# Comparing discrete echoes counts and intensity sums from ALS for estimating forest LAI and gap fraction

Svein Solberg

Norwegian Forest and Landscape Institute, 1431 Ås, Norway
svein.solberg@skogoglandskap.no

## Abstract

Effective leaf area index (*LAI*) of a forest is mathematically related to gap fraction, and may be estimated from the penetration rate of airborne laser scanning (ALS). The aim of this study was to compare the usefulness of four alternative ALS penetration rates for this purpose, and in particular to determine if any of them could produce un-biased estimates of gap fraction. This would be valuable in forests having a large fraction of small gaps. A 21 km$^2$ pine forest was covered with ALS, and ground measurements of gap fraction and *LAI* were done with LICOR's LAI-2000 at 20 plots within the area. The alternative penetration rate methods utilizes intensity data and multiple echoes. All alternatives were strongly related to gap fraction and *LAI*. However, none of the alternatives provided un-biased gap fraction estimates. The simple method of first echo counting, after being categorized as canopy and gap echoes, turned out to have the smallest deviation from gap fraction, with a slight underestimation. The methods of summing up intensities and counting of multiple echoes overestimated gap fraction, however, they were still strongly related to gap fraction and *LAI*, and may serve as supplementary methods in dense canopies with many small gaps, due to their higher sensitivity to small gaps.

*Keywords: laser scanning, ALS, LIDAR, gap fraction, LAI*

## 1. Introduction

ALS is intuitively, with its high number of evenly distributed pulses, a perfect tool for sampling based estimation of an area property such as gap fraction or canopy cover, where the return of each pulse is classified as either canopy or gap. Canopy cover may further be converted into *LAI* by inverting the gap fraction and log-transformation based on Beer-Lambert law (Chen et al. 1997). It has been demonstrated that such log-transformed, inverted gap fraction data from ALS are very strongly and linearly related to *LAI* from ground based measurements with R$^2$ values around 0.90 and higher (Solberg et al. 2005, Solberg et al. 2006). We may, however, cast a critical view on this apparently promising method: *LAI* in those studies were derived from the gap fraction as sampled from the first echoes of the laser beams, and it is a general concern that these first echoes underestimate gap fraction because the footprint size of the beam is too large to penetrate down small gaps (Lovett et al. 2003; Morsdorf et al. 2006). Hence, the method will apparently only work in forests where gaps smaller than the footprint size are infrequent, or eventually that they make up some constant fraction of the total gap fraction.

The ALS data have, however, another and less used attribute that may be useful here: the intensity. Lovett et al. (2003) have presented the idea of estimating gap fraction by summing up intensity data. The intensity of an echo depends on the size of the hit object and its reflectivity. If the reflectivity were fairly constant and similar for canopy and ground, then the intensity will represent mainly the size of the hit object. A large object would produce a compact return pulse and a high intensity value, and vice versa. Thus, we might derive unbiased estimates of canopy cover and gap fraction by replacing the simple counting of echo categories by summing up their intensities. Small canopy objects will the be given less weight compared to large objects, and in this way correspond to the fraction of cover within the footprint. Hence, although it may be correct that an entire laser beam can not penetrate down small gaps, the gap fraction might still be correctly recorded as a reduced intensity of an echo. Another option is also available: It seems likely that one would get higher penetration rates, eventually producing un-biased gap fraction estimates, by including the intermediate and last echoes of the pulses. These echoes represent the deeper penetration of the laser beam into the canopy.

The aim of this study was to compare the usefulness of four alternative ALS methods for mapping of gap fraction and LAI, and in particular to test a hypothesis that we can derive unbiased estimates of gap fraction from ALS data by using intensity data and eventually all echo categories.

## 2. Method

### 2.1 Study area and data sets

The data for this study is taken from a laser scanning campaign in 2005, which is a large and homogeneous data set, with three replicated sets of measurements and ALS acquisitions during the summer. The study area, data set and results are described in full detail by Solberg et al. (2006), and is briefly outlined here. The aim of that campaign was to map the defoliation caused by insect attack on the trees. ALS data together with ground based measurements of *LAI* were gathered. The study area was a 21 km$^2$ large area along a river and a relatively flat, sandy valley area along it. The area was dominated by Scots pine.

ALS data were gathered on May 10th; August 1st, and September 1st, with two pulses per m$^2$, and a maximum scan angle after pre-processing of ±12°. The technical properties of the scans were as follows:

- Aircraft: Piper PA31-310
- Mean flight altitude: 650 m above ground
- Laser scanner: Optech ALTM 3100C
- Wavelength: 1064 nm
- Pulse repetition frequency: 100 kHz
- Pulse width: 16 ns
- Pulse energy: 66 μJ
- Peak power: 4.1 kW
- Mean footprint diameter at the ground: 17 cm

The discrete return pulses were categorized as ‘only’; ‘first of many’; ‘intermediate’; and ‘last of many’, and in the further data handling here the two former categories were grouped together and termed ‘first’ echoes. Each echo had the variables *x*, *y*, *z*, *dz*, *intensity* and *echo category*, where *dz* represents the height above ground and all these variables were provided by the ALS acquisition company.

Twenty field plots were laid out over the area as a stratified, systematic sampling with four age classes (ranging from newly regenerated to old stands) and five replicates of each. At each of the three points of time *LAI* was measured at each plot using LICOR’s LAI-2000 plant canopy analyzer at five points of each plot (centre and 3m away towards the cardinal directions), and these measurements were tripled making a total of 15 measurements at each plot and each point of time. The measurements were done 1m above ground. Reference measurements at nearby open fields were done simultaneously and every 15 second, and every plot measurement was joined with the reference station measurement being closest in time. The *LAI* value obtained from LAI-2000 represents the hemi-surface area of the foliage objects, which is half their total surface area.

### 2.2 Calculus and statistics

ALS penetration rates were calculated for four alternative approaches, which was combinations of either using first echoes or using all echoes; and secondly, either doing echo counting or intensity summing. First, the echoes were classified into *canopy* echoes if they had *dz* values above 1m above ground, and *ground* echoes if less than 1m. For the echo counting approach the following model was used:

$$P = N_g / (N_g + N_c) \quad , \qquad (1)$$

where $P$ is the penetration rate, $N_g$ is the number of ground echoes, and $N_c$ is the number of canopy echoes. This was done on ALS data for five circles of various size (5; 10; 15; 20; and 25m) around each plot. In the case of using intensities the following model was used:

$$P = \sum I_g / (\sum I_g + \sum I_c) \quad , \qquad (2)$$

where $I_g$ and $I_c$ are the intensity of a ground echo and a canopy echo, respectively. The influence of reflectivity was not included in this equation, because data on reflectivity was not available. This implies that if this approach should turn out to be successful, it would mean that reflectivity was equal for all hit objects. It can also be noted that this ratio is not affected by the variation in the distance from the laser scanner within the scanned area.

The data from the LAI-2000 were first calculated into effective leaf area index, *LAI* using the following model provided by the producer:

$$LAI = 2\sum w_i \ln(GF_i^{-1}) / d_i \, , \qquad (3)$$

where $i$ is the ring number, $GF_i$ is gap fraction seen in the zenith angle direction of ring $i$; The terms $w_i$ and $d_i$ are ring specific factors provided by the hardware producer, representing the observed canopy volume, and the view path length, respectively, for that ring. A median LAI value was then calculated based on the replicate measurements done at each plot and each point of time. The median was used in order to exclude the influence from outlier results, which may be frequent in such data due to the sub-optimal weather conditions during measurements, such as direct sun light and partly clouded sky. *LAI* was calculated with two alternatives: using rings 1-4 only, and using all rings 1-5.

The data from the LAI-2000 were secondly calculated into foliage orientation, - in two alternative ways, - first by assuming a spherical foliage angle distribution corresponding to a mean tilt angle of 60°, and second by calculating the mean tilt angle (*MTA*) based on the LAI-2000 data. The projected fraction of the foliage area, $G(\theta)$ is equal to cos(MTA), and this is half the hemi-surface area for a spherical foliage angle distribution. *MTA* was calculated with the default method using the Fv2000 software which is shipped with the LAI-2000 hardware. As for *LAI*, *MTA* was calculated both for rings 1-4 and for all rings 1-5. It turned out that this produced a number of cases having the inappropriate *MTA* values 0 and 90°. In order to counteract these problems, each plot was provided with its median *MTA* value, being calculated across replicates and across repeated measurements during the season, however, where all values of 0 and 90° degrees were discarded from the median calculation.

Vertical gap fractions might be derived directly from the innermost ring, with its near-vertical view of ±12°. This gap fraction, however, suffers from large random errors as a very tiny bit of the canopy is seen with this ring, depending strongly on where the instrument is put, which could be under a tree or in a between-tree gap. The alternative and robust method applied here was to utilize the data from the other rings of LAI-2000. The random error is then considerably reduced as a much larger canopy volume is measured. Each ring has a gap fraction in its view direction, and these gap fractions were recalculated into one common, vertical gap fraction based on LAI and on foliage orientation:

$$LAI = \frac{1}{G(\theta)} \ln(GF^{-1}), \text{ and hence } GF = e^{-G(\theta) \cdot LAI} \quad . \tag{4}$$

After this preparation of ALS and LAI-2000 data, they were combined for modelling. A gap fraction corrective, $c$, was introduced for handling of systematic under- (or over-) estimation of gap fraction when using ALS penetration rate:

$$GF = P^c \quad , \qquad \ln GF = c \cdot \ln P \tag{5}$$

This model has two intuitively suitable properties. First, it is a relationship that meets the requirement that the ALS penetration rate and the gap fraction have to be equal in two cases; - when gap fraction is zero (a completely opaque canopy layer) and when it is one (a clear cut), while in-between here the relationship can be either linear or non-linear. Second, it works as a scaling factor in a model used for estimating *LAI* based on ALS penetration, and hence, allowing LAI to be strictly linearly related to $ln(P^{-1})$ even if the gap fraction estimate is biased:

$$LAI = \frac{1}{G(\theta)} \ln(GF^{-1}) = \frac{c}{G(\theta)} \ln(P^{-1}) \quad , \tag{6}$$

where the ALS penetration rate is an unbiased estimate of gap fraction if $c$=1; it underestimates gap fraction if $c$<1; and vice versa.

In this study no-intercept regression models are widely used, and such models don't have a trivial definition for the coefficient of determination ($R^2$), and I used the following formula in accordance with recommendations from Kvålseth (1985).

$$R^2 = 1 - \frac{n}{n-p} \frac{\sum (y-\hat{y})^2}{\sum (y-\bar{y})^2} , \tag{7}$$

where $n$ is the number of observations, $p$ is the number of model parameters, $\hat{y}$ is the model prediction for a given observation $y$, and $\bar{y}$ is the mean of all observations $y$.

## 3. Result

For analyzing the relationships between ALS penetration rates and LAI-2000 gap fractions it turned out that the optimal selection of data (the strongest relationships) were obtained by using ALS data from within a 15m radius around the plot centre, and by excluding the outermost ring of LAI-2000. This is reasonable as the forest stands are fairly small and stand edges are likely to be seen by the outermost ring in many cases. Hence, the outermost ring does not contribute with valuable data, but rather produce noise only, and is hence excluded from estimation of both *LAI* and *MTA*.

All the alternative ALS penetration rates were strongly correlated to gap fraction, having $R^2$

values ranging from 0.71-0.94. However, none of the four methods produced unbiased estimates of gap fraction. Compared to the simple method of counting first echoes, all the three other methods produced higher penetration rates, as expected. Not only had these gap fraction estimates higher values, but they were biased the opposite way, - overestimating gap fractions. The results were fairly similar for the two alternatives for foliage orientation. Assuming a spherical foliage angle distribution was apparently quite right: The more detailed estimation of *MTA* indicated slightly more horizontally oriented foliage, however, the difference was minor, and also it introduced new random errors causing lower $R^2$ values.

The two alternative counting methods, i.e. using either first or all echoes, were equally good in terms of strength of relationship, and also they were equally much biased, although they were biased in opposite ways. As seen in Fig. 1 the penetration rates calculated this way were generally close to the gap fractions (circle symbols and solid fit lines, Fig. 1).

In addition to being weaker related to gap fraction, the intensity based approaches suffered in general from a poorer fit than the counting methods. Inspection of residual plots revealed a tendency of non-linearity. The intensity based penetration rates represented an overestimation of gap fraction, particularly at higher values (triangles, Fig. 1). In particular, summing up intensities for all echoes produced much higher values than the gap fractions.

Table 1: Results of no-intercept log-log models for LAI-2000 gap fractions against ALS penetration rates (Eq. 5). The estimated slope is an estimate of the gap fraction corrective, *c*. Results of the various alternatives for data selection and handling are presented (echo categories used; counts versus intensity sums; and foliage orientation assumed to follow a spherical foliage angle distribution or estimated mean tilt angle *MTA*)

| Echoes used | Method | Foliage orientation | $c$ | $R^2$ | Residual check |
|---|---|---|---|---|---|
| First | Echo counts | Spherical | .77 | .94 | Ok |
| --“-- | Intensity sums | --“-- | 1.17 | .86 | Curved |
| All | Echo counts | --“-- | 1.12 | .94 | Ok |
| --“-- | Intensity sums | --“-- | 1.45 | .87 | Curved |
| First | Echo counts | MTA | .81 | .83 | Ok |
| --“-- | Intensity sums | --“-- | 1.22 | .72 | Ok |
| All | Echo counts | --“-- | 1.18 | .81 | Ok |
| --“-- | Intensity sums | --“-- | 1.51 | .71 | Ok |

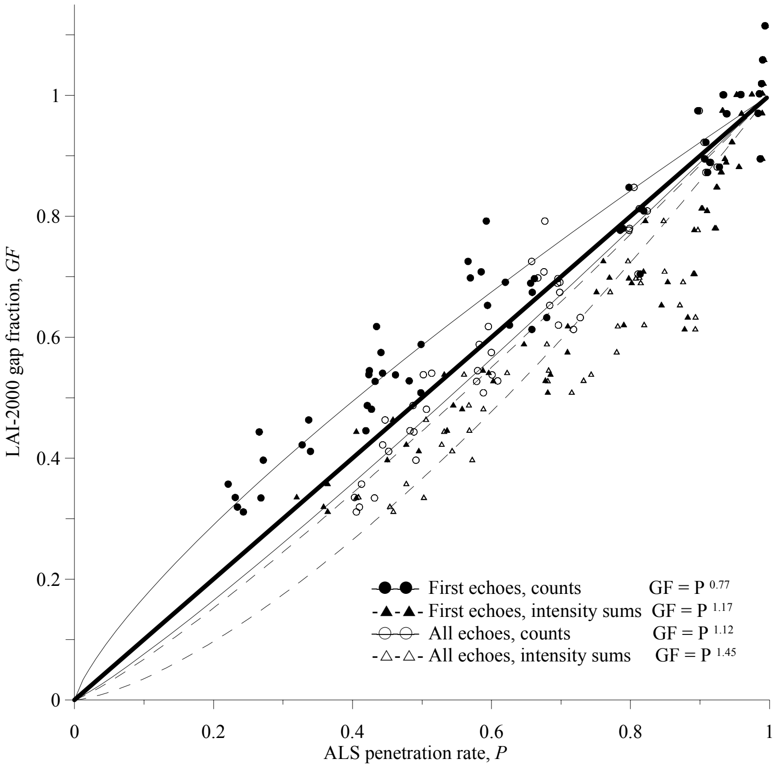

Figure 1: Gap fraction, *GF*, plotted against ALS penetration rates calculated in four alternative ways, based on rings 1-4 of LAI-2000 and assuming a spherical foliage angle distribution. The lines represent the fit functions (Table 1). A few *GF* values are above 1.0 which result from higher light intensities measured in a plot compared to the reference measurement in an open field nearby.

Two models of *LAI* as a function of ALS penetration rates (Eq. 6) are shown in Fig. 2, and is a reformulation of the results shown in table 1. The tendency of a non-linear relationship for the intensity-based method is shown in Fig. 2, right. At low *LAI* values, this method tend to underestimate *LAI*. In comparison, the echo counting method is straight linear.

*LAI*, as well as its changes from before to after insect defoliation, were now estimated based on each of the four alternative ALS penetration rates for the entire area with a 10x10m grid, producing a large data set of 139720 grid cells having four alternative *LAI* and *LAI* change values. The parameter estimates for the slopes in these LAI models were equal to $c/G(\theta)$. This corresponds to two times the *c* parameter estimate (Table 1), being 1.54; 2.34; 2.24; and 2.90, respectively for the four methods (counting first echoes, summing intensity of first echoes, counting all echoes, and summing intensities of all echoes). This is implicitly based on the assumption that foliage orientation is the same all over, which seems to be quite correct based on the results presented above (Table 1). All alternatives *LAI* variables were strongly correlated. However, again, the intensity based methods had a tendency of non-linearity, which tended to underestimate *LAI* at low *LAI* values and vice versa (Fig. 3, left), as shown above (Fig. 2, right). The two echo counting methods, however, were linearly related over the entire *LAI* range. All the *LAI* change variables were highly correlated, and despite its tendency of non-linearity, the intensity based method was the one being most strongly correlated to the *LAI* change based on echo counting ($R^2$=0.93, Fig. 3, right).

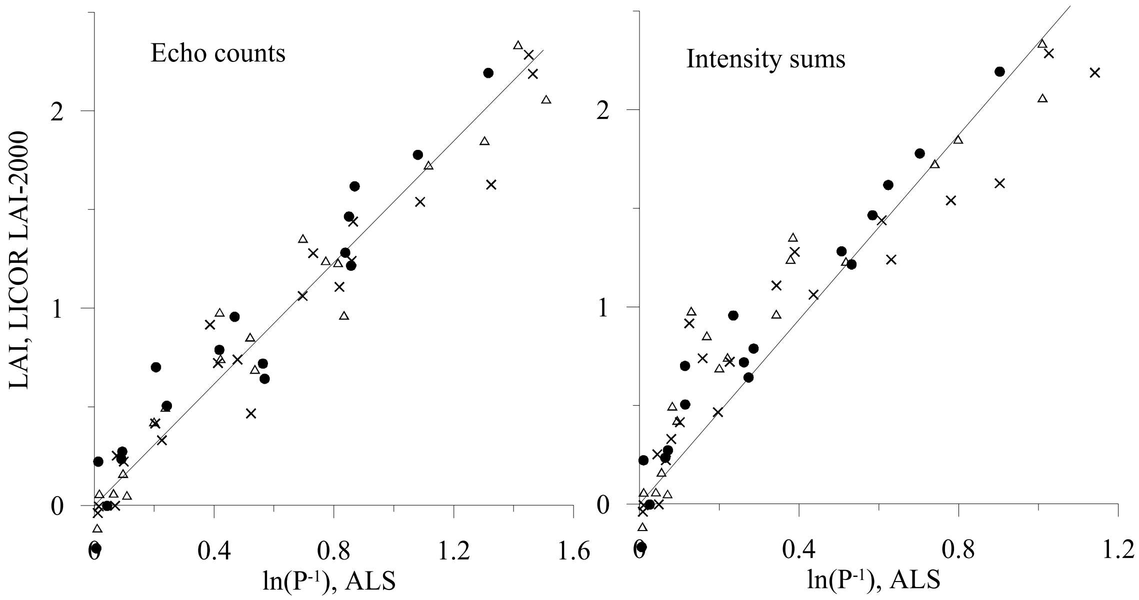

Figure 2: LAI from ground based measurements with LAI-2000 based on rings 1-4 plotted against log-transformed ALS penetration rate data ($P^{-1}$) from ALS data. The two alternative penetration rates are derived from using first echo counts (left) and first echo intensity sums (right).
Data are from three repetitions in time: ● May; x August; and Δ September.

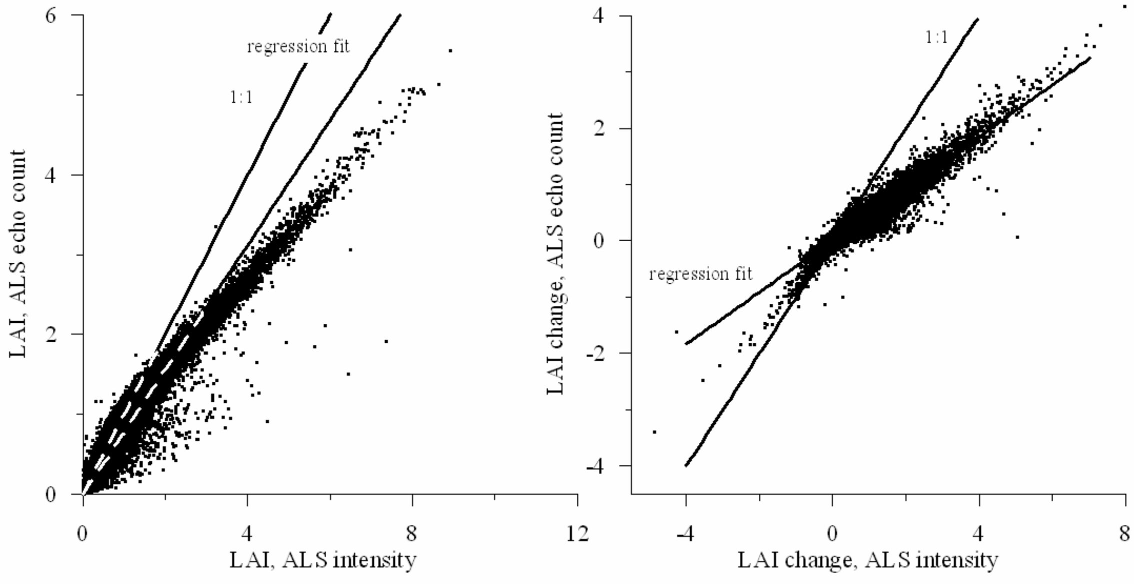

Figure 3: *LAI*, and its change during the summer, calculated for a 10x10m grid using the two ALS alternative methods echo counting and intensity summing. N=139720.

First echoes from the canopy had generally lower intensity than those from the ground, with mean values being 70 and 130, respectively. Defoliation caused both a higher fraction of echoes to penetrate the canopy layer and a reduced intensity of canopy echoes. The intensity of ground echoes was not affected.

## 4. Discussion

Using the standard method of counting first echoes was the most appropriate method in this data set. Compared to LAI-2000 measurements it underestimates gap fraction corresponding to a 20-30% underestimation of *LAI*. However; it is experienced in other studies that LAI-2000 normally underestimates *LAI* with some 15% because of light scattering in the canopy (Chen et al. 1997). It is likely that such a systematic error is present in this data set, since the LAI-2000 measurements often were carried out in sub-optimal weather conditions such as mid-day sunshine and partly cloudy. If we in the present data set adjust up the *LAI* data from LAI-2000 with 15%, then the method of counting first echoes would produce almost un-biased estimates of gap fraction. The parameter estimates of the gap fraction corrective, *c*, in Table 1 would then increase to 0.89 for a spherical foliage angle distribution, and to 0.93 if using the calculated mean tilt angles. For comparison, the other alternative penetration rates would then strongly overestimate gap fraction, with parameter estimates, *c*, ranging from 1.29 and upwards.

The results discussed above are not likely to depend much on footprint size and other acquisition settings. It is often claimed though, that the penetration of ALS downwards into the forest canopy depends on footprint size and pulse energy, and related factors such as flight altitude. However, a physical explanation for this is not evident and straight-forward: On one hand large footprints are intuitively less able to penetrate down small gaps as compared to small footprints. On the other hand, when large footprints hit the canopy they may be fragmented into a high number of non-detectable echoes, having their first detectable echo located deep down in the canopy, causing an apparent deeper canopy penetration. A careful review of a number of studies addressing the influence of footprint size and other ALS acquisition factors (Nilsson 1996; Lovell et al. 2003; Næsset 2004) suggests that when using first echoes (echo categories 'only' and 'first of many') the acquisition factors have generally minor influence on canopy penetration.

As expected, it is possible to estimate gap fraction and *LAI* from summing up echo intensities. The relationships with ground based measurements were strong, having $R^2$ values of 0.86-0.87. However, compared to the simpler method of counting echoes, it was a less appropriate method, having lower $R^2$ values; showing a non-linearity in the relationships; and having systematic overestimation of the gap fraction. This overestimation becomes particularly large if we assume that LAI-2000 overestimates *LAI*. This overestimation may have two causes: The intensity is likely corresponding to the size of the hit object, i.e. the uppermost canopy objects the laser beam hits. The fraction of the footprint that is continuing downwards through the canopy is then handled as if it continued all the way down to the ground, representing gaps. However, the photons may well hit canopy objects further down. The interpretation is visualized in Fig. 4. If using the echo counting method, an echo is weighed as if the laser beam hit an opaque object covering the entire footprint, which obviously represents an underestimation of the gap fraction within that footprint. And if using the intensity summing method, the echo is given a weight corresponding to the size of the uppermost hit branch, which represents an overestimation of the gap fraction. A second problem with the intensity based method is that the reflectivity may indeed be variable, and hence, the results obtained here may only be valid for this type of forest, - pine trees growing on soils covered mostly by reindeer lichens. Pine trees are found to have a particularly low canopy reflectance as compared to many other foliage objects because of multiple scattering within the shoots (Smolander and Stenberg, 2003). Laboratory experiments have demonstrated that the backscatter from the reindeer lichens *Cladina* and *Cladonia* with the NIR wavelength used in ALS is very high, eventually being exceptionally intense when the lichens have complex branching structures (Kaasalainen and Rautiainen, 2005).

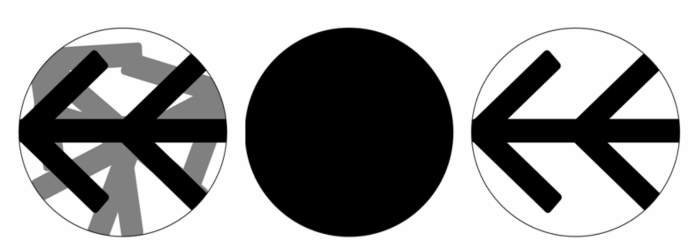

Figure 4: An imagined case where a circular laser beam footprint hits a canopy (black = uppermost canopy object, grey=canopy objects lower down and white=gap (left); discrete echo counted as a full reflection (black circle) causing underestimation of gap size (centre); and echo intensity corresponding to the size of the uppermost canopy object (black) causing overestimation of gap size as objects deeper down are taken as gap.

Using echo counting and including all echo categories (i.e. including intermediate and last echoes) produced close to un-biased estimates of gap fraction, and had as strong relationships to ground based measurements as using first echoes only. As an estimate of gap fraction it was biased, i.e. overestimating gap fraction. However, even if it could not produce un-biased estimates of gap fraction, it represents a valuable method alternative to be used in very dense canopies. *LAI* estimation based on first echoes may easily get saturated in ALS scans having few pulses per square meter and if one wants to map *LAI* with a high spatial resolution. The saturation problem here refers to cases when there are no ground echoes, and then *LAI* can not be estimated. By including the intermediate and last echoes the saturation problem will be much reduced. However, the usefulness of this approach may be limited by the minimum time distance needed to separate two echoes: If the canopy layer has a low surface height, there might not be enough time to separate multiple echoes. And there may be hardware-specific differences in this minimum time distance, which would cause inconsistencies in multi-temporal ALS data sets from different producers.

Finally, the methods based on intensity sums and counting of all echoes may be useful for forest health monitoring in forests having small gaps, due to their higher sensitivity to small gaps, and as they were strongly correlated to the other LAI change variables.

## 4. Conclusion

The aim of this study was to compare four alternative ALS penetration rates for estimation of gap fraction and LAI. None of the alternatives produced unbiased estimates of gap fraction. The best result was obtained by estimating gap fraction as the fraction of the first echoes that were classified as ground echoes. This produced slightly underestimated gap fraction values, however, they were strongly correlated to the ground measured gap fractions. The methods of summing up intensities and counting of multiple echoes overestimated gap fraction, however, they were still strongly related to gap fraction and *LAI*, and may serve as supplementary methods in dense canopies with many small gaps, due to their higher sensitivity to small gaps.

## Acknowledgements

The Norwegian research council, together with the fund Skogtiltaksfondet and the Ministry of Agriculture and Food, are acknowledged for funding the REMFOR project, in which this study is carried out.

## References

Chen, J. M., Rich, P.M., Gower, S.T., Norman, J.M. and Plummer, S., 1997. Leaf area index of boreal forests: Theory, techniques, and measurements. *Journal of Geophysical Research*, 102(D24), 29429–29444.

Kaasalainen, S. and Rautiainen, M., 2005. Hot spot reflectance signatures of common boreal lichens. *Journal of Geophysical Research-Atmospheres*, 110 (D20), D20102.

Kvålseth, T.O., 1985. Cautionary note about $R^2$. *The American Statistician,* 39/4: 279-285.

Lovell, J.L., Jupp. D.L.B., Culvenor, D.S. and Coops, N.C., 2003. Using airborne and ground-based ranging LiDAR to measure forest canopy structure in Australian forests. *Canadian Journal of Remote Sensing*, 29: 607-622.

Morsdorf, F., Kotz, B., Meier, E., Itten, K.I. and Allgower, B., 2006. Estimation of LAI and fractional cover from small footprint airborne laser scanning data based on gap fraction. *Remote Sensing of Environment*, 104(1):50-61.

Næsset, E., 2004. Effects of different flying altitudes on biophysical stand properties estimated from canopy height and density measured with a small-footprint airborne scanning laser. *Remote Sensing of Environment*. 91, 243-255.

Nilsson, M., 1996. Estimation of tree heights and stand volume using an airborne lidar system. *Remote Sensing of Environment* 56, 1–7.

Smolander S. and Stenberg, P., 2003. A method to account for shoot scale clumping in coniferous canopy reflectance models. *Remote Sensing of Environment*, 88(4), 363-373.

Solberg, S., Næsset, E., Aurdal, L., Lange, H., Bollandsås, O.M. and Solberg, R., 2005. Remote sensing of foliar mass and chlorophyll as indicators of forest health: preliminary results from a project in Norway. *Proceedings of ForestSat 2005* (H. Olsson, editor), 31 May – 3 June 2005, Borås, Sweden (Report 8a of the National Board of Forestry, Sweden), pp. 105-109.

Solberg, S., Næsset, E., Hanssen, K.H. and Christiansen, E., 2006. Mapping defoliation during a severe insect attack on Scots pine using airborne laser scanning. *Remote Sensing of Environment*, 102, 364-376.