Detecting Multi-layered Forest Stands Using High Density Airborne LiDAR Data

Jan-Peter Mund, Robert Wilke, Michael Körner and Alfred Schultz

Eberswalde University for Sustainable Development, Germany · jan-peter.mund@hnee.de
Landesforstbetrieb Brandenburg (LFE), Eberswalde/Germany

Full paper double blind review

Abstract

Since two decades, the use of terrestrial laser scanning (TLS) and Airborne Light Detection and Ranging (LIDAR) has become very prominent in analysing 3D forest structures (AKAY et al. 2009). The potential of full waveform analysis of high density Airborne LiDAR data (ALS) for the detection and structural analysis of multi-layered forest stands is not yet well investigated (JASKIERNIAK et al. 2011), although ALS data provide exact information on tree heights of multi-layered forest stands using particular laser pulse characteristics (GAULTON & MALTHUS 2010).

Since the mid-19th century, managed forests in Brandenburg have been dominated by Scots pine monocultures. In the last fifteen to twenty years many forest stands were converted into multi-layered mixed forests by silvicultural conversion of forests and natural succession (MLUR 2004). Today, the majority of forest stands in the federal state of Brandenburg remain dominated by Scots pine (*Pinus sylvestris*) in the canopy layer, while European beech (*Fagus sylvatica*) or Sessile oak (*Quercus petraea*) are predominant in the understorey.

In this study, we investigate and discuss the potential of full waveform high density airborne LiDAR data (ALS) for detecting, classifying, and stratifying discrete vegetation layers at forest stand level, based on 0.1ha investigation plots. Full waveform high density ALS data on each 5th percentile level was used in combination with binary logistic regressions to discover the structural layer type of multi-layered forest stands from normalized discrete ALS pulses. The results of the descriptive statistics of ALS point clouds and binary logistic regression models produce particular forest layer profile indices of understorey vegetation and canopy layer. Such parameters can further be used as variables for forest structure analysis algorithms, and can be empirically tested against stand characteristics. The validation of ALS data and model results is tested against empirical forest mensuration data of the "Datenspeicher Wald 2 (DSW 2-Forest inventory data)" and field survey reference points using error matrices.

We demonstrate that binary logistic regression analyses are functional for establishing a prediction model. The model was applied successfully on larger forest stands and forest areas, and can become useful for identifying and separating single from multi-layered forest stands using percentiles of total amounts of ALS return pulses on a 10x10m raster size with a high overall accuracy of 90%. The established model has the potential for a broad range of forest management applications, such as timber inventory evaluation, forest growth

GI_Forum – Journal for Geographic Information Science, 1-2015.

modelling, monitoring of vegetation dynamic and succession, as well as ecological classifications and the detection of deadwood in forest stands (KIM et al. 2009).

1 Introduction

Technologies like terrestrial laser scanning (TLS) and airborne laser scanning (ALS) are accepted and effective methods for the structural analysis of forest physiognomy (WOODS et al. 2008; YU et al. 2004). Standardized methods and applications of TLS cover individual tree observations, rapid collection of mensuration data from forest stands up to large forest areas, as well as observation and monitoring (BALTSAVIAS 1999, KRAUS 2002, WEHR 1999). Today ALS data is used for the standardized three-dimensional analysis of individual trees, larger forest areas, and managed forest stands (PIROTTI 2011, ZHANG et al. 2011). However, the potential of high-density airborne LiDAR data (ALS) for the detection and structural analysis of multi-layered forest stands is not yet well investigated (SCHMID et al. 2008, DRAKE et al. 2003, KIM et al. 2009).

Typical parameters and variables derived from TLS and ALS using simple laser pulse characteristics are stand heights and number of stems, leaf area index and volume, or biomass at stand level (LEFSKY et al. 1999, MEANS et al. 2000). Also, researchers investtigated ALS and full waveform laser data to calculate forest mensuration and inventory-related variables such as tree density and age structures (HEURICH 2008). Different methods and algorithms for a semi-automated extraction of 3D tree measurements and structures were established, but only a few studies explored statistical indices which represent the vertical profile of vegetation structure in LiDAR data (PIROTTI 2011). Most of these studies applied the unimodal Weibull distribution function to ALS point clouds to determine canopy models. ALS data provides exact tree height information of forest stands, but in a more complex, multi-layered forest ecosystem, typical spatial patterns of LiDAR point clouds may be highly variable between forest types and age classes, thus unimodal distribution functions do not fit well as a prediction model (WOODS et al. 2008). To monitor silvicultural conversion rates or to identify ecological succession processes, researchers collected standardized parameters at forest stand level. Information about its status and the progress of such sustainability measures are required for forest management plans and sustainable forest growth scenarios (JASKIERNIAK et al. 2011). However, collecting data about the forest structure, its biomass, and ecosystems services is still difficult and cost-intensive using traditional forest mensuration methods and terrestrial forest inventories HEURICH (2008).

Today, still the majority of forest stands in the federal state of Brandenburg are dominated by Scots pine (*Pinus sylvestris*) in the canopy layer, while European beech (*Fagus sylvatica*) or Sessile oak (*Quercus petraea*) are predominant in the understorey (MLUR 2004). It is still challenging and cost-intensive to collect exact measurements and spatial information about the silvicultural conversion process beyond forest mensuration or inventory sample points.

In this study we investigate, analyse, and discuss the potential of high density ALS data (> 25 points/m², see Table 1) for detecting, stratifying, and classifying discrete vegetation layers at forest stand level based on 0.1ha investigation plots. High density LIDAR point cloud data and binary logistic regression analyses are used to differentiate single and multi-layered forest structures at stand level. The tested methods and experiences may serve as a

guideline to set-up a semi-automated detection algorithm for multi-layered forest stands in the future. The results can be useful for a broad range of forest management applications such as timber inventory evaluation, forest growth modelling, monitoring of vegetation dynamic, and succession as well as ecological classifications.

2 Study Area

All investigation and validation plots of this study are located in the north east of Germany in the federal state of Brandenburg, south-east of Eberswalde in the Region of Spechthausen − Melchow (see Fig. 1). The high density ALS point cloud (>25 points/m²) was provided by the federal state forest agency of Brandenburg, LFE (KÖRNER 2014).

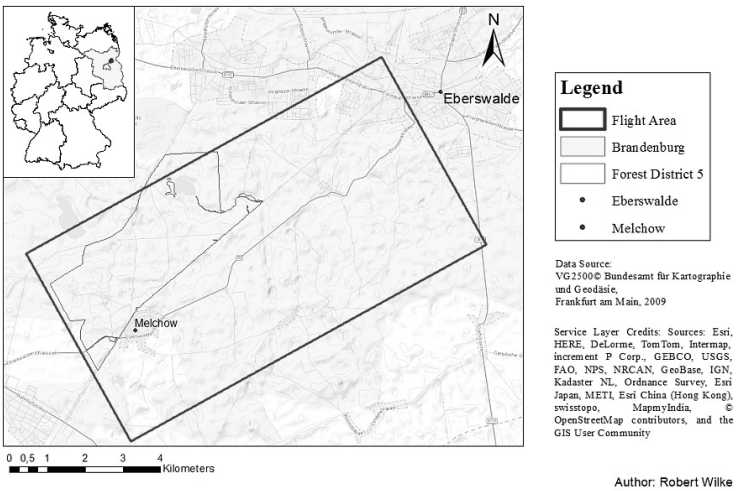

Fig. 1: Flight area and Forest district No 5 in the Eberswalde region (WILKE 2014)

The selected Brandenburg forest district No 5 is dominated by Scots pine forest stands of different age classes with a share of 66% (see Fig. 2 & Fig. 3). The implementation of sustainable forest management concepts and political pressure accelerated a continuous silvicultural restructuring of single layered forest (SCHRÖDER et al. 2014).

The silvicultural restructuring and sustainable forest management aims at producing an ideally mixed forest within the next 40 years, consisting of a multi-species and multi-layered forest structure, including deciduous tree species like European beech or Sessile oak (MLUR 2004). Since the year 2000, several forest stands in Brandenburg have been restructured, resulting in such a multi-layered forest structure (see Fig. 3). Inside the ALS survey area, a number of 10 ALS investigation and terrestrial validation plots (60m x 60m) were selected randomly. For evaluation purposes, an additional terrestrial survey and manual forest mensuration was carried out in March 2014 using a dGPS Survey Controller R4 Version 10.0 by Trimble.

Fig. 2: Typical single layered forest stand dominated by Pinus silv. (WILKE 2014)

Fig. 3: Multi-layered forest stands with Pinus silv. and Fagus silv. (WILKE 2014)

3 Data Used

The full waveform LiDAR data was provided by the Brandenburg Federal state forest management agency "Landesforstbetrieb Brandenburg (LFE)" in the framework of the EU Interreg IVa project ForseenPOMERANIA.

Table 1: Features of original LiDAR data collection

Date	Scanner type	Flight altitude	Flight speed	Point density
March 6, 2013	LMS-Q 680i-400 kHz	600m above ground	222 km/h	25 points/m²

Table 2: Information about the point record of two different investigation plot types

Plot type	Plot size [m²]	Number of point records	Point density/[m²] (all returns)	Point density/[m²] (last returns)
Single layered forest stands	3.608	17.508	49,2	26,4
Multi layered forest stands	3.604	164.610	45,7	22,6

The original survey data was collected on the 6th of March 2013 with a laser point density of 25 hits per square meter (Körner 2014). The raw data at coordinate and attribute level (x;y;z) was controlled and validated by the federal state agency for surveying and geo-basis information in Brandenburg ("Landesbetrieb für Landesvermessung und Geobasisinformation Brandenburg" LGB). Since the majority of ALS and TLS forest studies only focus on the unimodal Weibull distribution function, this study analysed single and multiple laser return pulses of up to 5 return levels in the standardized "las" format using the software LAStools produced by Rapidlasso GmbH 2014. According to GATZIOLIS & ANDERSEN (2008), SCHMID et al. (2008) and JASKIERNIAK et al. (2011) only multiple laser returns per

pulse can be used to analyse the 3D forest structure. For the first overview around 5 million points were displayed with the LAStools-tool LASview.

4 Methodology

Data pre-processing of ALS data was applied using tools such as spatial clipping and data compression in order to reduce the total amount of ALS raw data. Further standardized ALS pre-processing tasks like indexing and quality control of return pulses have been applied. In a second step a point data classification into two discrete classes – ALS canopy returns and ground level pulse returns – is applied, in order to eliminate wired z-values and separate the original point cloud into two vertical layers (KIM et al. 2009). The absolute height level is validated using dGPS measurements and handheld laser tree height measurements. Further conventional ALS data management steps such as processing of a canopy surface model (DSM), a normalized DSM, and calculating a hill shade model HSM on a 10×10cm raster with the LAStools tool LAS2dem were applied, according the guideline for managing and processing LiDAR data, published in 2008 by GATAZOLIS & ANDERSEN (2008).

The DSM model is used to survey the location of investigation plots in the forest, and compare them with the location of the investigation plots in the processed ALS point cloud. Structural and spatial variables and ground level are separated and extracted from the preprocessed ALS point cloud using the normalized surface model (nDSM).

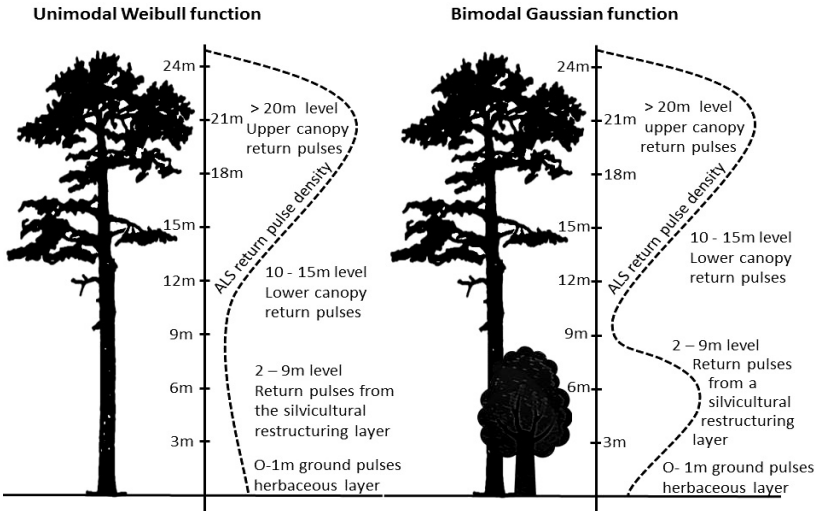

Fig. 4: Unimodal and bimodal function of ALS return pulses representing regrowth of younger trees and silvicultural restructuring process of forest stands (MUND 2015)

A prognosis and stand detection model is required for further data classification, and for the identification of multi-layered forest stands (WOODS et al. 2008). Descriptive statistics

using the Kolmogorov-Smirnov test show that all pre-processed ALS data is not normally distributed. The results indicate a binary logistic regression analysis to establish a prediction model for the detection of multi-layered forest stands. From a list of empirical investigation plots, all multi-layered forest stands are identified and separated from single layered forest stands applying a standardized data processing algorithm and a rule-based classification model. The selection of investigation and validation plots with a similar upper canopy level allows a comparison of results, and supports the calibration of the model. Characteristic ALS attributes and specific spatial structures of ALS return pulses i.e. return pulses with a certain level above ground are chosen as classifiers, separating data from the entire point cloud into different classes.

We compared the density of return pulses and time delay of ALS data for each class level, in order to determine the most suitable bimodal distributions function. Descriptive statistics were applied to select and establish variables for the detection of different layer types (see Table 2). Every vegetation layer influences the number of ALS points in relation to the respective height level (see. Fig. 4 & Fig. 6). The minimum, maximum, and each 5th percentile of height values is calculated for each investigation plot, and then resampled in a two-step approach. Variables for the model have been selected based upon the number of ALS points in relation to the height level. For the majority of all investigation plots, only 3-5 resampled percentiles (p35, p55, p95) contain sufficient ALS data representing main physiognomic forest structures (see Fig. 10 & Fig. 11). Relative level data of each percentile layers above ground characterizes the different forest stand types. The result is a vertically separated and resized point cloud divided into four vegetation height classes (see Fig. 5 & Fig. 6). The class levels differentiate ALS return pulse density, representing different stages of silvicultural forest restructuring from Scots pine monocultures to multispecies mixed forest stands.

A model to identify and predict multi-layered forest stands is established, applying a binary logistic regression. The following equation produces results for forest stands containing two or more tree layers (1) or only one tree layer (0).

$$P(y=1 \mid x_i) = \frac{e^{b_1 * x_1 + b_2 * x_2 + \ldots + b_n * x_n + b_0}}{1 + e^{b_1 * x_1 + b_2 * x_2 + \ldots + b_n * x_n + b_0}}$$

y = second forest stand layer (1 = yes | 0 = no)

b_i = coefficient and constant values

x_i = percentile of occurrence

Coefficients and constant values for the model are based upon relative percentile values from descriptive statistical tests and regression functions. In order to test the best fitting model, the analyses were conducted on two different raster sizes of investigation (5×5m and 10×10m). The model results are transformed into a raster map to provide forest stand characteristics for an entire forest district.

5 Results and Discussion

The results of descriptive statistics for ALS data of 10 investigation plots are separated between single and multi-layered forest stands and presented in Table 2. The validation of forest canopy heights in both n plot types investigated were tested against two different empirical data sets: 40 surveyed forest canopy reference points from a forest mensuration campaign and visualized qualitative attribute data from the Brandenburg DSW 2. The height accuracy of control points and the difference between measured height values and ALS data is about 0,02m. The average absolute error ranges from 0.01m to 0.02m and the standard deviation calculated for ALS data is 0.024m. Classified and height normalized ALS data from both single and multi-layered forest stands are plotted on 2D point clouds to visualize forest stand structure types (see Fig. 4; Fig. 5 & Fig. 6).

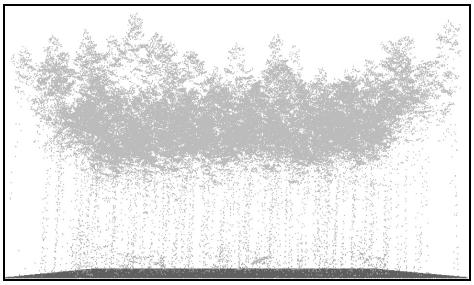

Fig. 5: Typical single layered forest stand

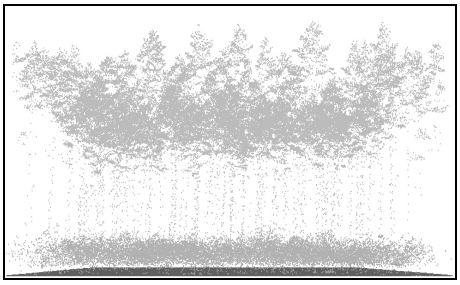

Fig. 6: Multi-layered forest stand with understorey

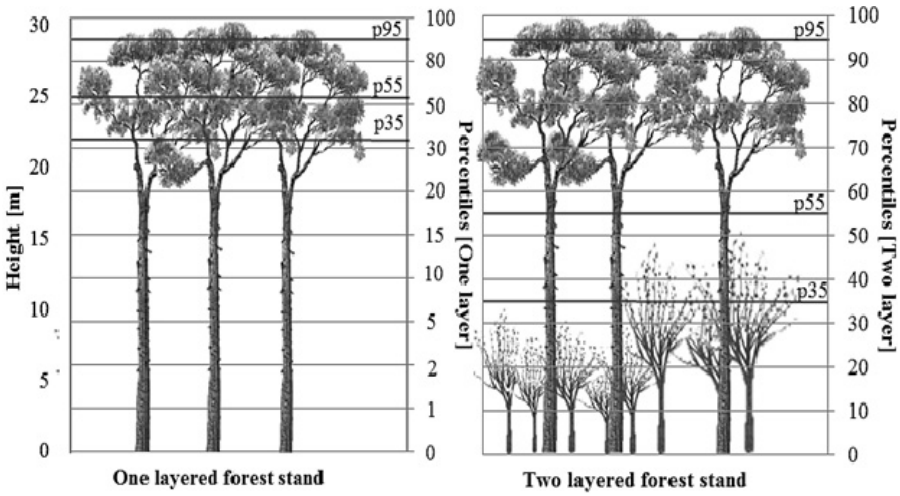

Fig. 7: Characteristics of percentiles p35, p55, p95 for investigation plots (WILKE 2014)

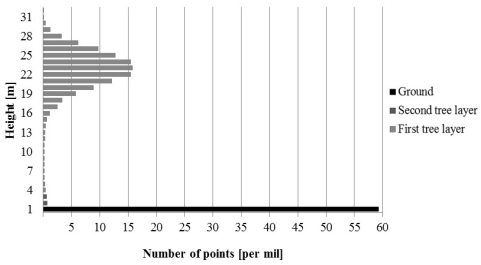

Fig. 8: Univariate distribution of ALS returns from single layered forest stand

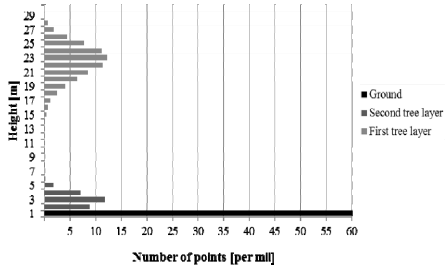

Fig. 9: Bivariate ALS returns from multi-layered forest stand with understorey

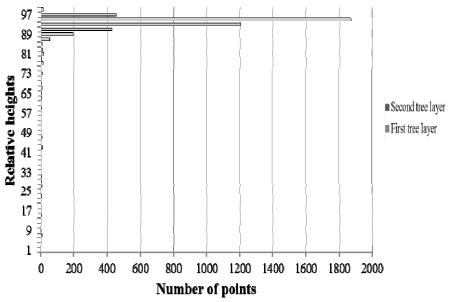

Fig. 10: Distribution of the relative percentiles for p95ra (10×10 m² raster)

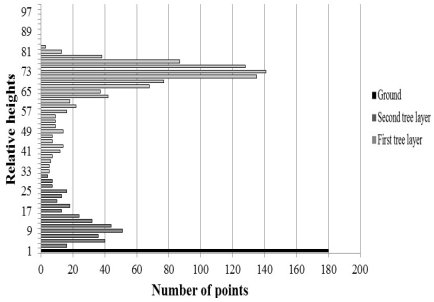

Fig. 11: Distribution of the relative percentiles for p35ra (10×10 m² raster)

Figures 8-11 give a good visual impression and provide quantitative results of the differences between point clouds of single and multi-layered forest structures. A single structured forest stand shows a proportion of 62.4% for the canopy layer to 1.2% of the secondary layer to the total amount of all ALS returns (see Fig. 8).

A multi-layered forest stand shows different proportions, with 38.4% returns from the canopy layer and 18.1% from the secondary or understorey layer (see Fig. 7 & Fig. 8). The ground points in both plots have a share of 35%.

In order to test and establish an independent prediction model using a binary logistic regression function, relative as well as absolute values of percentiles were tested to describe height differences between forest layers. Based upon analytical statistics of relative height values for three different percentile levels (p35, p55, p95, see Fig. 7; Fig. 10; Fig. 11), variables were selected and tested in two binary logistic regression functions on a 5×5m and 10×10m raster model. The application of a 5m² raster instead of the 10m² raster size produced a 50% higher amount of calculated values in total, which might result in an overestimation of predicted multi-layered forest stands. The predictor variables for both models were tested with the *Nagelkerke* R^2 test for the investigation and evaluation plots in detail.

The probability of predicting the structure of forest layers with the help of the predictor variables increased to 82% using the 10×10m raster size instead of the 5×5m raster model. Predictor variables gathered from a 10×10m model produced a 27% higher influence on the result than predictor variables out of the 5×5m raster model. As a consequence, the 10×10m model was used to identify and separate single from multi-layered forest stands in the Brandenburg forest district No 5. Fig. 7 shows the resulting 2D forest structure and Fig. 12 shows a map presenting spatial patterns of predicted variables for the 10×10 m² raster model. Single layered forest stands are displayed in light grey and multi-layered forest stands are coloured in dark grey, while the "No data" class contains forest stands with other tree species such as Norway spruce, Douglas fir and Red alder.

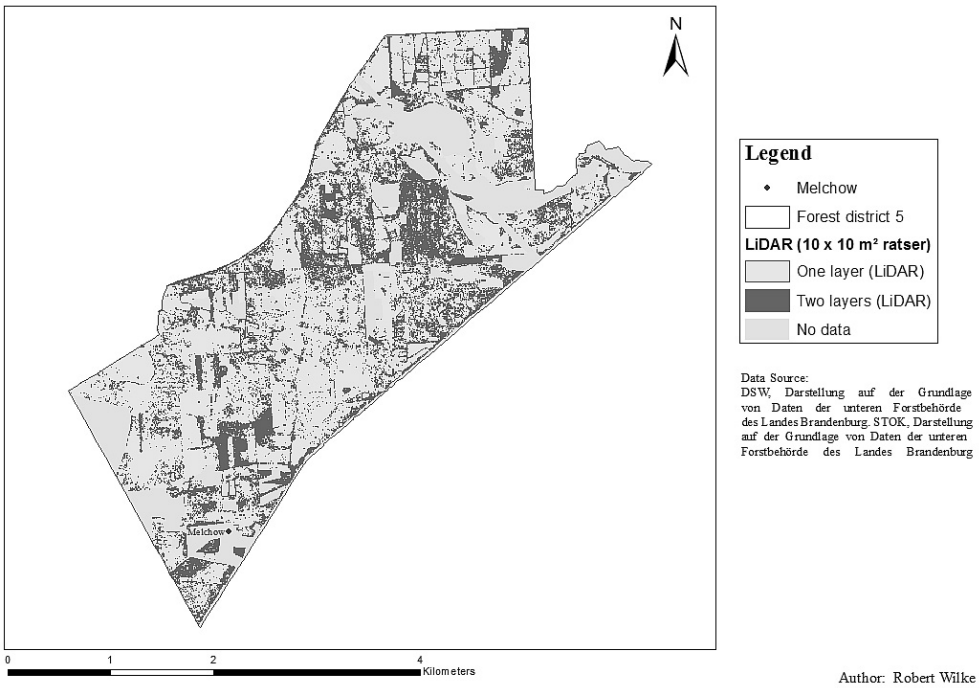

Fig. 12: Spatial pattern of multi-layered forest stands in Forest district No 5, based 10×10m prediction model calculated from a bilinear logistic regression (WILKE 2014)

A comparison of results from both prediction models for single and multi-layered forest stands with empirical data from the Brandenburg DSW 2 data base proves a high level of agreement. Modelling a raster size of 5×5m, the calculated overall accuracy was 65%, while the overall detection accuracy increased up to 90% when applying a raster model of 10×10m to extract the forest variables.

6 Conclusion

In this study, high density ALS data (> 25 pulses/m²) are used to create and test a prediction model to detect and separate single layered from multi-layered forest stands in a forest district in NE Brandenburg. We demonstrate that normalized ALS data and descriptive statistics of multiple ALS return pulses in relation to the height are functional for establishing a prediction model based on each 5th percentile level. Two prediction models of different raster sizes were tested to identify and separate single from multi-layered forest stands. Due to their specific morphology, some tree species like Norway spruce, Douglas fir, or Red alder are not suitable for this prediction model using percentiles. In forest conditions, these tree species produce cylindrical shaped crowns and often have branches down to near ground level. In the ALS 3D point cloud, only open-grown trees can be separated into a distinct trunk section and a crown section.

The established prediction model uses a binary logistic regression of ALS returns, and derives particular forest structure variables. The selected 10×10m raster model produced an overall high accuracy of 90% agreement with empirical validation data. The resulting prediction model has proven a clear potential to detect and separate multi-layered forest stands from single layered forest stands. Further generalization of normalized input data or increasing the raster model size could enhance the prediction model stability, as variations inside forest stands gain less influence on any individual raster.

Acknowledgement

The LiDAR data were provided by the Eberswalde forestry state centre of excellence (LFE) from their EU Interreg IVa project ForseenPOMERANIA. The ForseenPOMERANIA project dealt with the development of a transnational decision support system for remote sensing and model-based estimation and simulation of woody biomass in the forests of the POMERANIA region (SCHRÖDER et al. 2014, KÖRNER 2014).

References

AKAY, A. E., OĞUZ, H., KARAS, I. R., ARUGA, K. (2009), Using LiDAR technology in forestry activities. Environmental monitoring and assessment, 151 (1-4), 117-125.

DRAKE, J. B., KNOX, R. G., DUBAYAH, R. O., CLARK, D. B., CONDIT, R., BLAIR, J. B. & HOFTON, M. (2003), Above-ground biomass estimation in closed canopy Neotropical forests using lidar remote sensing: factors affecting the generality of relationships. Global Ecology and Biogeography, 12, 147-159.

GATAZOLIS, D. & ANDERSEN, H. (2008), A guide to LiDAR data acquisition and processing for the forests of the Pacific Northwest.

GAULTON, R. & MALTHUS, T. J. (2010), LiDAR mapping of canopy gaps in continuous cover forests: A comparison of canopy height model and point cloud based techniques. International Journal of Remote Sensing, 31 (5), 1193-1211.

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.

JASKIERNIAK, A, D., LANEB, P. N. J., ROBINSON, C. A. & LUCIEERA, A. (2011), Extracting LiDAR indices to characterise multilayered forest structure using mixture distribution functions. Remote Sensing of Environment, 115 (2), 573-585.

KIM, Y., YANG, Z., COHEN, W. B., PFLUGMACHER, D., LAUVER, C. L. & VANKAT, J. L. (2009), Distinguishing between live and dead standing tree biomass on the North Rim of Grand Canyon National Park, USA using small-footprint lidar data. Remote Sensing of Environment, 113 (11), 2499-2510.

KÖRNER, M. (2014), Gewinnung forstlicher Informationen durch flugzeuggetragenes Laserscanning. In: Landesbetrieb Forst Brandenburg (Hrsg.). Biomasseschätzung für Wälder mittels Fernerkundung und Modellierung. Eberswalder Forstliche Schriftenreihe, 56. LFB, 45-50.

LEFSKY, M. A., COHEN, W. B., ACKER, S. A., PARKER, G. G., SPIES, A. & HARDING, O. (1999), Lidar Remote Sensing of the Canopy Structure and Biophysical Properties of Douglas-Fir Western Hemlock Forests. Remote Sensing of Environment, 70 (3), 339-361.

MINITERIUM FÜR LANDWIRSCHAFT, UMWELT- UND RESSOURCENSCHUTZ (2004), MLUR Waldbau-Richtlinie 2004, Potsdam.

PIROTTI, F. (2011), Analysis of full-waveform LiDAR data for forestry applications: a review of investigations and methods. Forest – Biogeosciences and Forestry, 4 (3), 100-106.

SCHMID, K., WATERS, K. & DINGERSON, L. (2008), LIDAR 101: An introduction to lidar technology, data, and applications. NOAA Coastal Services.

WEHR, A. & LOHR, U. (1999), Airborne laser scanning – an introduction and overview. ISPRS Journal of Photogrammetry and Remote Sensing, 54 (2-3), 68-82.

WOODS, M., LIM, K. & TREITZ, P. (2008), Predicting forest stand variables from LiDAR data in the Great Lakes – St. Lawrence forest of Ontario, 84 (6), 827-839.

YU, X., HYYPPÄ, A., J., KAARTINENA, H. & MALTAMOBET, M. (2004), Automatic detection of harvested trees and determination of forest growth using airborne laser scanning. Remote Sensing of Environment, 90 (4), 451-462.

ZHANG Z., LIU, X., PETERSON, J., WRIGHT, W. (2011), Statistical analysis of airborne LiDAR data for forest classification in the Strzelecki Ranges, Victoria, Australia.