources Forêt Bois (LERFoB), Centre INRA de Nancy-Lorraine, 54280 Champenoux, France*

[2] *AgroParisTech, UMR 1092, Laboratoire d'Etude des ressources Forêt Bois (LERFoB), ENGREF, 14 rue Girardet, 54000 Nancy, France*

Highlights: We proposed a method to determine local curvature along the trunk. A sample of 42 beech poles whose the half are thinned in 2007-08 and were scanned during winter 2007-08 and 2013-14 by using T-LiDAR measurements. To obtain the motion induced by tree growth, we compare curvature profile between 2007 and 2013.

***Key words**: T-LiDAR, motion, curvature, trunk, beech*

Introduction

The shape of a tree is the result of its acclimation to different environmental factors all along its life. The main external drivers are the light availability which is crucial for fulfilling the photosynthetic needs, conditioning the tree development, and more particularly the primary growth of apical meristems. In addition to this architectural development, the secondary growth induces passive effects through the increase of the tree biomass and its unbalanced distribution in interaction with gravity.

Against gravity effects, trees can react by controlling actively their posture by modifying the nature and the amount of wood produced by the cambium inducing reaction wood and ring width anisotropy respectively. These different mechanisms may result in change of curvature of the neutral line of the woody axes contributing to an uprighting or a sagging movement of the stem [1-2].

However, curvature data are really difficult to measure precisely in field work, time-consuming and poorly informative by traditional ways. At tree scale, terrestrial LiDAR (T-LiDAR) allow to gain these informations rapidly and precisely.

This paper is part of a study on tree biomechanics which one aim is the assessing local curvature along the trunk. Curvature profile are collected for 42 beech poles which growth during 60 years under canopy. Trees are thinned for winter 2007-2008. To study shape trees we realised T-LiDAR scanning during winter 2007-2008 and 2013-2014.

Analysis

First, trees were isolated by Faro Scene program. 3D points clouds are analysed by a cooperative program call Computree (http://computree.onf.fr/?lang=fr). By creating a “tree-steps”, cylinders are adjusted along the trunk (thickness = 10 cm). Then, one circle per cylinder were fitted. Thanks to a manual step, we selected circles which corresponding to the trunk and the dominant axis.

We used a specific plugin developed by the Lerfob’s laboratory to calculate curvature circles from 5 adjusted circles on both sides of the adjusted circle at the trunk height of interest.

The curvature assessment implied the calculation of inverse of the curvature circle radius. For each curvature circles, the corresponding height was exported with circles radius.

Results

To illustrate the kind of result obtained after T-LiDAR scanning analyses, we present an example for a control tree (no thinned) in Figure 1.

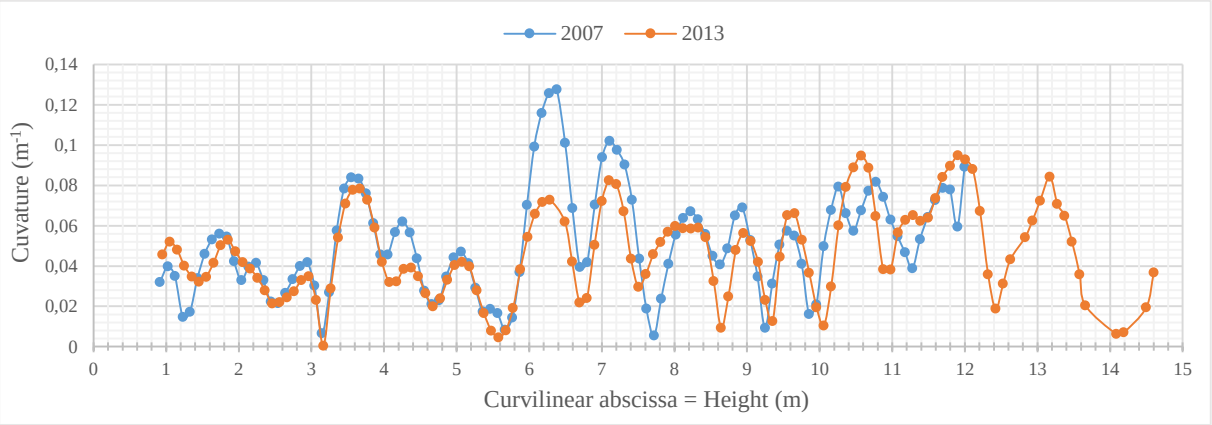


Figure 1: Example of curvature profile at 2007 and 2013 for a tree developed under canopy.

This tree are 18.50 m in 2007 and 19.60 m in 2013 tall. The method determine the curvature until 14.5 m for the last year. Because of T-LiDAR scanning quality, 2007 resolution is less precise than during the 2013 data acquisitions. Here, curvature values fluctuated between 0 and 0.14 m^{-1}.

This profile shows a difference in term of curvature intensity at 4.35, 6.30 and 7.20 m. The maximum of intensity difference is at 6.30 m. A wood sample was removed for futures measurements like materiel wood properties (wood reaction distribution or wood density measurements). In this way, we can identify wood properties which are modified specifically by curvature changes.

References

[1] Coutand, C., Fournier, M., & Moulia, B. (2007). The gravitropic response of poplar trunks: key roles of prestressed wood regulation and the relative kinetics of cambial growth versus wood maturation. Plant physiology, 144, 1166–80.

[2] Moulia, B., & Fournier, M. (2009). The power and control of gravitropic movements in plants: a biomechanical and systems biology view. Journal of experimental botany, 60, 461–86.

Detection of multi-layered forest development classes using airborne laser scanning

Rubén Valbuena[1], Matti Maltamo[1], Juho Heikkilä[2], Petteri Packalen [1]

[1] *University of Eastern Finland (UEF), School of Forest Sciences, PO Box 111. Joensuu, Finland*
rubenval@uef.fi, matti.maltamo@uef.fi, petteri.packalen@uef.fi
[2] *Finnish Forest Centre (Metsäkeskus), Aleksanterinkatu 18 A, FI-15111 Lahti, Finland*
juho.heikkila@metsakeskus.fi.

Highlights: Low-density national ALS surveys can reliably be employed to assess silvicultural development classes. However, shelterwood stands with shade-tolerant ingrowth showed high confusion with other mature forests, probably due to the lack of pulses reaching the understory. A rule-based approach using no field data obtained similar accuracies as the supervised classification.

Key words: *Lidar; forest structure; shelterwood; seed-tree; quadratic discriminant analysis.*

Introduction & Objectives

The main purpose of this research was to develop methods for automated detection of multi-layered development classes (DC) by airborne laser scanning (ALS) remote sensing. The methods developed in this research will be applied using the laser scanning data made publicly available by National Land Survey of Finland (NLS) and Suomen Metsäkeskus (SMK, Finnish Forest Centre) [1], consequently allowing consistent replication and potential country-wide upscaling.

Silvicultural development classes Finland

The types of silvicultural DCs which are multi-layered, and therefore the target of this research, are (codes denoting each class derive from their original names in Finnish language):

- Seed-tree stands (*S0*). Stands where few mature parent trees have been left after harvesting, and therefore seedlings grow along with them. It is used for natural regeneration of pine and birch, which is shade-intolerant and therefore mature seed trees (parent) have very low density.
- Shelterwood stands (*05*). It is used for natural regeneration of spruce, which is shade-tolerant and therefore the overstory has much higher density than in *S0*. Stands are usually even denser and surely more mature than *04* (see below). While containing trees at a mature overstory, which serves as shelter, it will also have a spruce-dominated understory of seedlings, and possibly saplings as well.
- Seedling stands with shelter trees (*Y1*). It is a multi-layered deciduous forest in which the understory has been established. While containing trees at a young overstory, which serves as shelter, it has and understory of saplings, possibly containing seedlings as well. *Y1* differs from *05* that in the degree of maturity of the shelter layer, and also that *Y1* regenerates deciduous and *05* does with spruce.

These are to be discriminated from those silvicultural DCs which are considered even-sized forest structures:

- Young thinning stands (*02*)
- Advanced thinning stands (*03*)
- Mature stands (*04*)

And, of course, they also have to be discriminated from non-forest areas and forest development classes which belong to the stand re-initiation phase at the start of a rotation in a silvicultutral management system based on clear-cut and planting:

- Seedling stands (*T1*)
- Sapling stands (*T2*)

Materials and Methods

Study area and remote sensing data

The research was conducted in a pilot study area in the region of North Karelia, surrounding Joensuu and extending to Outokumpu in the west and Pyhäselkä to the south. The total extent of the pilot included approximately 200,000 ha covered by forest. The ALS survey was appointed to Blom Kartta Oy (Finland). Laser data was acquired during May 2012 with an ALS60 system from Leica Geosystems (Switzerland). Flying at a height of 2300m above ground rendered an average density of 0.91 points per squared-meter. Country-wide laser data is being consistently acquired using broadly similar parameters [2].

Design of the field campaign

Field data acquisition for multi-layered development classes was adapted from the ordinary practice followed by SMK, which modifies plot size according to tree development (Figure 1). For Even-sized development classes, namely *02*, *03* and *04*, the field data was acquired over circular plots (sub-plot D1, Fig. 1). Plot size was reduced sampling stands, *T2* (sub-plot D2, Fig. 1). For seedlings stands, *T1*, species counts were carried out four distributed positions (sub-plot D3).

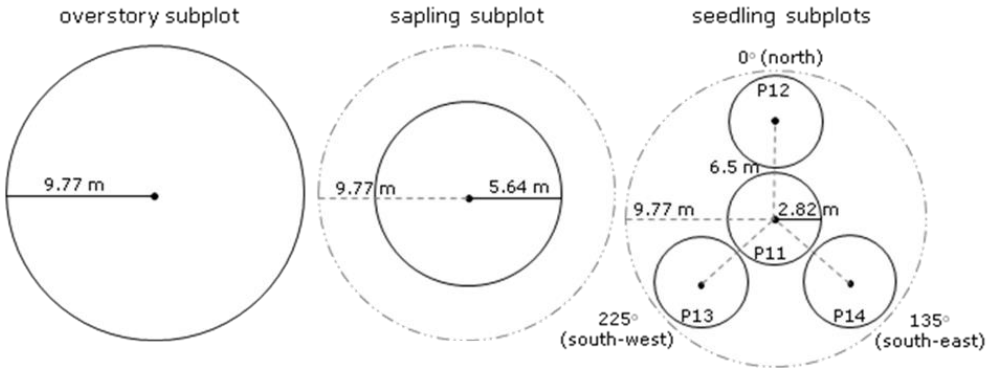


Figure 1: Field plot acqusition designed for multi-layered forest. Three sub-plots were measured using a same plot centre: D1 (left), D2 (middle) and D3 (right)

The adaptation to multi-layered development classes, namely *Y1*, *05* and *S0*, consisted principally on the possibility for measuring all these three plot types from a same plot centre (sub-plot D1, D2 and D3, Fig. 1). Plot centres were located by stratified random sampling, so that their positions were determined randomly within stands which belonged to the target DCs in the previous management plan. While stacking them out by high-grade GNSS, the diameters at breast height for each tree and its distance to plot centre were simultaneously measured. This way, the operator can determine on-the-go whether a tree belongs to its corresponding sub-plot. Following Valbuena et al. [3], all subplots were merged afterwards into one by direct repetition of the data acquired in the inner subplots.

Supervised classification

We conducted a supervised classification of all the possible DC (see above), using the ALS metrics as predictors. This approach was carried out as an alternative to using direct rules without any field data for training [4]. It was therefore possible to observe the increase in accuracy obtained when using a supervised classification, against the unsupervised approach by direct mathematical rules. After considering many alternatives for classifying the dataset (support vector machine, naïve Bayesian classifier, minimum volume ellipsoid estimator), a quadratic discriminant analysis (QDA) model based on maximum likelihood estimation was selected, since it provided highest accuracy.

Results and Discussion

As the supervised classification was calibrated with the field training plots, its accuracy was expected to improve the results obtained by a classification using direct mathematical rules [4], which we refer to as rule-based approach. The predictors were selected based on these previous findings obtained in a multivariate exploratory analysis. The field data protocol hereby presented was therefore employed in the selection of the variables, on a previous stage to the supervised classification. It allowed to calculate the Gini coefficient (GC) of tree size inequality [5], as well as the basal area larger than mean ($BALM$) [6]. The relations of these forest structure indicators with the target DCs [7] were first studied, and then the relations of these indicators with the ALS metrics [3]. It was then found that L-moments and canopy cover metrics have most explanatory power towards the targeted DCs, specially with regards to identifying multi-layered forests [4]. Consequently, we employed as predictors the first, second and third L-moments, and the proportion of ALS returns above their mean height.

Taking into account all the silvicultural DCs separately, the classification reached an overall accuracy of 66.9%, obtaining a coefficient of agreement $\kappa = 0.60$ (Table 1). However, there was no improvement with regards to multi-layered development classes, which were the forest areas actually targeted by this study. The supervised method obtained a 62.8% of correctly classified multi-layered-development classes (*S0*, *Y1*, *05*), whereas the rule-based method [3] reached 63.8%. Plots located at even-sized development classes were classified with a higher accuracy of 61.7%, whereas 75.4% of seedling and sapling stands were correctly identified.

Table 1: Contingency table of classification obtained by quadratic discriminant analysis including all the silvicultural development classes (codes are explained in the introduction).

Predicted	Observed								
	T1	*T2*	*02*	*03*	*04*	*05*	*Y1*	*S0*	totals
T1	35	52						6	13.7%
T2	1	152	4					4	23.7%
02		7	82	25	1	1			17.1%
03			12	83	17	3			16.9%
04				7	43	4		1	8.1%
05			11	19	16	20	2		10.0%
Y1			4	7	3	3	12	11	5.9%
S0		1						27	4.6%
Totals	5.3%	31.2%	17.1%	20.8%	11.8%	4.6%	2.1%	7.2%	679

To allow direct comparison of the supervised method against the rule-based method, classes in Table 1 were also be aggregated into two groups: even-sized (*T1*, *T2*, *02*, *03*, *04*) and multi-layered (*S0*, *Y1*, *05*). This aggregation showed an overall accuracy of 87.7% with a $\kappa = 0.57$ in discriminating multi-layered DCs from even-sized area. This result is very similar to that obtained by applying a simple rule based on the L-coefficient of variation [3], which by itself reached an overall accuracy of 87.0% and $\kappa = 0.53$ without the support of field data.

The maps obtained from maximum likelihood classification of all development classes showed good results that would, however, require segmentation procedures for their analysis at stand level. Mature dense forests showing large tree size inequality were well detected. The presence of highest values of tree size inequality in the presence of seed and retention trees, as well as stand edges, was easily observed in the resulting maps.

Although the supervised approach requires the use of field data, it had the advantage of yielding a classification calibrated to the pilot area and the actual target DCs used in silvicultural practice in Finland. The method is can therefore be replicated elsewhere by tailoring the definition of DCs to other silvicultural system, and training with local field plots. The field protocol provided can be employed for increasing the understanding on the relations between the ALS metrics and the characteristics of the diameter distributions that can be used to discriminate DCs [3]. However, once these relations are clear, field effort can be considerably reduced by simply visually classifying forest plots within either DC, using expert knowledge in the field.

The conclusions reached in this research were:

- The low density ALS survey available in the NLS data was demonstrably sufficient for discriminating multi-layered development classes from even-sized ones.
- Seed tree stands and, in general, forest areas where light availability is not a limitation, may be detected directly by simple rules [4], without the need of field data.
- Multilayered shelterwood mature stands, on the other hand, benefited from employing a discriminant analysis based on maximum likelihood classification, as the low-density nature of the NLS dataset may hamper the discrimination between *04* and *05*.

Acknowledgements

This research was funded by Suomen Metsäkeskus. Special thanks to Jussi Lappalainen (SMK), Heli Laaksonen (MML) and Aki Suvanto (Blom Kartta Oy) for their support.

References

[1] NLS - National Land Survey of Finland 2013. Laser scanning data (available online at maanmittauslaitos.fi). Visited in Sep. 2013.

[2] Ahokas, E., Hyyppä, J., Yu, X. & Holopainen, M. (2011) Transmittance of airborne laser scanning pulses for boreal forest elevation modeling. Remote Sensing 3: 1365-1379

[3] Valbuena, R., Maltamo, M., Martín-Fernández, S., Packalen, P., Pascual, C., & Nabuurs, G.J. (2013). Patterns of covariance between airborne laser scanning metrics and Lorenz curve descriptors of tree size inequality. Canadian Journal of Remote Sensing, 39, S18–S31

[4] Valbuena, R., Mehtätalo, L., Maltamo, M., & Packalen, P (2015) A rule-based method using L-moments for direct detection of shade-intolerant regeneration in Boreal forests from airborne laser scanning. Silvilaser 2015 Conference, La Grande Motte, France.

[5] Gini, C. (1921) Measurement of inequality of incomes. Economic Journal, 31,124–26

[6] Gove, J.H. & Patil, G.P. (1998) Modelling the basal area-size distribution of forest stands: a compatible approach. Forest Science 44 (2): 285–297.

[7] Valbuena, R., Packalen, P., Mehtätalo, L., García-Abril, A., & Maltamo, M. (2013) Characterizing forest structural types and shelterwood dynamics from Lorenz-based indicators predicted by airborne laser scanning. Canadian Journal of Forest Research, 43, 1063–1074

Towards UAV based laser scanning for forest mapping

Dr. Antoine Cottin[1], Sam Fleming1, Dr. Tristan Allouis[2]

[1]*Carbomap Ltd., Edinburgh, a.cottin@carbomap.com, s.fleming@carbomap.com*
[2]*L'Avion Jaune, tristan.allouis@lavionjaune.fr*

Highlights: Evaluation of the potential for a new kind of LiDAR sensor to retrieve topography and forest parameters. This study was conducted in different forest types : mountainous forest, tropical rainforest, and mixed temperate forest. The results were compared to traditional airborne survey, terrestrial laser scanning, and ground control points.

Key words: UAV, LiDAR, forest, management

Introduction

Interest in the use of Unmanned Aerial Vehicles (UAVs) has seen an unprecedented rise over the past year. Within the forest mapping industry, this is in part by the first entries to the market of purpose-built, full LiDAR systems, which provide new monitoring opportunities. The ease of use, and low cost of purchase and deployment, mean that UAVs have quickly become an attractive solution [2]. This combination of technologies provides very high point density LiDAR data which is collected at lower costs. Although UAVs are limited to a small-scale approach compared to traditional airborne surveys, this provides the opportunity for further cost reductions in data collection as key areas of interest can be targeted with rapid deployment. This solution therefore opens up LiDAR monitoring to small forest projects due to the improved economic feasibility. However, as the technology is so cutting edge, there is a requirement to fully test its capabilities. This work focuses on one of the recently released UAV-LiDAR instruments, YellowScan. The aim is to understand if and how UAV based Lidar is suitable for forest mapping. To do so, the amount of echoes that actually reached the ground is assessed for different type of forest, as well as the accuracy of the resulting digital terrain model. After computing a Canopy Height Model (CHM), some wildly used forest metrics such as tree height and basal area are retrieved and their accuracy assessed relatively to traditional airborne lidar, terrestrial laser scanning and ground truth.

Study Sites

This project tested the YellowScan system, discussed in the next section, on a number of different forest types:

1. Saint Jean-de-Couz, France. This forest area is located in the French Alps, and is predominantly composed of spruce and common beech tree species. This location has a range of different sized and shaped trees, with varying tree density. The location is under management by the Office National des Forêts (ONF).

2. Paracou Research Station, French Guiana. This tropical rainforest is highly diverse (typically more than 200 tree species per ha) and has a dense canopy and a mean canopy height of 30m. The relief of the area is mild. The research station is part of the labex CEBA and is managed by CIRAD.

3. Barnaby Slough, Washington State, USA. This location is a temperate mixed-species forest, with data collected during the leaf off season. The terrain at this location is relatively flat. The Skagit River System Cooperative manage the forest, and the study site is part of a floodplain restoration program.

Materials

YellowScan® LiDAR system

YellowScan [3] is an all-integrated ultra-light laser scanner system intended for UAVs and other ultra-light aircrafts, created by L'Avion Jaune company. With a weight of 2.2 kg, and compact dimensions (20 x 17 x 15 cm). It incorporates a laser scanner operating at 905nm, an inertial measurement unit of 0.25° accuracy, a bi-frequency GNSS receiver, an on-board computer and a battery for 3 hours operation. The system has a low power consumption (12 Watt), and operates at 40kHz. YellowScan can operate up to 100 m above ground level with a resolution of 10 cm. It has a typical scan angle of 100° and can record up to 3 echoes per pulse, allowing topographic information under vegetation cover to be extracted.

Airborne/UAV Surveys and Reference Datasets

During the early stage of development of the YellowScan system, surveys were carried out using manned aircrafts. For Saint Jean-de-Couz, the survey was conducted using an autogyro flown at 80 km/h at 75m AGL, which provided a density comprised between 5 pts/m² and 10 pts/m². The reference dataset at this location was collected using Terrestrial Laser Scanning, FARO Focus3D, which were collected at 5 plots. Auxiliary ground truth data, in the form of tree position, tree height, and Diameter-at-Breast-Height (DBH) for 1064 trees, was also provided.

At Paracou, the LiDAR system was attached to a manned helicopter flown at 40 km/h and 50m AGL, which resulted in a point density comprised between 10 pts/m² and 20 pts/m². A traditional airborne LiDAR survey was also conducted for this site, Riegl LMSQ560, which were used to provide an additional raw point cloud, which was subsequently classified and a Digital Terrain Model (DTM) generated. Ground control points were also collected, as an auxiliary source of ground truth data.

For the Barnaby Slough survey, the system was fully developed and functional, therefore it was mounted on an octocopter from Drone-Motion, flown at 20 km/h and 50 m AGL providing a density comprised between 30 pts/m² and 50 pts/m². The ground truth data for this location was in the form of 15 ground control points, collected using a Differential GPS system.

Methods

Laser penetration through the canopy

The ability of the YellowScan system to penetrate through the forest canopy was assessed at all three locations. This was achieved by classifying the raw point cloud using Lastools software into ground and non-ground echoes. A ratio between the number of ground echoes and number of emitted pulses is then retrieved.

Digital Terrain Model (DTM)

Once the ground layer was extracted from the classified point cloud, a Digital Terrain Model was generated. Due to a poor absolute accuracy in the ground control points for the Saint Jean-de-Couz and Paracou locations, only the accuracy of the DTM for the Barnaby Slough site was assessed. Indeed, the original goal of this survey was to provide an accurate topographic model for the local hydrological restoration project. The vertical accuracy of the LiDAR-DTM was assessed by making a comparison to corresponding ground control points.

Canopy Height Model (CHM) and metrics extraction

Once the DTM was extracted, the maximum height of the trees above the ground layer was computed and the CHM was generated. A segmentation algorithm was then applied to identify individual trees and retrieve their maximum height.

Individual tree heights have been manually compared to ground truth over the Saint Jean-de-Couz site, to overcome absolute positioning issue. Terrestrial laser scanning dataset was used in a similar way to assess the quality to tree height retrieval from YellowScan data.

Over Paracou site, a similar process has been applied to the Riegl data set and the results have been compared to YellowScan data.

Extraction of diameter at breast height (DBH) and basal are (BA) a have been tried, based on site-specific allometry. Results have been compared to ground truth over the Saint Jean-de-Couz site.

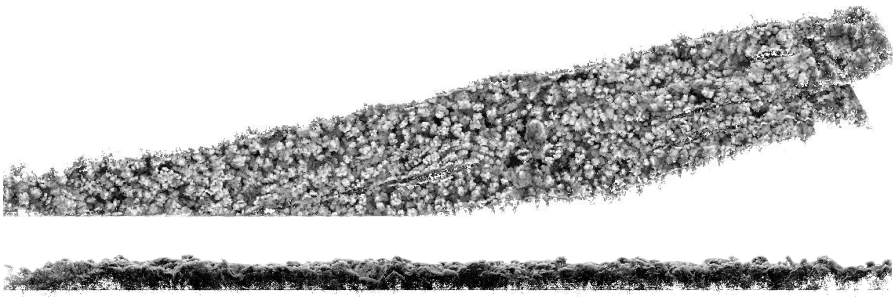


Figure 1: Canopy Height Model from the Paracou site. Upper image is top view, lowere image is right side view.The grey scale is from an Ambient Occlusion shader computed on the points cloud..

Results

Laser penetration through the canopy

For the Paracou location, ground points represent only 0.25% of the YellowScan hits compared to Riegl 5%. It implies one point every 2.1 meters.

Canopy penetration at the Barnaby Slough location is 24% representing approximately 9.6 ground points per square meter.

Canopy penetration at the Barnaby Slough location is 35% representing approximately 1 ground point per square meter.

The Paracou location returned the lowest amount of ground points. This is because of the almost complete canopy coverage, and highly complex forest.

Digital Terrain Model (DTM)

The achieved accuracy for the Paracou location had a total RMSE of 0.1m, a bias of 0.01m, and a standard deviation of 6.9m, in the vertical direction, compared to the DTM extracted from the airborne LiDAR campaign.

A comparison between the ground control points and the YellowScan LiDAR derived DTM showed that the accuracy for Barnaby Slough was 5.18cm ± 6.7cm.

Canopy Height Model (CHM) and tree metrics

A comparison of the CHM from YellowScan LiDAR data and the TLS data shows for the highest trees an average vertical difference of 0.47 m ± 2.36 m. A comparison of the tree's height between the YellowScan LiDAR data and the actual ground data from ONF shows that the mean elevation difference is 0.11 m ± 8.10 m.

For the DBH, ONF's ground data shows an average of 30.17 cm ± 18.25 cm whereas the estimate DBH from YellowScan LiDAR data shows an average of 77.98 cm ±15.99 cm. The BA from ONF data is 0.09 m^2 ± 0.10 m^2 versus 0.49 m^2 ± 0.18 m^2 for YellowScan LiDAR data.

Discussion

This study was limited by the absolute accuracy of YellowScan data at the early stage of development, when both Paracou and Saint-Jean-de-Couz site were surveyed. As a consequence, a detailed analysis of DTM accuracy has not been performed, despite the availability of ground control points.

UAV lidar, that require compromise between weight and power to be take on-board small UAVs, uses very low power laser. As a consequence, tropical forests are very challenging sites. The high foliage density of the forest itself limits the ease of ground identification. However the ground return was very low, its amount was enough to compute a rough DTM and extract tree heights at accuracy comparable to traditional airborne systems. For other sites, the amount of ground points was enough to compute a DTM at a resolution of 1m. This is particularly true when YellowScan was flow at low speed using a UAV. In this case, the density of ground points overcomes the density generally provided by classical airborne lidar. For the DBH and BA estimation, the results are generally over estimate compare to the ground data. These discrepancies can be attributed to; positioning inaccuracy, the number of echo per shot recorded, forest density, and the data processing workflow.

The next stages in a forest mapping workflow would be to extract other forest metrics from the data. After tree identification and crown segmentation, estimation of aboveground biomass and carbon would have been performed [4]. If multiple datasets over time are available, then the change in forest cover, from which changes in forest carbon, can also be measured. However, as no ground truth data was available for these more advanced metrics, we have not included them in the evaluation here.

Conclusion

YellowScan was able to collect terrain elevation with a vertical accuracy of less that 10cm, which shows that it is possible to extract a sufficient number of ground echoes for assessing forest metrics, even in one of the densest forests in the world.

It can also be concluded that the Multi-Rotor Copter UAV is well suited for small-scale, very high point density, surveys. However a Fixed-Wing UAV would be better suited to larger scale projects, where a greater coverage and lower cost per hectare are required, at the expense of a lower point density. However, even with the fixed wing the point density is higher than in traditional airborne LiDAR surveys. It is well known that LiDAR can be processed effectively for the extraction of individual trees [1], and the higher point density offered by UAV-LiDAR this process is expected to improve this process. The extraction of typical forest metrics was applied to the data used for this assessment, demonstrating that it is indeed possible, the next steps would be for a full evaluation of UAV-LiDAR for this purpose in comparison to traditional Airborne LiDAR, with rigorous ground truthing to quantitatively describe its performance.

References

[1] Clark, M.L., Roberts, D.A., Ewel, J.J., & Clark, D.B. (2011). Estimation of tropical rain forest aboveground biomass with small footprint lidar and hyperspectral sensors. Remote Sensing of Environment, 115, 2931–2942

[2] Cottin, A., Fleming, S., & Woodhouse, I.H. (2014). Pushing the Boundaries of UAV LiDAR for Forestry Applications. ELMF 2014, Amsterdam, 8-10 December 2014

[3] YellowScan, http://yellowscan.lavionjaune.com/home2/

[4] Allouis, T., Durrieu, S., Véga, C., & Couteron, P. (2013). Stem Volume and Above-Ground Biomass Estimation of Individual Pine Trees From LiDAR Data: Contribution of Full-Waveform Signals. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 6 (3), 924–934

Computation of tree volume from terrestrial LiDAR data

Jules Morel[1], Alexandra Bac[2], Cédric Véga[3]

1 : French Institute of Pondicherry, Pondicherry, India.
2: Laboratoire des Sciences de l'Information et des Systèmes, 163 AV. de Luminy, 13288 Marseille, France
3 : Institut National de l'Information Géographique et Forestière, Laboratoire de l'Inventaire Forestier, 11 rue de l'Ile de Corse, 54000 Nancy, France

Highlights: This paper introduces a new method for the detailed estimation of volumes of individual trees from terrestrial laser scanner data. This method is based on adapted compactly supported radial basis function models together with a deformable models approach.

Key words: *terrestrial lidar, deformable models, poisson surface reconstruction, forest plots, volume estimation.*

Introduction

Tree volume is a key parameter of forest inventories. Its estimation usually relies on specific volume equations based-on diameter at breast height (DBH) and height which were derived from localized and costly destructive sampling. Terrestrial lidar (TLS) technology allow to acquiring very high resolution and accurate 3D data of forest scenes [1], thus providing a promising way to acquire volume information from non-destructive sampling and at a low cost.

Methods developed to estimate tree volume from TLS data are generally based-on the fitting of cylinders along the trunk and the main branches [2-4]. Such methods consider the tree as a set of perfect cylinders, which is far from being true and might leads to substantial errors. Also these methods often require a pre segmentation of the point clouds in order to process each tree individually [4].

Recently, [5] introduced a Modified Hough Transform Approach (MHTA) to locate and reconstruct individual tree within a scanned forest plot. The method use principal component analysis to simultaneously reduce noise, simplify the 3D point cloud and compute approximate normals. Taking advantages of the computed normal, an original circle Hough transform is then applied to detect tree stems. While the method provide information about both tree position and structure, it still suffers from issues related to geometric fitting, and is not adapted to complex tree structures like those found with the tropics.

Here, we tackle this issue by introducing an innovative method dedicated to reconstruct trees. The method uses MHTA to detect and segment individual trees in forest plots. Then, a continuous wrapping surface is adjusted around the data points distributed on the trunk and on the big branches.

Preliminary work

The Poisson Surface reconstruction (PSR) [6] approximates a point cloud by an implicit surface whose gradient fits the normals of the points. The method was found to perform well on dense and continuous data such as the lower part of a trunk and the biggest branches of a well scanned tree (Figure 1- a). But its performance decreases with the point density, occlusion rate and the object size (Figure 1- b). Hence, the method as such is not appropriate for real data collected in larger plots in natural forested environment.

To overcome this flaw, our idea is to combine approximation of the points when the density is high enough together with constraints on the surface shape (cylindric-like shape and low curvature) in occluded areas. We propose to implement constrained deformable models on implicit functions to fill the occlusions by automatically reconstructing missing data.

Study Design

First approximation

The starting point is the set of cylinders produced by the MHTA. This algorithm allows to extract each tree from the scene in order to process them individually, plus, it provides a segmentation of the point cloud in bounding boxes, one for each cylinders. In each bounding boxes, points are approximated by a cylindrical quadric (elliptic cylinder, ellipsoid or hyperboloid). Then, we extend the local reconstruction to the whole space

by a compactly supported radial basis functions (CSRBF) partition of unity [7]; actually, we merge those local implicit approximations using such CSRBF functions. The zero levelset of the resulting merged implicit function is extracted with a marching cube-like algorithm [8]. Hence, the trunk and big branches are modelled by a continuous surface (a zero level set), visualized as a triangular irregular network (TIN) (Figure 1- c).

The surface produced is a good approximation in the occluded areas, but it is a coarse one close to data points and can be improved. Therefore, we use this first approximation to initialize a deformable model to push the surface towards data points while minimizing the curvature energy to guarantee the reconstruction in the occluded areas. Two paths have been explored. Both of them use the previous CSRBF model as an initialization. The first one is based on a constrained minimisation of energy on level set surfaces, inspired by the deformable model of [9]. The second one is based on the PSR where the resulting surface is constrained by our first approximation.

Deformable model

In this method, we define the energy of the surface as a Mumford-Shah functional. Its minimization guides our deformable model. In practice, CSRBF centers are set in the centers of the leaf of an octree structure refined along the first approximating surface and the around the data points. For each of these centers we compute a deformation vector that will lead to the velocity of deformation in the deformable model (see Figure 1 - d). The surface is then deformed iteratively.

Modified Poisson surface reconstruction

In the classic PSR, the surface arises from an indicator function whose gradient best approximates the vector field defined by the normals of the points. Our idea is to add the influence of our first approximation surface in the computation of the vector field. By computing the indicator function from this gradient field, we fit accurately the sample data while filling the holes in the occluded areas.

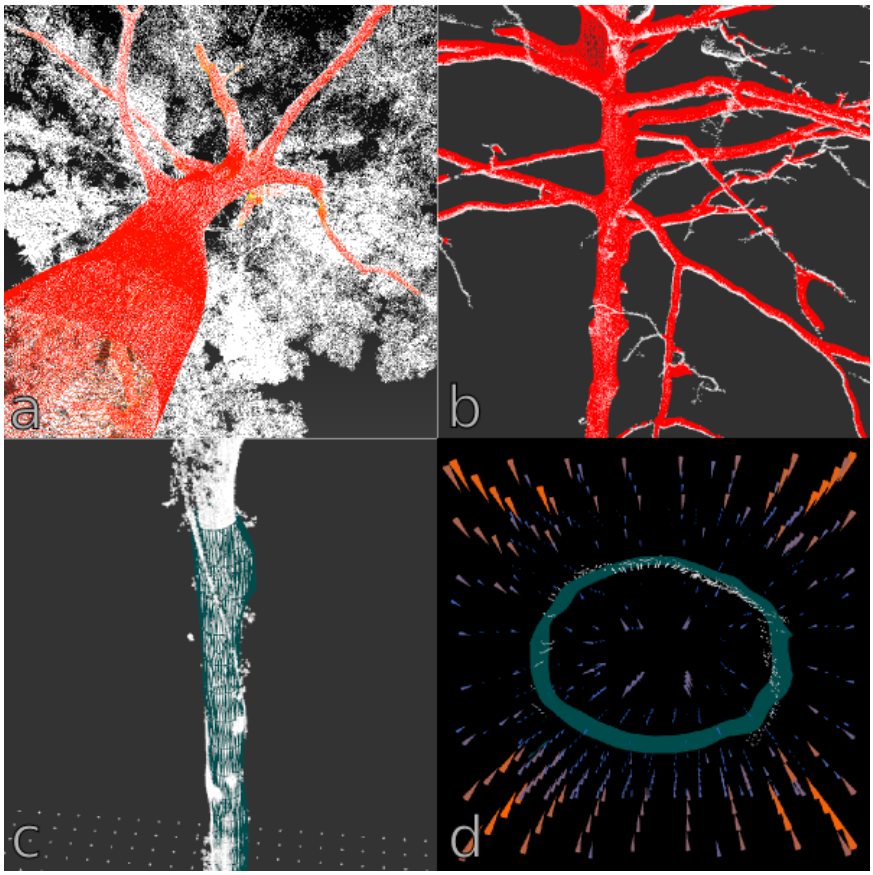


Figure 1: a) and b) provided example of traditional PSR reconstruction for leaf-on and leaf-off tree conditions, showing evidence in reconstruction quality due to leaves and low point densities. An example of CSRBF reconstruction is shown in c) and d) shows a sample of deformation vectors along a tree trunk.

Experimental and first conclusions

The method has been tested on data acquired in a tropical context (© Nicolas Barbier, UMR AMAP). The input of the algorithms was single trees, as the MHTA produces. At this stage of development, only a visual analysis of the result has been effective. As expected, our preliminary results demonstrated that the modified PSR is less sensitive to both the point density and object size than the original PSR did. A quantitative evaluation will be conducted using field data for the tree circumference and from simulation for the whole tree structure. Because is allow to approximate the real structure and the wooded part of the tree, the method has a high potential for tree volume computation, and the improvement of volume equations.

References

[1] Dassot, M., Colin, A., Santenoise, P., Fournier, M., & Constant, T. (2012). Terrestrial laser scanning for measuring the solid wood volume, including branches, of adult standing trees in the forest environment. Computer an electronics in Agriculture, 89: 86-93.

[2] Delagrange, S., Jauvin, C., & Rochon, P. (2013). PypeTree: A Tool for Reconstructing Tree Perennial Tissues from Point Clouds. Sensors, 14, 4271–4289.

[3] Lefsky, M.A., & McHale, M.R. (2008) Volume estimates of trees with complex architecture from terrestrial laser scanning. Journal of Applied Remote Sensing, 2, 023521.

[4] Pfeifer, N., B. Gorte, and D. Winterhalder. 2004. Automatic reconstruction of single trees from terrestrial laser scanner data. *Int. Archives of Photogrammetry and Remote Sensing*, Vol. XXXV, B5, pp. 114-119.

[5] Bac, A., Ravaglia, J. Morel, & Véga, C. (2015). Detection of single trees in T-LIDAR scans of dense forest areas. LSIS research report. 2013.

[6] Kazhadan M., Bolitho M., Hoppe H., 2006. Poisson surface reconstruction. SGP '06 Proceedings of the fourth Eurographics symposium on Geometry processing, p 61-70.

[7]Ohtake, Y., Belyaev, A., Seidel, H., 2004. 3D Scattered Data Approximation with Adaptive Compactly Supported Radial Basis Functions, in: Shape Modeling Applications.

[8] Bloomenthal, J., 1994. An implicit surface polygonizer. In Graphics Gems IV, Academic Press, 1994, pp. 324–349.

[9] Gelas A., Bernard O., Friboulet D., Prost R., 2007, Compactly Supported Radial Basis Functions Based Collocation Method for Level-Set Evolution in Image Segmentation. Transactions on image processing, vol. 16, no. 7

Algorithms and Tools for Point Cloud Based Vegetation Studies - The OPALS Forestry Package

Lothar Eysn[1], Gottfried Mandlburger[1], Markus Hollaus[1], Johannes Otepka[1] and Norbert Pfeifer[1]

[1]*Vienna University of Technology, Department of Geodesy and Geoinformation, Research Group Photogrammetry, Gußhausstraße 27-29, A-1040 Vienna, Austria (Lothar.Eysn, Gottfried.Mandlburger, Markus.Hollaus, Johannes.Otepka, Norbert.Pfeifer@geo.tuwien.ac.at)*

Highlights: The operationalization of five different applications in forestry within the scientific software package OPALS is presented. The OPALS forestry package enables handling and processing of large point cloud data in a research as well as in a production environment.

Key words: LiDAR; forest delineation; biomass estimation; scientific software, point cloud processing

Introduction

Recent developments in LiDAR and image matching technology allow precise and fine quantification, qualification and mapping of the resource forest. The increasing availability of 3D data derived from these technologies opens new opportunities for large area applications in the forestry domain. Different tasks as, for example, the delineation of forested areas, the evaluation of forest biomass, the detection of forest change or the extraction of single tree information are already close to an operational status, also for forests composed of a mixture of species..

While the acquisition of 3D data is more or less state of the art, software for processing these data is still developing. Efficient algorithms and software tools are needed to tackle the above mentioned tasks. In many cases, operational software tools and their corresponding algorithms originate from a scientific environment.

In this study, the scientific software package, OPALS (Orientation and Processing of Airborne Laser Scanning data) [1] for processing of 3D point cloud data is presented with focus on forestry and vegetation applications. This software is a means of sustainably providing new algorithms to researchers, but is also fit for efficiently handling and processing large data in a production environment. The study concentrates on the operationalization of five different tasks: 1) reconstruction of a canopy height model (CHM) optimized for forested areas 2) detection of single trees, 3) delineation of forested areas, 4) estimation of stem volume and 5) forest change detection based on multi-temporal data. All presented algorithms were developed in-house.

OPALS

The requirements to modern software are manifold. Mass production as well as demanding research should be ideally enabled in different scripting and programming languages. The software OPALS [2] is a modular, scientific software system consisting of small well defined components (modules). The aim of the software is to provide a complete processing chain for processing LiDAR data including: waveform decomposition, georeferencing, quality control, point cloud classification, topographic modeling and several fields of applications like for example forestry and hydrology. Each module is available as (i) a command line program,(ii) a Python module, and (iii) a C++ class (via DLL linkage). While the command line executables allow straight forward execution of the modules from within the operating systems command prompt, the latter enable embedding OPALS functionality in high level programming and scripting languages, allowing for the construction of complex work flows. Above all, the python integration has proven its worth as Python combines the advantages of an easy to use scripting and a full featured programming language.

A central data administration unit (OPALS Data Manager, ODM) is available for efficient data management. Tile-wise kd-trees are built up allowing high performance spatial queries (k-nearest neighbors and fixed distance queries). Additionally, the ODM also features an administration scheme for storing arbitrary attributes on a per point basis. The attributes may either stem from the initial sensor data analysis (e.g. amplitude, echo width, etc.) but may also be derived during data processing (e.g. local 3D metrics, surface normal vectors, etc.). Thus, the additional information system is highly dynamic and can, therefore, be used to transfer information between different modules without the need for external storage of attributes.

Land cover dependent surface model

Reliable modeling of the forest canopy is needed when it comes to forestry related applications. As the canopy height is very important for many forestry applications, it must be ensured that the derived digital surface models (DSM) follow the uppermost part of the canopy but also the terrain in forest gaps. Within the OPALS

forestry package, a land cover dependent approach described in Hollaus, *et al.* [3] is implemented for the derivation of a DSM. This approach uses the strengths of different algorithms for generating a DSM by using surface roughness information to combine two DSMs, which are calculated based (i) on the highest point within a raster cell and (ii) on moving least squares interpolation with a plane as functional model (i.e. a tilted regression plane is fitted through the k-nearest neighbors). As the surface roughness is high in forests, the created models correctly follow the uppermost canopy within the forest as the highest point per cell method is used. Additionally, local height jumps at forest gaps and forest borders are preserved and show sharp edges. In contrast, for the modelling of the open terrain and for cells with no points the moving least squares interpolation is used, which enables smooth results. Finally, a CHM is calculated by subtracting the heights of a digital terrain model (DTM) from the DSM.

Single tree detection

The identification of single trees and their parameters is an important task for analyzing large forested areas. The single tree detection can be performed based on rasterized LiDAR data or directly in the 3D point cloud. In OPALS forestry, both of these options are available. Focal statistics using different kernel sizes and shapes (i.e. circular, diamond) can be applied to rasterized LiDAR data (i.e. CHM), to detect local maxima in the model. Height thresholds and a forest mask limit the detected positions to forested areas. Alternatively, based on the 3D point cloud, local point cloud metrics can be calculated from the data for each point. This means, that the distribution of a selected point attribute (i.e. height above terrain) within a local neighborhood is analyzed. Statistical features (i.e. max, quantiles) are derived and stored as additional attributes of the current point. The point neighborhood can be defined as an infinite vertical cylinder (2D), a bounded vertical cylinder (2.5D) or a sphere (3D). The selected statistical features are calculated by aggregating the signed attribute deviations of the feature point and its neighbors from a reference model. Different reference models (i.e. horizontal or tilted plane, raster model) are handled. For the single tree detection, the normalized point cloud is analyzed for the statistical feature "Max" using a 2.5D neighborhood, and a horizontal plane with z=0 as a reference model (Figure 1).

Forest delineation

As presented in Eysn, *et al.* [4], forested areas can be automatically delineated based on LiDAR data by subsequently checking four clearly defined geometrical criterions (area, height, width and crown coverage) of a forest definition. The method is implemented using the Python bindings of OPALS. The following steps are carried out: The minimum height criterion is applied by height thresholding of rasterized CHMs. Buildings and other artificial objects are previously removed from the scene by using pre-derived point cloud metrics and image morphology. For the crown coverage calculation the 'tree triples' approach as described in Eysn, *et al.* [4] is used, where single trees are detected, crown sizes are estimated and a Delaunay triangulation is used to connect the trees. The minimum area criterion is applied by using standard GIS-queries on all areas fulfilling the previously checked criterions. The minimum width criterion is applied by using morphologic operations (open, close) based on the intermediate result fulfilling the criteria height, crown coverage and area. Because the width criterion can change the given areas, an iterative process of checking minimum area and width is applied.

Stem volume estimation

For estimating the growing stock, the method described in Hollaus, *et al.* [5] is available in the OPALS forestry package. The method assumes a linear relationship between the growing stock and the LiDAR derived canopy volume, stratified according to different canopy height classes to account for height dependent differences in the canopy structure. The estimation process consists of two general steps, where (i) the linear model is calibrated based on forest inventory (FI) data and LiDAR data within well known areas (i.e. forest inventory plots) and (ii) the calibrated model is applied to large areas. For the calibration process, information within canopy height classes, in detail the canopy volume within a certain height layer ($v_{can,i}$), is obtained by processing zonal statistics on the point cloud or rasterized CHM models. The locations of the zones are defined by polygons (i.e. shape file). For each zone, the proportion of LiDAR points or raster cells within the corresponding canopy height class is analyzed and linked to the $v_{can,i}$. Using the given FI data and the obtained $v_{can,i}$ values, the model coefficients are calibrated using the software package R [6]. Statistical measures and graphical representations are derived to enable an insight about the quality of the calibration. Additionally, stratified models can be calibrated to account for local forest changes. For example, crown cover classes can be used to calibrate separate stem volume models. Finally, area-wide maps can be calculated by checking the local heights of the CHM and multiplying them with the corresponding coefficients. The resulting maps can be processed in different spatial resolutions (i.e. 1x1m or 10x10m). Using zonal statistics, the information can be aggregated to existing forest stand polygons.

Forest change detection

Multi-temporal point cloud data can be used to monitor forest change, which can be, e.g., a change in forested area, harvested areas inside the forest or a change in growing stock. Changed forested areas can be

identified by comparing different forest delineations of different epochs. The identification of harvested areas can be performed by comparing the DSM heights of different epochs. Based on the estimated stem volume, derived from multiple datasets, the changes in growing stock can be assessed. A mandatory initial step is to test and enhance the quality of the input data. Especially height differences of stable objects between multiple surface models originating from errors in the georeferencing may occur. In OPALS, least square matching (LSM) of smooth surfaces (i.e. open terrain or buildings) can be used to minimize these errors between multiple models. For differentiating between exploitation and forest growth the area of interest is classified into areas with an (i) increased (= forest growth) and (ii) decreased (= exploitation) surface height. As for each LiDAR data set small differences in the tree crown representation within the DSMs can occur, morphologic operations and a minimum mapping area are applied to a DSM difference map. For each classified area (exploitation, forest growth), the change for the assessed growing stock is analyzed separately. The transferability of the calibrated models from one acquisition time to the other is tested by applying the calibrated models to the LiDAR data, which were not used for the calibration. Finally, the derived growing stock maps are validated with the corresponding FI data.

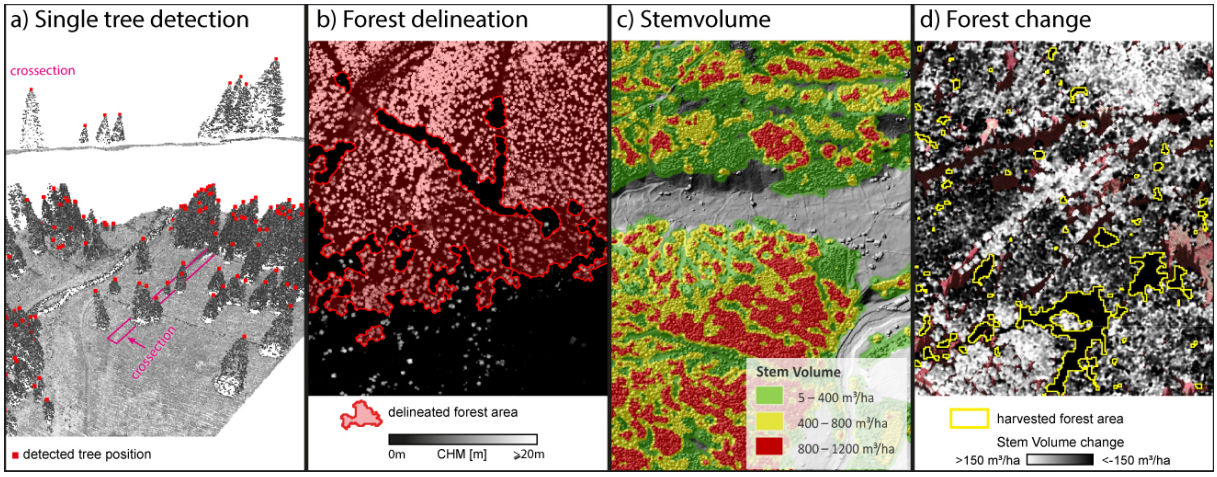


Figure 1: a) Detected single tree positions based on the 3D point cloud, b) Delineated forest areas based on LiDAR data c) aggregated stem volume map and d) Detected forest changes based on matching data

Conclusion

The basic concept of the comprehensive LiDAR software OPALS representing a sustainable basis for future research work and operational use was presented. A series of forestry related research topics were presented, considering both, their theory and the embedding into OPALS. All processes are fully automated, take care of the data handling and therefore enable operational use. The methods were already successfully tested and applied to large areas as i.e. Voralberg, Austria (2600 km²) or Canton Bern, Switzerland (6000 km²). The forest delineation shows a high performance and delivers robust and comprehensible results. The method can be applied to different forest definitions, which is significant when the results are used for obligatory reporting. The presented single tree detection enables processing of different point cloud metrics using the powerful attribute scheme of OPALS. The stem volume estimation and forest change detection opens up new possibilities for a sustainable forest management. The OPALS forestry package will be expanded with new tools in the future, to ensure that research driven methods will be available for operational use.

Acknowledgements

This work was partly funded by the EU FP7 “Advanced_SAR” project.

References

[1] OPALS Orientation and Processing of Airborne Laser Scanning Data. http://geo.tuwien.ac.at/opals

[2] Pfeifer, N.; Mandlburger, G.; Otepka, J.; Karel, W., OPALS–A Framework for Airborne Laser Scanning Data Analysis. Computers, Environment and Urban Systems 2014, 45, 125-136.

[3] Hollaus, M.; Mandlburger, G.; Pfeifer, N.; Mücke, W. In Land Cover Dependent Derivation of Digital Surface Models From ALS Data, ISPRS Comm. III Symp. PCV2010, Saint-Mandré, France, 2010; pp 6-6.

[4] Eysn, L.; Hollaus, M.; Schadauer, K.; Pfeifer, N., Forest Delineation Based on Airborne LIDAR Data. Remote Sensing 2012, 4, 762-783.

[5] Hollaus, M.; Wagner, W.; Schadauer, K.; Maier, B.; Gabler, K., Growing Stock Estimation for Alpine Forests in Austria: A Robust LiDAR-Based Approach. Canadian Journal of Forest Research 2009, 39, 1387-1400.

[6] R: A Language and Environment for Statistical Computing. http://www.R-project.org

A new automatic coarse-to-fine registration method of terrestrial LiDAR measurements in forest areas

Yiming Chen, Wuming Zhang, Hongtao Wang, Donghui Xie, Guangjian Yan

State Key Laboratory of Remote Sensing Science, Beijing Key Laboratory of Environmental Remote Sensing and Digital City, School of Geography, Beijing Normal University, 100875 Beijing, China.
cym_bnu@mail.bnu.edu.cn

Highlights: A new coarse-to-fine registration method is proposed. The backsighting orientation approach is used for the coarse registration. In the fine registration, corresponding stem centers are extracted to calculate the rigid-body transformation parameters. This method overcomes the difficulty in placing artificial targets and improves the scanning efficiency in dense forest areas.

Key words: *Terrestrial Laser Scanning, Registration, Point cloud processing, Coarse-to-fine*

Introduction

Terrestrial LiDAR Scanning (TLS) is an active technology which can provide both horizontal and vertical information. In the last few years, it has been widely used in different forestry applications. However, during the scanning, there are some constraints caused by Terrestrial LiDAR (T-LiDAR) system and complex environmental conditions. Obstructions caused by different vegetation elements frequently occur during the scanning, especially in natural forest stands with high stem density. This can lead to the partial observations and missing features from individual scans. To mitigate these effects and measure whole 3-D structure at a forest site, TLS data acquired from multiple scans are required. The registration procedure becomes vitally important.

In the forest areas, three main challenges need to be considered for registration. 1) Most T-LiDAR instruments are heavy (for example, the total weight of the scanner and accessories used in our experiments is larger than 15kg). In addition, natural forest sites are usually hard to access. It is better not to carry any other instruments to the forests and use only TLS system for accurate registration. 2) GPS usually performs badly with low spatial accuracy under a dense forest canopy. As a result, it is unreliable for accurate registration purpose. 3) Few natural tie points can be found in the forest. Therefore, artificial reflectors are widely used for the registration. However, setting up these reflectors is time-consuming. As the working hours of the TLS system are limited by the environmental conditions (such as the weather and daytime hours) and the instrument battery, we should improve the efficiency in data acquisition, especially in placing reflectors.

Some reflector-free registration methods were proposed to address the difficulties in placing artificial reflectors. Most of these methods extracted natural features (such as ground surface points, stem centers, and tree surface) as tie points. Then the iterative closest point (ICP) approach was commonly used to align adjacent scans. However, in order to obtain initial values for the ICP approach, some field work using the total station or reflective targets was still needed. In addition, some manual steps were also reported to extract tree stems.

In this paper, an automatic coarse-to-fine registration method is proposed. The main objectives of this method are (1) improving the efficiency of TLS in natural forest stands by placing only one artificial reflector; and (2) aligning multiple scans in an automatic and robust way with high accuracy. In the coarse registration, the backsighting orientation approach is applied. Unlike other approaches, only the TLS system is needed for the backsighting orientation. In addition, only one reflector is used as the backsighting target and located at each scan position. Compared with the reflector-based methods using at least four reflectors, our method is much more time-saving. In addition, the coarse registration procedure also provides the initial transformation parameters for determining tie points in the fine registration step. The fine registration step extracted corresponding stem centers as tie points and performed the rigid-body transformation for the alignment. To evaluate our method, we compared it with the reflector-based method. Results show that the proposed method offers the high accuracy of 1.5cm.

Materials and Data acquisition

Study Area

Our experiment was conducted in the northeast of China in August 2014. The study area is located in the Saihanba National Forest Park, Chengde, Hebei province (42° 23.87′ N, 117° 19.05′ E). Our scanning plot is a larch forest plot with few shrubs on the ground, and we chose a 45m by 45m square region for scanning.

Data acquisition

In this research, a Riegl VZ-1000 terrestrial laser scanner was used. The Zenith angle ranges from 30 ° to 130 ° and the azimuth angle ranges from 0 ° to 360 °. At this forest site, the whole plot was covered by ten scans. In order to reduce the accumulated registration error, all the scans in a forest site formed a closed ring. For the backsighting orientation approach in the coarse registration, we placed one reflector at each scan position. The scan angle stepwidth was set to 0.03 °. Thus, the point spacing was 2.4cm at a range of 45m. To test the registration accuracy, we also placed at least four corresponding reflectors in every two adjacent scans for the reflector-based registration method. It should be noted that these reflectors were not used in our coarse-to-fine registration method. They were only used for the result comparison with the reflector-based method.

Methods

Coarse registration

Considering the difficulty in setting up artificial reflectors, we apply the backsighting orientation approach in the coarse registration step. A reflector is placed at the target scan position to represent the scan location. The coordinates of the backsighting target (reflector) are extracted from the reference scan. Thus, three translation parameters are acquired. The rotation angles around the X and Y axes are recorded using the built-in inclination sensor of the scanner. Then the remaining degree-of-freedom (rotation angle around the Z axis) is calculated using the difference between the northing angles (angle against the north direction) of the target scan and the reference scan. In this paper, the northing angle is measured by the internal GPS receiver and the internal digital compass of the scanner. Compared with other backsighting orientation methods, our coarse registration only uses the TLS system for the backsighting orientation. No other measuring instrument is needed.

Fine registration

The objective of the fine registration is determining tie points and use them for the rigid-body transformation. Owing to the environmental conditions in the forest, few natural features can be directly found. Therefore, we calculate the corresponding stem center locations as tie-points. The fine registration procedure contains three main steps: 1) Ground points filtering; 2) Tie point determination and 3) Rigid-body transformation.

In the first step, we use the adaptive TIN models method which was proposed by Axelsson [1]. The ground points are filtered by the TIN (DEM), and all the remaining points are regarded as non-ground points.

In the second step, stem centers are extracted as tie points. Therefore, the stem fitting approach is needed. We assume that the trees are straight up in the forest. Thus, in a horizontal layer, all the points of a stem form a circular. Then in each point cloud, we slice the horizontal layer at 1.3m above the ground. The slicing is based on the TIN model to ensure the extraction of the same portion of each corresponding stem even if there is a slope at the site. To obtain enough points for stem fitting, the horizontal layer is sliced with a thickness of 5cm in each scan. After that, all the sliced points are segmented into several components to present the "potential" stems using the connectivity analysis. The least square circle fitting method is then applied to extract stem circles and calculate the stem center positions and diameters. For each potential stem, the standard deviation of the distances between all the points to the calculated stem center is obtained using Equation (1). If the standard deviation (*std*) exceeds the threshold, the potential stem will be regarded as wrong stem and discarded automatically. Finally, the average z-coordinate of the points in each stem is calculated as the z-coordinated of the stem center.

$$std = \sqrt{\frac{(\sqrt{(x_1 - X_0)^2 + (y_1 - Y_0)^2} - R)^2 + (\sqrt{(x_2 - X_0)^2 + (y_2 - Y_0)^2} - R)^2 + \ldots + (\sqrt{(x_n - X_0)^2 + (y_n - Y_0)^2} - R)^2}{n}} \quad (1)$$

where R represents the calculated stem radius, (X_0, Y_0) represents the calculated stem center location, (x_k, y_k) (k = 1, 2, 3, …, n) represents the points of a potential stem after the slicing.

After the coarse registration procedure, the corresponding points from every two scans are roughly aligned using the calculated transformation parameters. Based on the coarse alignment, it is easy to determine corresponding stem centers in two adjacent scans using a nearest neighbor searching approach.

In the rigid-body transformation step, the least square method is applied for the fine registration. Using the Equation (2), the rigid-body transformation parameters (three translation vectors and three rotation angles) are calculated with the corresponding stem center pairs in both the target scan ($(X_i, Y_i, Z_i)_{tar}$) and the reference scan ($(X_i, Y_i, Z_i)_{ref}$). Applying these transformation parameters to all scans, the fine registration procedure is done.

$$\begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix}_{ref} = R \begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix}_{tar} + \begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{bmatrix} \quad (2)$$

Where R represents a standard 3×3 rotation matrix, $[\Delta X \quad \Delta Y \quad \Delta Z]^T$ are the three translation vectors.

Results and Discussion

In this study, we took the local coordinate system of the scan position one (Scan 1) as the reference coordinate system. The registration order of all the scans reversed the scanning order, which was "Scan 10 to Scan 1, Scan 9 to Scan 10, … , Scan 2 to Scan 3". In order to evaluate the registration accuracy, we also placed at least four reflectors in every two adjacent scans for the reflector-based registration method.

Results

In our experiment site, we used three corresponding reflectors in every two adjacent scans to show the coarse registration results. For each corresponding reflector, the position errors of its center points in two scans are calculated as the registration accuracy. Considering all the reflectors at the site, the mean position errors in x-, y- and z- directions are 0.129m, 0.187m and 0.012m. The 3-D position error is 0.240m, and the standard deviation is 0.182m. Results show that position errors in the horizontal coordinates are much larger than the vertical coordinate. As we used the location of the backsighting reflector to represent each target scan's position and to calculate the three translation vectors, the location accuracy of the coarse registration is high. In addition, the rotation angles around the x and y axes are precisely recorded by the built-in inclination sensor of the scanner. Therefore, the larger horizontal position error is mainly caused by the inaccurate rotation angle around the z axis.

Similar to the coarse registration, we calculated the position errors as the fine registration accuracy. The mean position errors in x-, y- and z- directions are 0.005m, 0.005m and 0.011m. The 3-D position error is 0.015m, and the standard deviation is 0.006m. We also compared our proposed method with the reflector-based method. Using the reflector-based method, the mean accuracy is 0.006m, and the standard deviation is 0.002m. Although the accuracy of our method is a bit lower than the reflector-based method, we highly improve the efficiency in the scanning. We also compared our results with Ni's reflector-free method [2]. In his method, the 3-D position error is 8.5cm. Therefore, our method shows higher registration accuracy than Ni's method.

Discussion

In the forest area, the environmental conditions are complicated which made the forest site difficult to access. Considering the total weight of our TLS system, it is hard to set up the instrument and carry it among all the scan positions. Therefore, we should carry as few instruments as possible. In this paper, the proposed registration method only uses the TLS system for relative registration between every two scans. Moreover, the weather needs to be considered. The wind and rain commonly affect the scanning. Therefore, the efficiency needs to be improved in data acquisition. In the reflector-based registration method, at least four corresponding reflectors are needed. However, at a natural forest site, the obstructions occur frequently. It is hard and time-consuming for setting up these reflectors. At this experiment site, if we use the reflector-based method, it will take more than 20 minutes in setting up and placing reflectors. Whereas, using our coarse-to-fine method, we only need to place one reflector at the location of each scan position. Even at a dense forest site, it will take less than 5 minutes.

Conclusion

In this paper, an automatic, single-reflector registration method using a coarse-to-fine strategy with Terrestrial LiDAR data is proposed. The coarse registration uses the backsighting orientation approach which requires only one reflector at each scan position. In addition, compared with the conventional method, we only use the TLS system for the backsighting orientation. In the fine registration procedure, stem center positions are calculated as tie-points to obtain the rigid-body transformation parameters. To test our method, experiments were conducted in the northeast China. The registration accuracy of our method is 1.5cm.

The main objective of this research is to make the data acquisition more efficient. As only one reflector is needed at each scan position, this method avoids the challenge of placing many artificial reflectors. In addition, it is also an automatic registration method. Results show that our registration method provides much better accuracy than other reflector-free methods. Although the accuracy is a bit lower than the reflector-based method, the scanning efficiency is highly improved. Therefore, our coarse-to-fine registration method is more feasible for forest measurements considering both the scanning efficiency and registration accuracy.

Acknowledgements

This work is supported by the National Basic Research Program of China (973 Program) Grant No. 2013CB733402 and the Open Research Fund of Key Laboratory of Digital Earth Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences No. 2014LDE015. This work is also supported by National Natural Science Foundation of China Grant No. 41171265, 41331171 and 40801131.

References

[1] Axelsson, P. (2000). Dem generation from laser scanner data using adaptive tin models. International Archives of Photogrammetry and Remote Sensing, 22, 111–118.

[2] Ni, W., Sun, G., Guo, Z., & Huang, H. (2011). A method for the registration of multiview range images acquired in forest areas using a terrestrial laser scanner. International journal of remote sensing, 32, 9769-9787.

Modeling Individual Tree Height from LiDAR data in a Tropical Rainforest

Wan S. Wan Mohd Jaafar[1], Iain H. Woodhouse[1], Carlos A. Silva[2]

[1]*School of Geosciences, University of Edinburgh, EH8 9XL, Edinburgh, United Kingdom.*
Email: W.B.Wan-Mohd-Jaafar@sms.ed.ac.uk
[2]*University of Idaho, 875 Perimeter Drive, Moscow, ID 83844, United States.*

Highlights:
Forest plays a key role in the global carbon cycle. An accurate estimation of total biomass and its component is critical for understanding the carbon cycle in the forest ecosystem. Individual tree detection is important to accurately extract tree structural attribute of trees; tree height is critical for biomass estimate.

Key words: *individual tree detection, tropical rainforest, LiDAR*

1. Introduction

Tropical forests are known for their complex stand structure and abundant variety in species composition ([1]). The chosen study site is Pasoh Reserve Forest at Peninsular Malaysia. This site contains mixed species and was dominated by trees from the dipterocarp family which is lowland dipterocarp is the most dominant one. Lowland dipterocarp forest is one of the most species-rich communities in the world, with more than 200 species per hectare ([2], [3]). This type of forest is typically characterised by a complex three-dimensional structural. Therefore, individual tree detection is seen as the most relevant approach to extract the structural attribute of the tree in tropical rainforest. Light Detection and Ranging (LiDAR) is the best remote sensing technology for modelling the forest's structure and for retrieving several forest attributes. The precision and accuracy of LiDAR based estimations are strongly influenced by site conditions and by the characteristics of the sensors. LiDAR forest inventory methodology based on individual tree detection has been widely studied, but is not widely used in practice, due to assumed problems related to tree detection under various forest condition especially in a dense tropical forest. Unlike most of the published algorithms that detect individual trees from a LiDAR-derived raster surface, at this stage we worked directly with the LiDAR point cloud data and field inventory information to separate individual trees and to estimate individual tree metrics for individual tree biomass modelling. Later we will compare the capability of the current technique with other common segmentation methods in determining locations of individual trees by using raster surface models like Canopy Height Model (CHM).

2. Methods

2.1 Study Area

The Pasoh Reserve Forest is located on peninsular Malaysia approximately 70 km southeast of Kuala Lumpur. The Pasoh Reserve Forest study site (2.98 N 102.31 E) is located about 8 km from the town of Simpang Pertang, approximately 70 km southeast of Kuala Lumpur (140 km by road). Pasoh Forest has an area of approximately 140 km2, mainly covered with lowland dipterocarp forest and with hill dipterocarp forests in its north-eastern boundary. The core area which is approximately 600 ha within the reserve is still covered with old growth forest; most of the surrounding area has been logged in the past, providing many examples of regenerating lowland forest. Pasoh Forest is one of the most complex dense species-rich communities in the world with 340,000 trees (≥1 cm diameter at breast height) consisting of 818 species.

2.2 Field Measurement

The fieldwork runs at Pasoh Reserve Forest, Negeri Sembilan, Malaysia from 21st October - 23rd October 2014 with 9 field members. Six main forest parameters have successfully been collected; accurate individual tree position (x, y), Diameter at Breast Height (DBH), Total Height, Bole height, Crown Diameter and Tree Species. To minimise the uncertainties associated with biomass measurements, the layout of forest sampling must follow the standard measurement guideline from Malaysia National Forest Inventory (MNFI). Stratified random sampling of square shape were established to measure trees with DBH >10cm. There are 142 individual tree measurements recorded

from the 8 plots with 25m x 25m measurement. Out of this number, three trees were excluded from analysis due to the condition of being broken at the first branch. We used GIS software ArcMap 10.1 to calculate crown cover by creating a polygon layer using the field measured crown diameter and plotting it in the form of shape files with correct x and y location for each of individual tree measured in the field. Each of the tree polygon has been assigned with a unique identification number (ID) and has been projected to the same projection of LiDAR data which is UTM_Zone_48N.

2.3 LiDAR Operations

The LiDAR data was obtained from a private Malaysian airborne company. IGI LiteMapper-5600 system with lidar sensor Riegl Q560 was used with a 60^0 scan angle from nadir. The processed data was received as a point cloud format in LAS and ASCII XYZ format with a projection of UTM_Zone_48N. The average point spacing of LiDAR data used for this study was 8.8 points per square meter. The value of Root Mean Square achieved was 0.092 meter and proved that the accuracy of LiDAR data is in between tolerance of 0.15m.

2.3.1 Data Pre-processing

A Digital Terrain Model (DTM) was created in two steps from the discrete return LiDAR data: first, the data was filtered to remove the above-ground returns (using algorithm adapted from [4]), and secondly the DTM was created by calculating the average elevation from the remaining (ground) LiDAR returns within a cell (cells that contain no points were filled by interpolation using neighbouring cells). The DR point clouds were then normalized against the ground surface height and extracted for each plot using the coordinates of the lower left and upper right corners. All point cloud data processing was performed using the FUSION software [5].

2.3.2 Individual tree extraction – clipping point cloud data with crown polygon from forest inventory

We choose an automatic method using Web-LiDAR 3D ClipTree application [6] which is available on the web at http://forest.moscowfsl.wsu.edu:3838/csilva/LiDAR3DClipTree/ to delineate each individual tree by clipping the LiDAR point cloud by the crown polygon created using the field measurement. Figure 1 shows the main display of the Web-LiDAR ClipTree application. The application has five major displays: (1) settings menu; (2) summary of the LiDAR metrics; (3) Profile viewer of LiDAR data; (4) Top LiDAR data viewer; and (5) 3D LiDAR data viewer. Individual tree LiDAR metrics were then computed using the rMetrics function in the rLiDAR package [7]. The generated metrics from LiDAR were used to model the individual tree heights.

Figure 1: Main display of the Web-LiDAR forest inventory application: 3D ClipTree tool

3. Results and Discussion

The automated clipping works by assigning the point cloud data fall within the crown polygon according to the x and y position. Each of the point cloud data has been assigned with the unique ID similar to the polygon layer where it was overlapped. Each set of measurement was used to compute the individual tree metrics using R statistics. After an individual tree had successfully been extracted, we validated the LiDAR derived height with the field measured height for each of the individual tree detected. Figure 2 shows the scatterplot of individual tree heights from LiDAR compared with field height measurement.

The results show a good correlation and the value of Root Mean Square achieved, proving a good tolerance.

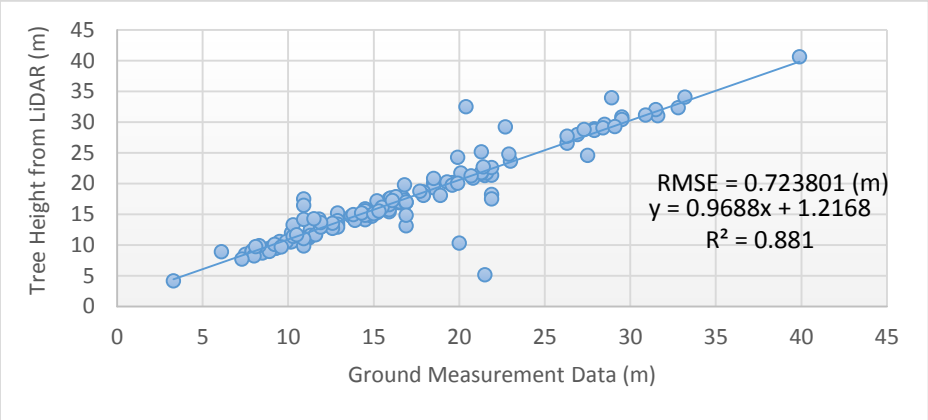


Figure 2: Scatterplot individual tree heights from LiDAR compared with ground measurement

3.1 LiDAR Metrics

Discrete data based metrics are descriptive structure statistics calculated from the height normalized LiDAR point cloud measurement in three-dimensional space. In this study, 24 metrics for each tree within 25 x 25 m area of the 8 plots were calculated and this can be summarized into 3 different categories; (i) Height description (Minimum Height (h_{min}), Maximum Height (h_{max}), Mode Height (h_{mod}), Median Height (h_{med}), Standard Deviation of Height (h_{sd}), Variance of Height (h_{var}), Coefficient of variation of heights (h_{cv})); (ii) Height percentile values (h_5,h_{10},h_{20},h_{25},h_{30},h_{40},h_{50},h_{60},h_{70},h_{75},h_{80},h_{90},h_{95},h_{99}) and (iii) Canopy-related metrics (Kurtosis Height ($h_{kurtosis}$), Skewness Height ($h_{skewness}$), Canopy cover (CC)). Discrete data-based metrics were computed using all the LiDAR returns.

4. Conclusions

This study intends to use different approaches on utilizing LiDAR for individual tree detection over dense tropical rainforests. Individual tree detection has significant implications in forestry. We believe that, once accurately segmented, tree structural attributes such as tree height, crown diameter, canopy height and volume and species type can be derived. These are the most important parameter used to estimate biomass. By accurately extracting forest parameter this can generate an accurate biomass result and will provide reliable guidelines for biomass study in tropical rainforest.

References

[1] Steiniger, M.K. (2000). Satellite estimation of tropical secondary forest above ground biomass: Data from Brazil and Bolivia. International Journal of Remote Sensing, 21 (6-7), 1139-157.

[2] Symington, C. F. (1943). Foresters' manual of Dipterocarps. Malay. For. Rec. No. 16. Penerbit Universiti Malaya, Kuala Lumpur.

[3] Wyatt-Smith, J. (1964). A preliminary vegetation map of Malaya with descriptions of the vegetation types. J. Trop. Geogr. 18: 200–213.

[4] Kraus, K., and N. Pfeifer. (2001). Advanced DTM generation from LIDAR data, International Archives of Photogrammetry and Remote Sensing, Volume XXXIV-3/W4, 22–24 October, Annapolis, Maryland, USA: 23–30.

[5] Mcgaughey, R.J. FUSION/LDV: Software for LiDAR Data Analysis and Visualization [Computer program]. Washington: USDA, Forest Service Pacific Northwest Research Station. (2014). 150 p. Available: http://forsys.cfr.washington.edu/fusion/FUSION_manual.pdf. Assessed on: 5 May, 2015.

[6] Silva, C.A, Hudak, A.T, Crookston, N.L. (2014). Web-LiDAR forest inventory: 3D ClipTree application. Available: http://forest.moscowfsl.wsu.edu:3838/csilva/LiDAR3DClipTree/ Accessed on: 1 May 2015.

[7] Silva, C.A., Nicholas L., Andres T., Lee A. Package rLiDAR. Available: http://cran.r-project.org/web/packages/rLiDAR/rLiDAR.pdf . Accessed on: 19 May. 2015.

Identification of 16 individual tree species with ALS in a complex southern boreal forest of Canada

David Hernandez[1], Benoît St-Onge[1], Bastien Ferland-Raymond[2], and Chhun-Huor Ung[3]

1. *Department of Geography, University of Quebec at Montreal, Canada*
2. *Consulting statistician, Québec, Canada*
3. *Formerly with the Canadian Wood Fiber Center, Natural Resources Canada*

Highlights: Mixed boreal forests have complex species assemblages. We tested an individual tree species identification approach for discriminating between 16 tree species (10 needleleaf, 6 broadleaf). Among three classifiers (random forest, kNN, and SVM), random forest produced the best kappa: 0.902 for needle/broad leaves, and 0.561 at the species level.

Key words: *Species, delineation, airborne laser scanning, random forest, kNN, SVM.*

Introduction

Despite the rich spectral information contents of optical images, radiometric standardization problems encountered in correcting for the varying sun-object-sensor geometry greatly complicates the retrieval of species information from this type of remote sensing data. In recent years, the potential of ALS for identifying the species of individual trees has been demonstrated (see for example the review by Vauhkonen et al 2014). However, most studies were performed in forests having limited tree species diversity, for example in Scandinavia or Western Europe. In contrast, the forests found in the mixed boreal zone of Canada often exhibit complex species assemblage (e.g. > 12 species locally), as well as complicated 3D structures (mix of large and small trees). What is more, the accuracy and level of detail on species composition in current Canadian forest stand maps is deemed insufficient. Therefore, there is a need to further explore the discrimination of species using ALS in these types of forests.

Our objective was to identify the species of individual trees within a southern boreal forest exhibiting a large diversity (16 species from 10 genera). We also assessed the performance of three different classification methods (random forest, support vector machine, and kNN), as well as the effect of merging individual species in various groups (e.g. needleleaf and broadleaf).

Methods

Study region and data

The study region is the "Valcartier Forest", located near Quebec City, Canada (71°30'W, 46°54'N). It comprises 16 typical species of the mixed boreal region: *Betula alleghaniensis* (BOJ), *Betula papyrifera* (BOP), *Picea glauca* (EPB), *Picea mariana* (EPN), *Picea abies* (EPO), *Picea rubens* (EPR), *Acer rubrum* (ERR), *Acer saccharum* (ERS), *Fagus grandifolia* (HEG), *Larix laricina* (MEL), *Populus tremuloides* (PET), *Pinus strobus* (PIB), *Pinus resinosa* (PIR), *Tsuga canadensis* (PRU), *Abies balsamea* (SAB), and *Thuya occidentalis* (THO). Field observations of species were carried out in plots or pure tree patches and the centre of these spatial units were geolocated with a GNSS receiver.

A lidar survey was performed in summer of 2013 with a first return density of approximately 7 returns/m^2 Discrete XYZ coordinates and intensities were recorded for up to four returns per pulse. Raster DTMs, DSMs, and canopy height models (CHM) were generated through interpolation. The absolute Z value of lidar returns was converted to height above terrain by subtracting the corresponding DTM value to produce *XYH* points. The intensity values were standardized according to their approximate range that was estimated based on the nominal flight altitude, DSM, and pulse incident angle.

Analyses

Automated individual tree delineation was first performed on the lidar CHM using an in-house application developed in C++. It implements a method that involves the following steps: 1) adaptative gaussian filtering (σ proportional to local CHM height), identification of local maxima within a circular window of adaptative radius, 3) inter-tree valley detection using a circular Laplacien filter, and 4) region growing from local maxima until stopping conditions are attained (e.g. a strong valley is met, maximum radius for given H is reached, etc.). An interpreter then chose well delineated crowns that fell within the field inspected plots or patches to create a crowns sample with known species. A total of 429 trees were sampled, with 8 to 58 trees per species (average of

27 trees/species). For each crown, the CHM pixels and the height and intensity of the normalized returns (XYH and standardized intensity) were extracted. From these, 21 metrics involving either height or intensity were computed. The metrics were designed to capture different aspects of the crowns: chiefly shape, porosity, and reflectance. The best subset of 11 variables was selected with the goal of minimizing correlation between variables and removing those with very low variability. This subset was composed for example of (due to space limitation only a few variables are described): the slope of the crown profile, the ratio of median 1st return height over the maximum height, the standard deviation of the vertical distance between the 1st and 2nd returns (pulse-wise), the average deviation between the original first return height and the smoothed raster crown, the standard deviation of the average height measured respectively on four concentric annuli centred on the local maximum, the maximum intensity within the crown, the average difference of the intensity between 1st and 2nd returns (pulse-wise), etc.

Three classification methods were compared using R: random forest (RF), *k*-nearest neighbours (kNN, with $k = 5$), and support vector machine (SVM). 60% of the sample was used for training, and 40% for verification. 1000 different training/verification subsets were created randomly and Cohen's kappa coefficient was evaluated for each subset and classification method. We hereafter report on the average kappa of each method.Furthermore, the classifications were done for different levels of species groups: 1) needleleaf vs broadleaf, 2) all species taken separately, 3) genus groups (e.g. all spruces, all pines, etc.), and 4) species groups designed to minimize confusion while preserving a relatively large number of groups.

Results

The classification for level 1 (needleleaf vs broadleaf) yielded the following average kappa values, respectively for RF, kNN, and SVM: 0.902, 0.900, and 0.836. At the species level (2), the average kappa values were markedly lower, being respectively of 0.561, 0.527, and 0.471. Table 1 gives the average number of observations for each cell of the full confusion matrix for the individual species classification (because they result from averaging 1000 different subsets, the numbers may not be integers). As expected, confusion occurs mainly among the two broad level 1 groups, but is less frequent between them. Confusion is worse among broadleaf trees. Some coniferous trees stand out as being easily distinguishable from all other trees, notably *Picea glauca* (EPB), *Larix laricina* (MEL), and *Thuya occidentalis* (THO). There are some instances where, for this level of classification, confusion occurs between a deciduous and a coniferous tree (e.g. EPN with BOP). Confusion among level 1 groups is not necessarily within genera. For example, there is much more confusion between *Acer saccharum* (ERS) and *Betula Allghaniensis* (BOJ, different genera) than between ERS and *Acer rubrum* (ERR, same genus).

Table 1: Confusion matrix of the random forest classification of the 16 species (see the "Study region and data" section for explanation of the species codes).

	BOJ	BOP	EPB	EPN	EPO	EPR	ERR	ERS	HEG	MEL	PET	PIB	PIR	PRU	SAB	THO
BOJ	5.1	1.8	0.0	0.0	0.0	0.0	0.5	5.4	0.5	0.0	1.3	0.1	0.0	0.0	0.0	0.0
BOP	2.3	3.2	1.6	2.3	0.0	0.0	0.0	2.6	0.1	0.0	0.2	0.0	0.3	0.0	0.3	0.1
EPB	0.0	0.3	7.5	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.1	1.1	0.8	0.0	1.3	0.0
EPN	0.0	1.3	0.2	9.5	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.1	0.1
EPO	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.9	0.0	0.0	2.2	0.0
EPR	0.6	0.4	0.0	0.9	0.0	0.9	0.0	0.1	0.1	0.0	0.3	0.9	0.4	0.1	1.6	1.0
ERR	0.3	0.0	0.0	0.1	0.0	0.0	2.7	0.3	0.7	0.0	0.4	0.0	0.2	0.0	0.0	0.0
ERS	2.8	0.3	0.0	0.5	0.0	0.2	0.0	17.4	0.2	0.0	1.5	0.0	0.0	0.0	0.1	0.0
HEG	1.7	0.7	0.0	0.2	0.0	0.0	1.0	0.9	1.3	0.0	0.8	0.0	1.5	0.0	0.6	0.0
MEL	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	3.4	0.0	0.1	0.2	0.0	0.7	0.0
PET	1.5	0.1	0.0	1.0	0.0	0.1	0.0	2.2	0.7	0.0	5.5	1.4	1.4	0.2	0.7	0.0
PIB	0.6	0.0	0.1	0.6	0.2	0.1	0.0	0.0	0.0	0.2	0.9	5.2	2.1	0.2	3.0	0.1
PIR	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.1	0.2	0.0	0.3	1.0	13.1	0.0	0.7	0.0
PRU	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.9	0.0	0.0	0.4	0.1	0.0	1.7	0.3	0.0
SAB	0.0	0.3	0.3	0.4	0.2	0.2	0.0	0.3	0.1	0.4	0.4	1.5	0.3	0.0	13.8	0.0
THO	0.0	0.0	0.1	0.3	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.1	2.4

This situation is further evidenced by the results of level 3 classification in which species are grouped by genus. The respective kappa values for RF, kNN, and SVM are 0.555, 0.536, and 0.459, i.e. even worse than those resulting from the individual species classification. Finally, species between which large levels of

confusion existed were pooled into multi-species groups that may include also multiple genera (but never a mix of needleleaf and broadleaf species). This lead to two super-groups: EPR, EPB, EPO, PIR, PIB, and SAB (total of 3 needleleaf genera), and ERS, BOJ, BOP, HEG (total of 3 broadleaf genera). The classification was therefore performed for two super-groups and six individual species. The kappa values respectively for SF, kNN and SVM were 0.716, 0.714, and 0.594. These values are approximately midway between those of the needleleaf/broadleaf classification, and the individual species classification, but are much more precise in terms of number of species/groups (8 vs 2).

Discussion and conclusion

The results first indicate that ALS data can accurately discriminate between coniferous and deciduous trees at the individual tree level (kappa = 0.902). It should be remembered that the kappa coefficient does not represent the raw percentage of correct classification, but rather a score in which the effect of correct classifications resulting from mere chance was removed. A kappa value of 0.902 therefore indicates near perfect classification. Although similar results were obtained in the recent past, we have showed in the present study that such a high accuracy can also be obtained when the species composition is complex. Of course, the same level of performance cannot be achieved at the individual species level when a high number of species is present. Our analysis shows that the confusion between species is not necessarily between species of a same genus. This owes to the sometimes important intra-genus variations in crown architecture, leaf shape or size, foliage reflectance, etc. For example, *Picea glauca* (EPB) has a sharply arched crown envelope, with a rather dense foliage, and limited branch protuberance, while *Picea abies* (EPO) has a more ample profile, with protruding, isolated, and arched branches. In general, tree species with an idiosyncratic architecture will be easily distinguishable from other species. This brought us to propose a compromise by performing classifications at mixed levels, i.e. both individual species and species groups. This approach offers a relatively high species "resolution" while still having an acceptable performance. Indeed, a kappa of 0.716 (mixed levels) is deemed "strong", as opposed to a kappa of 0.561 (individual species) considered as "moderate". A mixed level classification may not however be optimal with regards to the information needs of the forest industry or ecological science. The wood quality of the different species make them amenable to different uses and mill processing. For example, the wood quality of *Abies balsamea* (SAB) is significantly lower than that of the *Picea* genus. Unfortunately, confusions with other species brought us to merge SAB with a few *Picea* species in the mixed level approach. The same applies to ecological science. If, for a given biodiversity assessment, a particular rare species needs to be mapped but is easily confused with other species, ALS would not be very useful in identifying it.

It should be noted that our analysis is solely based on well delineated crown samples that were selected by a human interpreter. We expect that tree crown delineation errors would somewhat corrupt the estimation of tree shape, porosity, and possibly reflectance. Therefore, the performance levels reported above would certainly drop in a scenario where the classification is applied to all delineated crowns of a study area. Furthermore, we have mainly selected the crowns of fully grown trees. Knowing that tree shape also varies with size, there is no guarantee that a particular classification method calibrated on taller trees will also work on shorter trees.

We may not have yet fully exploited the potential of standard ALS point clouds for species classification, but the above results lead us to think that further improvements would not be dramatic. Data fusion or sensor enhancements however have the potential for significant performance gains. Extracting multispectral information from optical images for each lidar first return is one avenue. What is possibly an easier approach is to include topographic derivatives, such as the topographical wetness index, to help separate species that have similar architectures, but different ecological niches. Finally, the advent of multispectral lidar, such as Optech's Titan with bands in the green, near infrared, and middle infrared, will certainly help in increasing the performance of species identification.

Acknowledgements

We thank the Canadian Wood Fibre Centre (Natural Resources Canada) for the grant that made this study possible, as well as the authorities of the Canadian armed forces of the Valcartier Garrison who contributed to the acquisition of lidar data and to the field data collection.

References

[1] Vauhkonen, J., et al. (2014). Tree species recognition based on airborne laser scanner and complementary data sources. In M. Maltamo, E. Naesset & J. Vauhkonen (Eds.), Forestry Applications of Airborne Laser Scanning (pp. 135-156). Springer.

A comparison of spatialisation methods for the aggregation of LiDAR forest estimates at the compartment level

Jean-Matthieu Monnet[1,2], Alain Munoz[3]

1: Irstea, UR EMGR, France
2: Univ. Grenoble Alpes, F-38402 Grenoble, France
3: ONF, Département Recherche Développement et Innovation, France

Highlights: From the plot to compartment level, prediction error decreases from 15 to 6.4% for basal area, 26 to 7.7% for stem density and 6.5 to 3.4% for dominant diameter. The major criterion for spatialisation is to respect the calibration plot size, whereas for aggregation the issue of compartment borders depends on the forest parameter.

Key words: *forest inventory, airborne laser scanning, spatialisation, compartment.*

Introduction

Long term forest management planning is usually based on statistical inventories, which are not meant to provide information at the compartment level. The area-based approach is now a widely used method to derive continuous maps of forest parameters from airborne laser scanning (ALS) data and field calibration plots [1]. The accuracy of estimates is frequently based on cross-validation with plot-level data, whereas compartments are frequently the basic unit of forest management. The calibrated model first has to be used on a partition of the area of interest into a set of spatial entities for which the model is supposed to be valid. The most straightforward way is to apply a grid which pixel size is similar to plot size. Then to aggregate the pixel values into the forest compartments, the case of borders has to be handled carefully as border pixels may represent a significant proportion of the forest surface. The objective of this article is to compare different possibilities for the spatialisation step (spatial entity size, shape and spacing) and different border exclusion thresholds for the aggregation at the compartment level. The study is based on a dataset of 35 fully-callipered compartments (total 380 ha).

Material

The study area is located in the Jura mountains (France). The stands are uneven-aged, and dominated by silver fir (*Abies alba*), Norway spruce (*Picea abies*) and European beech (*Fagus sylvatica*). The ALS data were acquired with a LMS-Q560 full-waveform scanner. The final pulse density is 9.3 m^{-2}.

For the calibration of estimation models, 139 nested plots of 17 m radius are available in the 580 ha of public forests. 26 were excluded because harvesting occurred during the one year lapse between the ALS flight and field measures. Only trees with diameter at breast height (DBH) larger than 17.5 cm are considered. For each plot, three forest parameters are calculated: basal area, dominant diameter (diameter of the 8 largest trees) and stem density. For the validation, 35 compartments representing a total of 380 ha were fully-callipered. The three forest parameters are calculated for each compartment.

Methods

ALS models

ALS metrics are computed for each plot based on the normalised point cloud extracted over its spatial extent. Metrics include the usual height and density metrics, as well as tree metrics derived from a preliminary segmentation of the canopy height model.

A prediction models is calibrated for each forest parameter by fitting ordinary least squares regression with at most three ALS metrics as independent variables. Selection is based on the adjusted R^2. Models are automatically checked for linear model assumptions, and manually checked for dependencies to other forest parameters. Accuracy is evaluated by computing the root mean square error *rmse* in leave-one-out cross validation.

Spatialisation

Spatialisation scenarios are based on three parameters for the partition of the area of interest into a set of spatial entities, depending on the spacing, shape and size. The following scenarios are tested: disks of 17 m radius (surface and shape identical to calibration plots); squares of 20 m side (different surface and shape);

squares of 30 m side (different shape but similar surface). For square shapes, spacing is chosen equal to their side length. For disks, spacings of 10, 20, 30 and 40 meters are tested. Spacing modifies the amount of overlap between adjacent entities regarding the data used for metrics calculation.

Aggregation

Aggregation of ALS estimates in one compartment is done by computing the mean value of entities which centres lie inside the compartment, but at a minimum distance from the compartment border. Distance thresholds of 0, 10 and 20 m are tested. For validation the *Bias* and *RMSE* are computed by using the full-calliper data.

Results and discussion

Table 1 presents the accuracy of ALS estimates at the plot and compartment levels (spatialisation 17 m disk, 20 m spacing, aggregation with a 10 m threshold). In a study carried in Scandinavian forests [2], the standard deviation of plot-level error was 14.8 to 23.3% for basal area and 25.3 to 26.2% for stem density. At the compartment level, those values decreased to respectively 8.7 to 13.6% and 14.3 to 29.3%. The standard deviation of error was thus reduced by a factor of approximately 2. In our case it is also around 2 for basal area and dominant diameter, and 3 for stem density. The factor of surface increment between the plot and compartment levels is between 30 and 200, depending on compartment surface, compared to 16 in the Scandinavian study. This might explain why the compartment-level accuracy is better. Another explanation could be the uneven-aged structure, which should reduce the spatial correlation of errors. It is noteworthy that compartments with the largest errors are also those which differ most from the uneven-aged structure.

Table 1: Accuracy of ALS estimates at the plot and compartment levels.

Parameter	Plot level (n=113)			Compartment level (N=35)				
	R^2	*rmse*		R^2	*Bias*		*RMSE*	
Basal area ($m^2.ha^{-1}$)	0.76	4.5	15%	0.85	-0.4	-1.3%	1.9	6.4%
Stem density (ha^{-1})	0.57	72	26%	0.76	-7.4	-2.9%	20	7.7%
Dominant diameter (cm)	0.88	3.3	6.5%	0.91	-1.2	-2.5%	1.7	3.4%

Duplat and Perrotte [3] estimate that in the case of a full-calliper inventory 95% of the compartments have their basal area values within [-15%, 10%] of the true value. If those errors have a Gaussian distribution, then the distribution of differences between the full-calliper and ALS values is not significantly different from it. In this forest, the compartment-level accuracy of the ALS-based inventory is thus similar to a full-calliper inventory.

Regarding the spatialisation and aggregation parameters, it turns out that the compartment-level bias and error are higher in absolute value when the 20 m square entities are used, i.e. when the calibration plot size is not respected. Shape has only a limited effect as the 30 m square entities yield accuracies similar to disks of same surface and spacing for all forest parameters. The errors for disks with smaller or greater spacing are in general slightly higher. The distance thresholds for the aggregation step have different effects on bias and error, depending on the forest parameter. These differences may be due to the fact that the metrics selected in the models have different sensitivity regarding heterogeneous areas.

With compartments of a surface around a few hectares, the proportion of border entities is significant, especially when the perimeter to surface ratio is high. In the case of an uneven-aged forest, the entities located across two forested compartments should be less of a problem than for two adjacent, regular stands of different ages. However, the borders between forest and non-forest types should probably be tackled differently. One possibility could be to apply land-cover dependant strategies for the integration of border values. The other one would be to constrain the partition into the spatial entities with compartment borders, but the issue of respecting the surface criterion might turn out tricky.

Conclusion

The accuracy of ALS-based estimates for basal area, dominant diameter and stem density is more than twice better at the compartment level than at the plot level, and for basal area reaches a level similar to a full-calliper inventory. For the spatialisation step, it is crucial to respect the size of calibration plots, whereas for the aggregation the issue of compartment borders has to be tackled differently depending on the forest parameter.

References

[1] White, J., Wulder, M., Varhola, A., Vastaranta, M., Coops, N.C., Cook, B., Pitt, D. & Woods, M. (2013). A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach. 50 p.

[2] 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.

[3] Duplat, P. & Perrotte, G. (1981). Inventaire et estimation de l'accroissement des peuplements forestiers. Office National des Forêts. 432 p.

Low cost and accurate forest biomass estimation using a portable terrestrial laser scanner

Akira Kato[1)], Matt Bradford[2)], L. Monika Moskal[3)], Koji Kajiwara[4)], Yoshiaki Honda[4)]

1) Graduate School of Horticulture, Chiba University, Japan
2) Commonwealth Scientific and Industrial Research Organization Land and Water, Tropical Forest Research Centre, Australia
3) Precision Forestry Cooperative, School of Environmental and Forest Sciences, College of the Environment, University of Washington, USA
4) Center for Environmental Remote Sensing, Chiba University, Japan

Highlights: A low cost portable terrestrial laser system was developed to estimate tropical forest biomass for REDD+ project. The portable laser has the limited coverage when it is compared with high grade sensor. To make it useful in practical field survey, we developed a field sampling scheme and an algorithm to estimate tropical forest biomass accurately.

Key words: *Tropical forest, REDD+, Biomass, TLS, Simulation, Monitoring*

Introduction

Forest biomass estimation of REDD+ project requires low cost and accurate monitoring technology. However, the cost of 3D data acquisition is expensive if airborne laser is taken over rural area in developing countries. To obtain low cost 3D data, a portable terrestrial laser scanner systems have been developed [1][2][3]. To make it useful in practical field survey, the capability of portable sensors is tested for the tropical biomass estimation. And the optimum sensor location and the optimum number of scan are required to estimate a large scale tropical forest biomass.

In this study, we used a large scale forest inventory data collected in 25 ha and different sampling designs (random, systematic, and stratified sampling) are applied to find the feasible sampling scheme and estimate a large scale tropical forest biomass accurately.

Data and methodology

Our study area is Robson`s Creek, Australia. There is the 25 ha supersite established by CSIRO Land and Water, tropical forest centre. Field data was collected and trees within the 25 ha were all measured [4]. The portable sensor used for this study was SICK LMS511 (SICK Inc.). To understand the limited capability of the portable laser systems, the higher grade sensor (RIEGL VZ400, RIEGL Inc.) was also used at the same site and compared with the portable sensor. The sensor setting is shown in Table 1.

Table 1: Two different terrestrial laser sensors used for this study

Sensor Name	RIEGL VZ400	SICK LMS511
Wavelength	1550 nm	905 nm
Max. distance	600 m	40 m
Angle	Horizontal angle 360°	Horizontal angle 360°
	Vertical angle 100°	Vertical angle 150°

Field validation requires the allometric equation of tropical forest biomass. The allometric equation used for this study only needs diameter at breadth height (d.b.h) as an input parameter [5]. To derive the d.b.h. accurately from 3D data, the d.b.h height is identified from the digital terrain model (DTM). To get the d.b.h. height, DTM was initially created from the minimum height value of laser returns in each 50cm x 50cm mesh grid. And the spatial filter was applied to smooth the surface and fill the gap of no laser returns.

The individual stem location is identified using the sliced 3D data at d.b.h height. Based on the stem location, the 3D laser data around d.b.h height (1.3 m height) is extracted. And the stem shape is automatically recognized using radial basis functions [6]. And the stem diameters are measured from the stem circumference. The results from two different sensors (higher grade sensor and the portable sensor) are compared to get their accuracy in d.b.h measurement.

Three different sampling designs were applied and simulated using 25 ha field data to identify the optimum location of sensors and the optimum number of scan. The error was obtained by the difference between the 25 ha

biomass estimated by field based measurement and by laser based measurement from the limited number of samples. As the number of sample increases, the error decrease. When the error reaches 5 % of the entire biomass estimation, the optimum number of samples is found.

In considering practical use of the portable terrestrial sensor, the stratified sampling technique is developed for this study. The definition of the stratified sampling in this case is "large tree based sampling technique". The large tree means the tree height above 20 m. When the large tree is found in the woods, the plot centre is established at the large tree. And terrestrial laser sensor is set within 10 m radius from the plot centre. If the sensor is set right next to the big tree (the plot centre), the obscured area becomes large. Therefore, terrestrial laser sensor is located in three positions equally spaced to make a triangle to minimize the obscured area. After the data is collected in three positions, the 3D data is merged together.

Result and discussion

The main purpose of this research is to get the optimum location and number of scan using the portable terrestrial laser to estimate tropical forest biomass. In Figure 1, the blue dots represent the sliced laser data around d.b.h height and the red dots represent the shape recognition from our method. The red dots are lined up exactly on blue dots to recognize the stem circumference accurately. This technique improves the d.b.h measurement when it is compared with the conventional method using the circle fitting technique.

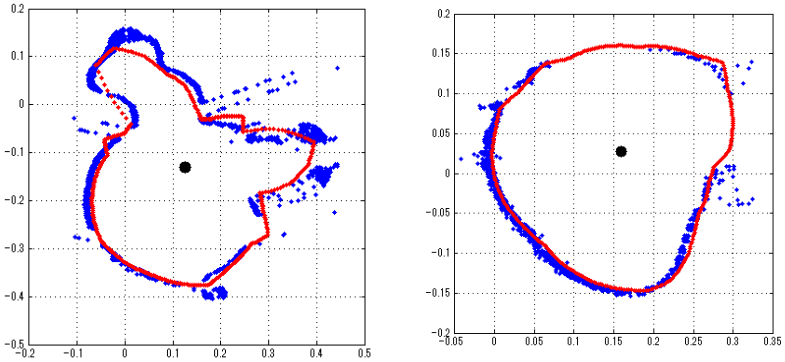
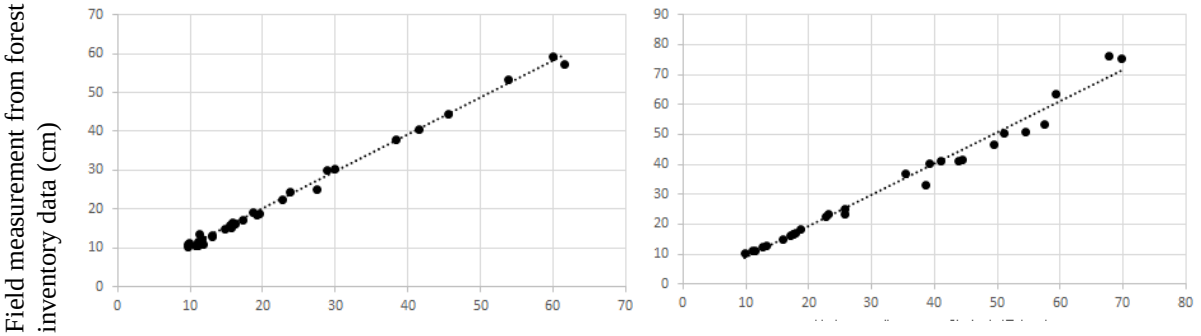


Figure 1: The result from automatic stem shape recognition (blue dots: sliced laser data, red line: automatically recognized stem circumference)

With this improved stem recognition technique, the d.b.h measurement of two different laser sensors is compared. For the d.b.h measurement, the higher grade sensor has 1.14 cm and the portable sensor has 2.77 cm Root Mean Square Error (RMSE) respectively. Therefore, the portable sensor can provide good quality of data for d.b.h. measurement.



DBH measurement from the high grade sensor (VZ400, left) the portable sensor (LMS511, right) (cm)

Figure 2: Comparison of d.b.h. measurement between the higher grade and portable sensor

The simulation result from "large tree based stratified sampling" is shown in Figure 3. This result shows that the stratified sampling using the large tree (tree height above 20 m) has less than 5% errors at around 30 plots. Thirty plots are the optimized number found by this simulation. It is easier to find the large trees in woods and come back to the same location in the future. The large tree based sampling technique is a new approach and is useful method in finding location when it is compared with the conventional random and systematic sampling design.

As a result, the plot establishment using the portable laser sensor requires three scan positions for one plot. A large tree like tree height above 20 m is a good location to become the plot centre. And the minimum thirty plots are needed to estimate large scale tropical biomass accurately with 5 % error (Figure 4).

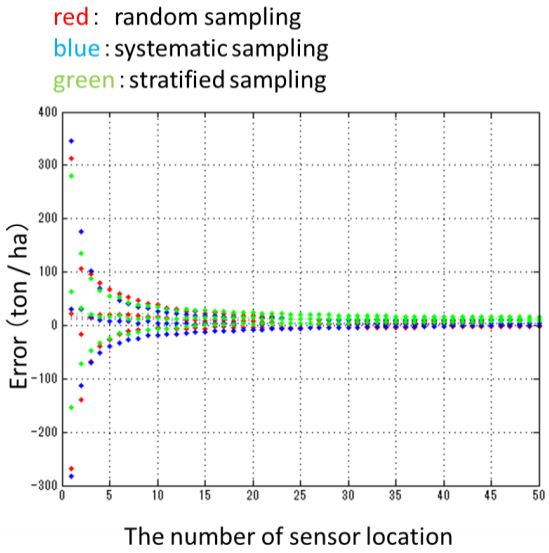


Figure 3: Simulation result from different sampling designs. All sampling designs approach within 5% error at 30 plots.

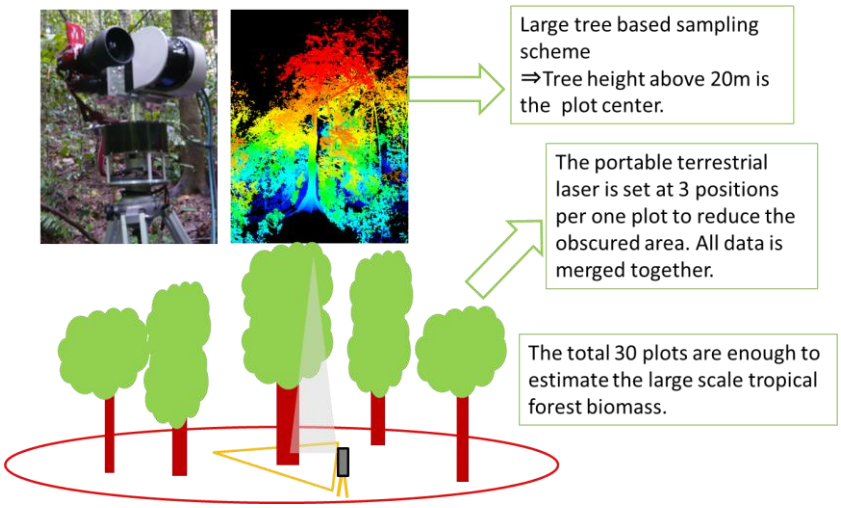


Figure 4: Plot establishment using the portable terrestrial laser.

Reference

[1] Kelbe, D, Romanczyk, P., &van Aardt, J. (2012) Automatic extraction of tree stem models from single terrestrial lidar scans in structurally heterogeneous forest environments. 12th International conference on LiDAR applications for assessing forest ecosystems, SilviLaser

[2] Jaakkola, A., Hyyppä , J., Kukko, A., Yu, X., Kaartinen,H., Lehtomäki, M., Lin, Y. (2010) A low-cost multi-sensoral monile mapping system and its feasibility for tree measurements. ISPRS Journal of Photogrammetry and Remote Sensing 65(6), 514-522.

[3] Yang, X., Strahler, A., Schaaf, C.B., Jupp, D.L.B., Yao, T., Zhao, F., Wang, Z., Culvenor, D.S., Newnham, G.J., Lovell, J.L., Dubayah, R.O., Woodcock, C.E., Ni-Meister, W. (2013) Three-dimensional forest reconstruction and structural parameter retrievals using a terrestrial full-waveform lidar instrument (Echidna). Remote Sensing of Environment 135, 36-51.

[4] Bradford, M.G., Metcalfe, D.J., Ford, A., Liddell, M.J., McKeown, A. (2014) Floristics, stand structure and aboveground biomass of a 25-ha rainforest plot in the wet tropics of Australia. Journal of Tropical Forest Science 26(4), 543-553.

[5] Brown S (1997) Estimating biomass and biomass change of tropical forests: a primer. FAO Forestry Paper 134. FAO

[6] Kato, A., Moskal. L.M., Schiess, P., Swanson, M.E., Calhoun, D., Stuetzle, W. (2009) Capturing tree crown formation through implicit surface reconstruction using airborne laser data. Remote Sensing of Environment 113(6), 1148-1162.

Mapping regional forest aboveground biomass using lidar remote sensing with a calibrated global forest canopy height map

Nian-Wei Ku and Sorin C. Popescu

Department of Ecosystem Science and Management, Texas A&M University

Highlights: The study validated and calibrated a global forest canopy height map with airborne lidar data in the Southeast US and created the biomass map in the East Texas forests. The investigations show the calibrated global forest canopy height map can be used to extend forest aboveground biomass maps to regional-scale.

Key words: *forest biomass, airborne lidar, forest canopy height, map validation, upscaling*

Introduction

This study employed airborne lidar remotely sensed data and a global canopy height map to estimate forest biomass from local to regional scales. The first objective of this study is validating the global forest canopy height maps with airborne lidar data in the Southeast United Sates. As a second objective, we upscaled the local aboveground biomass estimates to the extent of East Texas forests. Three global forest canopy height maps have been released in the past few years. Lefsky (2010) created the first global forest canopy height map in 2010 and Simard et al. (2011) generated another height map in 2011. Los et al. (2012) have also created a vegetation height product in 2011. Simard's global forest canopy height map is the primary map for the study. The map integrated the data of the Geoscience Laser Altimeter System (GLAS) with Shuttle Radar Topography Mission (STRM), climatic and other ancillary data with the regression tree method, Random Forest, to generate a global forest canopy height map. Moreover, Simard' map covered large area of forests and involved more mosaic crops, open forest, and saline flooded forests. Therefore, Simard's map reveals not only the tall trees in forests, but also the short woody plants on the ground (Simard et al., 2011).

Bolton et al. (2012) have initially investigated the accuracy of Lefsky and Simard's maps over Canada with airborne lidar data. They concluded that the relationship between airborne lidar data and the global canopy height maps can be used to develop approaches for more accurate aboveground biomass estimates. For the second objective, the calibrated global canopy height map was used to upscale the forest aboveground biomass from local scales to larger extents in Texas. The validation and calibration procedure of the first objective will be applied to the canopy height map over East Texas forests to generate a forest aboveground biomass map. Primary interest of this study is in generating a forest aboveground biomass map for East Texas.

The specific objectives of the study are 1) validate and calibrate a global forest canopy height map with lidar-derived canopy height information on Southeast United States forests, and 2) upscale the forest aboveground biomass estimates from local to regional scale by using the calibrated global forest canopy height maps in East Texas.

Material and methods

The airborne lidar data is collected between 2010 and 2012 from the following sources: (1) NASA's Goddard's LiDAR, Hyperspectral & Thermal Image (G-LiHT) program; (2) National Ecological Observatory Network's (NEON) prototype data sharing program; (3) NSF Open Topography Facility; and (4) Spatial Sciences Laboratory (SSL) in the Department of Ecosystem Science and Management at Texas A&M University. The airborne-scanned areas are located at the states of Alabama, Arkansas, Florida, Maryland, Mississippi, North Carolina, South Carolina, Tennessee, Texas, and Virginia. Therefore, the airborne-scanned areas cover a wide variety of vegetation types across the Southeast United States and the airborne lidar data is used to validate and calibrate the global canopy height map.

Airborne lidar data is preprocessed to acquire lidar-derived metrics (Table 1) and generate four different levels of point cloud data. Because short vegetation with high density airborne lidar points cloud data potentially affect the statistical calculation of forests, we assume that the open ground is less than 1 m, understory in the forests is less than 3 m, and shrubs are less than 5 m. The four levels of point cloud data will be 1) point cloud data with ground, 2) greater than 1 m, 3) greater than 3m, and 4) greater than 5m. Each point cloud data level is processed individually and result in seven airborne lidar-derived metrics.

Simard's global forest canopy height map can be downloaded at the NASA Jet Propulsion Laboratory website (link: http://lidarradar.jpl.nasa.gov/). The airborne lidar-derived metrics in various locations will be used to validate the accuracy of global forest canopy map. The root mean square error (RMSE) and the coefficient of

determination (R^2) are used for examining the discrepancies of canopy height between the airborne lidar-derived metrics and the global forest canopy height map, and assisted to calibrate the global forest canopy height map.

Table 1. The list of lidar-derived metrics from airborne lidar point cloud data

I. Point cloud data with ground	II. Point cloud height ≥ 5 m
a.Maximum Height	a.Maximum Height
b.Mean Height	b.Mean Height
c.50th percentiles Height	c.50th percentiles Height
d.75th percentiles Height	d.75th percentiles Height
e.90th percentiles Height	e.90th percentiles Height
f.95th percentiles Height	f.95th percentiles Height
III. Point cloud height ≥ 3 m	IV. Point cloud height ≥ 1 m
a.Maximum Height	a.Maximum Height
b.Mean Height	b.Mean Height
c.50th percentiles Height	c.50th percentiles Height
d.75th percentiles Height	d.75th percentiles Height
e.90th percentiles Height	e.90th percentiles Height
f.95th percentiles Height	f.95th percentiles Height

The second objective extends the aboveground biomass estimates from local scale to a regional scale and result in an aboveground biomass map in East Texas refer to the global forest canopy height map. The materials include global forest canopy height map, vegetation type map which was produced by LANDFIRE program of US Department of Interior and US Department of Agriculture, Forest service, and local-scale aboveground biomass maps in East Texas. Then the vegetation type map will assist to distinguish the different type of vegetation in East Texas. Nelson et al. (2009) used Landsat scenes and unsupervised clustering technique to generate a 23-class land cover map. Sequentially, they combined the classification map and the local vegetation zone maps for samplings and deriving regional-scale aboveground biomass estimates. According to Nelson's methodology, the vegetation type map in East Texas is used to distinguish the vegetation type on the global forest canopy height map. The different vegetation type affects the coefficients and parameters in the regression equations for scaling up the aboveground biomass estimates.

The global forest canopy height map and local-scale aboveground biomass maps are correlated to each other and generated the aboveground biomass estimate regression equations. The local-scale aboveground biomass map in East Texas is resampled at 1 km spatial resolution to match the cell size of the global forest canopy height map. The relationship between pixel values of the 1 km local-scale aboveground biomass map and the height statistics of global forest canopy height map is used to develop the regional-scale aboveground biomass estimation equation. This equation is applied to the extent of East Texas using the global forest canopy height map to derive the aboveground biomass map of East Texas.

Preliminary Results and discussions

Results for preliminary investigations into the validation of the global canopy height map using airborne lidar data are still undergoing in the progress. We computed various height metrics from airborne lidar at 1x1 km spatial resolution to match the global canopy height map and generated differential maps and error plots for validation. Both differential maps and error plots show differences between mean, maximum, 50th percentile, 75th percentile, 90th percentile, and 95th percentile heights. Some preliminary results indicate error patterns of the global canopy height model that are dependent on actual canopy height. Quantifying such errors will allow us to calibrate the global canopy height map and improve estimates of biomass at regional scale.

Conclusions

In summary, the study investigations will utilize lidar data collected from multiple platforms to estimate aboveground biomass at multiple scales. The methodology developed for this study could be used to extend forest biomass maps to larger national and continental scales.

References

Lefsky, M. A. (2010). A global forest canopy height map from the Moderate Resolution Imaging Spectroradiometer and the Geoscience Laser Altimeter System. Geophysical Research Letters, 37(15), L15401.

Los, S., Rosette, J., Kljun, N., North, P., Chasmer, L., Suarez, J., et al. (2012). Vegetation height products between 60° S and 60° N from ICESat GLAS data. Geoscientific Model Development, 5, 413-432

Simard, M., Pinto, N., Fisher, J. B., & Baccini, A. (2011). Mapping forest canopy height globally with spaceborne lidar. Journal of Geophysical Research: Biogeosciences (2005–2012), 116(G4)

Bolton, D. K., Coops, N. C., & Wulder, M. A. (2012). Preliminary investigation of the accuracy of global canopy height products over Canada using airborne lidar data. SilviLaser 2012, SL2012-062

Integrating lidar-derived canopy structure metrics into evapotranspiration modeling

Cameron K. Houser

Forest Resources and Environmental Conservation, Virginia Tech, Blacksburg, Virginia

Highlights Understanding the dynamics of evapotranspiration is essential for monitoring and predicting landscape water use and vegetative drought stress. Surface characteristics, such as canopy structure, are important but difficult to estimate modelling parameters that can influence ET estimates. Current methods of estimating canopy characteristics for evapotranspiration modelling include using remote sensing-derived fractional cover estimates and land surface classifications to assign parameters such as roughness length. Given that roughness length is a height-based parameter, Lidar can be used to characterize canopy structure and its relationship to forest-level evapotranspiration.

Key words: **Evapotranspiration, Lidar, Canopy structure**

Introduction

Evapotranspiration (ET) modeling using remotely-sensed thermal data is an effective method for estimating water use because it can be used across varied spatial scales, it is cost-effective useful in areas of the world that are lacking detailed meteorological data [1]. High accuracy models exist for corn and soy bean agricultural fields and there is interest in developing methods to apply such models to managed forests [2]. Canopy resistance and height parameters in evapotranspiration modeling are often either experimentally calibrated value or a constant derived based on vegetation type. The disadvantage of parameterizing models based on experimental methods is that the results are only locally valid while, parameterizing by vegetation cover type ignores variation in forest structure [3]. Additionally, efforts to account to for variation in canopy structure often assume a "Big Leaf" model in which the source of sensible and late heat are assumed to be at the same height and temperature [4].

Lidar sensors have been used to successfully characterize and monitor structural vegetation characteristics across landscapes [5]. The integration of lidar data into evapotranspiration modeling will allow for a more dynamic account of variability in canopy heterogeneity. Therefore, implementing lidar-derived estimates of canopy structure is the necessary next step incorporating canopy-level heterogeneity in to evapotranspiration modeling.

Research

The goal of this research is to use lidar data taken with NASA's G-LiHT (Goddard's Lidar, Hyperspectral,Thermal) sensor and, in conjunction with field-based and eddy-covariance data, incorporate forest structural characteristics into the DisALEXI (Disaggregated Atmospheric Land Exchange Inverse) evapotranspiration model.

Study Area

The study site is located in the lower coastal plains region of North Carolina near Plymouth, NC. The site is predominantly composed of managed loblolly pine plantation forests, followed by mixed forests. Within the study area, there are two Ameriflux eddy covariance flux towers which cover structurally different areas of forest (i.e. clear cut and mature forests).

Methods

In order to integrate Evaluate and process lidar data collected from G-LiHT (Goddard's Lidar, Hyperspectral, Thermal) sensor (Fig. 1A) and derive canopy-related lidar metrics such as the Canopy Height Model (CHM; Fig.1B) for the site. The lidar-derived metrics are then scaled and formatted for incorporation into DisALEXI (Fig.1C).

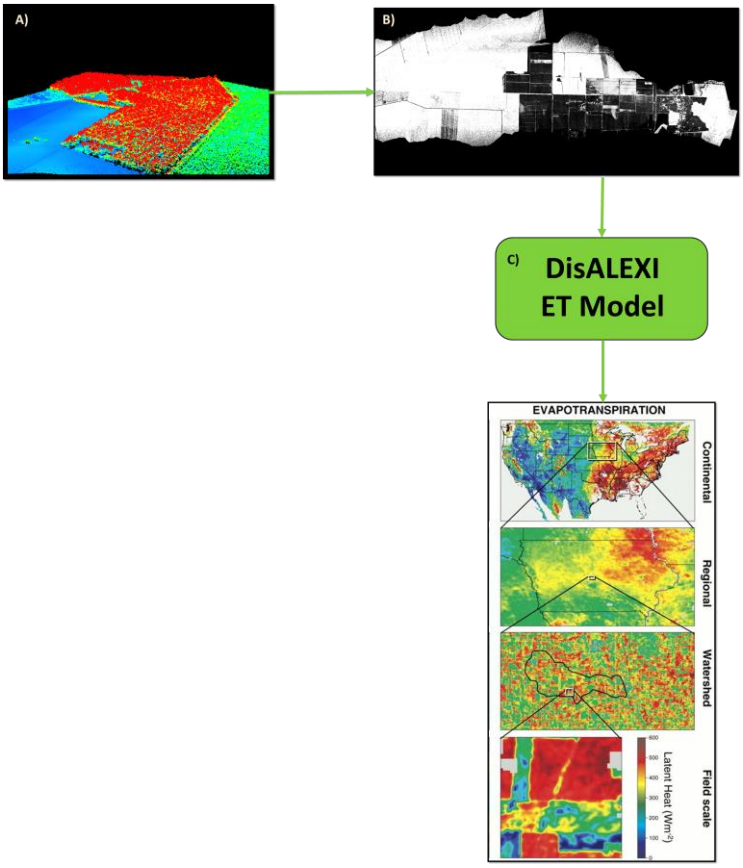


Figure 1: **2** A conceptual workflow for the integration of lidar data into DisALEXI. (Fig. 1 D adapted from Anderson et al. 2007)

Summary

By incorporating lidar data, we can produce accurate stand-level estimates for forests because the integration of 3D structural data into evapotranspiration modelling will allow for a more dynamic account of variability in canopy heterogeneity. Additionally, incorporating realistic values of forest height and structure will allows us to to better characterize the importance of forest structure to evapotranspiration dynamics.

References

The following format for journals [1] books [2] and chapters in an edited book [3] should be used:

[1] Van der Geer, J., Hanraads, J. A. J., & Lupton, R. A. (2010**)**. The art of writing a scientific article. Journal of Scientific Communications, 163, 51–59.

[2] Strunk, W., Jr., & White, E. B. (2000). The elements of style. (4th ed.). New York: Longman, (Chapter 4).

[3] Mettam, G. R., & Adams, L. B. (2009). How to prepare an electronic version of your article. In B. S. Jones, & R. Z. Smith (Eds.), Introduction to the electronic age (pp. 281–304). New York: E-Publishing Inc.

[1] Anderson, M., Kustas, W. (2008). Thermal remote sensing of drought and evapotranspiration. Eos (Washington. DC)., 89, 233–234.

[2] Choi, M., Kustas, W. P., Anderson, M. C., Allen, R. G., Li, F., Kjaersgaard, J. H. (2009). An intercomparison of three remote sensing-based surface energy balance algorithms over a corn and soybean production region (Iowa, U.S.) during SMACEX. Agricultural and Forest Meteorology, 149, 2082–2097.

[3] Schaudt, K. J., & Dickinson, R. E. (2000). An approach to deriving roughness length and zero-plane displacement height from satellite data, prototyped with BOREAS data. Agricultural and Forest Meteorology, 104(2), 143–155.

[4] Fisher, J. B.; Debiase, T. a.; Qi, Y.; Xu, M.; Goldstein, A. H. (2005) Evapotranspiration models compared on a Sierra Nevada forest ecosystem. Environmental Modelling and Software, 20, 783–796.

[5] Popescu, S. C., & Wynne, R. H. (2004). Seeing the trees in the forest: Using lidar andmultispectral data fusion with local filtering and variable window size for estimating tree height. Photogrammetric Engineering and Remote Sensing, 70(5), 589–604.

Multi-footprint airborne LiDAR in forest vegetation

Ilkka Korpela

University of Helsinki, Department of Forest Sciences

Highlights: Multispectral visible-NIR LiDAR data are a tempting option as directional reflectance properties of vegetation scatterers are spectrally variant, when sampled across a wide spectrum. Multifootprint LiDAR measures structure. This study examines the option of using waveform data from pulses of maximal eye-safety irradiance. An interesting normalization of amplitude data was carried out.

Key words: *Divergence, Radiometric normalization, Waveform, Canopy structure*,

Introduction

Airborne LiDAR data have become an integral part of forest inventories. Efficient, customized and reliable sensors attract customers. An example is the multispectral LiDAR. Titan by Optech uses λ of 532, 1064 and 1550 nm. The sparse set of bands may well suffice for many forestry applications, as directional reflectance across far-away bands is less correlated. Vegetation is dark in the visible range so that sensitive receivers are needed, because the pulse power is limited by eye-safety concerns.

No airborne LiDAR or camera sensors are optimized for forest RS, and multi-purpose sensors are used instead. Simulations have highlighted how certain sensor properties influence the information available. Examples include the pulse length in LiDAR or the spectral response functions of multispectral cameras. Simulations are inexpensive compared to experimental research, but constrained by the inherent inaccuracy of the needed geometric-optical vegetation models as well as that of the atmospheric effects.

Airborne LiDAR is about 'sampling the scene with time-stamped photons'. The circular illuminated area, the footprint, has a Gaussian power distribution. For a fixed pulse power, the received signal is independent of divergence in large, flat and homogenous targets. When the beam divergence increases, the at-sensor signals weaken for twigs or leaves, as the irradiance incident on them decreases with increasing footprint. Eye-safety limits the maximal footprint energy that can enter a binocular aperture. Output power and energy have to be controlled, when acquisition height or the beam divergence change. The λ influences eye-safety as well. The maximum irradiance near the footprint center for a binocular aperture is approximately equal for footprints that are clearly larger than the aperture. For fixed divergence, the dependency of the maximum peak pulse power and the eye-safe range is rather linear until the power and range are such that the footprint becomes smaller than the binocular aperture.

The volumetric canopy is scanned over a range of zenith angles. A canopy point has a certain sky view factor comprised of gaps. A more exact description of the local directional illumination accounts for multiple scattering and the transmission of light in the foliage. The canopy can also be considered as a volume with a directional differential backscatter cross-section in addition to the directional function of illumination incident, which has a direct and diffuse component. Thus, the high directionality of LiDAR view-illumination geometry is altered by multiple scattering (returning pulses elongate always), which contributes up to 50% in NIR, in favorable foliage geometry. Of all photons emitted a fraction reaches the receiver aperture and all have different travel time. The speed and sensitivity of the receiver determines if the signal is meaningful or covered by noise. A stable transmitter is needed for accurate interpretation aswell. For fixed pulse power and receiver sensitivity, a sensor with smaller beam divergence delivers more observations from the forest floor and the maximal signals for small targets are stronger as the irradiance near the footprint center is higher. The wide pulses are more likely to result in transmission losses, i.e. weak backscattering that does not trigger the waveform recording or remains in the noise. However, if the footprint irradiance can be set to the maximum eye-safety level for any footprint diameter, the same energy ultimately reaches the forest floor as with the narrow pulses, and the wider pulses will deliver more backscattered energy and detectable echoes. This makes the signal distributions to vary according to footprint diameter, which is interesting if one considers multi-footprint data for the estimation of canopy structure.

This experimental study looks at the influence of eye-safe variable-size footprints on waveform-recording LiDAR data. The investigations are confined to boreal forests in Finland and the main rationale was the exploration for usable information for tree species discrimination.

Materials

We experimented in Hyytiälä (62°N, 24°E). Permanent plots have accurately positioned trees. The status of trees was verified in high-density LiDAR and image data of 2013. The area is 600 ha and is systematically covered by over 1500 aerial images and 18 airborne LiDAR datasets since 1946/2004.

Individual trees were manually measured for an accurate and up-to-date 3D treetop position and crown shape, using aerial images and dense LiDAR data from 2012 (Fig. 1). Only trees that were discernible in these data were included in the single-tree analyses. Most dominant, co-dominant and intermediate trees belong to this set, especially as the tree positions were known to the operator.

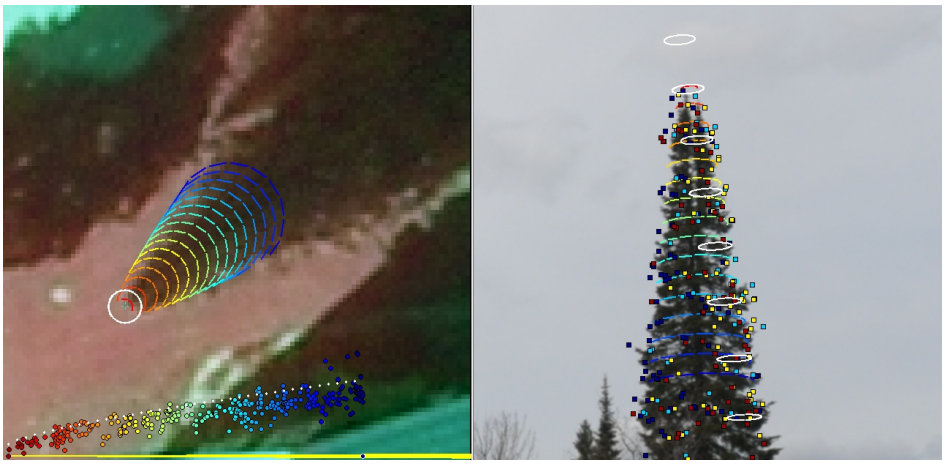


Figure 1: Example of simultaneous treetop positioning and estimation of the crown envelope. The white circle in the aerial image shows the image position of the treetop, and the resulting 'epipolar cylinder' is shown in the terrestrial image with the circles. The cylinder intersects with the LiDAR points that are superimposed in the terrestrial view.

The WF data were acquired with a Leica ALS60 sensor (Tables 1, 2). The pulse peak power of the ALS60 is adjustable at 1% intervals, i.e. 20 dB in total. The pulse power in ALS60 is set by accounting for an average range, desired signal-to-noise ratio, an expected reflectivity of targets and the binocular eye-safety. The receiver front-end implements an automatic gain control (AGC), which varies 0–4 dB in response to scene brightness. The gain was fixed in the 2013 dataset. The influence of the AGC can be normalized, using a stored voltage value of each pulse. The needed large dynamic range of LiDARs is realized by the transmitter in ALS60 in contrary to many other sensors, which may have several receivers (after the photodiode) that differ in gain. The 2012 multi-footprint data was acquired by adjusting peak power to reach the same at-sensor signal for the same well-defined targets from varying altitudes, while the 2013 multi-range data was acquired, using fixed power and receiver gain. ALS60 is primarily a discrete-return (DR) sensor and waveform (WF) storage calls for an optional digitizing module. It is important to note that the storage is started by a threshold-exceeding signal that triggers the first DR echo. The WF includes some 30 samples that precede first echo, so that the 256-sample WF can capture height differences of about 34 meters. The WFs (amplitude data) were normalized for the AGC and atmospheric attenuation (0.2 dB/km) , then mildly low-pass filtered, and attributes such as the peak amplitude, FWHM, total energy, length, and rise speed were determined for continuous sequences that raised above the apparent noise. There could thus be several sequences in a WF, but the first was the most interesting as it comprises least by apparent backscatter cross-section. The (above-the-noise) length of the continuous sequences depended on the signal strength, while the FWHM width measure was less affected. The shorter 7.8 ns pulses, which are used when the PRF is above 120 kHz, have a sharp ascend, and the FWHM (WF shape, receiver response, rise of the pulse) depended on the signal strength, which is a phenomenon not observed for the 'slower' 10 ns pulses.

Results and Discussion

Receiver linearity, i.e. the linearity of the amplitude values w. r. t the at-sensor irradiance, was examined first as it is crucial. The overlapping multi-range, fixed power data of 2013 were used for this, and the signals from well-defined targets of varying reflectance from the three heights were contrasted. The radar equation states that the signal strength should vary with the second power of range ratios. Figure 2 shows that the sensor differed slightly from perfect linearity. This observation, if it is true for all Leica ALS50/60 sensors, explains why earlier tests for the effect of range have failed. The results apply only to peak amplitude, and the non-linearity may be due to the photodiode response (unlikely), signal-dependent band-width differences of transimpedance amplifiers, or a non-linear response and limited band-width of the oscilloscopes used for digitizing the WF. In addition, the non-linearity may change when the AGC is operating as it is band-limited amplifier. Example of the effect of the footprint size in surfaces that comprise different

level and scale of reflectance variation are shown in table 3. The radiometric normalization that accounted for differences (deviations from the nominal, Table 2) in pulse power and atmospheric attenuation as well as the AGC, resulted in peak amplitude data were that were within 5% from all acquisition heights. This meant that the objective of acquiring 'maximally eye-safe' data was successful. The WF attribute histograms and the measures for canopy-induced transmission losses differed in canopies representing different species, age and density. If these differences can be harnessed to assist for example species identification remains an interesting future topic in multi-footprint LiDAR.

Table 1. Sensor parameters.

Property/Sensor	ALS60 + WDM65
Wavelength, nm	1064
WF of transmitted pulse	No
Divergence, mrad, $1/e^2$	0.22
Front-end amplifier	Automatic Gain Control
Triggering of WF-recording	Discrete-return circuit
WF-recording sequence	Single, 128 or 256 samples
WF Sampling rate, ns	1 or 2
Discrete-return echoes	on-the-fly; 1, 2, 3 or 4
Echo attributes	range, XYZ, intensity
Acquisition height, m	500, 1000, 2000, 2700

Table 2. Characteristics of the LiDAR acquisitions. In dataset names 2012_#, # stand for nadir range in hectometers.

Dataset	2012_ (05, 10, 20, 27)	2013b_ (07, 08, 09)
Purpose	Multi-footprint data	Linearity tests
Date, local	July 5, 2012	June 15–16, 2013
Altitude, m	500, 1000, 2000, 2700	700, 800, 900
WF density per m^2	5.3, 5.4, 3.2, 2.2	4.7, 4.1, 3.6
Emitted pulse, FWHM, ns	7.8, 10.1, 10.3, 10.4	7.8
Nominal power, % of max	3, 14, 55, 100	15
Footprint, $1/e^2$, m	0.11, 0.22, 0.45, 0.59	0.15, 0.18, 0.20
AGC	on	off

Table 3. Coefficient of variation of peak amplitude. Surfaces are in decreasing order of reflectance (0.45→0.04).

Surface	Footprint, cm			
	11	22	44	59
Hay, < 50 cm	7.0	5.8	4.6	3.6
Short grass	4.9	4.3	4.0	3.5
Mire surface	9.7	8.3	6.3	5.4
Fine sand	4.5	3.5	3.0	2.8
Asphalt, old	7.4	7.3	6.9	7.2
Asphalt, new	5.1	4.9	5.0	5.1
Gravel road	11.4	9.7	7.7	6.9
Bitumen	9.4	9.4	9.0	8.6

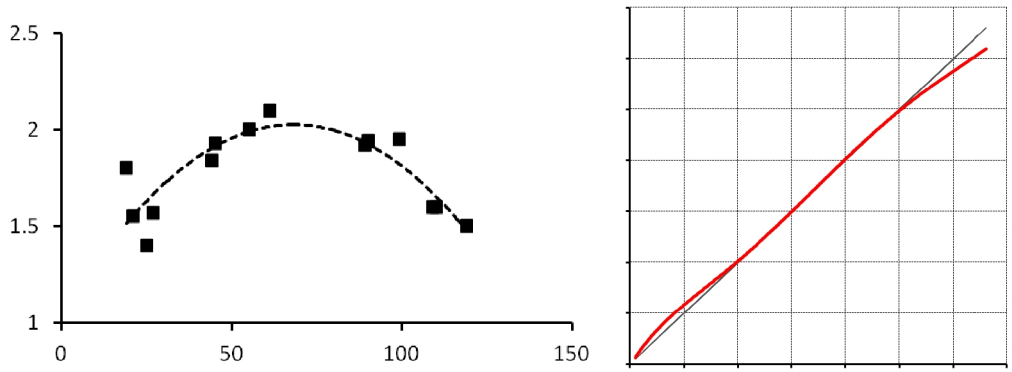


Fig 2. Left: Estimated power terms of range correction for the sensor as a function of the amplitude, i.e. scene brightness. Right: An illustration of nonlinearity that produces the response in the left sub-figure. Horizontal axis has the at-sensor radiance and the vertical axis is the observed amplitude.