Estimation of canopy cover in dense mixed-species forests using airborne lidar data

Tauri Arumäe & Mait Lang


To link to this article: https://doi.org/10.1080/22797254.2017.1411169

© 2017 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

Published online: 18 Dec 2017.

Article views: 78
Estimation of canopy cover in dense mixed-species forests using airborne lidar data

Tauri Arumäe1,2 and Mait Lang3,4

1Institute of Forestry and Rural Engineering, Estonian University of Life Sciences, Tartu, Estonia; 2Forest survey management division, State Forest Management Centre, Tallinn, Estonia; 3Department of Remote Sensing, Tartu Observatory, Tõravere, Tartumaa, Estonia

ABSTRACT
Airborne laser scanning (ALS) data and digital hemispherical photos (DHP) from 93 sample plots in Laeva test site, Estonia, were used to study effects of phenology and scan angle on the ALS-based canopy cover (CC) estimates. The relative share of first returns (P1/A) for 6185 forest stands was analysed. The CC was calculated using different height thresholds and echoes, and was compared with the CC estimates based on DHP (CCDHP) and crown model (CC_Rcrown). The first of many echoes-based canopy cover estimate (CC_ALS1,3,1) saturated at values greater than 80%. The strongest correlation of CC_DHP was found with CC_ALS1,3,1 using all echoes and a 1.3 m height break (R² = 0.81, RMSE = 11.8%). Correcting the estimate for view nadir angle did not improve the correlation of CC_ALS1,3,1 with CC_DHP. The CC_Rcrown had a weak correlation (R² < 0.25) with CC_ALS and with CC_DHP. The P1/A was not influenced by tree species composition, but by phenology, stand relative density and forest height; however, CC_ALS was not dependent on stand height. Foliage phenology had a substantial effect on CC_ALS and CC_DHP. In dense mixed-species forests, we recommend to use all returns for canopy cover estimation.

Introduction
Airborne laser scanning (ALS) has become a widely used remote sensing method for assessing forest structure variables in operational forestry (Andersen, McGaughey & Reutebuch, 2005). The main applications of ALS are forest height map construction (Arumäe & Lang, 2016a; Næsset, 1997a), wood volume and biomass estimation (Bouvier, Durrieu, Fournier, & Renaud, 2015; Ferraz, 2015; Ferraz et al., 2016; Guerra-Hernández et al., 2016; Næsset, 1997b; Patenaude et al., 2004, Popescu, Zhao, Neuenchwander, & Lin, 2011), carbon stock monitoring (Bright, Hicke, & Hudak, 2012) and biodiversity mapping (Müller & Vierling, 2014; Smith, Anderson, & Fladeland, 2008). Vertical canopy cover (CC) is the key factor for most of these estimates and when combined with other forest structure parameters it is used for leaf-area index (LAI) mapping (Korhonen & Morsdorf, 2014; Solberg et al., 2009) or could be used for planning of thinnings in forest management (Vastaranta et al., 2011). Forest canopy structure is additionally described with mean tree height (H), crown length and the ratio of crown length to H, crown cover, effective CC and angular canopy closure. Distinguishing between different CC estimates is essential for forest structure studies (Jennings et al., 1999; Korhonen, Korpe, Heiskanen, & Maltamo, 2011). The CC, as defined by Jennings et al. (1999) and as used in this study, is the share of ground covered by the vertical projection of the canopy and is commonly expressed as a percentage. The simplest method for measuring vertical CC is using the Caijanus tube (Korhonen, Korhonen, Rautiainen, & Stenberg, 2006; Rautiainen, Stenberg, & Nilson, 2005). Canopy closure, on the other hand, is defined as the proportion of the sky hemisphere obscured by the vegetation canopy when viewed from a single point (Jennings et al., 1999). Canopy closure can be estimated from digital hemispherical photos (DHP) or measured using the LAI-2000 plant canopy analyser (Jonckheere, Nackaerts, Muys, & Coppin, 2005).

The CC and crown cover (ratio of the total area of crown vertical projections divided by the sampling area) are also estimated using crown radius models and stand density (Spurr, 1948). Tree crown radius is commonly estimated using stem diameter at breast height (Davies & Pommerening, 2008; Spurr, 1948) and can be used for tree competition assessment (Purves, Lichstein, & Pacala, 2007). The relationship between stem diameter at breast height, tree crown radius and CC can be used to estimate the mean tree size from ALS data (Ferraz, Saatchi, Mallet, & Meyer, 2016). However, crown radius models may yield different crown cover estimates, depending on the definition of crown radius, species and age (Kandare, Orka, Dalponte, Næsset, & Göbakken, 2017).

CONTACT Tauri Arumäe tuauri.arumae@rmk.ee Institute of Forestry and Rural Engineering, Estonian University of Life Sciences, Kreuzwaldi 5, Tartu 51014, Estonia
© 2017 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
Commonly, tree crowns are modelled as solid geometrical shapes – a cone, ellipsoid (Kuusk & Nilson, 2000) or convex irregular polygon (Mongus & Žalik, 2015). If gaps in the tree crown are accounted, then the result is an effective CC estimate (Duncanson, Cook, Hurt, & Dubayah, 2014).

A DHP records all the radiation penetrating through tree crown or plant canopy, taking into account all the gaps inside the tree crowns, and the share of sky pixels in the total number of pixels is the effective canopy closure. For the conversion of canopy closure into CC, the view zenith angle should not exceed 20° (Korhonen, Korpela, Heiskanen & Maltamo, 2011) and the estimates should be corrected for within-crown gap fraction (Nilson & Kuusk, 2004). The sources of uncertainties in the ALS data-based estimates of CC (CC$_{ALS}$) are related to the scanning setup where the view nadir angle (VNA) is usually in the range of 0° ≤ VNA ≤ 30° and the angular dependence of the observations is characteristic of canopy closure. There are no clear rules of how to account for within-crown gaps. The pulse footprint of the scanners is commonly larger than gaps inside crowns. For example, Leica ALS50-II at a 2400-m flight altitude and beam divergence of 0.22 mrad (Leica, 2007) has a pulse footprint of about 50 cm in diameter. The canopy itself is also semi-transparent on near-infrared (NIR) wavelengths used in most of the topographic mapping oriented ALS devices, so defining crown gaps is problematic. A laser scanner can also retrieve echoes from overlapping crowns, which would allow estimating crown cover, when the overlapping areas of the crowns are taken into account separately.

The most common method for the CC$_{ALS}$ estimation is using a height break or threshold and calculating the share of the first and first of many echoes above the height break in the total number of echoes (Korhonen, Korpela, Heiskanen & Maltamo, 2011). The threshold is commonly set at a few metres above the ground corresponding to the live crown base height (Smith et al., 2009). However, it is found that the CC$_{ALS}$ estimate based on the first and first of many echoes using the, e.g. Leica ALS50-II scanner will result in the CC$_{ALS}$ estimate saturation, especially in dense deciduous forests (Lang, 2010; Lang, Arumäe, & Anniste, 2012). The CC$_{ALS}$ is influenced also by scanning parameters (Keränen, Maltamo, & Packalen, 2016) and the ratio of the first echo count to all echoes ($P_{1/A}$) which characterizes pulse splitting and is affected by phenology stages, especially for deciduous forests (Wasser, Day, Chasmer, & Taylor, 2013). In NIR spectral region, broadleaved deciduous tree species have a higher reflectance than needle leaf species (Kuusk, Lang, & Kuusk, 2013) which can also influence the $P_{1/A}$. As a result, CC is at best estimated from ALS data with a root mean square error of 10% (Ferraz et al., 2015). Similar results are shown for canopy closure estimates (Mooser, Roubinek, Schleppi, Morsdorf, & Jonas, 2014). In practical applications where CC$_{ALS}$ or its analogues are used for ALS-based wood volume models (Arumäe & Lang, 2016b; Bouvier et al., 2015) an error of 10% in CC$_{ALS}$ causes about a 15% difference in predicted wood volume.

The aim of this study was to test different CC$_{ALS}$ estimation methods in dense deciduous broadleaf-dominated hemi-boreal forests stands. The ALS data were from two phenological phases – before bud swelling with leaves off (bBS) and after the final leaf unfolding stage (aFLU) with leaves fully developed. CC estimates from DHP (CC$_{DHP}$) were chosen as the reference for other methods. Additional tests were carried out using CC estimates based on tree position data and the crown radius model (CC$_{Crown}$). The influence of scan angle, $P_{1/A}$ and plot location on the CC$_{ALS}$ estimates was investigated.

**Materials and methods**

**Study site**

The 15 × 15 km test site is located in south-eastern Estonia, near Laea (Lang, Arumäe, Lükk, & Sims, 2014). The terrain is rather flat. Half of the area is covered by forests. The forests are of mixed species and multi-layer structure is common, with a dense understory layer of Padus avium Mill. and Corylus avellana L. Dominant tree species are silver birch (Betula pendula Roth), Norway spruce (Picea abies L.), trembling aspen (Populus tremula L.), black alder (Alnus glutinosa L.) and Norway spruce in the lower layer. The most common site types are the Aegopodium and Filipendula (Löhmus, 2004) and based on FAO-UNESCO, the soil types are mainly fertile Calc Eutric Gleysols and Eutri Histic Gleysols. The forest height can reach up to 37 m (typical height in mature forests is 25–30 m) and the basal area is up to 40 m$^2$/ha according to forest inventory (FI) data.

The first dataset in the tests was FI data for 6185 stands. The FI data were used for studying the influence of deciduous species fraction on $P_{1/A}$. The majority of the stands were inventoried in 2013. The average stand size was 2.0 ha and dominating tree species are silver birch and trembling aspen with a common second layer of Norway spruces. The second dataset was based on 93 sample plots extracted from the Estonian Network of Forest Research (ENFR) database (Kiviste et al., 2015). According to Kiviste et al. (2015) all the trees on these plots were calipered, tree positions were mapped and model trees were selected for crown base height and height measurements. The rest of the tree heights were estimated using diameter-height
models based on sample tree measurements. Dominating tree species in the ENFR plots are silver birch and trembling aspen, with an average age of 59 years for silver birch and 65 years for trembling aspen. The average height of the forests was 23 m and basal area average was 25 m²/ha. The average relative density (also known as stand stocking index) was 81%.

**ALS data**

The ALS data were collected in spring (06 May 2013) and summer (13 July 2013) by the Estonian Land Board using the scanner Leica ALS50-II. The ALS point density for bBS was 0.5 points/m², the flight altitude was 2400 m and ALS pulse footprint diameter on the ground was 50 cm. Point density for aFLU dataset was 2 points/m², flight altitude 1800 m and pulse footprint diameter on the ground was 40 cm in diameter. The bBS and aFLU datasets were combined in one analysis to increase the range of $CC_{ALS}$ and to use the bBS data as a substitute for possible defoliation effects. The VNA did not exceed 28°. The ALS data were processed using FUSION/ LDV freeware (McGaughey, 2014).

The $CC_{ALS}$ was calculated using different height breaks ($z$) with a 1-m step starting from 1.3 m and ending at the live crown base height ($H_{LCB}$) to reduce the forest understory vegetation influence. The $CC_{ALS}$ was calculated as follows:

$$CC_{ALS,z,E} = \frac{100 \cdot P_E(h_p > z)}{P_E}$$

where

- $P$ – the number of echoes,
- $h_p$ – the pulse return height from the ground,
- $E$ – selection of the first or first of many ("1") or all ("A") echoes.

The ALS-based live crown base height ($H_{LCB,ALS}$) was estimated with the model (2) taken from Arumäe and Lang (2013), where $H_{LCB,ALS,0}$ is calculated using point cloud height distribution mode value ($H_{Mode}$) and standard deviation ($H_{Stdv}$) excluding the points with $h_p \leq 1.3$ m.

$$H_{LCB,ALS,0} = H_{Mode} - \frac{H_{Stdv}}{2}$$

Correlation between the measured $H_{LCB}$ and $H_{LCB,ALS,0}$ was strong ($R^2 = 0.79$). However, $H_{LCB,ALS,0}$ overestimated $H_{LCB}$ on average by 6.8 m and therefore a linear correction model (3) was applied.

$$H_{LCB,ALS} = 0.69 \cdot H_{LCB,ALS,0} - 1.3$$

We tested also the influence of VNA correction on $CC_{ALS}$ depending on the selection of echoes and ALS measurement geometry. To study the $CC_{ALS}$ dependence on scanning angle, first 38 of the ENFR plots were selected where the scan angle was large (18–28°). The sample plots occurred on the overlapping area of scan swaths and, as the result of the flight plan, they were scanned from two opposite directions ($ALS_{Left}$; $ALS_{Right}$). The relationship between $CC_{ALS,Left}$ and $CC_{ALS,Right}$ was analysed before and after the VNA correction with the model by Korhonen and Morsdorf (2014)

$$CC_{ALS} = CC_{ALS,1.3.,1} - 0.0253 \cdot \theta_{scan} \cdot F_{max},$$

where

- $CC_{ALS,1.3.,1}$ – the CC estimate using the first or first of many echoes,
- $\theta_{scan}$ – mean scan angle (°),
- $F_{max}$ – height of the highest echo above the digital terrain model (m).

In pulse splitting analysis, both canopy and ground returns were included to test the influence of tree species on $P_{UA}$. The impact of foliage phenology on the occurrence of returns within the canopy and the corresponding $P_{UA}$ was analysed in the ENFR plots including only the pulse returns with $h_p > 1.3$.

**Digital hemispherical photographs**

The DHP measurements were carried out in the summer of 2013. On 32 out of the 93 ENFR plots, DHP measurements were carried out also in spring at the time before bud swelling. Twelve photos per plot were taken following the VALERI protocol (Validation of Land European Remote Sensing Instruments, http://www.avignon.inra.fr/valeri/). Three sampling points were marked in all four cardinal points, with a 4-m distance between each sample. For hemispherical photos, we used a Nikon D5100 with the Sigma 4.5 mm F2.8 EX DC HSM Circular Fisheye lens and a Canon EOS 5D with the Sigma 8 mm f:3.5 EX DG Fisheye lens. The HSP software (Lang, Kodar, & Arumäe, 2013) was used for DHP processing. The $CC_{DHP}$ for each plot was then calculated as an average of $CC$ estimates from 12 single photos. Pixels from the view zenith angle of less than 9° were used to measure $CC$ (corresponds roughly to the first ring of LAI-2000).

**Crown radius models**

Two crown radius ($R_{Crown}$) models were initially tested – the first model was published by Jakobsons (1970) and is based on stem diameter at breast height ($d$). The second $R_{Crown}$ model (model 14 from Lang, Nilson, Kuusk, Kivist, & Hordo, 2007) is based on tree height and $d$. Using the calculated $R_{Crown}$ values, the $CC$ ($CC_{R_{Crown}}$) estimate was calculated by merging all the crown shapes. The $CC_{R_{Crown}}$ using
Jakobsons (1970) model was 32–42% greater compared to the \( CC_{R_{\text{Crown}}} \) calculated using model by Lang et al. (2007). Additionally, the average CC for all 93 plots with Jakobsons’ (1970) \( R_{\text{Crown}} \) model was 50% greater than with the model by Nilson, Lang, Kuusk, Anniste, and Lükk (2000) that relates \( CC = 100\cdot(0.8987+0.044) \) where \( T \) is stand relative density. Therefore, Jakobsons’ (1970) model was abandoned.

Tree crown vertical projections on the ground can be represented as concentric rings of a particular radius in the case of equal spacing of the trees. In natural stands, however, distances between trees vary and, due to the competition for light, branches grow more likely towards open space. As a result, the overlaps between tree crown projections decrease and CC increases. Since tree positions in the ENFR plots were known, a model from Lang and Kurvits (2007) was used to adjust crown projections according to the neighbours of each tree, while the area of each crown projection was fixed to that of the circle with the predicted crown radius. This model was able to draw more realistic crown shapes, and a visual comparison of the adjusted crown projections with orthophotos showed a good agreement. For each sample plot, all crown projections were then merged into one polygon using QGIS to calculate the \( CC_{R_{\text{Crown}}} \). After the crown shape modification, CC increased (mean \( CC_{R_{\text{Crown}}} \) by 3% and the maximum difference was 16%) compared to the circular crown projections-based CC. We assumed that this also increases the correlation with \( CC_{\text{ALS}} \). We did not apply edge correction (see, e.g., Lilleleht, Sims, & Pommerening, 2014) and a small underestimation of CC near a sample plot border is still possible. To study the edge effect, a 2-m wide buffer was excluded from outside the sample plots. The parts of crown projections within the decreased sample plot were then extracted and \( CC_{R_{\text{Crown}}} \) was estimated again. The \( CC_{R_{\text{Crown}}} \) estimates were compared with the aFLU \( CC_{\text{ALS}} \).

**Manipulation of plot centre location**

To estimate the stability of \( CC_{\text{ALS}} \), the centre positions of the 93 ENFR plots were randomly dislocated for 100 times within a radius of 10 m, which corresponds roughly to the estimated maximum error in the ENFR plot coordinate measurements. The \( CC_{\text{ALS},1.3_\Lambda} \) was calculated for each dislocated point cloud sample and variation was analysed.

**Error estimates**

The mean error of estimate (MEE) was calculated as

\[
\text{MEE} = \frac{\Sigma (X - Y)}{N} \quad (5)
\]

and the root mean square error (RMSE) was calculated as

\[
\text{RMSE} = \sqrt{\frac{\Sigma (X - Y)^2}{N}} \quad (6)
\]

where \( X \) is the argument, \( Y \) is the dependent variable and \( N \) is the number of observations.

**Results**

The shifting of the centre coordinate showed that the \( CC_{\text{ALS},1.3_\Lambda} \) estimates are rather stable concerning errors in ALS point cloud sample locations. The standard error of \( CC_{\text{ALS},1.3_\Lambda} \) for the 93 ENFR plots was 0.15% and the average interquartile range of \( CC_{\text{ALS},1.3_\Lambda} \) was 2.2%. The largest range of \( CC_{\text{ALS},1.3_\Lambda} \) estimates was 23%, nine sample plots had a \( CC_{\text{ALS},1.3_\Lambda} \) range of larger than 10% and the average range of \( CC_{\text{ALS},1.3_\Lambda} \) was 6.1%.

There was a substantial influence of phenology on the \( CC_{\text{ALS}} \) estimates. For bBS, the average \( CC_{\text{ALS},1.3_\Lambda} \) of the ENFR plots was about 20% smaller compared to the average of aFLU conditions (Table 1).

The correlation of \( CC_{\text{DHP}} \) with the first echoes-based \( CC_{\text{ALS},1.3_\Lambda} \) was weaker (\( R^2 = 0.75 \), \( \text{RMSE} = 20.7\% \)) than with \( CC_{\text{ALS},1.3_\Lambda} \) (\( R^2 = 0.81 \), \( \text{RMSE} = 11.8\% \). Figure 1, Table 1). When aFLU and bBS measurements were tested separately, the smallest RMSE was found for \( CC_{\text{ALS},1.3_\Lambda} \) and \( CC_{\text{DHP}} \) of aFLU flight and the \( R^2 \) corresponded to a moderate correlation.

The change in correlation between \( CC_{\text{ALS,2}} \) and \( CC_{\text{DHP}} \) when the height break was raised from \( z = 1 \) to \( z = 5 \) m was not significant. Raising the \( z \) even higher, up to 10 m, somewhat improved the predictive power of the first echoes, but the increase in \( R^2 \) was small. The \( CC_{\text{ALS,LCB}} \), estimated at \( z = H_{\text{LCB,ALS}} \), had surprisingly only a weak correlation (\( R^2 = 0.15 \)) with \( CC_{\text{DHP}} \) for aFLU dataset and no correlation for bBS measurements (\( R^2 = 0 \)). This was most likely due to the ALS-based \( H_{\text{LCB}} \) model (2) not being able to predict the \( H_{\text{LCB}} \) for bBS dataset. The \( H_{\text{Mode}} \) for bBS data occurred near the minimum height threshold of 1.3 m following the mode value of the pulse return height distribution and the \( CC_{\text{ALS,LCB}} \) was substantially overestimated. This was probably the result of a dense forest understory and second layer of spruces in many stands that caused pulse returns from below the upper tree layer and flattened the height distribution of pulse returns and changed the position of \( H_{\text{Mode}} \). Using the measured \( H_{\text{LCB}} \),

<table>
<thead>
<tr>
<th>Phenophase</th>
<th>( CC_{1.3} ) using the first and first of many echoes (%)</th>
<th>( CC_{1.3} ) using all echoes (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>bBS</td>
<td>Mean ( \text{RMSE} ) ( \text{MEE} ) ( R^2 )</td>
<td>Mean ( \text{RMSE} ) ( \text{MEE} ) ( R^2 )</td>
</tr>
<tr>
<td>aFLU</td>
<td>74.0 34.8 –31.7 0.64 59.7 19.3 –17.3 0.62</td>
<td></td>
</tr>
<tr>
<td>Combined</td>
<td>93.9 14.0 –11.9 0.32 83.5 7.4 –1.5 0.37</td>
<td></td>
</tr>
</tbody>
</table>

a: bBS: before bud swelling; aFLU: after final leaf unfolding.
for z instead of the point cloud-based $H_{LCB,ALS}$ improved the correlation of $CC_{ALS,LCB}$ with $CC_{DHP}$ ($R^2 = 0.22$) when tested for aFLU dataset. We found also that $CC_{DHP}$ was almost always smaller than $CC_{ALS}$ (Table 1, Figure 1). Since the gap fraction estimation from DHP with the LinearRatio (Cescatti, 2007) method that is used in HSP software is an unbiased technique (Lang, Kodar & Arumäe, 2013, Lang et al., 2017), the $CC_{ALS}$ is probably overestimated. However, in our study we did not correct $CC_{DHP}$ for within the crown gap fraction which would increase aFLU $CC_{DHP}$ by a few percentages and thus established a good correlation between the mean $CC_{ALS,1.3,A}$ and $CC_{DHP}$.

The pulse splitting within the canopy showed also a difference between aFLU and bBS datasets. The pulses were splitting less ($p$-value < 0.01) and the share of the first returns increased using aFLU data ($P_{1/A} = 0.82$) compared to the bBS dataset ($P_{1/A} = 0.76$). This is also one of the reasons for the systematically overestimated $CC_{ALS}$ for bBS data (Figure 1), since the number of returns from the canopy increased.

The comparison of $CC_{ALS,1.3,1}$ in the ENFR plots that were scanned twice indicated that at a large VNA, the CC estimates for sample plots may vary substantially when calculated from repeated scans taken from different directions – the maximum differences ranged up to 8%. The VNA correction improved the $CC_{ALS,1.3,1}$ estimation (Figure 2) precision as appeared from the analysis of overlapping scan swaths. The VNA correction also decreased $CC_{ALS}$ by 13% on average, but the scatter between the repeated measurements remained considerable (Figure 2(b)). The VNA uncorrected dataset (Figure 2(a)) showed a slightly weaker correlation between $CC_{ALS,Left}$ and $CC_{ALS,Right}$ ($R^2 = 0.81$, RMSE = 4.2%) compared to the VNA corrected dataset (Figure 2(b); $R^2 = 0.89$; RMSE = 3.8%); however, the change in $R^2$ was statistically insignificant when tested using Sheskin’s (2000) statistical test.

The $CC_{ALS}$ of all ENFR plots was then corrected for the VNA and compared to $CC_{DHP}$ (Figure 3). The scan angle correction decreased the $MEE$ of the relationship between $CC_{ALS,1.3,1}$ and $CC_{DHP}$ from −12.8 to −4.7; however, the change in $R^2$ was not significant according to the statistical test (Sheskin, 2000). The scan angle correction caused a systematic underestimation of $CC_{ALS,1.3,A}$ compared to $CC_{DHP}$ with $MEE$ increasing from −2.8 to 5.3, but the increase in $R^2$ was again not statistically significant.

The percentage of deciduous tree species in the dominant layer of trees had no correlation with the first echo ratio $P_{1/A}$ ($R^2 = 0$) using the aFLU dataset. Whereas $P_{1/A}$ had a negative moderate correlation with stand relative density ($R^2 = 0.35$) and a moderate negative correlation with the ALS-based 80th height percentile ($H_{P80}; R^2 = 0.56$). The VNA had an influence on $P_{1/A}$. The share of pulses giving only a single echo increased with VNA ≥ 10° and the $P_{1/A}$ at VNA ≥ 16° was 2% larger compared to the near nadir $P_{1/A}$ ($p$-value < 0.01 based on 379 stands scanned at nadir and 420 stands scanned at VNA>16°).

In contrary to the expected relationship between $CC_{ALS}$ and tree crown-based CC, the concentric ring-based $CC_{RCrown}$ showed no correlation with $CC_{DHP}$ or $CC_{ALS}$. The neighbourhood corrected $CC_{RCrown}$ showed a weak correlation with $CC_{DHP}$ ($R^2 = 0.14$) and also with $CC_{ALS,1.3,A}$ ($R^2 = 0.22$) and $CC_{ALS,1.3,1}$ ($R^2 = 0.14$). Our test with the excluded 2-m-wide buffer from outside the sample plots where the crown model may have lacked the influence of external competition did not show a marked increase in the correlations with other CC estimates. The 2-m buffer zone exclusion increased the average of $CC_{RCrown}$ by 2.7% and increased the $R$ of the $CC_{RCrown}$ relationship with $CC_{DHP}$ ($R^2 = 0.18$), $CC_{ALS,1.3,A}$ ($R^2 = 0.24$) and $CC_{ALS,1.3,1}$ ($R^2 = 0.16$), but statistically the influence

---

**Figure 1.** First echoes-based canopy cover (CC) calculated at 1.3-m height break $CC_{ALS,1.3,1}$ saturates in dense forests in summer (aFLU) (a) compared to all echoes-based $CC_{ALS,1.3,A}$ (b).
was insignificant ($p$-value > 0.05). The correlation of $CC_{RCrown}$ with $CC_{ALS}$ based on the measured $H_{CB}$ was weak ($R^2 = 0.02$), but the mean values of the two CC estimates (67.4% and 69.7%, correspondingly) were closer than $CC_{RCrown}$ and $CC_{ALS,1.3}$ (67.4% and 83.6%, correspondingly).

**Discussion**

The comparison of $CC_{ALS}$ and $CC_{DHP}$ in the best-case scenario had an RMSE of 11.8%, which is comparable to the results of similar studies (Ahmed, Franklin, Wulder, & White, 2015; Smith et al., 2009). The relatively large RMSE is most likely the result of measurement errors and comparison of different variables as CC. For example, DHP is measuring light that is penetrating the canopy and the tree crowns are accounted as semi-transparent, as opposed to the $CC_{RCrown}$ which accounts crowns as solid shapes, but ALS-based measurements use the NIR spectral region where also foliage is semi-transparent. Also, the ALS observations are not points, but samples with an area defined usually by the laser beam divergence and flight altitude (Baltsavias, 1999). Sample plot position errors also propagate into $CC_{ALS}$ and decrease the reliability of the CC estimates. Although we found that $CC_{ALS,1.3}$ was usually not substantially influenced by the positioning errors of sample plots, this was probably due to the ENFR plots being well-positioned in homogeneous parts of forest stands. On the other hand, $CC_{ALS}$ calculated for a sample plot using ALS point clouds from different flight paths had an uncertainty of up to 8% in CC units in the forests.

Somewhat surprisingly, there was only a weak correlation between $CC_{RCrown}$ and $CC_{ALS}$, This is most likely because $CC_{RCrown}$ and $CC_{ALS}$ are different by their definition or is because of the used $R_{Crown}$ models, which were not able to predict accurately tree crown radius in the dense stands. One reason might be also the influence of understory trees, for which there was no data to calculate $CC_{RCrown}$ estimates.

**Figure 2.** Comparison of $CC_{ALS}$ estimated from the opposite direction flights using the first echoes (a) without VNA correction and (b) using Korhonen and Morsdorf’s (2014) VNA correction model.

**Figure 3.** Influence of view nadir angle correction on $CC_{ALS,1.3}$ estimates using the first echoes (a) and using all echoes (b).
However, there was not much of an improvement in correlation when $CC_{ALS}$ was estimated at measured $H_{L_{CB}}$ threshold to exclude the influence of understory vegetation. Only the mean value of $CC_{RCrown}$ was in a better agreement with $CC_{ALS,L_{CB}}$ compared to $CC_{ALS,1.3}$. The correlation of $CC_{RCrown}$ with $CC_{DHP}$ and $CC_{ALS,L_{CB}}$ did not show a significant improvement after adjusting crown shapes, however, the $CC_{RCrown}$ was estimated according to the competition of neighbouring trees. It is possible that the tree crown shapes vary in nature much more than predicted by the model, and in semi-naturally developed forests, tree stems are not strictly vertical due to competition during their growth. As a result, the $CC$ can be much larger in many plots compared to the estimates based on the crown projections and assumed vertical stems at tree stump locations.

In the dense forests, the $CC_{ALS}$ was estimated using the regular height break $z = 1.3$ m and saturated using the first or first of many echoes at CC greater than 80%, whereas the $CC_{ALS,z=A}$ did not saturate. There was not much of an influence on $CC_{ALS}$ from the increase in $z$ until 10 m above the ground or until the $H_{L_{CB}}$ level was reached. Raising the threshold higher than 1.3 m did improve the predictive power of the first echoes, but overall, the relationship between $CC_{ALS}$ and $CC_{DHP}$ was still weak and the change in the determination coefficient was statistically insignificant. Using the ALS-based $H_{L_{CB}}$ as a threshold introduces additional errors in CC caused by the errors in $H_{L_{CB,ALS}}$ estimation. On the other hand, the lack of sensitivity of $CC_{ALS}$ to the change in $z$ is probably caused by the upper layer dense canopy in the forests (images can be found in the appendix in Lang et al., 2014), which captures most of the pulse energy and thus relatively fewer echoes from the lower layers down to the ground vegetation are triggered. On the other hand, there is a clear negative correlation between the leaf area index of upper and lower canopy layers in forests (Chianucci, Puletti, Venturi, Cutini, & Chiavetta, 2014; Kodar et al., 2011). In our test, DHPs were taken always at 1.3 m above the ground and a denser forest understory in the case of a smaller CC of the upper layer can explain the weak correlation between the values of $CC_{ALS}$ calculated using $H_{L_{CB}}$ and $CC_{DHP}$.

In addition to the selection of pulse returns and discrimination threshold for the CC estimation, there are also sampling problems and issues related to the interaction between the lidar pulse and forest canopy. The laser scanners measure at various VNAs and this raises the question of representativity of higher-order echoes in the case of small sample plots in tall forests. With the VNA increase, the second and higher-order returns are displaced horizontally by $h_{diff} \times \tan(VNA)$, where $h_{diff}$ is the vertical distance between the first and subsequent returns of the lidar pulse. For example, at $VNA = 25^\circ$ and $h_{diff} = 21$ m, the displacement in the horizontal direction is 9.8 m. Therefore when a pulse is sent out at a large VNA, the first echo is triggered from a plot or grid cell location and the following echoes would be received in a shifted location, but their occurrence is still determined by the part of the canopy that triggered the first echo. This makes the $CC_{ALS}$ estimate dependent on the grouping of trees and sensitive to variation in forest height. However, there was a positive correlation between the VNA and canopy $P_{1/A}$ at larger scan angles ($VNA > 15^\circ$) meaning fewer returns per pulse at the view angles. The reasons for this may include the increase in footprint size at larger scanning angles (Heritage & Large, 2009; Korhonen, Korpela, Heiskanen & Maltamo, 2011) and the fact that at large scan angles the probability of seeing crown sides increases and so does the path length through the canopy. The increased path length within the forest canopy causes more scattering which flattens the distribution of the received energy leaving less chance for the scanner internal software to distinguish more than just the first return from emitted pulses.

Additionally to the VNA, the species composition is known to influence the point cloud height distribution and therefore the estimates of $CC_{ALS}$ (Wasser, Day, Chasmer, & Taylor, 2013). The higher reflectance of deciduous broadleaf tree species in NIR should, in theory, result in a greater $P_{1/A}$ as the pulses are being split less due to a stronger signal from the top of the canopy. However, based on the 6185 forest stands, the fraction of broadleaf trees in the upper layer had no influence on $P_{1/A}$, but $P_{1/A}$ was influenced by the structural features of the forest – stand density and height. The correlation between $P_{1/A}$ and forest height is partially explained by the fact that the Leica ALS50-II scanner can register up to four echoes from a single pulse, with a minimum distance of greater than 3.5 m between the echoes. Therefore, the splitting of pulses can only occur in forests where the tree height exceeds the multiple value of the spacing distance. The $P_{1/A}$, and therefore $CC_{ALS}$, is also influenced by the scanner automatic gain control (AGC) (Vain, Yu, Kaasalainen, & Hyppa, 2010) which regulates the intensity of the emitted pulses, but the effect was not studied in this research. Similarly to AGC, which regulates the emitted energy, the flight altitude and, in turn, the footprint size has an effect on echo registration (Gaveau & Hill, 2003). The increased pulse splitting found on ENFR plots during bBS compared to aFLU phenophase may also be one reason why all returns-based $CC_{ALS}$ estimates were still systematically greater compared to $CC_{DHP}$ in spring, while there was a good agreement in summer. Similarly to Straatsma and Middelkoop (2006) we found that there were fewer second or higher-order returns per pulse in summer triggered by the
canopy compared to leafless conditions in spring. This is related to the higher NIR reflectance of the foliage compared to leafless branches, which creates a better-defined mode value into the distribution of returned photons. However, \( CC_{ALS} \) and \( CC_{DHP} \) both showed a consistent increase in CC in accordance with the foliage phenology. The phenology-driven change in \( CC_{ALS} \) may be an indicator of the share of deciduous species in the estimation of species composition in forest stands.

To account for the complex influence of the VNA on \( CC_{ALS} \) estimates, we applied the VNA correction model published by Korhonen and Morsdorf (2014). In general, there was a small increase in correlation between the \( CC_{ALS} \) and \( CC_{DHP} \) and a decrease in RMSE, but the improvement was statistically insignificant. The VNA correction of \( CC_{ALS,1.3_A} \) resulted in a systematically smaller estimate compared to \( CC_{DHP} \) and is therefore not recommended for \( CC_{ALS,1.3_A} \). This indicates that the VNA correction models for ALS-based CC estimates are dependent on the selection of pulse returns, since the model was originally constructed for the first returns of laser pulses. For future studies, a VNA correction model also for all returns-based \( CC_{ALS} \) should be developed.

Conclusion

We analysed discrete-return ALS data from dense deciduous broadleaf-dominated mixed forests for CC estimation using references obtained from tree crown models and DHP. The uncertainties in the ALS-based CC estimates due to the errors in sample plot location and sampling of point clouds from different scans at large VNAs are roughly in the same range and reach up to 8–10% in CC units in the stands. Correction of VNA effects systematically decreased CC estimates, but did not decrease variability and there was no improvement in the correlation with CC estimated from DHPs. A VNA effect correction model developed for the first returns yields biased CC values when applied to all returns-based CC estimates. There is not much of an influence on ALS-based CC estimates when selecting a threshold height from the range between 1.3 m above the ground and the level of the live crown base in the dense forests.

There was a good agreement between the CC estimated from ALS data and DHPs, while both estimates had only a weak correlation with the CC based on crown radius models. The weak correlation was probably related to the different meaning of the variables and failure of crown models to account for single tree crown plasticity in the dense forests. The increase in foliage density decreases the number of returns per pulse triggered from the forest canopy. This effect may be important for ALS-based defoliation estimates which will be thus less sensitive to an actual decrease in foliage density. In general, the number of returns per pulse in the aFLU phenophase ALS data was not dependent on the proportion of broadleaf trees but only on forest height and stand relative density. Finally, for the estimation of CC using ALS data in the dense forests, we recommend using the first returns and VNA effect correction or all returns without the correction.

Acknowledgments

The authors would like to thank the Estonian Land Board for the airborne lidar data. Data acquisition was financed by the Estonian State Forest Management Centre. Ave Kodar helped during field measurements. Data analysis was supported by the Ministry of Education and Research grant IUT21-4. The Estonian Network of Forest Research Plots was supported by the Estonian State Forest Management Centre and the Estonian Environmental Investment Centre. The authors are also thankful for the helpful comments from the anonymous reviewers.

Disclosure statement

No potential conflict of interest was reported by the authors.

References


Arunäe, T., & Lang, M. (2016b). ALS-based wood volume models of forest stands and comparison with forest inventory data. Foresty Studies, 64, 5–16. doi:10.1515/fsmu-2016-0001


Lang, M., Kodar, A., & Arumäe, T. (2013). Restoration of above canopy reference hemispherical image from below


