Geo-information Science and Remote Sensing

Thesis Report GIRS-2018-34

Mapping the shrub layer in a forest using LiDAR

Testing a method to predict the presence of a shrub layer using the Actueel Hoogtebestand Nederland point clouds

Gersom Zomer

2 July 2018

WAGENINGEN
UNIVERSITY & RESEARCH

100years
1918 — 2018

Mapping the shrub layer in a forest using LiDAR

Testing a method to predict the presence of a shrub layer using the Actueel Hoogtebestand Nederland point clouds

Gersom Q. Zomer

Registration number
940212989040

Supervisors:

Dr. Harm Bartholomeus
Dr. ir. Jan den Ouden

A thesis submitted in partial fulfilment of the degree of Master of Science
at Wageningen University and Research Centre,
The Netherlands.

2 July 2018
Wageningen, The Netherlands

Thesis code number: GRS-80436
Thesis Report: GIRS-2018-34
Wageningen University and Research Centre
Laboratory of Geo-Information Science and Remote Sensing

Acknowledgements

This study would not have been possible without the help and supervision of Harm Bartholomeus and Jan den Ouden who provided me with valuable input over the course of this project. I would also like to thank the State Forestry Service for providing me with the field data used in this study. Lasty I would like to thank Judith Poelman for providing valuable input for my thesis and supported me throughout.

Abstract

In this study the usability of the AHN2 ALS-LiDAR data-set for the prediction of a shrub layer was investigated. Two LiDAR structure indices, the Undergrowth Return Fraction (URF), and the Undergrowth Cover Density (UCD), were proposed. The shrub layer was described using a binomial parameter that indicated whether a shrub layer was absent or present. The binomial parameter was derived from field measurements. The presence of a shrub layer was predicted with the URF and UCD in a logistic regression for six forest species compositions. The URF performed slightly better than the UCD in the logistic model but the overall predictive power of the model was low with a maximum R^2 of 0.44. The species composition of the shrub layer has a significant effect on the UCD but not the URF. Furthermore, the results also indicate there is a difference in detectability of evergreen species and deciduous species when leaf-off ALS-LiDAR is used. The low model performance for deciduous species combined with the apparent difference in detectability between evergreen and deciduous species indicate that leaf-off LiDAR data might not be suitable for the purpose of shrub layer detection, thus limiting the usability of the AHN2 for shrub cover prediction.

Keywords: ALS-LiDAR, AHN, leaf-off, Shrub layer, Forest structure,

1 Introduction

Forests are important ecosystems and provide habitat for a large number of species (Hermy & Bijlsma, 2010). Forest structure is an important indicator of ecological quality, as a forest rich in structure generally supports more niches and thus increases species richness (Bazzaz, 1975). More specifically, high structural diversity due to the presence of shrubs and deadwood is an important habitat requirement for birds (Hanzelka & Reif, 2016; James & Wamer, 1982; Martinuzzi et al., 2009), arthropods (Müller & Brandl, 2009; Roberson et al., 2016; Vierling et al., 2011) and mammals (Ewald et al., 2014). In conclusion, mapping and monitoring forest structure is important, as it provides valuable ecological information for nature management.

The structure of a forest is the product of all vegetation elements within a forest stand. These vegetation elements include: living trees, shrubs, standing and lying deadwood, and herbaceous plants. The number and density of these vegetation elements and how they relate to each other in the vertical and horizontal space can be described in terms of structural variation or structural richness. A forest is considered structurally rich if more vegetation elements are present and there is a large variety in the composition of these structure elements. The importance of different structure elements and the scale at which the structure variation should occur, differs per organism (Tews et al., 2004). Quantification of the forest structure allows forest managers to monitor changes in the ecological value of forest stands, as well as provide insight into the areas where the structure can be enriched to achieve management goals.

The quantification of forest structure has been the focus of many studies, which employed a multitude of different methods. Generally a parameter is derived or calculated from field measurements and used as an index of structure. These indices either describe forest structure in the horizontal or vertical dimension, where vertical indices describe structure variation over height, while horizontal indices describe structure variation over distance. Good examples of vertical structure indicators are the vegetation cover values in different height strata (Williams & Marsh, 1998) and the foliage height diversity (FHD), which is obtained by counting vegetation contacts with a pole (Ding, Liao, & Yuan, 2008). The number of plant species and their respective canopy cover in a plot is a prime example of a horizontal structure indicator (Bazzaz, 1975). A combination of different structure indices is often used to predict the abundance of a certain species in so called habitat structure-species diversity (HS-SD) relationships (Simonson et al., 2014).

The recent introduction of Light Detection And Ranging (LiDAR) has sparked renewed interest in forest structure assessment, as this enables the direct measurement of vertical forest structure for large areas (Dubayah & Drake, 2000). The advent of LiDAR saw the introduction of many new structure indicators which mainly focus on the canopy. Some notable LiDAR structure indicators include: mean height or the standard deviation of the mean height (Müller & Brandl, 2009), density or height related percentiles (Naesset & Naesset, 2003; Zellweger et al., 2014), fractional vegetation cover values (Latifi et al., 2016), and vegetation density (Næsset, 1997).

While plenty of LiDAR-based forest structure assessment studies have focussed on the forest canopy structure, fewer work has been done on the assessment of sub-canopy vegetation layers such as secondary tree layers or shrubs. This is largely due to the difficulty an Arial Laser Scanner (ALS) has to penetrate to lower vegetation layers, which require high scanning densities for accurate mapping (Hamraz et al., 2017). ALS-LIDAR has been

used to map the undergrowth of a forest with varying success.

Martinuzzi et al. (2009) predicted understorey shrub presence or absence for 20 by 20 m pixels with a random forest model. Understorey shrubs were deemed present if the total cover of all shrub species, excluding tree seedlings, exceeded 25%. The model achieved an 83% classification accuracy using the proportion of ground returns, the proportion of returns from 1–2.5 m and the terrain slope. Latifi et al. (2016) used LiDAR to predict the density of the first, second and third tree layer as well as the density of the shrub and herb layers. The final regression model for the shrub cover used the vegetation density and the number of returns below 0.5 m divided by the number of returns below 5 m and yielded an $R^2 = 0.37$. Wing et al. (2012) presented a new LiDAR index called the Understorey LiDAR Cover Density (ULCD) and showed a strong relation with the field measured shrub cover in interior Ponderosa Pine (*Pinus ponderosa*) stands. A weighted regression model with the ULCD, standard deviation of canopy height and proportion of points between 30 – 40 m predicted the shrub cover with an $R^2 = 0.74$. Hill & Broughton (2009) predicted the presence or absence of an understorey in deciduous forests, using first and last LiDAR returns from data acquired in early spring (leaf-off) and mid-summer (leaf-on). The model using the leaf-on and leaf-off data achieved a 77% classification, and the model using only leaf-off data achieved a 72% classification accuracy. However all these studies required extensive field measurements and used leaf-on LiDAR-data.

Ideally, shrub cover and other structural parameters of forest stands can be derived from data that are already available, without the need for new measurements to be made. This could be achieved by using information from nationwide ALS-LiDAR height surveys. The validation of the structure assessment can be done from field data recorded in regular forest stand surveys. The Actueel Hoogtebestand Nederland (AHN) or Height Database Netherlands is an ongoing nationwide ALS-LiDAR mapping effort to create a high-fidelity Digital Evelation Model (DEM) of the Netherlands. The AHN provides open ALS-LiDAR data and offers an unique opportunity for forest structure assessment.

The aim of this study is to investigate whether the shrub layer can be predicted for fully developed deciduous, coniferous, and mixed forests using ALS-LiDAR data from the AHN. To answer if the structure of the understorey of a forest can be mapped using ALS-LiDAR, the AHN point clouds will be combined with available field data. Three research questions are adressed. The first question that is addressed is: (1) Can the presence of a shrub layer be determined from the AHN? The Dutch forests are highly heterogenous and have a large spatial variation in species composition. Therefore, the second question is defined as: (2) Does the effectiveness of structure assessment vary between different tree species? The AHN2 data is aqcuired under leaf-off conditions, which differs from prior studies. A third research question is therefore defined as: (3) Does the vegetation species composition affect shrub layer prediction using ALS-LiDAR in leaf-off conditions?

2 Material and Method

2.1 Study area

For this study, data from forest surveys in two different locations in the Veluwe nature area were used. The Veluwe is the largest forested area in the Netherlands located in the province of Gelderland. The first location is the forest area around the village of Kootwijk (52°11'05.8"N 5°46'12.7"E), located at the western edge of the Veluwe (fig. 1). The area is dominated by conifer-

ous species, including: Scots pine (*Pinus sylvestris*), Douglas Fir (*Pseudotsuga menziesii*), and Japanese Larch (*Larix kaempferi*). Most of these stands were planted at the start of the 20th century as part of the large reforestation campaign by the State Forestry Service. The second location are the forests near the village of Vierhouten (52°19'55.5"N 5°49'46.5"E), located at the northeast corner of the Veluwe (fig. 1). This area shows a larger variety in species composition with a larger presence of broadleaf species, including European Beech (*Fagus sylvatica*) and Common Oak (*Quercus Robur*); alongside Scots pine and Douglas Fir. The study areas were selected because they were surveyed by the State Forestry Service around the same time the LiDAR data was acquired.

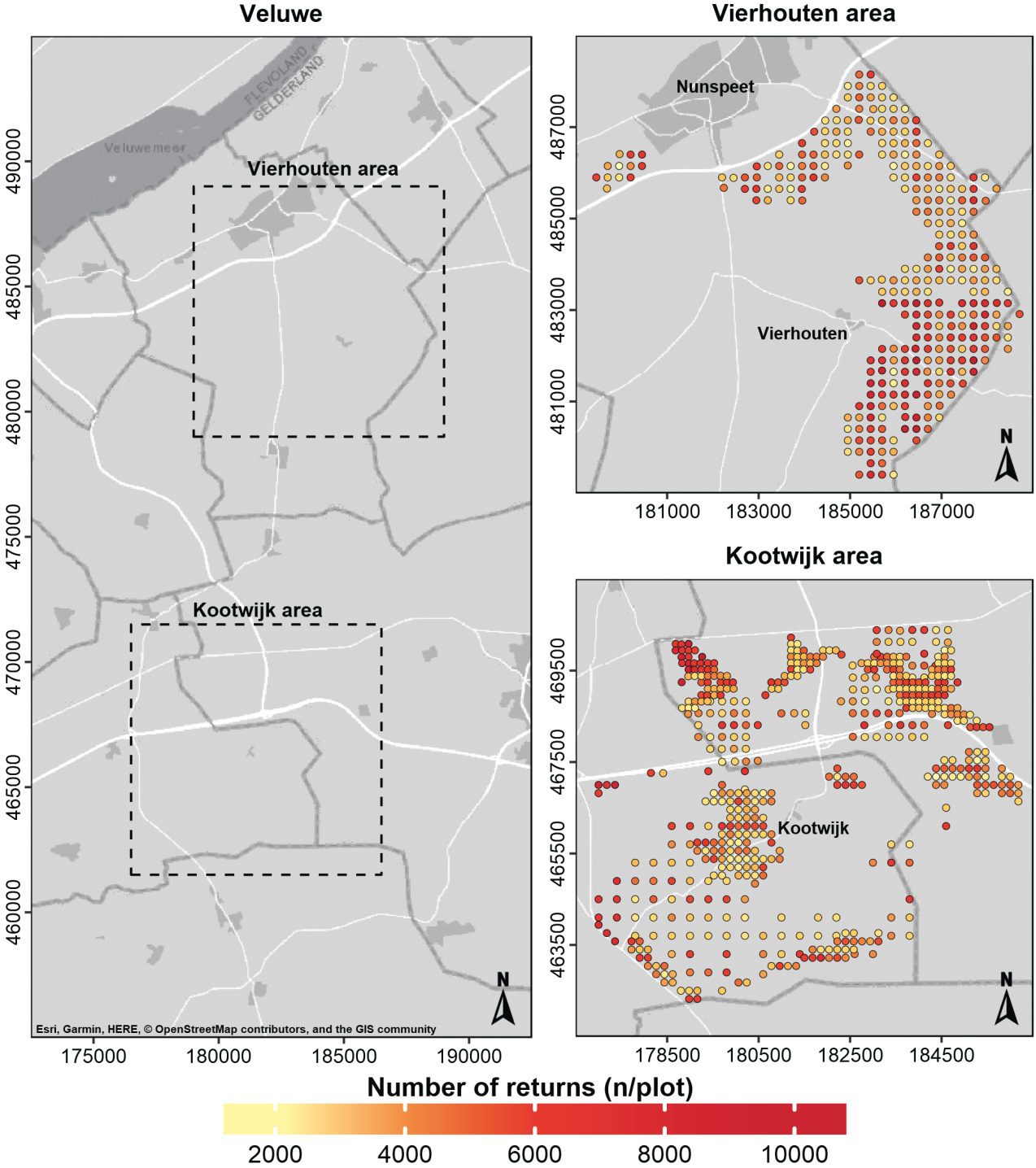

Figure 1: *An overview of the study area and the forest inventory plots with their respective AHN-LiDAR scanning density in number of points per plot. The coordinates are in Rijksdriehoek (m).*

2.2 Valdidation data

Field observations from 825 plots that are part of the State Forestry Service forest inventory system or Systeem Houtmeetkundige Inventarisatie (SyHI), were used as ground truth. The SyHI-method consists of regular stand surveys on an evenly spaced grid of circular plots, and is used by the State Forestry Service for stand monitoring. Every year a different forest on the Veluwe is surveyed. The radius of a plot depends on the forest density and varies between 5 and 20 m. The radius is chosen in such a way that at least 20 trees fall inside the plot (Berg, 1996). For each tree within a plot the Diameter Breast Height (DBH) and tree species is recorded, to get insight into stand composition, stand development and tree growth. The number of seedlings as well as the density and composition of the shrub layer are recorded using a fixed radius of 8 m (Silve, n.d.). The field observations used in this study were done in different years, with the area around Kootwijk being measured between 2009 – 2010, and the area around Vierhouten being measured in 2011.

There are three different structure parameters the State Forestry Service recognizes in the undergrowth: the shrub cover (%), the number of recent seedlings (n/ha) and the number of established seedlings (n/ha). The shrub cover is expressed as the percentage of a plot covered by the crown projection of the shrubs in a plot, where a shrub is defined as a woody plant with a height between 50 cm and 5 m and a DBH smaller than 5 cm. Recent seedlings are young trees which are larger than 50 cm and have a DBH smaller than 5 cm, while established seedlings are young trees with a DBH between 5 and 8 cm. Simply put, the recent seedlings are part of the shrub layer, while established seedlings form the youngest cohort of trees.

Beside the undergrowth parameters the surveys also include a stand based structure indicator, the forest development stage. Six developmental stages are recognized: the open phase, the juvenile phase, the dense phase, the pole phase, the tree phase, and the deterioration phase (Verheyen et al., 2010). The forest development stages are a description of the forest vegetation structure during its development, and give an indication of the quantities in which certain forest structure elements are present (fig. 2), providing valuable information about the expected forest structure.

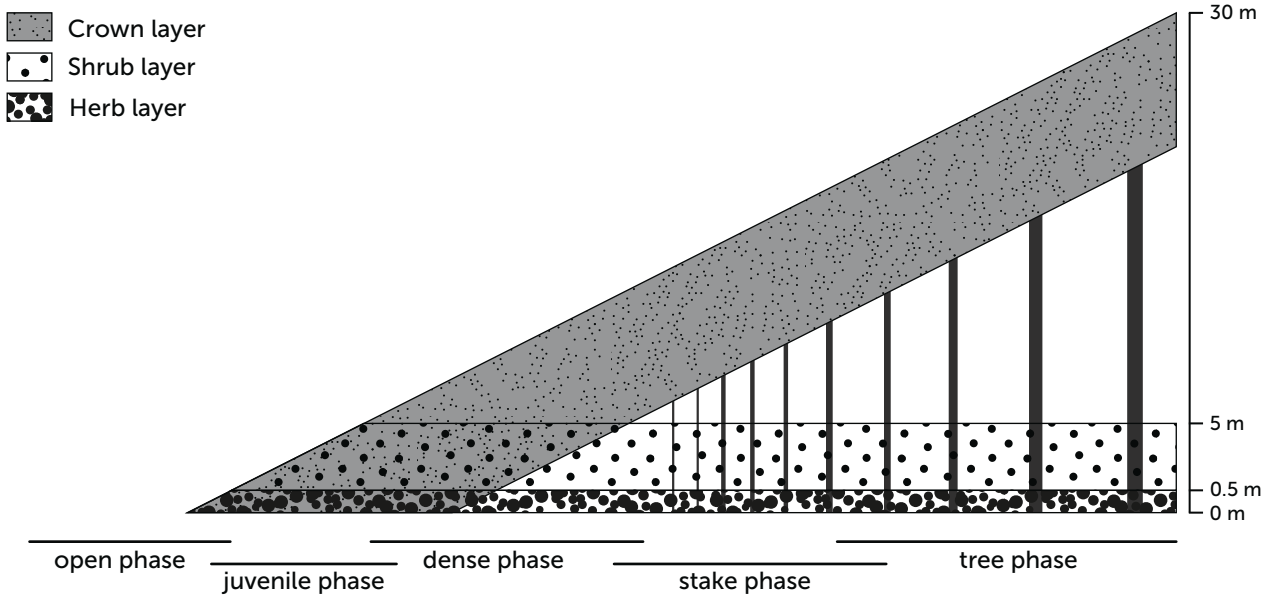

Figure 2: *Conceptual structural model of a forest thorugh different development stages.*

From the SyHI-surveys the forest development stage, the dominant tree species, the dominant tree height, the number of seedlings, the shrub cover percentage and the shrub composition are used in this study.

2.3 LiDAR data

The point clouds used for this study were recorded as part of the second iteration of the AHN also named AHN2. The AHN is an ongoing effort to create a high precision and resolution DEM of the Netherlands. Currently, data for the third iteration, called AHN3, is being acquired and expected to be completed in 2019. The current version available for the entire Netherlands is the AHN2 for which the data was acquired between 2007–2012 with an minimum density of 6–10 points per m^2, to achieve a DEM with a maximum stochastic and systematic error of 5 cm (van der Zon, 2013).

The AHN data is made available as pre-processed point clouds containing X, Y, and Z-coordinates. Other parameters, such as the return intensity, scan angle, and return number, are not published. The LiDAR data is split in two distinct point clouds: a point cloud containing only ground level returns, which is defined as the border between ground and air, and ground and water (van der Zon, 2013); and a point cloud containing all non-ground returns. The latter dataset thus contains points describing the vegetation in forested areas.

The AHN-LiDAR data was acquired using the helicopter based FLI-MAP 400 system (Ludikhuize & Stroeven, 2008), and was recorded between 1 December 2010 and 31 March 2011 (van der Zon, 2013).

2.4 Methodology

2.4.1 Lidar Data Processing

The LiDAR data was processed in order to develop the model and compare the LiDAR data to the field validation data. First, the AHN point clouds were projected to the Rijksdriehoek system (EPSG:28992). Next, the point clouds were clipped to the extent of the plots using a radius of 35 m for the ground returns point clouds and a radius of 28 m for the non-ground returns point clouds. The ground points were clipped to a larger plot radius to ensure ground level information was present for all the vegetation points, which is needed for point cloud normalization.

Cloud point normalization is a process in which height of the ground level is subtracted from all the non-ground returns, making the height of all non-ground returns relative to the ground level. The normalization was accomplished by converting the ground point clouds to a DEM with a raster size of 50 cm using LAStools. The point cloud normalization was then done using the 'lidR' library for R.

After the normalization was completed all points below 50 cm were removed to ensure that no returns from herbs were included in the model. Finally, the points in the ground and non-ground clouds were classified based on their distance from the plot centre. The classes used were 0–8 m, 8–18 m, and 18–28 m, which were used in further analysis.

2.4.2 LiDAR indices

There is a large variation in the number of LiDAR returns between the different SyHI-plots, in some cases the number of returns differs more than 5000 points (fig. 1). The variation in the number of LiDAR returns between plots is caused by two factors. First, a LiDAR pulse can be partially reflected by vegetation before hitting the forest floor, resulting in multiple discrete returns for each pulse. Second, some plots are located at overlapping flight lines, increasing point density. Flight lines overlap to ensure good LiDAR coverage at the edges of a scan swath. fraction based LiDAR structure indices were used for the shrub layer model, to allow for comparison of plots despite the differences in LiDAR returns.

Two different LiDAR indices were defined and used for the shrub layer model. Namely, the Undergrowth Cover Density (UCD) and the Undergrowth Return Fraction (URF). The URF is a measure of the fraction of LiDAR returns present between 0.5–5 m, in respect to the total number of ground and non-ground returns in a plot and is obtained with the following formula:

$$URF = \frac{n_{shrub}}{n_{total}}$$

where n_{shrub} is the number of LiDAR returns between 0.5 and 5 m in a plot, and n_{total} is the total number of ground and non-ground returns combined in a plot. The UCD is a measure of the fraction of Lidar returns present between 0.5-5 m in relation to the number of ground returns and is obtained by the following formula:

$$UCD = \frac{n_{shrub}}{n_{shrub} + n_{ground}}$$

where n_{shrub} is the number of LiDAR returns between 0.5-5 m in a plot, and n_{ground} is the number of LiDAR ground returns in a plot.

The concept behind the URF and UCD can be explained by comparing the LiDAR vegetation return profile of three different conceptual stands: one stand without shrub layer (fig. 3a) one stand with a dense shrub layer (fig. 3b), and a stand with a dense shrub layer and a low overstorey (fig. 4). By calculating the fraction of LiDAR vegetation returns for different height strata the vertical distribution of vegetation elements can be visualized. When more vegetation elements are present in a stratum, more LiDAR pulses are expected to reflect from that stratum, leading to a higher return fraction. These return profiles can be compared to make differences in the vertical distribution of vegetation elements apparent. So, a stand with a dense shrub layer (fig. 3b) is expected to have a higher fraction of returns in the first five meters, than a stand without a shrub layer.

The return fractions are negatively correlated to the height of a stand. Overall return fractions decrease if a stand increases in height and thus the number of strata increases. When two conceptual stands with identical shrub layers and different tree heights are compared, the return fractions, and specifically the URF, are different (figs. 3b & 4). The number of ground returns is directly affected by the height and density of vegetation present.

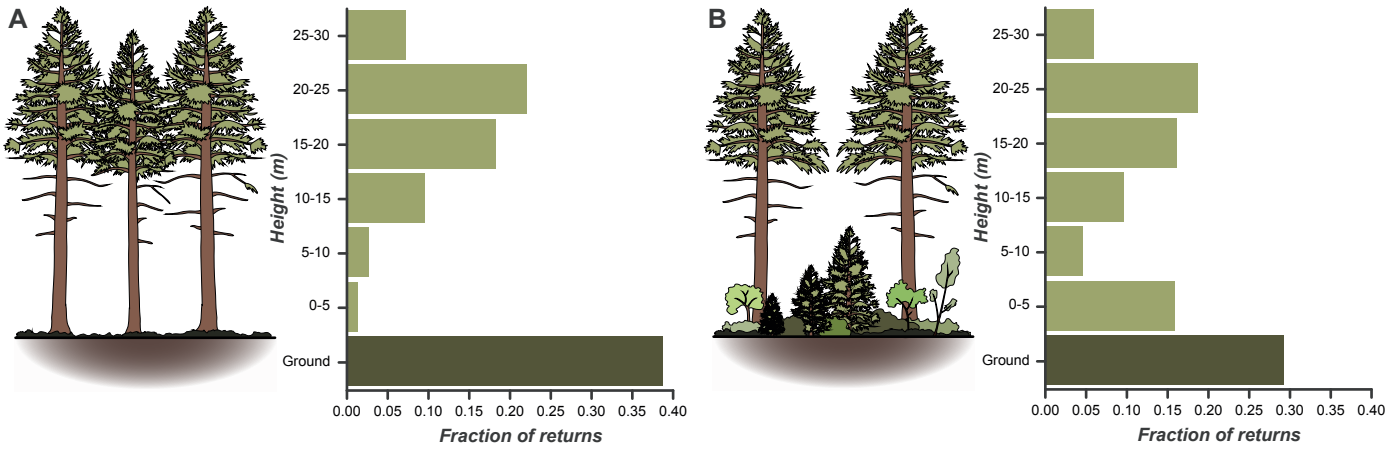

Figure 3: *Conceptual presentation of a coniferous forest stand with the expected fraction of LiDAR returns from different height classes. (A) situation when a shrub layer is absent, with an URF of 0.02 and an UCD of 0.05; (B) situation when a shrub layer is present, with an URF of 0.16 and an UCD of 0.36.*

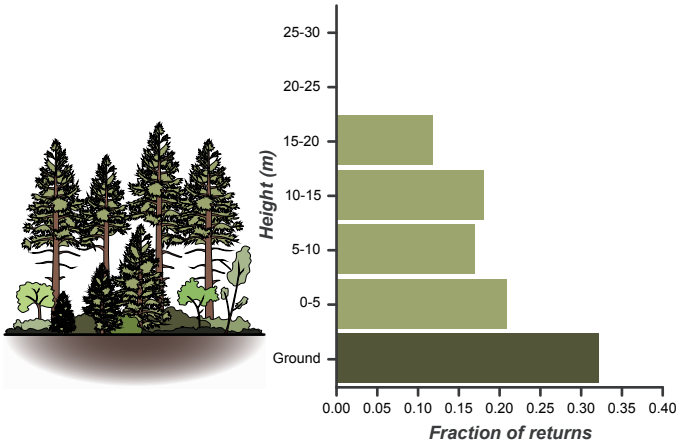

Figure 4: *Conceptual presentation of a coniferous forest stand with the expected fraction of LiDAR returns from different height classes when a shrub layer is present under a low overstorey, with an URF of 0.21, and a UCD of 0.38.*

By using the number of ground returns the UCD aims to remove the effect of stand height on URF by using the number of ground returns, to standardize for vegetation height.

2.4.3 Selection of field data

A proper separation of canopy and shrub layer is a large concern for LiDAR-based shrub cover models. A good separation of the shrub and canopy is most likely in the tree phase. Therefore, only the plots that were in the tree phase were selected for analysis. In other phases a shrub layer is either absent (pole phase) or not properly separated from the canopy layer (open, juvenile and dense phase) (fig. 2).

Another major concern is a mismatch of the plot locations in the field measurement with the plot locations in the LiDAR point clouds. While each plot has its own set of coordinates, it is uncertain as to whether the provided coordinates are accurate. When SyHi-measurements are made, the plot centre is determined using compass and tape measurer. It is unclear for which plots the accuracy of the coordinates is low and unknown how large the offsets are.

In order to minimize the potential effect of errors induced by a mismatch between field and LiDAR data, the plots were filtered based on the variation in forest structure of their direct surroundings. This spatial variation filtering consists of a number of steps. First, the URF and UCD were calculated for each plot for a radius of 8, 18 and 28 m. Second, the difference (Δ) between the URF for a radius of 8 and 18 m (URF Δ_{8-18m}), and 8 and 28 m (URF Δ_{8-28m}), were calculated for each plot by subtracting the larger radius from the smaller radius. Third, the UCD Δ_{8-18m}, and UCD Δ_{8-28m} were calculated similarly. Fourth, the difference between the field observed tree height and LiDAR derived tree height (Δ_{height}) was calculated for each plot by subtracting the LiDAR tree height from the field tree height, where the LiDAR derived height was defined as the highest vegetation return for a radius of 8 m. Fifth, the interquartile range (IQR) or mid-spread, which is the range between the 25% and 75% percentiles of a dataset, for Δ_{8-18m}, Δ_{8-28m} of the URF, UCD, and Δ_{height}, are derived and used to calculate the upper (Q_3 + 1.5 × IQR) and lower (Q1 - 1.5 × IQR) boundaries (fig. 5). Finally, all plots of which the URF Δ_{8-18m}, URF Δ_{8-28m}, UCD Δ_{8-18m}, UCD Δ_{8-28m}, or Δ_{height}, fell outside the previously calculated boundaries were removed from the analysis. This approach ensured that the plots with the largest spatial variation were removed, while some spatial variation was still allowed.

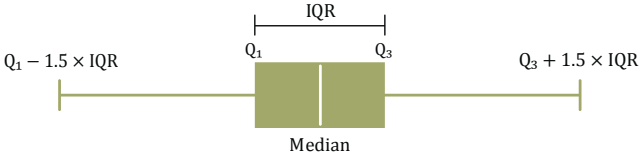

Figure 5: *Example of a boxplot with different terms, where Q_1 is the first quantile, Q_3 is the third quantile and IQR is the interquartile range.*

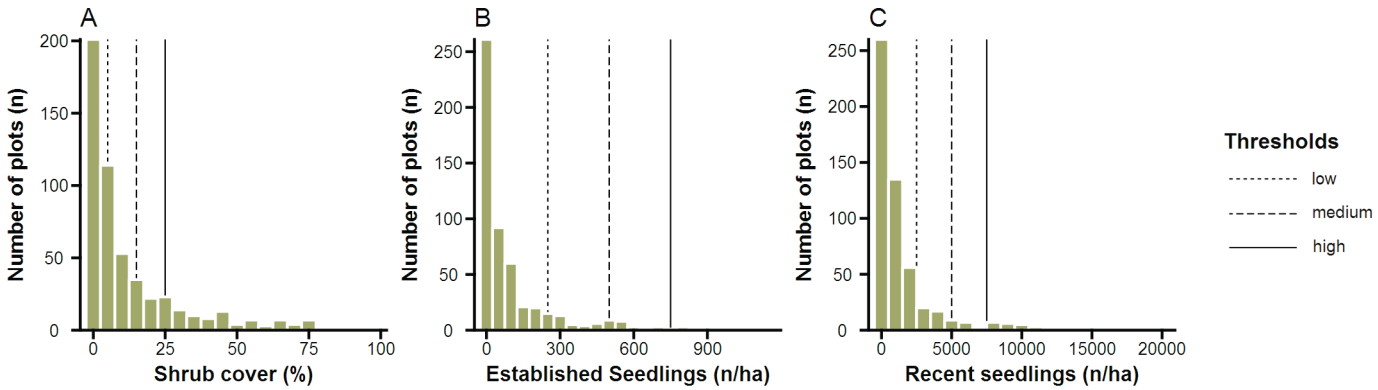

Figure 6: *Histograms of the field parameters used to determine the S_{layer} for all plots in the tree phase (n = 570) the. (A) Histogram of the shrub cover with a bin size of 5; (B) Histogram of the number of established seedlings with a bin size of 50; (C) Histogram of the number of recent seedlings with a bin size of 1000. the thresholds used for the S_{layer} are plotted as lines.*

2.4.4 Model discription

First a binominal parameter, called S_{layer} was defined that indicated whether a shrub layer was present (1) or absent (0). The S_{layer} was derived from the field data and is a description of the shrub layer based on the shrub cover (%), the number of recent seedlings (n/ha), and the number of established seedlings (n/ha). The S_{layer} was used because the number of LiDAR vegetation returns between 0.5–5 m was expected to increase with shrub cover. However, the number of LiDAR vegetation returns from 0.5–5 m was also expected to increase when forest regeneration or young established trees were present in the undergrowth.

The inclusion of forest regeneration and small trees in the S_{layer} should increase model performance. Furthermore, the inclusion of small trees and forest regeneration into the shrub layer model is also appropriate from an ecological perspective. Tree seedlings and young trees are important structure elements in the undergrowth of a forest.

Whether the S_{layer} was 1 (present) or 0 (absent) was determined using thresholds for the shrub cover, number of established seedlings, and number of recent seedlings. A shrub layer was present when one of the three parameters in a plot exceeded a threshold. Three different variants of the S_{layer} derived using different thresholds, were used in the model (Table 1). The thresholds were selected based on the distribution of the structure parameters (fig. 6). The main reason for the use of three variants of the S_{layer} was the possibility that different tree species had different optimal thresholds for the binominal parameter.

Table 1: *Threshold combinations for S_{layer}*

	Shrub cover	Established seedlings	Recent seedlings
S_{layer} low	5%	250	2500
S_{layer} medium	15%	500	5000
S_{layer} high	25%	750	7500

Finally, the logistic regression models predicted the S_{layer} low, S_{layer} medium or S_{layer} high (dependent variable), based on the value of either the URF or UCD (independent variable). A logistic regression model predicts the probability of the dependent variable being 1 or 0, based on one or more independent variables. The logistic regression was fitted independently for six different canopy layer tree species compositions. The fitted compositions were: Mixed compositions, European Beech (*Fagus sylvatica*), Common Oak (Quercus robur), Douglas Fir (*Pseudotsuga menziesii*), Japanese Larch (*Larix kaempferi*), and Scots Pine (*Pinus sylvestris*). The goodness of fit of these compositions was examined

for UCD and URF to investigate the effectiveness of the proposed shrub cover prediction model for different tree species.

To compare model performance and fit the Tjur (2009) pseudo-R^2 value was calculated for each model. The Tjur pseudo-R^2 is similar to an R^2 of a linear regression model. Both have a range from 0–1, and in both cases a higher value indicates a better fit. The Tjur pseudo-R^2 is obtained by calculating the mean of the predicted probabilities for both groups and taking the difference between them. The use of a pseudo-R^2 allowed for comparison of the performance URF and UCD for the same species.

2.4.5 Vegetation composition

A large concern was the difference in detectability between deciduous and evergreen species by ALS-LIDAR in leaf-off conditions. Leaf-off refers to the leave-less state of deciduous species in mid-winter, which were the conditions when the AHN-LiDAR scans were made. In leaf-off conditions, the only part of a deciduous plant that can be detected by LiDAR are the woody structures, while in evergreen plants both woody structures and leaves can be detected. This is expected to create a difference between the detectability of deciduous and evergreen species in leaf-off conditions.

Two test were performed to ensure there was no effect of species composition of the shrub layer and the overstorey on the shrub layer model. The first test compared the difference in URF and UCD values for two species composition of the shrub layer. The second compared the difference in occlusion effect of the overstorey for different species compositions.

The effect shrub composition had on the detectability by ALS-LiDAR was tested by selecting all plots with a shrub cover above 10% (n = 138) and splitting the plots into two groups based on the composition of the shrub layer, one group containing plots with a shrub layer predominantly made up of evergreen species, and another group containing plots with a shrub layer predominantly made up of deciduous species. If the shrub layer in a plot consisted for more than 50% of evergreen species it was put in the first group; the rest was put in the second group. The means of the URF and UCD for both groups were compared with a Wilcoxon-Mann-Whitney test.

The effect canopy species composition had on the detectability of the shrub layer was tested by comparing the means of the Overstorey Return Fraction (ORF) of the canopy species compositions used in the shrub cover model. The ORF was defined as the fraction of returns above 5 m in a plot and is obtained using the following formula:

$$ORF = \frac{n_{canopy}}{n_{total}}$$

were n_{canopy} is the number of vegetations returns above 5 m and n_{total} is the total number of returns including ground returns in a plot. A Kruskal-Wallis test was used, to examine if there was a significant difference in the average ORF between the fitted species compositions. A Dunn (1964) post-hoc test using the Hochberg & Benjamini (1995) method for p-value adjustment was used to determine which overstorey canopy compositions differ significantly from each other.

3 Results

3.1 Data filtering

Before the model was tested, the plots were selected on the forest development phase and the six dominant species that were tested for the model, reducing the number of plots from 825 to 529. These 529 plots were used to calculate the filtering thresholds of the URF $\Delta_{8\text{-}18m}$, URF Δ_{8-28m}, UCD Δ_{8-18m}, UCD Δ_{8-28m}, and Δ_{height}. After the spatial variation filtering was completed 426 plots remained, with a small part of the plots removed with each filtering step (Table 2).

The distribution of the URF Δ_{8-18m}, URF Δ_{8-28m}, UCD Δ_{8-18m}, and UCD Δ_{8-28m} were similar (fig. 8). The absolute differences of the URF were smaller compared to the UCD. This was caused by the fact that the UCD had higher values across all stand heights. The thresholds calculated for the Δ_{height} were -4.38 m for the lower bound and 5.64 m for the upper bound. Furthermore, the Δ_{height} showed a large spread with values of up to +10 and -15 m being observed (fig. 7). No clear spatial patterns in the Δ_{height} could be observed (fig. 9).

Table 1: *The number of plots after each filtering step.*

Filtering used	Plots remaining
None	825
Selected species	783
Tree phase	529
Δ_{height}	496
URF $\Delta_{8\text{-}18m}$	448
URF $\Delta_{8\text{-}28m}$	439
UCD $\Delta_{8\text{-}18m}$	429
UCD $\Delta_{8\text{-}28m}$	426

3.2 URF and UCD

The URF and UCD both describe the shrub layer and both range from 0 to 1. However, they responded differently to stand characteristics. First, the range of observed UCD values was higher than the range of observed URF values. Second, the URF was negatively related to the maximum height of the plot, while the UCD did not appear to be related to tree height at all (fig. 10a & c). Third, the UCD was positively related to the vegetation density, while the URF appeared to be unaffected (fig. 10b & d). The vegetation density was calculated as the percentage of vegetation LiDAR returns in a plot (Næsset, 1997). The relation between tree height and the URF is caused by a decrease in the absolute value of the return fractions from individual strata, as vegetation height increases. The UCD is calculated using the number of ground returns and as the number of ground returns approach 0 the UCD approaches 1. An increase in the amount and density of vegetation elements in a plot causes fewer LiDAR pulses to reach the forest floor. Therefore, a higher vegetation density leads to an increase of the UCD.

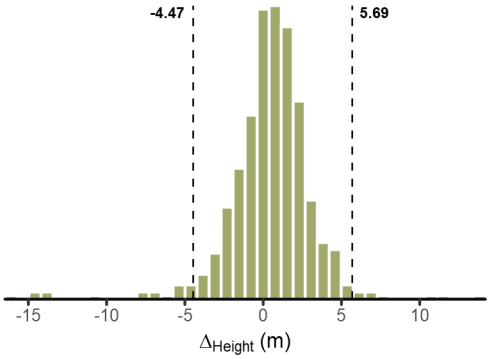

Figure 7: *Histogram of the Δ_{height}. The thresholds used for filtering are plotted as dotted lines.*

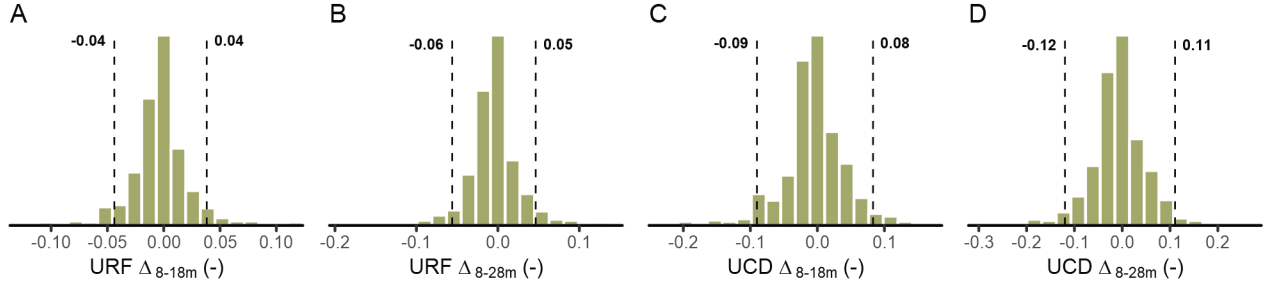

Figure 8: *Histograms of the (A) URF $\Delta_{8\text{-}18m}$, (B) URF $\Delta_{8\text{-}28m}$, (C) UCD $\Delta_{8\text{-}18m}$, (D) URF $\Delta_{8\text{-}28m}$. The thresholds used for filtering are plotted in each histogram as dotted lines.*

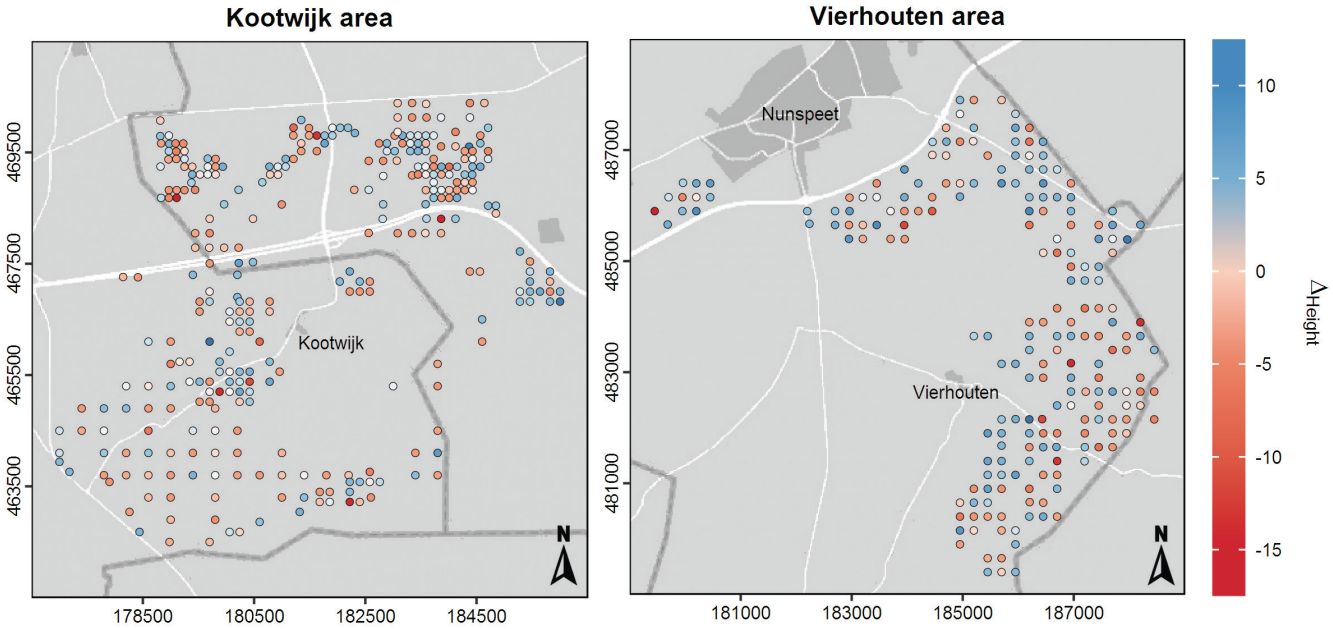

Figure 9: *The remaining forest inventory plots after data selection with their respective Δ_{height} in m. The coordinates are in Rijksdriehoek (m).*

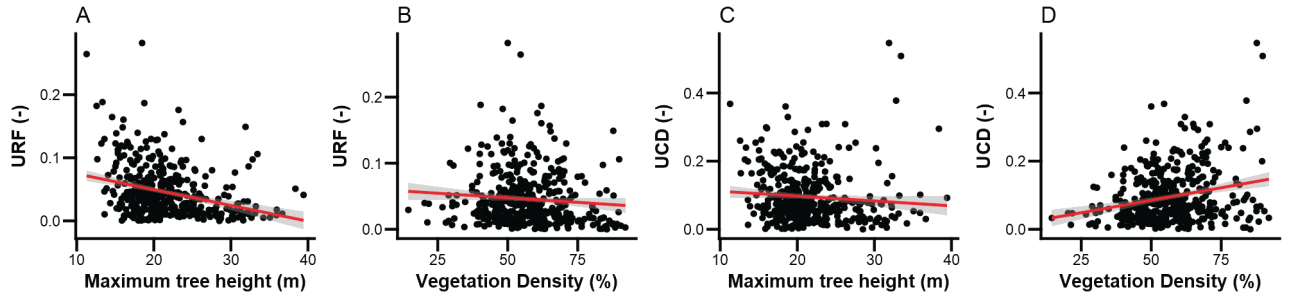

Figure 10: *Scatter plot of the URF and the UCD against the maximum tree height (A & C) and the vegetation density in a plot (B & D).*

Table 3: *Significance of the URF and the Tjur pseudo-R^2 of the binominal regression model, separated by main tree canopy species. S_{layer} low, medium, and high indicate the different threshold settings of the S_{layer} parameter (see Table 1). Empty cells or missing species indicate that a logistic model could not be fitted or was insignificant.*

	S_{layer} low		S_{layer} medium		S_{layer} high	
	Sig.	R^2	Sig.	R^2	Sig.	R^2
Common Oak	.041	.138	-	-	-	-
Mixed Composition	.001	.123	.004	.078	.013	.058
Scots Pine	.000	.103	.047	.024	-	-
Japanese Larch	-	-	.030	.352	.030	.352
Douglas Fir	.012	.298	.006	.435	.006	.379

Table 4: *Significance of the UCD and the Tjur pseudo-R^2 of the binominal regression model, separated by main tree canopy species. S_{layer} low, medium and high indicate the different threshold settings of the S_{layer} parameter (see Table 1). Empty cells or missing species indicate that a logistic model could not be fitted or was insignificant.*

	S_{layer} low		S_{layer} medium		S_{layer} high	
	Sig.	R^2	Sig.	R^2	Sig.	R^2
Mixed composition	.000	.120	.007	.062	.009	.062
Scots Pine	.000	.091	.031	.028	-	-
Japanese Larch	-	-	.034	.269	.034	.269
Douglas Fir	.024	.241	.010	.401	.014	.284

3.3 Shrub cover model

A logistic regression was fitted separately for each canopy species composition, using either the URF or UCD to predict one of the three S_{layer} parameters (see Table 1). Since a binominal regression does not provide the significance for the overall model, the significance of the independent variable was used to determine whether a model was significant. From the 36 models that were fitted 21 were significant (Tables 3 & 4).

The models where the URF was used to predict the shrub layer was significant for 11 of the 16 cases, while the UCD was significant for 10 of the 16 cases. The model was insignificant or could not be fitted for European Beech due to an insufficient number of plots where a shrub layer was present (less than 2). In general no shrub layer was present in Beech stands.

The species compositions of Douglas Fir and Japanese Larch achieved the best results in the logistic model with a maximum pseudo-R^2 of 0.435 for Douglas Fir and a maximum pseudo-R^2 of 0.352 for Japanese Larch. However, the pseudo-R^2 values for both species are still relatively low, and highlight that the overall model performance was poor.

The poor performance of the logistic regression is illustrated by the strong overlap of S_{layer} groups (absent, present), with the respect to the URF and UCD (figs. 11 & 12). This was especially true for Scots Pine stands, where for an URF of zero the probability that a shrub layer is present was 50% (fig. 11). The same is true for stands with a mixed composition where the probability that a shrub layer was present is 25% when the URF was zero (fig. 11). In case of a good fit the probability should

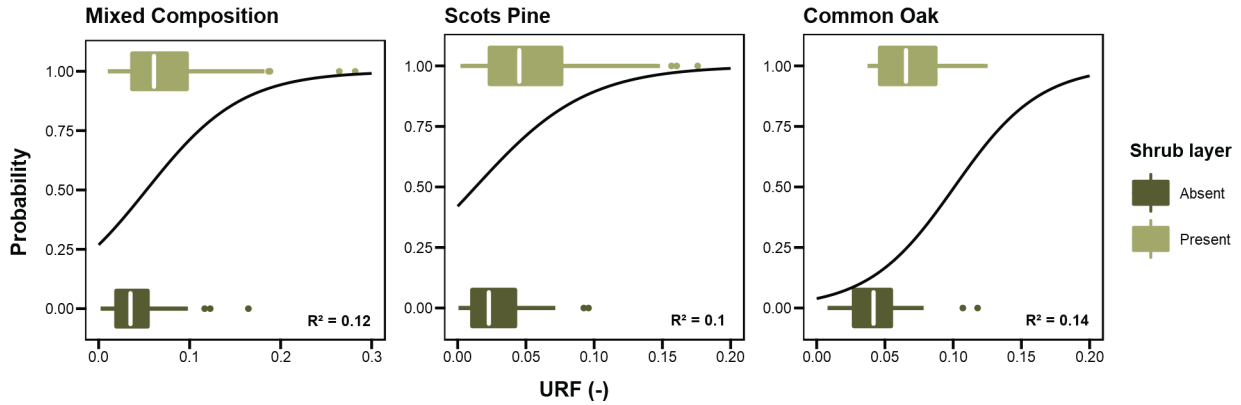

Figure 11: *Boxplots of the URF of SyHI-plots grouped by their presence or absence of a shrub layer for three canopy species compositions. The predicted probabilities of logistic regression are plotted as a black line. The thresholds of S_{layer} low were used to create the groups. The Tjur pseudo-R^2 is plotted in the lower right corner.*

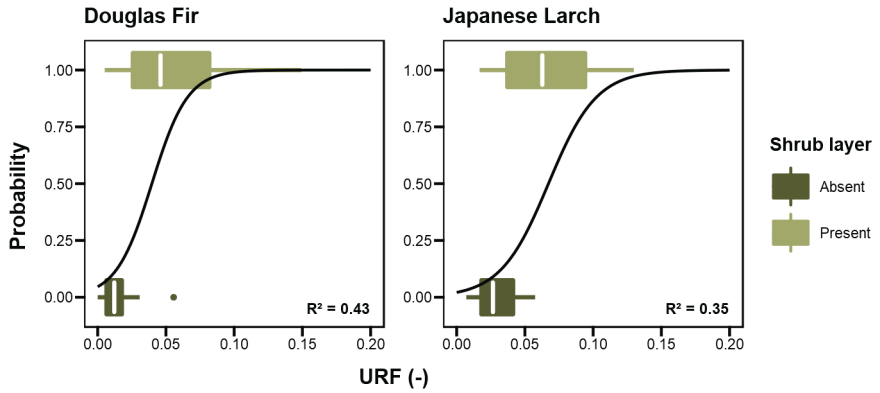

Figure 12: *Boxplots of the URF of SyHI-plots grouped by their presence or absence of a shrub layer for Douglas Fir and Japanese Larch. The predicted probabilities of logistic regression are plotted as a black line. the thresholds of S_{layer} medium were used to create the groups. The Tjur pseudo-R^2 is plotted in the lower right corner.*

follow a S-shape, similar to the predicted probabilities for Douglas Fir (fig. 12) and the values for the groups should sparsely intersect. But even for stands dominated by Douglas, which exhibited a relatively decent fit compared to other compositions, there was still quite some overlap of the S_{layer} groups.

3.4 Vegetation composition

A Wilcoxon-Mann-Whitney test was used to determine whether the URF and UCD were affected by the shrub layer composition. The results indicated that the average value of the UCD was greater for plots with a shrub layer consisting mostly of evergreen species (Median = 0.195), compared to plots dominated by a deciduous shrub layer (Median = 0.115), W = 2534, p < .000. The difference between the URF for plots dominated by an evergreen shrub layer and plots dominated by a deciduous shrub layer was not significant. Even when the threshold for an evergreen shrub layer was raised from 50% of the composition to 100% of the composition, the differences for the URF remained insignificant (fig. 13).

A Kruskall-Wallis test was used to compare the means of the ORF for different dominant plot species. The test indicated that there was a significant difference in the mean of the ORF between different tree species. (H_5 = 105.75, p < .000) with a mean rank of 372.82 for Douglas Fir, 216.31 for Scots Pine, 213.31 for Japanese Larch, 191.01 for mixed compositions, 135.92 for Common Oak, and 125.17 for European Beech.

A Dunn (1964) post-hoc test using the Hochberg & Benjamini (1995) method for p-value adjustment resulted in three groups of tree species for which the ORF differed significantly from each other. One group containing Douglas Fir; one group containing Japanese Larch, Scots Pine, and Mixed compositions; and one group containing Common Oak and European Beech (fig. 14). Douglas Fir had a much higher ORF than other species indicating that relatively fewer LiDAR-pulses reached the shrub layer. The remaining species have smaller differences in the ORF, but coniferous species and mixed stands still had a higher ORF than the deciduous species.

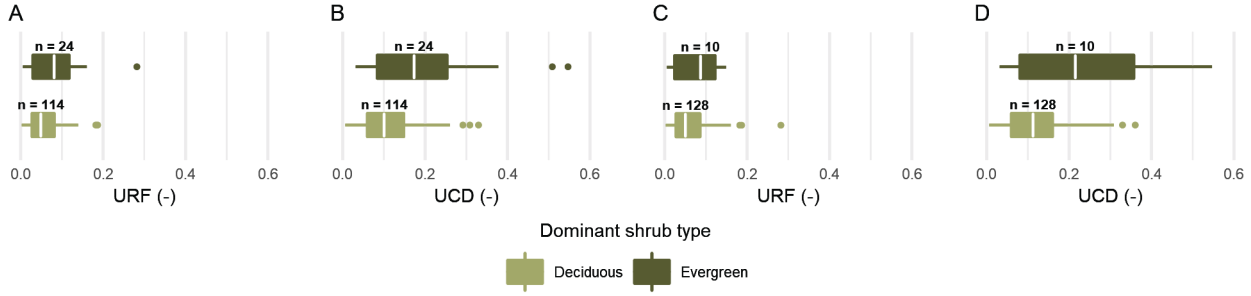

Figure 13: *. Boxplots of the URF and UCD split by shrub composition, with a 50% evergreen shrub composition threshold (A, B); and a 100% evergreen shrub composition threshold (C, D).*

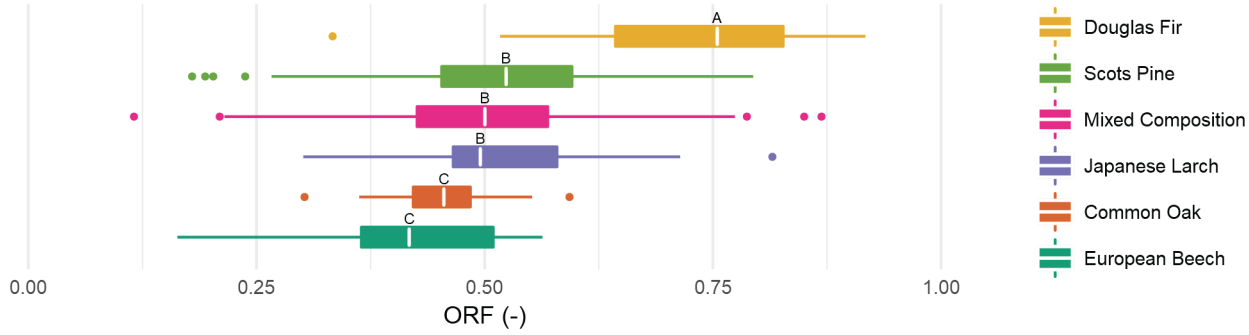

Figure 14: *Boxplots of the ORF for the six different plot stand compositions. The letters indicate the groups of species with significantly different ORF values than other species groups.*

4 Discussion

4.1 Model performance

This study aimed to assess whether the AHN LiDAR data could be used to predict the presence of a shrub layer in forest stands. The results show a poor predictive power of the URF and UCD for the presence or absence of a shrub layer. The UCD and URF both performed differently in the logistic model. The URF could be significantly fitted for more species than the UCD. Furthermore, the models using the URF achieved higher pseudo-R^2 values than the models using the UCD.

The UCD yielded insignificant model results for all deciduous species and this is likely caused by the fact that for a high number of ground returns, the sensitivity of the UCD to returns from the shrub layer is low. This implies that for an open vegetation structure, the predictive power of UCD is low, which is enhanced by the use of leaf-off data. This is further illustrated by the dependence of the UCD on overall vegetation density (fig. 10). Indeed, the shrub cover model was only significant for the four most dense species compositions (fig. 14 & Table 4).

The UCD was based on the ULCD which Wing et al. (2012) used to predict the shrub cover in interior Ponderosa Pine (Pinus ponderosa) stands. The ULCD is calculated by selecting all points between the maximum and minimum shrub height measured in the field, and filtering these points based on the LiDAR return intensity values. The filtered points from the shrub layer are then divided by themselves plus the number of ground points. The UCD on the other hand takes all the points between 0.5 and 5 m and does not include the return intensity filtering. In his study, Wing et al. (2012) reported an R^2 of 0.74 between the ULCD and the field measured shrub cover which is a stark contrast with the results of this study. This contrast can be explained by the fact that Wing et al. (2012) used leaf-on LiDAR data and high precision field measurements explicitly collected for the goals of his study.

When comparing the URF and UCD indices, the results indicate that sensitivity of the URF to changes in the shrub layer decreases as stand height increases (fig. 10). The UCD, on the other hand, performs better when the vegetation layer is very dense or very tall, as mentioned previously. Therefore, the use of the URF is recommended over the UCD when leaf-off LiDAR data is used. The relatively high number of ground returns, which are obtained in leaf-off conditions limits the range of the UCD, reducing its performance in comparison to the URF (Tables 3 & 4). How the URF and UCD perform in leaf-on conditions is not clear and should be further tested.

Hill & Broughton (2009) examined the difference between 'leaf-off' and leaf-on ALS-data in its ability to map the forest undergrowth in deciduous forests. To accomplish this, a combination of leaf-on and leaf-off data was used to predict the presence of a shrub layer. The leaf-off data was acquired in early April, when the shrubs had fully leaved-out but the trees just began budburst, which is different from the true leaf-off conditions of the AHN. The leaf-on data was acquired mid-summer when both the over and understorey had fully leaved-out. Hill & Broughton concluded that there was no large difference in the prediction accuracies of leaf-on and 'leaf-off' data, but recommended to further test how well the shrub layer can be detected in true leaf-off conditions.

The AHN in contrast is acquired in leaf-off conditions, which is in mid-winter. Model performance is low compared to the values of Hill & Broughton (2009), though it is unclear to what extent this can be contributed to the use of lead-off data. Results do indicate that the use of the UCD is less suitable when using lead-off data, as

previously mentioned. In this study, both deciduous and evergreen species are included and the results suggest there is an inherit difference in the detectability of deciduous and evergreen species when using true leaf-off data (fig. 13). The detectability of evergreen species is higher, supporting the notion that using leaf-off data limits the model performance. This is further discussed in section 4.2 and 4.3.

More research is needed for to the determine the usability of the AHN for shrub layer prediction. The inclusion of return number and return intensity in the upcoming AHN3, also acquired under leaf-off conditions, might increase the potential of the AHN for shrub layer prediction.

4.2 Overstorey species effect

The secondary goal of this study was to assess if there was a difference between tree species in the shrub layer prediction model. The results show that there was a definite difference between tree species, highlighted by the large variation in the pseudo-R^2 values of species per S_{layer} variant with some species performing better overall. Additionally, the six species achieved an optimal fit with different shrub layer definitions (S_{layer}). Tree species have an inherent difference in tree height, stand structure and vegetation density under similar growing conditions (Muys et al., 2010). These physiological differences between tree species are believed to be a large factor in the variation in the goodness of fit of species in the shrub cover prediction models.

Model performance was relatively good for both Douglas Fir and Japanese Larch. which are characterised by a low variety in vertical structure with straight trunks and few low hanging branches. On the other hand, Scots Pine and Common Oak typically have a much larger variation in vertical structure, with uneven trunks and a higher presence of low hanging branches. Furthermore, there were large differences in the number of trees per hectare between species. The species that performed poorly in the model also had a higher tree density, with large standard deviations (fig. 15). Not all returns between 0.5-5 m originate from shrubs or small trees, some returns also originate from the trunks of overstorey trees. If the tree density increases the number of trunk returns also increases, inflating the URF and UCD. If the tree density is low this inflation effect on the URF and UCD is reduced increasing model accuracy.

Lastly, a significant difference in the density of the overstorey (ORF) between tree species was observed (fig. 14). The density of the overstorey affected the performance of a species in the logistic model using the UCD. Douglas fir, which had the highest ORF also achieved the best fit with the UCD. However, Scots Pine, which had the second highest ORF, performed poorly in models using the UCD. Striking is the fact that the model using the UCD was insignificant for Common Oak which had the lowest ORF.

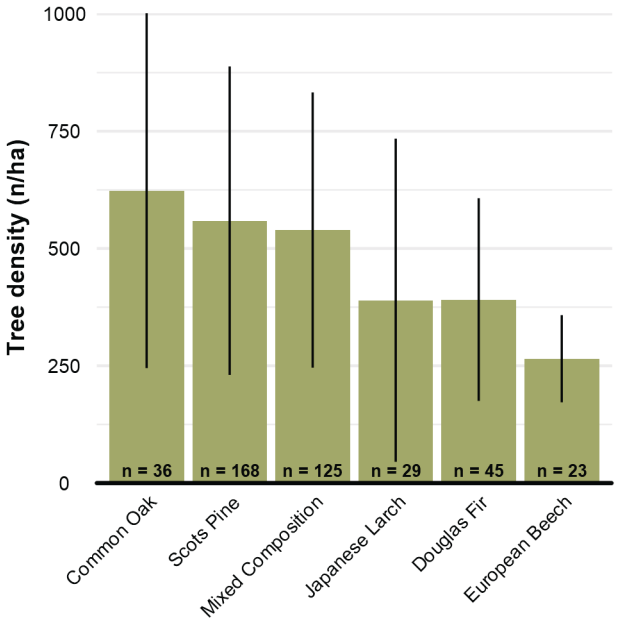

Figure 15: *Bar graph of the mean number of trees per hectare for different species, with error bars representing the standard deviation.*

How the URF and UCD exactly respond to the inherent differences in the vertical structure and the number of trees per hectare between tree species is not clear. More research is needed into the response of fraction-based LiDAR indices in respect to different tree species to examine how the overstorey species composition affects LiDAR based forest structure assessment.

4.3 Shrub layer species effect

The final goal of this study was to assess if there was an effect of composition of the shrub layer on shrub layer prediction using leaf-off LiDAR data. The results of the Wilcoxon-Mann-Whitney test showed that the UCD was significantly higher for shrub layers which were dominated by evergreen species (fig. 13).

Further inspection of the 138 plots that had a shrub cover above 10% made clear that 11 of the 24 plots of which the shrub layer was dominated by evergreen species (> 50%) were dominated by Douglas Fir trees. A similar pattern was found for the plots that had an evergreen shrub layer of 100%: 7 of 10 plots had a Douglas overstorey. This indicates that a shrub cover dominated by evergreens is likely to have a Douglas Fir overstorey, considering that the total number of Douglas fir plots was 13 (out of the 138 plots with a shrub cover above 10%).

As shown earlier, the UCD increases with vegetation density (fig. 10) and Douglas Fir had a high average vegetation density, which is highlighted by the high average ORF (fig. 14). This leads to significantly higher values for the UCD for plots which are for the most part dominated by Douglas Fir. Therefore, the significant relation between UCD and shrub layer composition is not necessarily caused by a difference in detectability of the shrub layer, as it is related to the overstorey species that mostly dominates stand with an evergreen shrub layer.

4.4 Model remarks

Aside form the use of leaf-off LiDAR data there are a number of other factors that are thought to have influenced the model performance. First, the binominal S_{layer} parameter was used to circumvent the sharp distinction between a shrub and a tree in the field data. The inclusion of young trees and forest regeneration in the model did not solve all definition problems. Problems still arise with trees that have a DBH just over 8 cm and are no longer classed as established seedlings. A large part of

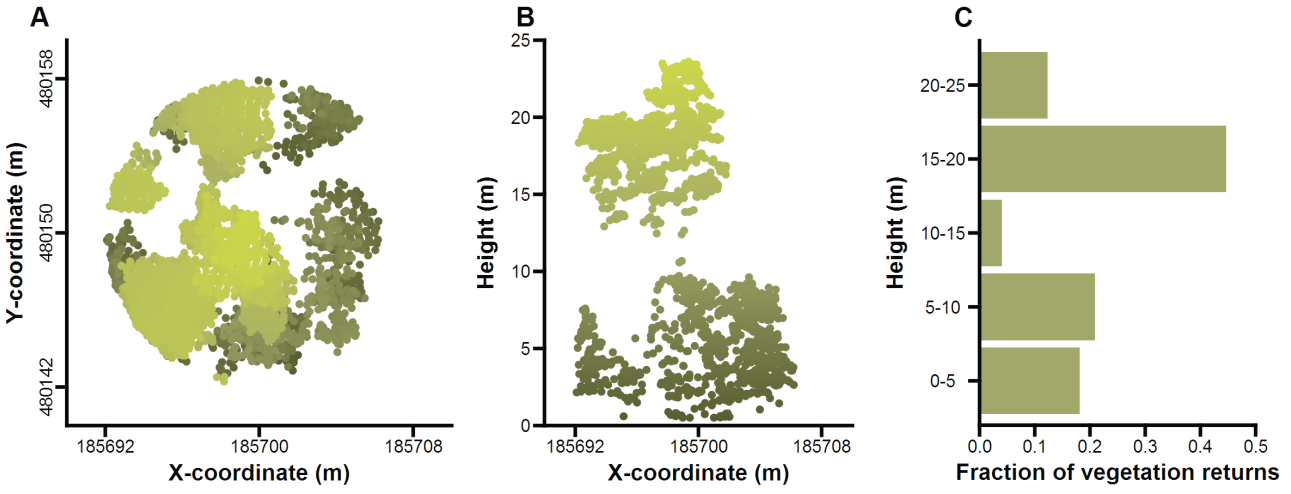

Figure 16: *Plot of the LiDAR vegetation returns of a Scots Pine stand with 0% shrub cover and 38 established seedlings. The high number of returns between 0.5–5 m are likely caused by young trees. With (A) the top view of the plot, (B) the side view of the plot and, (C) the vertical return fractions of the plot. The coordinates are in Rijksdriehoek (m).*

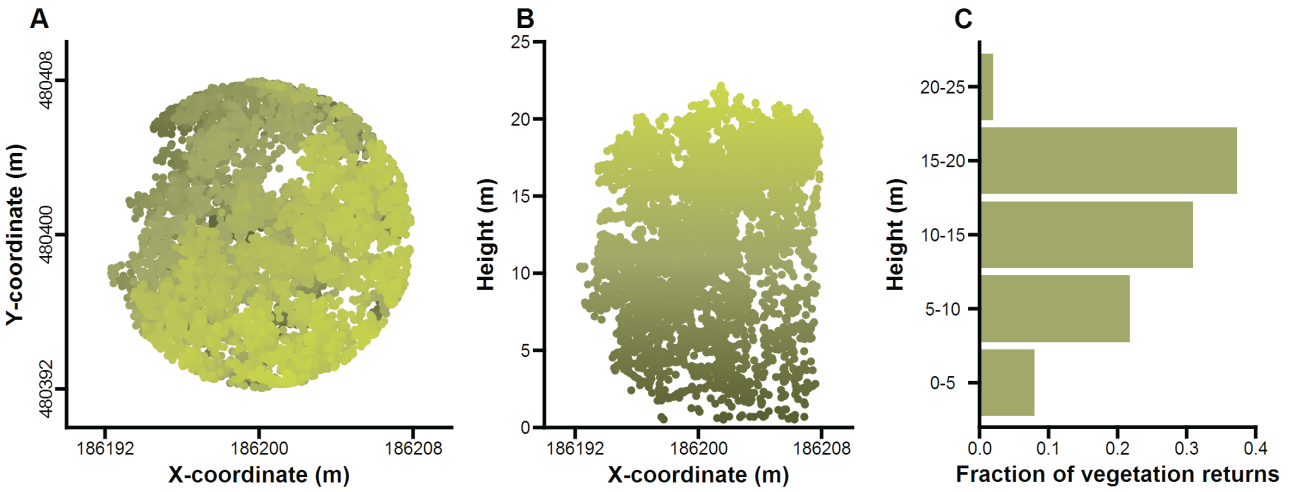

Figure 17: *Plot of the LiDAR vegetation returns of a Common Oak stand with a shrub cover of 0% and 0 established seedlings. The high number of returns between 0.5–5 m are likely caused by low hanging branches. With (A) the top view of the plot, (B) the side view of the plot and, (C) the vertical return fractions of the plot. The coordinates are in Rijksdriehoek (m).*

the branches and structure of these trees is located between 0.5–5 m increasing the number of returns. There are multiple examples of stands were this is the case, and close inspection of a LiDAR plot of such stands illustrates this clearly (fig. 16).

Second, the presented model is unable to differentiate between a shrub layer and low hanging branches. Low hanging branches can increase the number of returns in the lowest strata causing an overestimation of the shrub layer when they are present in a plot. A similar problem arises when the distribution of vegetation elements is very homogenous in the vertical direction. In this case the understorey and overstorey seamlessly blend into one another (fig. 17). The effect of low hanging branches on the prediction of the shrub layer is not the same for all species though. Deciduous species are generally more prone to this effect, because they do not shed their lower branches as rapidly as coniferous species.

Third, it remains unknown whether the field data describes the forest at the given coordinates, due to the inherent difficulty of finding exact coordinates in the field. While the presented method tries to compensate for a possible mismatch of field data and LiDAR data through the use of spatial filtering, it is unclear to what extent the spatial filtering was successful and if all plots with a data mismatch were removed. It is also possible that the employed filtering method inadvertently removed plots where the location of the field and LiDAR data matched. In fact, the spatial filtering method used in this study does not check how the field data relates to the LiDAR data, but instead opted to remove all plots that have a dissimilar forest structure in their direct surroundings. This approach ensured that plots which could potentially be large outliers were removed, but did not ensure the removal of plots with a data mismatch. That the spatial filtering was only partially successful is highlighted by the large differences for some plots between the field measured tree height and the LiDAR derived tree height. Although some variation could originate from measurement errors or tree felling between the moment of field measurements and LiDAR data acquisition, a mismatch of datasets remains the most probable cause.

Finally, the data used for this research was not specifically acquired with the goal of undergrowth prediction and monitoring in mind, and the data could therefore lack the needed precision and detail for accurate shrub layer prediction. This does not necessarily mean that it is impossible to

accurately map the undergrowth of a forest with open data, but clearly illustrates the high data requirements needed to accurately model forest undergrowth. The use of high resolution or multiple LiDAR data sets acquired at the right time, such as used by Hill & Broughton, (2009), or the use of highly accurate field measurements of the shrub layer, which were used by Wing et al. (2012), appear to drastically improve understorey model predictions.

5 Conclusions

This study investigated whether the shrub layer of a forest could be predicted using ALS-LiDAR data from the AHN. The results showed a poor explanatory power of the proposed LiDAR indices. The main reasons for the poor predictive power of the proposed shrub cover model are believed to be: (1) the use of leaf-off LiDAR data which appears to limit the detectability of deciduous shrubs and (2) the mismatch between the LiDAR data and the used validation data. This does not mean that the AHN itself is unsuitable for undergrowth modelling, but highlights the data requirements needed for accurate shrub layer modelling. The use of leaf-on data is recommended when further research into the prediction of the shrub layer with ALS-LiDAR is attempted.

6 References

Bazzaz, F. A. (1975). Plant Species Diversity in Old-Field Successional Ecosystems in Southern Illinois. Ecology, 56(2), 485–488. http://doi.org/10.2307/1934981

Berg, J. van de. L. B.-S. as bergetal1996. pd. (1996). SYHI, houtvoorraadinventarisatie bij het Staatsbosbeheer. Nederlands Bosbouwtijdschrift, 68, 145–148.

Ding, T. S., Liao, H. C., & Yuan, H. W. (2008). Breeding bird community composition in different successional vegetation in the montane coniferous forests zone of Taiwan. Forest Ecology and Management, 255(7), 2038–2048. http://doi.org/10.1016/j.foreco.2008.01.056

Dubayah, R. O., & Drake, J. B. (2000). Lidar remote sensing for forestry. Journal of Forestry, 98(6), 44–46. Retrieved from https://search.proquest.com/docview/220790008?accountid=27871

Dunn, O. J. (1964). Multiple Comparisons Using Rank Sums. Technometrics, 6(3), 241–252. http://doi.org/10.1080/00401706.1964.10490181

Ewald, M., Dupke, C., Heurich, M., Müller, J., & Reineking, B. (2014). LiDAR remote sensing of forest structure and GPS telemetry data provide insights on winter habitat selection of european roe deer. Forests, 5(6), 1374–1390. http://doi.org/10.3390/f5061374

Hamraz, H., Contreras, M. A., & Zhang, J. (2017). Forest understory trees can be segmented accurately within sufficiently dense airborne laser scanning point clouds. Scientific Reports, 7(1), 6770. http://doi.org/10.1038/s41598-017-07200-0

Hanzelka, J., & Reif, J. (2016). Effects of vegetation structure on the diversity of breeding bird communities in forest stands of non-native black pine (Pinus nigra A.) and black locust (Robinia pseudoacacia L.) in the Czech Republic. Forest Ecology and Management, 379, 102–113. http://doi.org/10.1016/j.foreco.2016.08.017

Hermy, M., & Bijlsma, R. J. (2010). Bosbeheer en biodiversiteit. In B. Muys, J. den Ouden, G. M. J. Mohren, & K. Verheyen (Eds.), Bosecologie en Bosbeheer (pp. 218–233). 317, , : ACCO.

Hill, R. A., & Broughton, R. K. (2009). Mapping the understorey of deciduous woodland from leaf-on and leaf-off airborne LiDAR data: A case study in lowland Britain. ISPRS Journal of Photogrammetry and Remote Sensing, 64(2), 223–233. http://doi.org/10.1016/j.isprsjprs.2008.12.004

Hochberg, Y., & Benjamini, Y. (1995). Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society. Series B (Methodological), 57(1), 289–300. http://doi.org/10.2307/2346101

James, F. C., & Wamer, N. O. (1982). Relationships between Temperate Forest Bird Communities and Vegetation Structure. Ecology, 63(1), 159–171. http://doi.org/10.2307/1937041

Latifi, H., Heurich, M., Hartig, F., Müller, J., Krzystek, P., Jehl, H., & Dech, S. (2016). Estimating over- and understorey canopy density of temperate mixed stands by airborne LiDAR data. Forestry, 89(1), 69–81. http://doi.org/10.1093/forestry/

cpv032

Ludikhuize, B., & Stroeven, C. (2008). Inwinning AHN-2 met nieuw FLI-MAP 400-systeem, 2–4. Retrieved from http://info.fugro.nl/info/archief/2008_06/juni234.pdf

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. http://doi.org/10.1016/j.rse.2009.07.002

Müller, J., & Brandl, R. (2009). Assessing biodiversity by remote sensing in mountainous terrain: The potential of LiDAR to predict forest beetle assemblages. Journal of Applied Ecology, 46(4), 897–905. http://doi.org/10.1111/j.1365-2664.2009.01677.x

Muys, B., Ouden, J. den, & Verheyen, K. (2010). Groei. In J. den Ouden, B. Muys, G. M. J. Mohren, & K. Verheyen (Eds.), Bosecologie en Bosbeheer (pp. 75–91). 63, , : ACCO.

Næsset, E. (1997). Estimating timber volume of forest stands using airborne laser scanner data. Remote Sensing of Environment, 61(2), 246–253. http://doi.org/10.1016/S0034-4257(97)00041-2

Naesset, E., & Naesset, E. (2003). Scandinavian Journal of Forest Research Practical large-scale forest stand inventory using a small-footprint airborne scanning laser Practical Large-scale Forest Stand Inventory Using a Small- footprint Airborne Scanning Laser. Scand. J. For. Res, 19, 164–179. http://doi.org/10.1080/02827580310019257

Roberson, E. J., Chips, M. J., Carson, W. P., & Rooney, T. P. (2016). Deer herbivory reduces web-building spider abundance by simplifying forest vegetation structure. PeerJ, 4, e2538. http://doi.org/10.7717/peerj.2538

Silve. (n.d.). Forest monitoring using SyHI / Woodstock. Retrieved November 14, 2017, from https://www.silve.nl/content/MonitoringWithWoodstock.htm

Simonson, W. D., Allen, H. D., & Coomes, D. A. (2014). Applications of airborne lidar for the assessment of animal species diversity. Methods in Ecology and Evolution, 5(8), 719–729. http://doi.org/10.1111/2041-210X.12219

Tews, J., Brose, U., Grimm, V., Tielbörger, K., Wichmann, M. C., Schwager, M., & Jeltsch, F. (2004). Animal species diversity driven by habitat heterogeneity/diversity: the importance of keystone structures. Journal of Biogeography, 31(1), 79–92. http://doi.org/10.1046/j.0305-0270.2003.00994.x

Tjur, T. (2009). Coefficients of Determination in Logistic Regression Models—A New Proposal: The Coefficient of Discrimination. Source: The American Statistician, 63(4), 366–372. http://doi.org/10.1198/tast.2009.08210doi.org/10.1198/tast.2009.08210

van der Zon, N. (2013). Kwaliteitsdocument AHN-2. Retrieved from http://www.ahn.nl/binaries/content/assets/hwh---ahn/common/wat+is+het+ahn/kwaliteitsdocument_ahn_versie_1_3.pdf

Verheyen, K., Ouden, J. den, Muys, B., & Aerts, R. (2010). Populatiedynamiek. In J. den Ouden, B. Muys, G. M. J. Mohren, & K. Verheyen (Eds.), Bosecologie en Bosbeheer (pp. 103–121). 63, , : ACCO.

Vierling, K. T., Bässler, C., Brandl, R., Vierling, L. A., Weiß, I., & Müller, J. (2011). Spinning a laser web: Predicting spider distributions using LiDAR. Ecological Applications, 21(2), 577–588. http://doi.org/10.1890/09-2155.1

Williams, S. E., & Marsh, H. (1998). Changes in small mammal assemblage structure across a rain forest/open forest ecotone. Journal of Tropical Ecology, 14, 187–198. http://doi.org/10.1017/S0266467498000157

Wing, B. M., Ritchie, M. W., Boston, K., Cohen, W. B., Gitelman, A., & Olsen, M. J. (2012). Prediction of understory vegetation cover with airborne lidar in an interior ponderosa pine forest. Remote Sensing of Environment, 124, 730–741. http://doi.org/10.1016/j.rse.2012.06.024

Zellweger, F., Morsdorf, F., Purves, R. S., Braunisch, V., & Bollmann, K. (2014). Improved methods for measuring forest landscape structure: LiDAR complements field-based habitat assessment. Biodiversity and Conservation, 23(2), 289–307. http://doi.org/10.1007/s10531-013-0600-7