Title:

Key structural features of Boreal forests may be detected directly using L-moments from airborne lidar data

Authors:

Rubén Valbuena*[(1)], Matti Maltamo[(1)], Lauri Mehtätalo [(2)], Petteri Packalen [(1)]

Affiliations:

(1) University of Eastern Finland, School of Forest Sciences, PO Box 111. Joensuu, Finland; rubenval@uef.fi; matti.maltamo@uef.fi; petteri.packalen@uef.fi

(2) University of Eastern Finland. School of Computing. PO Box 111. Joensuu, Finland; lauri.mehtatalo@uef.fi

*Corresponding author.

Abstract

This article introduces a novel methodology for automated classification of forest areas from airborne laser scanning (ALS) datasets based on two direct and simple rules: L-coefficient of variation $Lcv = 0.5$ and L-skewness $Lskew = 0$, thresholds based on descriptors of the mathematical properties of ALS height distributions. We observed that, while $Lcv > 0.5$ may represent forests with large tree size inequality, $Lskew > 0$ can be an indicator for areas lacking a closed dominant canopy. $Lcv = 0.5$ discriminated forests with trees of approximately equal sizes (even tree size classes) from those with large tree size inequality (uneven tree size classes) with kappa κ = 0.48 and overall accuracy OA = 92.4%, while $Lskew = 0$ segregated oligophotic and euphotic zones with κ = 0.56 and OA = 84.6%. We showed that a supervised classification could only marginally improve some of these accuracy results. The rule-based approach presents a simple method for detecting structural properties key to tree competition and potential for natural regeneration. The study was carried out with low-density datasets from the national program on ALS surveying of Finland, which shows potential for replication with the ALS datasets typically acquired at nation-wide scales. Since the presented method was based on deductive mathematical rules for describing distributions, it stands out from inductive supervised and unsupervised classification methods which are more commonly used in remote sensing. Therefore, it presents an opportunity for deducing physical relations which could partly eliminate the need for supporting ALS applications with field plot data for training and modelling, at least in Boreal forest ecosystems.

Key words

Airborne laser scanning; L-moments; Gini Coefficient; L-coefficient of variation; forest structure; tree size inequality; shade-tolerance.

1. Introduction

Airborne laser scanning (ALS) can be a valuable tool for studying structural properties of forests (Lefsky et al., 1999a; Drake et al., 2002; Frazer et al., 2005; Maltamo et al., 2005; Valbuena et al., 2016a). The relationships of ALS to forest structure can be employed to analyse asymmetric competition among trees (Kellner & Asner, 2009), and hence forest growth conditions (Stark et al., 2010). In fully-stocked forests (Gove, 2004) light resource pre-emption drives asymmetric competition processes, leading to mortality of the least competitive trees (Weiner, 1990). These are forests with closed canopies and structural properties yielding shady areas, i.e. oligophotic zones (sensu Lefsky et al., 2002), under the dominant tree crowns. In turn, detecting forest areas with light resource availability, which are characterized by large euphotic zones (sensu Lefsky et al., 2002), can be key to monitoring forest disturbance and regeneration. Several metrics derived from ALS height distributions have potential for describing these key characteristics related to forest structure (Zimble et al., 2003). For this reason, studies on ALS-based forest structure characterization by statistical inductive methods, which relate ALS metrics to field attributes empirically, are commonplace (Hall et al., 2005; Lefsky et al., 2005; Dalponte et al., 2008; Pascual et al., 2008; Disney et al., 2010; Jaskierniak et al., 2011; Ozdemir & Donoghue, 2013; Valbuena et al., 2014).

Size hierarchy among trees growing in the vicinity influences competition processes in the forest community (Weiner, 1990; Valbuena et al., 2012). Knox et al. (1989) suggested the Gini coefficient (GC) (Gini, 1921) as a consistent descriptor of tree size inequality, and hence a reliable indicator of competition conditions in the forest (Cordonnier & Kunstler, 2015). For this reason, in the context of ALS estimation, the GC of tree sizes has been used as a basis for stratifying the forest area into homogeneous structural types (Bollandsås & Næsset, 2007; Valbuena et al., 2013a). Furthermore, Knox et al. (1989) also suggested the inclusion of skewness as a complement to the GC in describing forest structural properties. For this reason, Valbuena et al. (2013a) included asymmetry in their analysis of forest structural properties, to study relations of relative dominance between different strata in the forest vertical profile.

While Bollandsås & Næsset (2007) employed stand register data from previous inventories for carrying out their stratification, it would be advantageous if the same remote sensing material could be used for wall-to-wall predictions of forest structure indicators and classifications into forest structural types (Lefsky et al., 1999b; Drake et al., 2002). In particular, Ozdemir & Donoghue (2013) and Valbuena et al. (2013b; 2016a) obtained predictions of the GC of tree size inequality with reliable accuracy. As previous research has concentrated on the forest response (Lefsky et al., 1999a; Valbuena et al., 2013a), and on its analysis and estimation by a wide range of different statistical methods – such as analysis of variance (Zimble et al., 2003), canonical correlation (Lefsky et al., 2005), parametric (Hall et al., 2005) and non-parametric (Valbuena et al., 2014) modelling, histogram thresholding (Maltamo et al., 2005), or finite mixtures (Jaskierniak et al., 2011) –, the next question to answer would be: do the ALS metrics have, by themselves, capacity to discriminate among forest structural types, making no use of statistical methods linking field data to ALS metrics?.

Moments are quantitative measurements of probability density distributions employed to summarize their properties. The most conventional are the product moments, expected values of the powers of a random variable which lead to the use of mean, variance and skewness as measures for location, scale and shape. These descriptors of ALS return height distributions are metrics commonly employed as auxiliary variables in forest assessment (e.g., Næsset, 2002; White et al., 2013). Alternatively, Frazer et al. (2011) and Ozdemir & Donoghue (2013) recently drew the attention towards the L-moments, a set of statistics known by their sample efficiency (i.e., reliability at low sample sizes) and robustness to outliers, compared to conventional moments (Hosking, 1990). Consider a sample order statistic $X_{\mathrm{k:r}}$ – the k^{th} smallest observation in a sample of size r –, which is a many-to-one transformation of a random sample of size r, and therefore a random variable. The L-moments are based on its expected values $E(X_{\mathrm{k:r}})$ (Appendix A). Moreover, L-moment ratios have the advantage of being bounded by finite intervals (Hosking 1989), making them comparable among ALS distributions differing in their mean height. The L-coefficient of variation (Lcv) and the L-skewness ($Lskew$) are two types of L-moment ratios (Appendix A.2). Lcv is the ratio of the second ($L2$) to the first ($L1$) L-moments:

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

where $E(X)$ is the expected value of X. In the case of ALS metrics, the variable X is the height of ALS returns. The Lcv is mathematically equivalent to the GC (Appendix A.3), and therefore the same properties apply to both of them. For instance, they are scale-invariant, and for positive random variables their values are bounded within the [0, 1] interval (Hosking, 1989). Also, Valbuena et al. (2012) showed that an asymptote at $GC = 0.5$ represents the case of maximum entropy among tree sizes in the forest. On the other hand, $Lskew$ is the ratio of the third ($L3$) to the second ($L2$) L-moments:

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

In the case of $Lskew$, its theoretical bounds are [-1, 1] (Hosking, 1989). The value of $Lskew = 0$ corresponds to a symmetric distribution, while positive or negative values denote the type of asymmetry for the distribution of ALS heights. This article employs these mathematical properties of L-moments for describing ALS height distributions, in contrast to inductively researching explanatory potential in relation to field data attributes.

The aim of this research was to develop simple methods for explaining key features related to forest structure from few L-moment ratios of ALS returns. Lcv and $Lskew$ were used for detecting tree size inequality and light availability, and they were utilized for an automated classification of forests from ALS datasets, which was applied directly without the use of field data. The idea builds upon the hypothesis that two deductive mathematical rules, $Lcv = 0.5$ and $Lskew = 0$, may be used to classify the forest area into two groups, based solely on the ALS height distributions. We studied whether such classifications would be sound in terms of explaining properties of size inequality among trees growing in vicinity (even or uneven tree sizes) and competitive conditions for light in the forest community (oligophotic or euphotic). We compared the reliability of the rule-based method to results obtained from a supervised classification. This article discusses suitable applications for this rule-based method.

2. Materials

2.1. Study area and ALS data

The research was conducted in a 252,000 ha study area including approximately 200,000 ha of the Boreal forest ecosystems typically found in the region of North Karelia (Finland), which consists of forests dominated by Scots pine (*Pinus sylvestris* L.) Norway spruce (*Picea abies* (L.) Karst.) or Birch species (*Betula* ssp.) with various degrees of admixtures also with other deciduous trees (such as *Alnus* ssp., *Populus* ssp. etc). The ALS data were acquired by Blom Kartta Oy (Finland) during May 2012 with an ALS60 system from Leica Geosystems (Switzerland). A flying height of 2,300 m above ground rendered an average density of 0.91 pulses per squared-meter. Country-wide laser data are being consistently acquired using broadly similar parameters (National Land Survey of Finland; NLS, 2013). Methods may therefore by consistently replicated throughout the country, bringing potential for upscaling the results obtained at national-level.

Heights above ground for individual ALS returns were calculated by subtracting the digital terrain model provided by the NLS. We considered that, as seedlings and saplings were included in field mensuration (Valbuena et al., 2016b), their influence in laser pulse interception had to be accounted for in ALS metric computation. Consequently, just a very small height threshold of 0.1 m was used, only with the intention to mask out the influence of the ground. Sample estimates of L-moments and their ratios (Wang, 1996) were computed from the heights of all the ALS returns located within each cell over a regular grid covering the entire study area. The spatial resolution of this grid was 16 m × 16 m, a customary practice in Finland that makes cell size roughly coincident in with the area of field plots operationally established and measured by Finnish Forest Centre (SMK, Suomen Metsäkeskus).

2.2. Field dataset used for validation

Field data for validation of the methods were partly acquired by University of Eastern Finland (UEF), partly provided by SMK. A total of $N = 244$ plots were acquired in a stratified random sampling fashion with approximately equal per-stratum sample sizes (Valbuena et al., 2016b). The strata employed were the forest development classes commonly used in operational management in Finland (per-stratum sample sizes were $n = 31$, unless specified): *Seedling*, *Sapling*, *Young*, *Advanced*, *Mature*, *Shelterwood*, *Seed-tree* ($n = 29$), and *Multi-storied* ($n = 29$). SMK's stand register data based on previous inventories was employed for the initial randomization of field plot locations. Valbuena et al. (2016b) provides details about acquisition protocol and processing of field data. Appendix B details the criteria used to assign a development class to for each field plot, a task carried out independently by experienced SMK personnel.

3. **Methods**

3.1. The rule-based method for stratifying forests based on ALS data

We used a deductive approach to thresholding using the L-moment ratios. The rules were deduced from their mathematical properties, as opposed to using inductive, supervised, data-driven optimization or classification:

- The value $Lcv = 0.5$ was used because it represents maximum entropy of tree sizes (Valbuena et al. 2012); also recall that $Lcv = GC$ (see Appendix A.3). Since Lcv describes the relative dispersion of ALS heights, we postulated that Lcv could be used as descriptor for structural properties related to tree size inequality, and hypothesised that this threshold could be suitable for discriminating forests with trees of approximately equal sizes – even tree sizes – ($Lcv < 0.5$) from those with high tree size inequality – uneven tree sizes – ($Lcv > 0.5$).
- The value of $Lskew = 0$ was chosen because it represents a symmetric distribution of ALS heights, and distinguishes plots with positive or negative skewness (Hosking, 1989). Being a descriptor of asymmetry, we postulated that $Lskew$ could be used as descriptor for structural properties related to competitive dominance and light availability characteristics (Valbuena et al., 2013a), and hypothesised that this threshold could be useful for discriminating oligophotic zones ($Lskew < 0$) from euphotic ones ($Lskew > 0$).

We classified forests throughout the scanned area according to these rules directly, avoiding the use of field data in the training stage of the classification. The capacity of these rules to describe structural features of the forest was validated by comparing the classifications at field plot locations to the known development classes determined at the field plots. For that purpose, the development classes were aggregated into the target forest structural properties: even/uneven tree sizes and oligophotic/euphotic.

3.2. Aggregation of development classes

With the intention to study the hypothesised relationship between these thresholds of L-moment ratios for ALS height distribution and their related structural properties of forests, we aggregated the forest development classes according to their structural properties. In even-aged silviculture, the succession of development classes usually follows this a basic chronosequence of even-sized forest types: *Seedling*, *Sapling*, *Young*, *Advanced* and *Mature* stands. Silviculture based on natural regeneration yields more complex uneven-sized structural types: *Shelterwood*, *Seed-tree*, and *Multi-storied* stands. In Finland, *Shelterwood* stands are forest areas attaining regeneration of shade-tolerant species under the shade casted by a closed dominant *Mature* canopy (Appendix B). This is the oligophotic zone (Lefsky et al., 2002), which in the context of Eurasian Boreal forests corresponds to regeneration areas for Norway spruce (note: there are many different types of shelterwood management systems and, although in Finland this term is used specifically for shade-tolerant regeneration – Appendix B –, in other countries it may refer to regeneration of shade-intolerant species too, e.g. Valbuena et al., 2013a). Other oligophotic areas are those which have reached the stem exclusion stage – *Young*, *Advanced* and *Mature* stands –, limiting light availability under the dominant canopy (Zenner, 2005). On the other hand, *Seed-tree* stands are areas where few parent trees provide seeds for natural regeneration which recruits in the understorey generating *Multi-storied* stands (Appendix B). These, as well as *Seedling* and *Sapling* stands, belong to the euphotic zone (Lefsky et al., 2002), where the absence of a closed dominant canopy brings enough light to the ground as to allow the growth of shade-intolerant species. Accordingly, to test the capacity of the $Lcv = 0.5$ and $Lskew = 0$ rules to discriminate forest areas according to their respective hypotheses, the development classes where aggregated as:

(1) First criterion. Inequality among tree sizes ($Lcv = 0.5$):

- Even tree size forest structural types: *Seedling*, *Sapling*, *Young*, *Advanced* and *Mature* stands. Characterized by low relative dispersion in tree sizes (Valbuena et al., 2013a).
- Uneven tree size forest structural types: *Shelterwood*, *Seed-tree* and *Multi-storied* stands. Characterized by high relative dispersion in tree sizes (Valbuena et al., 2013a).

(2) Second criterion. Relative dominance of overstorey over the understorey ($Lskew = 0$):

- Oligophotic (forest structural types with a closed dominant canopy not allowing shade-intolerant regeneration): *Young*, *Advanced*, *Mature* and *Shelterwood* stands. Characterized by negative asymmetries (Valbuena et al., 2013a).
- Euphotic (forest structural types with canopy openness allowing shade-intolerant regeneration): *Seedling*, *Sapling*, *Seed-tree*, and *Multi-storied* stands. Characterized by positive asymmetries (Valbuena et al., 2013a).

3.3. Comparison against supervised classification

In order to compare the rule-based method with more common data-driven methodologies based on inductive statistical inference, we contrasted the results against those obtained by a supervised classification. For that purpose, we employed the results obtained in Valbuena et al. (2016b) from a support vector machine (SVM) classification which employed the same field plot dataset at the training stage as the one used for accuracy assessment in the present study. SVM is becoming increasingly popular for classification of ALS data (Dalponte et al., 2008; García et al., 2011), since it is suitable for operating with big datasets and complex relationships of covariance. SVM is a hard classifier which calculates hyperplanes between classes under a cost function defined as a combination of maximizing distances from training samples to the hyperplanes while minimizing the error of misclassified samples. Using package *e1071* in R statistical environment (Meyer et al., 2014a) and a SVM C-classification method, Valbuena et al. (2016b) computed predictions of all the above-mentioned development classes separately which, in the present study, we aggregated into the established criteria: inequality (even and uneven tree size classes) and dominance (oligophotic and euphotic), as detailed above. It may be worth noting that, in contrast to the rule-based method which avoided the training stage, the SMV predictions were obtained by an error minimization method using field data support and the explanatory capacity of many more ALS metrics (Valbuena et al., 2016b: Table 2).

3.4. Accuracy assessment.

Field data plots were only used for assessing the accuracy of the rule-based method. Relationships among L-moments of ALS heights were observed in scatterplots which depicted the development class to which each plot belonged, observing the role of different development classes in these relationships. Development classes were grouped as described above, and the capacity of the $Lcv = 0.5$ and $Lskew = 0$ rules to describe those grouping characteristics was assessed with the help of contingency matrices. The degree of misclassification was evaluated by the final overall accuracy (OA) and per-class user's (UA) and producer's (PA) accuracies, which were all calculated following Olofsson et al.'s (2013) estimators for stratified random sampling as:

(3) $$OA = \sum p_{ii} ;$$

(4) $$UA = \frac{p_{ii}}{p_{i\cdot}} ;$$

(5) $$PA = \frac{p_{jj}}{p_{\cdot j}} ,$$

calculated from the proportions of the total area for each predicted (i) and observed (j) class. Given the stratified random sampling design, and to adjust the accuracy estimates to account for the unequal sampling intensities for each class, these proportions were weighted according to the share of area for each class (A_j) with respect to the total (A_t) (Olofsson et al., 2013), as observed from the SMK's stand register dataset employed in the initial stratified random sampling (Appendix B):

(6) $$p_{ij} = \frac{A_j}{A_t}\frac{n_{ij}}{N},$$

where n_{ij} was the number of plots observed for class j and predicted to be class i, and N the total number of plots. Similarly, Cohen's (1960) kappa coefficient (κ) was also calculated from these weighted proportions p_{ij}, employing the sample estimator for stratified random sampling suggested by Stehman (1996). Routines implemented in R-packages *vcd* (Meyer et al., 2014b) and *diffeR* (Pontius & Santacruz, 2015) were employed for these tasks. Results were compared with those resulting from grouping supervised SVM predictions, which were obtained in a leave-one-out fashion (Valbuena et al., 2016b). It is worth stating that the study design complied with Westfall et al.'s (2011) recommendations for stratified estimation.

4. Results

4.1. L-coefficient of variation of ALS heights

First, we studied the relation between the Lcv of ALS heights and the forest development classes observed at field plots. From Eq. (1), the rule $Lcv = 1/2$ can be represented in the $L2 \sim L1$ relation (dashed line in Fig. 1) as:

(7) $$L2 = \frac{L1}{2}.$$

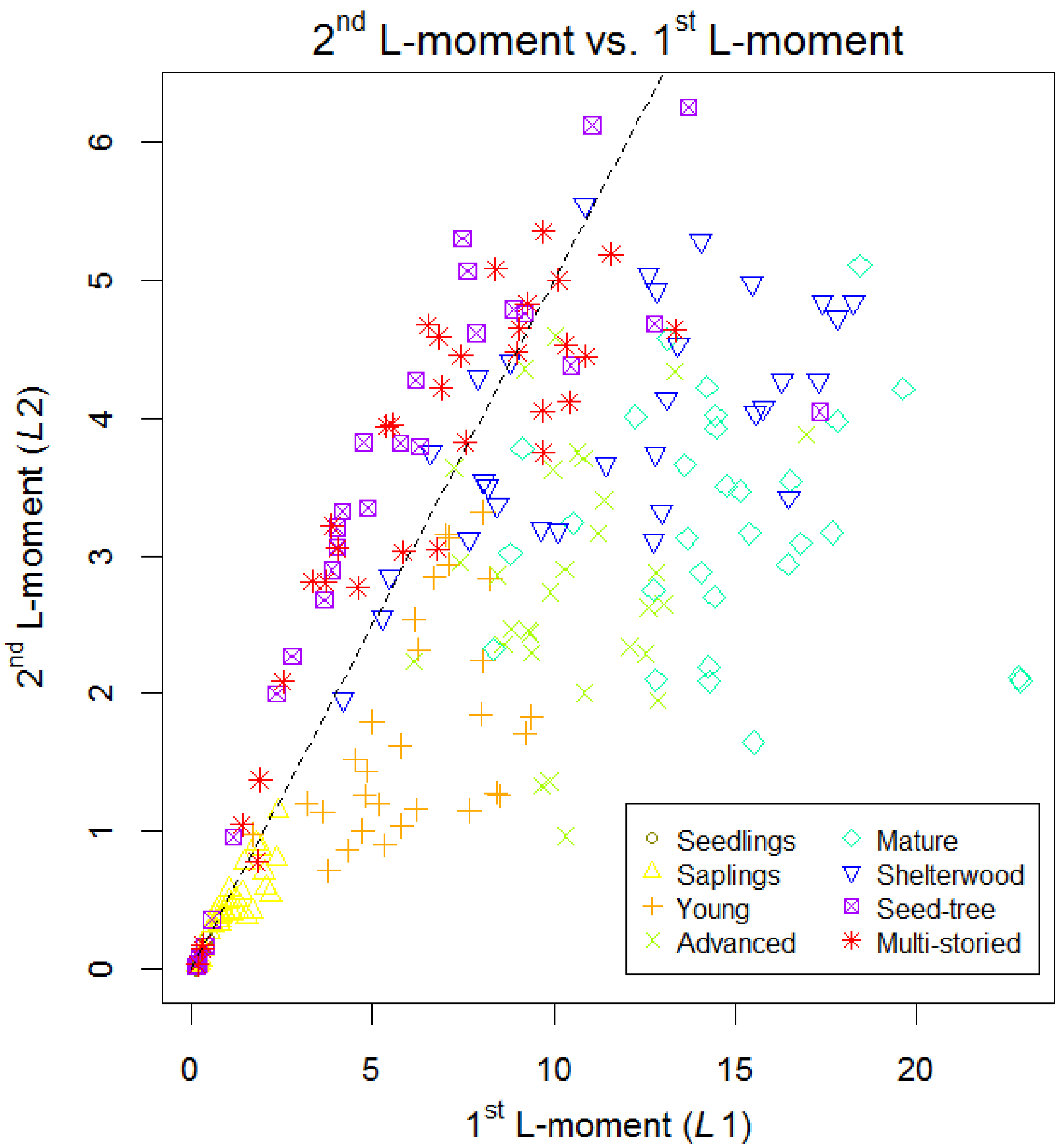

Figure 1. Relationship between the first and the second L-moments of ALS heights (i.e, L-coefficient of variation).

The $Lcv = 0.5$ threshold in Eq. (7) is depicted in Fig. 1 with a dashed line. Thus, Fig. 1 shows how the different forest development classes distribute themselves at either side of this threshold, using ALS metrics only. We observed that *Seed-tree* and *Multi-storied* stands, which usually present large values of relative dispersion in tree sizes ($GC > 0.5$), also had wide dispersion in their ALS returns being mainly above the threshold at $Lcv > 0.5$ as well. This rule, however, failed to identify forest areas with regeneration of shade-tolerant species recruited in the understorey under a closed dominant canopy. These correspond mainly to the *Shelterwood* development class, which fell largely under $Lcv < 0.5$. Fig. 1 shows that *Shelterwood* areas were difficult to discriminate from *Mature* forests, and hence they were likely to be misclassified by this rule as being even tree size forest types. Fig. 1 also shows the lack of independence of $L2$ from $L1$, since the spread of $L2$ values is larger for increasing $L1$. This demonstrates the advantage of the Lcv ratio, which normalizes the values of dispersion in $L2$, making them comparable among distributions differing in the mean ALS height ($L1$, see Eq. A3 in Appendix A).

Concerning the classification results, using the $Lcv = 0.5$ rule for discriminating even tree size (*Seedling*, *Sapling*, *Young*, *Advanced* and *Mature*) versus uneven tree size classes (*Shelterwood*, *Seed-tree* and *Multi-storied*) (Table 1), obtained an overall accuracy of 92.4% and a coefficient of agreement $\kappa = 0.48$. A total of 92.7% of the even-sized plots were correctly classified by this rule, with only few omission/commision errors. Most uncertainty was on the identification of uneven tree size forests, due to the inability for the $Lcv = 0.5$ rule to identify *Shelterwood* areas (Fig. 1), as this rule only classified 24.4% of those areas as being uneven-sized.

Table 1. Direct rule $Lcv = 0.5$. Contingency matrix of classification of even-sized versus uneven-sized development classes.

	Observed		
Predicted	even-sized	uneven-sized	Totals
even-sized	139	48	187
uneven-sized	11	46	57
Totals	150	94	244

4.2. L-skewness of ALS heights

The next step was to observe the capacity of $Lskew$ to incorporate additional information about forest structure with regards to the relationships of relative dominance among the trees. Using the rule $Lskew = 0$ in Eq. (2) gives

(8) $$L3=0.$$

Therefore the rule is demonstrated directly by the zero value on the y-axis of the $L3 \sim L2$ relation (horizontal dashed line in Fig. 2). In Fig. 2, we also observed a strong dependency of $L3$ on $L2$, since the spread of $L3$ values expands while $L2$ increases. This also illustrates the advantages of the $Lskew$ ratio, which normalizes the $L3$ values of asymmetry, making them comparable among distributions of differing dispersion of ALS heights (hence, of different mean ALS height as well).

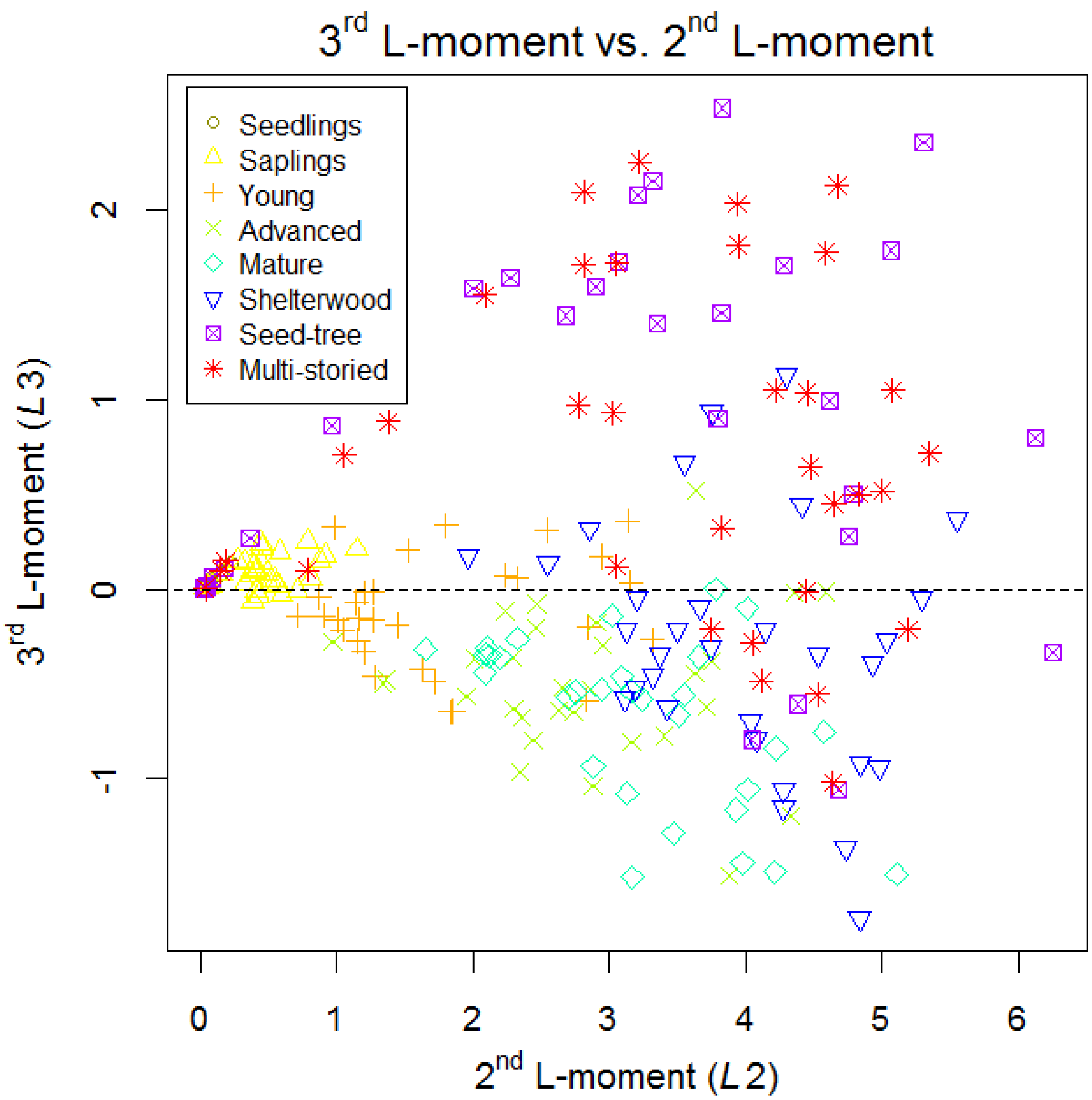

Figure 2. Relationship between the second and third L-moments of ALS heights (i.e., L-skewness).

The utility of analysing the asymmetry of the ALS height distributions was clear, as $Lskew$ was associated with the capacity of penetration of the laser pulses, and therefore with the openness of the canopy. Positive skewness ($Lskew > 0$) was observed when there were large proportions of ALS returns with relatively lower heights, which indicates few dominant trees allow the laser beam to reach lower areas underneath an open upper canopy. On the other hand, negative skewness ($Lskew < 0$) was observed when a closed dominant canopy backscatters most returns from the higher strata, and only few of them are returned from the understorey.

Regarding the discrimination of oligophotic (*Young*, *Advanced Mature* and *Shelterwood*,) and euphotic (*Seedling*, *Sapling*, *Seed-tree* and *Multi-storied*) areas of the forest (Table 2), the overall accuracy obtained was 84.6% and κ = 0.56. These accuracies were quite large, considering a method making no use of field data, an indication that $Lskew$ may be a good proxy for the degree of canopy closure.

Table 2. Direct rule $Lskew = 0$. Contingency matrix of classification of oligophotic (closed canopies) versus euphotic (open canopies) areas.

	Observed		
Predicted	oligophotic	Euphotic	Totals
oligophotic	102	17	119
euphotic	19	106	125
Totals	121	123	244

4.3. Comparing rule-based versus supervised method

Figure 3 shows a joint representation of both rules: $Lcv = 0.5$ and $Lskew = 0$, respectively represented by vertical dotted and horizontal dashed lines. It therefore illustrates how these measures of relative dispersion and asymmetry may be selected or combined in pursue of different objectives for classifying forest structure and development directly from the distribution of ALS returns. Furthermore, we also compared all results with those obtained by a supervised classification carried out with this same subsample dataset. Tables 3 and 4 are contingency matrices for the aggregation of development classes (according to section 3.2) predicted by the supervised SVM classification. For direct comparison, Table 5 includes a summary of results obtained by all the compared methods.

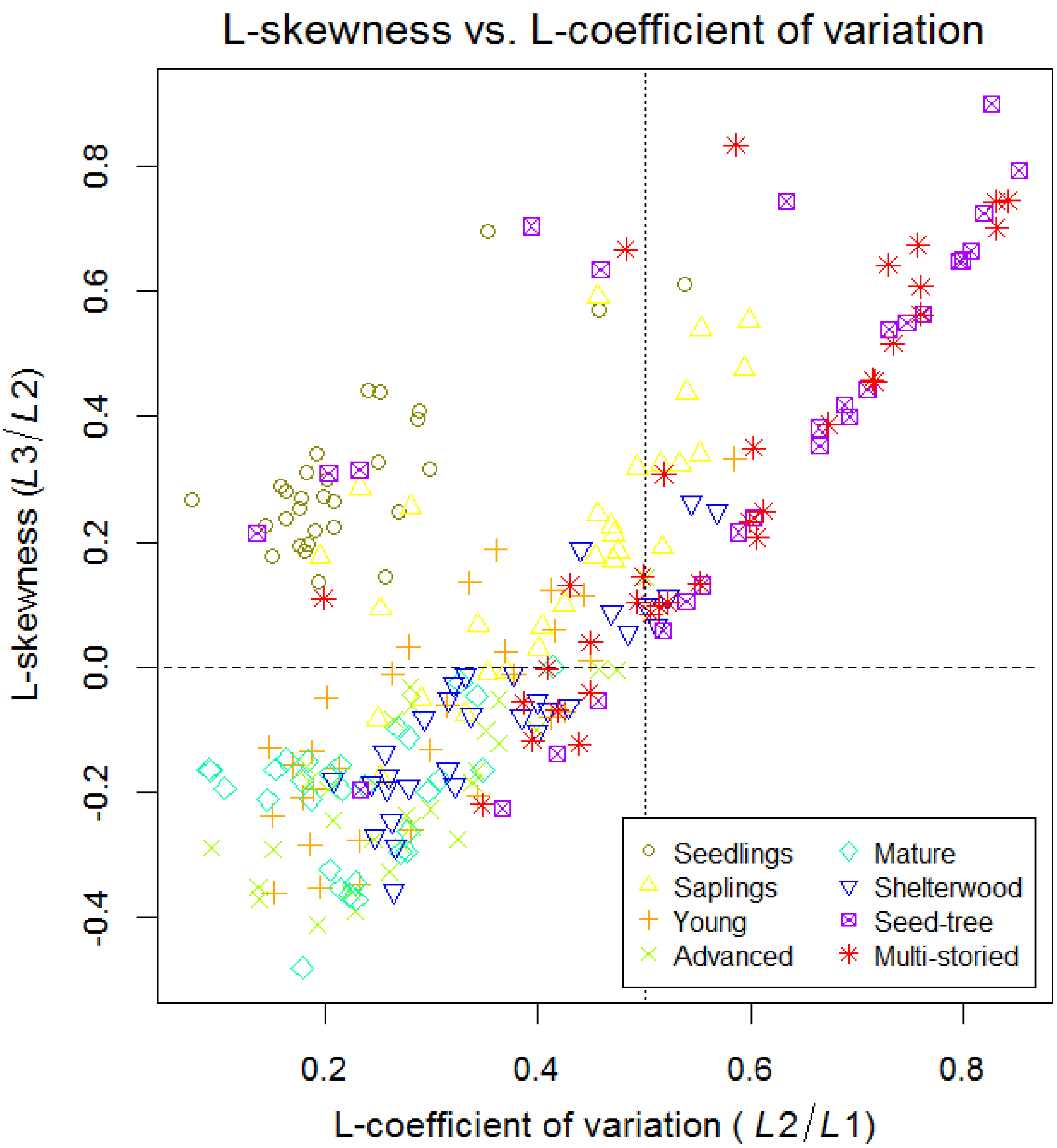

Figure 3. Relationship between the L-coefficient of variation and L-skewness of ALS heights.

Table 3. Supervised classification. Aggregated classes from Valbuena et al. (2016b). Contingency matrix of classification of even-sized versus uneven-sized development classes.

	Observed		
Predicted	even-sized	uneven-sized	Totals
even-sized	131	15	146
uneven-sized	19	79	98
Totals	150	94	244

Table 4. Supervised classification. Aggregated classes from Valbuena et al. (2016b). Contingency matrix of classification of oligophotic (closed canopies) versus euphotic (open canopies) areas.

	Observed		
Predicted	oligophotic	Euphotic	Totals
oligophotic	114	10	124
euphotic	7	113	120
Totals	121	123	244

Table 5. Comparison of accuracy results.

Stratification	**Rule-based classification**	**Supervised classification***
Even vs. Uneven Tree Size	$Lcv = 0.5$	SVM
Overall accuracy (OA)	92.4%	87.3%
kappa (κ)	0.48	0.34
Even tree size omission (PA)	92.7%	87.3%
Even tree size commission (UA)	99.6%	99.8%
Uneven tree size omission (PA)	48.9%	84.0%
Uneven tree size commission (UA)	4.2%	4.1%
Oligophotic vs. Euphotic	$Lskew = 0$	SVM
Overall accuracy (OA)	84.6%	93.8%
kappa (κ)	0.56	0.80
Oligophotic omission (PA)	84.3%	94.2%
Oligophotic commission (UA)	96.8%	98.3%
Euphotic omission (PA)	86.2%	91.9%
Euphotic commission (UA)	52.9%	76.8%

*aggregated from Valbuena et al. (2016b).

Regarding the results obtained from the supervised classification, it can be observed that the classification of forest areas into even and uneven tree sizes (Table 3) reached an overall accuracy 87.3% and κ = 0.34, whereas oligophotic versus euphotic (Table 4) obtained overall accuracy of 93.8% and κ = 0.80. Differences between the rule-based method and the supervised approach were not so large if taking into account the simplicity and lack of involvement of field data in the former one. User's accuracies obtained by the SVM classification were very similar to those yielded by the rule-based method (Table 5), which demonstrates that they are mainly due to differences in the proportions of area that each development class has from the population, and not differences between the two methods. The success of the $Lcv = 0.5$ threshold in classifying the even and uneven tree size forests and $Lskew = 0$ for segregating the oligophotic and euphotic areas of forest was remarkably good if compared to the supervised classification, which did not obtain much greater accuracies. The comparison of user's and producer's accuracies against the supervised classification however highlighted the two major differences: the rule-based method increased the errors due to omission of uneven-sized areas and commission of euphotic areas (Table 5).

5. Discussion

5.1. L-coefficient of variation may identify tree size inequality

Our prior presumption was that forests with trees of approximately equal sizes – i.e., even tree size classes –, since they would backscatter most ALS returns from a single canopy stratum, could be directly detected by low values of the Lcv of their ALS heights. Our results corroborate this presumption, since 92.7% of the even tree size plots were correctly classified by this rule (blue colour in Fig. 4 examples). Fig. 3 shows that most uncertainty in even tree size areas – those containing trees of approximately equal sizes – was due to *Sapling* stands, whereas not one single plot belonging to either *Advanced* or *Mature* development classes showed values of $Lcv > 0.5$. The low rate of omission errors implies that this rule could be used as a rather conservative and simple method when the purpose is to predict even tree size forest areas.

398

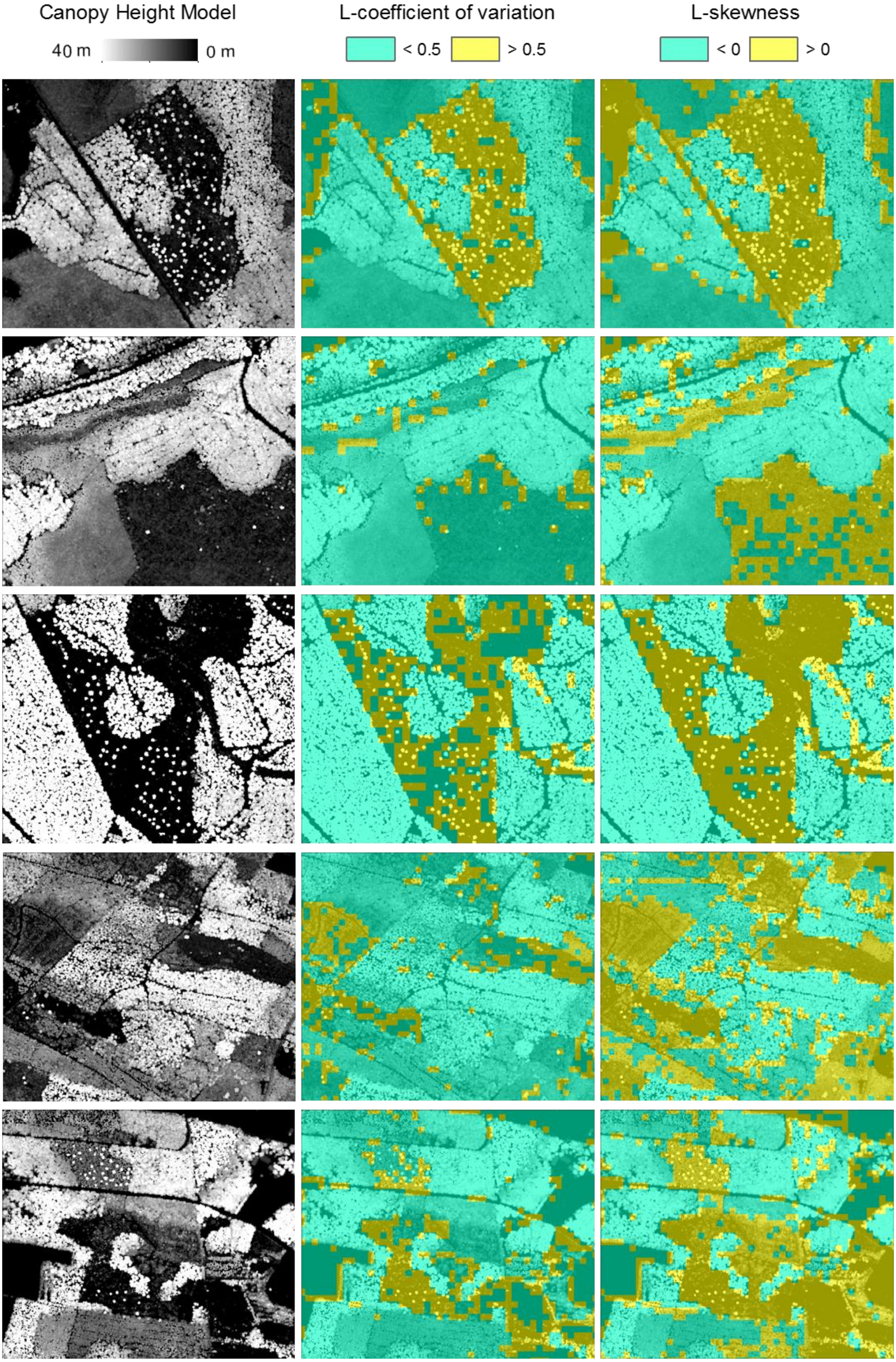
Canopy Height Model
40 m
0 m
L-coefficient of variation
< 0.5
> 0.5
L-skewness
< 0
> 0

Figure 4. Examples of resulting maps of forests stratified with rule-based method. Left: canopy height model (CHM). Middle: areas with $Lcv > 0.5$ in yellow (uneven tree sizes) and $Lcv < 0.5$ in blue (even tree sizes). Right: areas with $Lskew > 0$ in yellow (euphotic) and $Lskew < 0$ in blue (oligophotic). The reference CHM was made from the same ALS dataset, courtesy of Aki Suvanto (Blom Kartta Oy).

On the other hand, it was also expected that in the presence of structurally heterogeneous forests with more inequality of sizes among its trees, the ALS returns would also show a more spread pattern as they backscatter along the full vertical profile of the canopy, showing higher values of Lcv. In view of our results, that was the case for *Seed-tree* and most *Multi-storied* areas, although not for *Shelterwood* stands. We therefore propose that the direct rule $Lcv > 0.5$ may be used as an indicator of great tree size inequality only when regeneration is achieved by shade-intolerant species, and therefore it has been enabled by forest disturbance (Knox et al., 1989; Kellner & Asner, 2009). In other words, a correspondence between the GC of tree sizes (Valbuena et al. 2013a) and the Lcv of ALS heights may only happen when the large value of GC is due to the presence of a gap in the canopy, which allows a large proportion of the laser footprint to get through and disperse its corresponding returns along the vertical profile of the canopy (Stark et al., 2012). This highlighted the importance of employing an additional metric discriminating areas with a large euphotic zone from those where regeneration occurs in the oligophotic zone (Lefsky et al., 2002; Fig. 5). Whether or not more ALS metrics are required for fully describing the structural properties of forests, it is worth noting the recurrence of Lcv as a variable selected by many different automated methods tested in our previous studies, and therefore the role of Lcv in predicting structural attributes related to tree size inequality

(Valbuena et al., 2013b; 2014; 2016a) and forest development (Valbuena et al., 2013a; 2016b) seems clear.

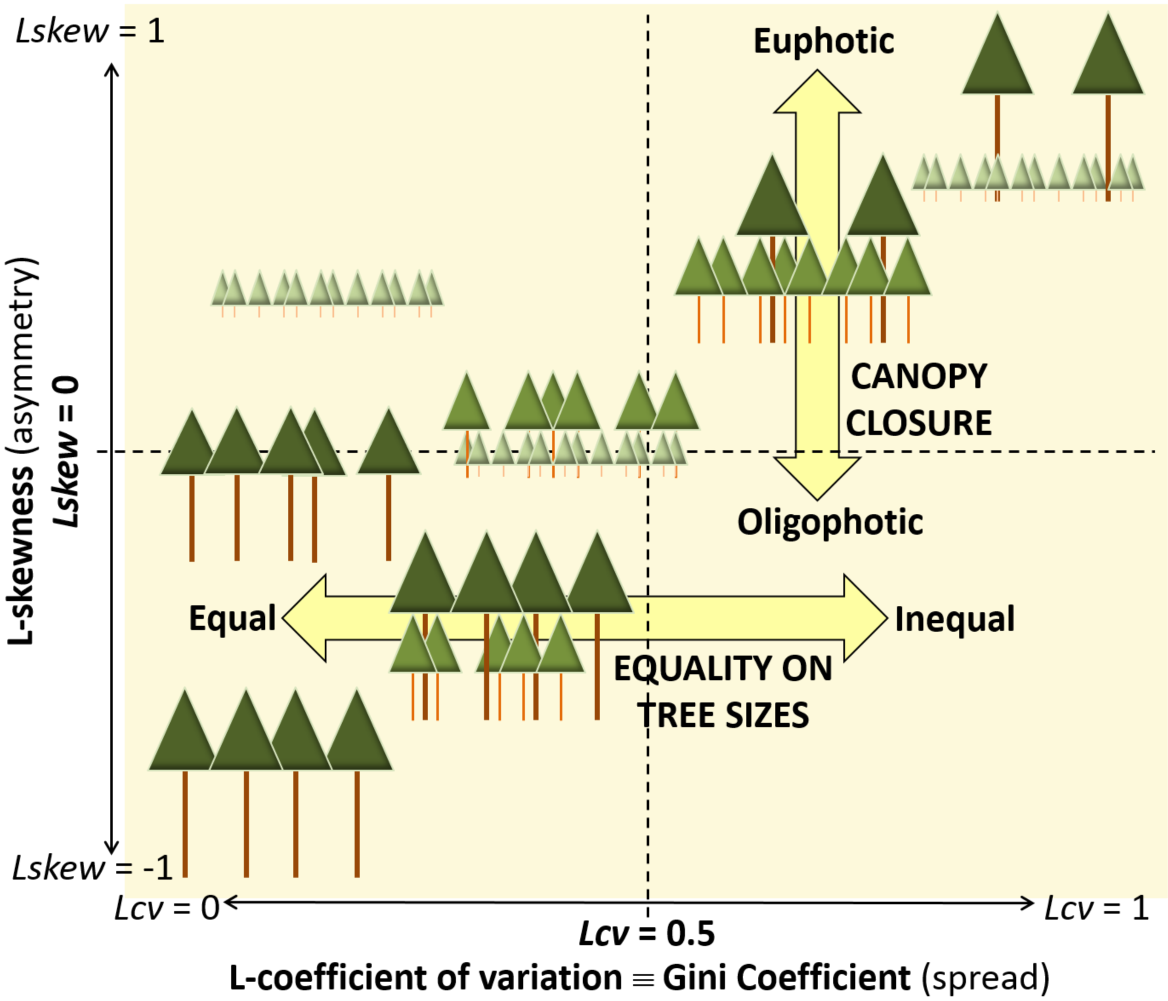

Figure 5. Schematic diagram representing the patterns of ALS return distribution that can be found in different types of forest structures, and how they are described by ratios of L-moments: L-coefficient of variation and L-skewness. Compare to Fig. 3 and Valbuena et al. (2013a: Fig. 4).

Exploring the reasons why only 24.4% of *Shelterwood* stands were classified by the $Lcv > 0.5$ rule as being uneven-sized, it could be taken into account that this development class was also the one showing most error in the SVM classification (Valbuena et al., 2016b). The fact that a supervised method, which used the explanatory potential of many other metrics as well, still failed to reliably identify *Shelterwood* areas may be an indication that the limitation is due not to the metrics but rather to the original ALS data. Due to the low-density nature of this national dataset (NLS, 2013), the laser footprint probably detects very infrequently the presence of understory under closed dominant canopies. In that case, scan density would need to be increased for this task. We considered the advantages of testing the rule-based method with this type of ALS dataset since, due to its simplicity, could have potential for replication at national scales. Further research should, however, employ datasets of larger densities to clarify whether Lcv could then show better capacity for detecting regeneration of shade-tolerant species. If direct replication of the rule-based method is to be envisaged, the effect of other flight parameters in these L-moment ratios, such as scanner device or maximum scanning angle (Næsset, 2004; Disney et al., 2010), should also be object of future investigations.

5.2. L-skewness may identify fully closed canopies

The threshold derived from the asymmetry measure of L-moments, $Lskew = 0$, was demonstrably practical with regards to discriminating oligophotic from euphotic areas. $Lskew < 0$ denotes areas where most ALS returns were backscattered from a closed dominant canopy which only allows small proportions of the laser footprint – and the light resource – to reach the understorey. Conversely, $Lskew > 0$ was observed whenever there were large proportions of ALS heights with relatively lower heights, and it was therefore related to the presence of only few returns backscattered from upper areas in the canopy, which indicates that the dominant trees allow the laser beam – and thereby the light resource – to reach lower areas underneath an open canopy. This can be relevant with regards to findings by Drake et al. (2002) and Lefsky et al. (2005), who found the degree of canopy closure to be one of the most relevant covariates in the relation between biomass and ALS heights.

It may be worth noting that the $Lskew > 0$ rule was capable for practically delineating *Seedling*, *Sampling* and *Seed-tree* stands directly (Fig. 4). Although the method was carried out at pixel-level, the resulting maps identified entire stands sharply. The rule-based stratification by $Lskew > 0$ was therefore fairly insensitive to the within-stand variation that usually makes difficult to discriminate stands, especially *Seed-tree* areas, by standard area-based procedures in remote sensing. These type of problems usually require more complex analyses at object-level – representing stands –, which involve segmentation procedures with subjective steps, parameters determined by trial-and-error, or manual delineation (e.g., Pascual et al., 2008). In contrast, the rule based method offers a simple procedure to determine *Seedling*, *Sampling* and *Seed-tree* stands directly.

5.3. Synergies between the rules

Overall accuracies obtained by the rule-based methods were, respectively, 92.4% and 84.6% which we considered a remarkable achievement for a rule-based method not requiring field support for training and that they were comparable to the results obtained by the supervised classification (87.3% and 93.8%, respectively; Table 5). As a rule of thumb, it may be affirmed that $Lskew > 0$ characterizes canopies not fully closed (areas not having reached stem exclusion), whereas those areas which also had values of $Lcv > 0.5$ presented high inequality among tree sizes driven by forest disturbance (Fig. 5). In our results in Fig. 3, values of wide dispersion $Lcv > 0.5$ occurred only in the presence of positive skewness $Lskew > 0$. This was also corroborated out of the sample, as pixels with $Lcv > 0.5$ also had $Lskew > 0$ as well (Fig. 4). This demonstrates that, in these low-density datasets, the variance of ALS heights only increases as a cause of openness in the canopy and an increase of the euphotic zone (Lefsky et al., 2002), possibly due to forest disturbance, which leads to positive skewness in the distribution. As a consequence, the maps obtained with $Lcv > 0.5$ were expanded by the $Lskew > 0$ rule (Fig. 4), extending the areas of large tree size inequality towards those simply presenting potential for growth with no limitation from light resource. In turn, negatively skewed $Lskew < 0$ ALS height distributions (Fig. 2) are indicative of forests with large oligophotic zone (Lefsky et al., 2002) and therefore can only allow the regeneration of shade-tolerant species. It is worth commenting that uneven tree size and euphotic forest areas stand out of a general relationship between first moments of ALS heights and forest attributes related to mean diameter (Lefsky et al., 2002, 2005), and therefore we suggest that one potential use of the rule-based method could to decrease the signal-to-noise ratio when obtaining ALS-assisted estimations in heterogeneous forest areas.

5.4. Practical benefits and further research needs

In this article, we applied deductive science (Appendix A) to infer that L-moments from the distribution of ALS returns can have a direct relationship to forest structural characteristics at the community level, namely tree size inequality and canopy closure (Fig. 5), in addition to the already well-known fact that ALS height relates to tree height (e.g., Lefsky et al., 2005; Maltamo et al., 2005; Miura & Jones, 2010). The main benefit of these research findings is on increasing our understanding (Fig. 5) of how ALS explains key structural features related to forest structure (Gove, 2004; Valbuena et al., 2012) and tree competition (Weiner, 1990; Cordonnier & Kunstler, 2015). These can be relevant to enhance the potential of ALS for describing light availability conditions (Lefsky et al., 2002), forest disturbance characteristics (Kellner & Asner, 2009), or tree growth (Stark et al., 2010) and regeneration (Valbuena et al., 2013a). Further research should clarify the role of different flight configurations, scanners systems or scanning density (Næsset, 2004; Disney et al., 2010) in the relationships between ALS L-moments and forest structural characteristics.

The resulting classification could be used e.g. in stratification of a forest area for the field data collection of an ALS inventory campaign, since Hawbaker et al. (2009), Maltamo et al. (2011) and Gobakken et al. (2013) demonstrated that a field sampling strategy based on *a priori* knowledge extracted from the ALS itself may be advantageous. In the presence of within-stand heterogeneity (e.g., Valbuena et al., 2013a), L-moments could be valuable for delineating microstands (van Aardt et al., 2006). There are potential applications in guiding future forest management operations directly from ALS datasets, once unveiling the relationship between GC and silvicultural alternatives (Pukkala et al., 2016) and thereby to L-moments of ALS returns. For ecosystem studies, there is potential for studying canopy structure, e.g., discrimination of single- and multi-layered forests, and other traits relevant to old-growth forests (Lefsky et al., 2002; Miura & Jones, 2010). We encourage further research to exploit the potential of L-moments in forest estimation and other applications.

6. Conclusions

We developed a rule-based classification deduced from L-moments summarizing the relative dispersion and skewness of ALS heights. Classification by two simple deductive mathematical rules, L-coefficient of variation $Lcv > 0.5$ and L-skewness $Lskew > 0$, was carried out directly on the ALS return cloud, omitting training stages making use of field plot data. Lcv was related to tree size inequality, while $Lskew$ provided information on the degree of closure of the dominant canopy. These provide relevant information about competition conditions in different areas of the forest, which can be deduced directly from ALS datasets. Our conclusions, however, may apply only to Boreal ecosystems, were light availability and its interception by the dominant canopy is the competitive process that limits forest growth. Some of the accuracies obtained were remarkably large, being a direct classification using no field data support, and they were comparable to those obtained by a supervised classification. Two flaws of the rule-based method were the omission of uneven-sized forest with shade-tolerant regeneration and commission errors for the euphotic areas, to be solved by further research perhaps making use of datasets with higher density. These rules can be executed directly over ALS datasets, providing an unambiguous procedure with multiple applications.

Acknowledgements

This research was funded by Suomen Metsäkeskus (SMK Finnish Forest centre). Special thanks to Juho Heikkilä and Jussi Lappalainen (SMK), Heli Laaksonen (NLS), and Aki Suvanto (Blom Kartta Oy) for their support at different stages of this study. Rubén Valbuena would like to thank The Finish Society of Forest Sciences for awarding a IUFRO Grant which sponsored his travel to present this work at Silvilaser 2015 Conference in La Grande Motte (France). Special issue editors are thanked for inviting this communication and three anonymous reviewers for their helpful and constructive comments.

References.

Asner, G.P., & Mascaro J. (2014). Mapping tropical forest carbon: Calibrating plot estimates to a simple LiDAR metric. *Remote Sensing of Environment, 140*, 614-624

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

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20* (1): 37–46

Cordonnier, T., & Kunstler, G. (2015). The Gini index brings asymmetric competition to light. *Perspectives in Plant Ecology, Evolution and Systematics*, *17* (2), 107-115.

Dalponte, M. Bruzzone, L., Gianelle, D. 2008. Fusion of hyperspectral and LIDAR remote sensing data for classification of complex forest areas. *IEEE Transactions in Geosciences and Remote Sensing*, *46* (5), 1416-1427.

David, H.A., & Nagaraja, H.N. (2003). *Order Statistics* (third edition). Wiley Series in Probability and Statistics. New York: John Wiley.

Disney, M.I., Kalogirou, V., Lewis, P., Prieto-Blanco, A., Hancock, S., & Pfeifer, M. (2010). Simulating the impact of discrete-return lidar system and survey characteristics over young conifer and broadleaf forests. *Remote Sensing of Environment*, *114*, 1546-1560

Drake, J.B., Dubayah, R.O., Knox, R.G., Clark, D.B., & Blair, J.B. (2002) Sensitivity of large-footprint lidar to canopy structure and biomass in a neotropical rainforest, *Remote Sensing of Environment*, *81* (2–3), 378-392.

Frazer, G.W., Wulder, M.A., & Niemann, K.O. (2005). Simulation and quantification of the fine-scale spatial pattern and heterogeneity of forest canopy structure: A lacunarity-based method designed for analysis of continuous canopy heights, *Forest Ecology and Management*, *214* (1–3), 65-90.

Frazer G.W., Magnussen S., Wulder M.A., & Niemann K.O. (2011). Simulated impact of sample plot size and co-registration error on the accuracy and uncertainty of LiDAR-derived estimates of forest stand biomass. *Remote Sensing of Environment*, 115, 636-649.

García, M., Riaño, D., Chuvieco, E., Salas, J., Danson, F.M. 2011. Multispectral and LIDAR Data Fusion for Fuel Type Mapping Using Support Vector Machine and Decision Rules. *Remote Sensing of Environment*, *115* (6), 1369-1379.

Gini, C. (1921) Measurement of inequality of incomes. *Economic Journal, 31*, 124–26.

Gobakken, T., Korhonen, L. & Næsset, E. (2013). Laser-assisted selection of field plots for an area-based forest inventory. *Silva Fennica, 47* (5), 943.

Gove, J.H. (2004) Structural stocking guides: a new look at an old friend. *Canadian Journal of Forest Research*, *34*,1044–1056.

Hall, S. A., Burke, I. C., Box, D. O., Kaufmann, M. R. & Stoker, J. M. (2005). Estimating stand structure using discrete-return lidar: An example from low density, fire prone ponderosa pine forests. *Forest Ecology and Management, 208* (1-3), 189-209.

Hawbaker, T.J., Keuler, N.S., Lesak, A.A., Gobakken, T., Contrucci, K., & Radeloff, V.C. (2009). Improved estimates of forest vegetation structure and biomass with a LiDAR-optimized sampling design. *Journal of Geophysical Research*, *114*, G00E04.

Helmert, F.R. (1876) Die Berechnung des wahrscheinlichen Beobachtungsfehlers aus den ersten Potenzen der Differenzen gleichgenauer director Beobachtungen. *Astronomische Nachrichten*, *88*, 127–132.

Hosking, J.R.M. (1989). Some theoretical results concerning L-Moments. *Research Report RC14492.* IBM research, Yorktown Heights.

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

Jaskierniak, D., Lane, P.N.J., Robinson, A., & Lucieer, A. (2011). Extracting LiDAR indices to characterise multilayered forest structure using mixture distribution functions. *Remote Sensing of Environment*, *115*, 573-585

Kellner, J.R. & Asner, G.P. (2009) Convergent structural responses of tropical forests to diverse disturbance regimes. *Ecology Letters*, *12*, 887–897

Kleiber, C. (2005) The Lorenz curve in economics and econometrics. Invited paper, Gini-Lorenz Centennial Conference, Siena (Italy).

Knox, R.G., Peet, R.K. and Christensen, N.L. 1989. Population dynamics in loblolly pine stands: Changes in skewness and size inequality. *Ecology*, *70*, 1153-1167

Lefsky, M. A., Cohen, W. B., Acker, S. A., Spies, T. A., Parker, G. G., & Harding, D. (1999a). Lidar remote sensing of biophysical properties and canopy structure of forest of Douglas-fir and western hemlock. *Remote Sensing of Environment, 70*, 339– 361.

Lefsky, M.A., Harding, D., Cohen, W.B., Parker, G., & Shugart, H.H. (1999b). Surface Lidar Remote Sensing of Basal Area and Biomass in Deciduous Forests of Eastern Maryland, USA. *Remote Sensing of Environment*, *67*, 83-98.

Lefsky, M.A., Cohen, W.B., Parker, G., & Harding, D., (2002). Lidar remote sensing for ecosystem studies. *Bioscience*, *52*, 19-30.

Lefsky, M.A., Hudak, A.T., Cohen, W.B., and Acker, S.A. (2005). Patterns of covariance between forest stand and canopy structure in the Pacific Northwest. *Remote Sensing of Environment*, *95*, 517-531

Maltamo, M., Packalen, P., Yu, X., Eerikainen, K., Hyyppa, J., & Pitkanen, J. (2005). Identifying and quantifying structural characteristics of heterogeneous boreal forests using laser scanner data. *Forest Ecology and Management, 216*, 41−50.

Maltamo, M., Bollandsås, O.M., Næsset, E., Gobakken, T., & Packalén, P. (2011). Different plot selection strategies for field training data in ALS-assisted forest inventory. *Forestry, 84*, 23-31

Miura, N. & Jones, S.D. (2010). Characterizing forest ecological structure using pulse types and heights of airborne laser scanning, *Remote Sensing of Environment, 114* (5), 1069-1076.

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A. & Leisch, F. (2014a) e1071: Misc Functions of the Department of Statistics, TU Wien. R package version 1.6-4. http://ugrad.stat.ubc.ca/R/library/e1071/html/00Index.html. Visited in Jan. 2014.

Meyer, D., Zeileis, A. and Hornik, K. (2014b) vcd: Visualizing Categorical Data. R package version 1.3-2. https://cran.r-project.org/web/packages/vcd/index.html Visited in Jan. 2014.

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

Næsset, E. (2002). Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment*, *80*, 88-99.

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* (2), 243-255.

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

Pascual, C., García-Abril, A., García-Montero, L.G., Martín-Fernández, S., & Cohen, W.B. (2008) Object-based semi-automatic approach for forest structure characterization using lidar data in heterogeneous *Pinus sylvestris* stands. *Forest Ecology and Management*, *255* (11), 3677-3685.

Pontius, R.G., & Santacruz, A. (2015). diffeR: Metrics of Difference for Comparing Pairs of Maps. R package version 0.0-4. http://CRAN.R-project.org/package=diffeR. Visited in Apr. 2016.

Pukkala, T., Laiho, O., & Lähde, E. (2016). Continuous cover management reduces wind damage. *Forest Ecology and Management, 372*, 120-127

Robbins, H.E. (1944). On the Expected Values of Two Statistics. *The Annals of Mathematical Statistics*, *15*, 321-323.

Stark, S.C., Leitold, V., Wu, J.L., Hunter, M.O., de Castilho, C.V., Costa, F.R., McMahon, S.M., Parker, G.G., Shimabukuro. M,T,, Lefsky, M.A., Keller, M., Alves, L.F., Schietti, J., Shimabukuro, Y.E., Brandão, D.O., Woodcock, T.K., Higuchi, N., de Camargo, P.B., de Oliveira, R.C., Saleska, S.R., & Chave, J. (2012). Amazon forest carbon dynamics predicted by profiles of canopy leaf area and light environment. *Ecology Letters*, *15* (12), 1406-1414.

Olofsson, P., Foody, G.M., Stehman, S.V., & Woodcock, C.E. (2013) Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation, *Remote Sensing of Environment*, *129*, 122–131.

Stehman, S.V. (1996) Estimating the Kappa coefficient and its variance under stratified random sampling. *Photogrammetric Engineering & Remote Sensing*, *62* (4), 401–402.

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

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

Valbuena, R., Maltamo, M., Martín-Fernández, S., Packalen, P., Pascual, C., & Nabuurs, G.J. (2013b). Patterns of covariance between airborne laser scanning metrics and Lorenz curve descriptors of tree size inequality. *Canadian Journal of Remote Sensing, 39*, S18–S31.

Valbuena, R., Vauhkonen, J., Packalen, P., Pitkänen, J., Maltamo, M. (2014) Comparison of airborne laser scanning methods for estimating forest structure indicators based on Lorenz curves. *ISPRS Journal of Photogrammetry & Remote Sensing, 95*, 23–33

Valbuena, R., Eerikäinen, K., Packalen, P. & Maltamo, M. (2016a). Gini coefficient predictions from airborne lidar remote sensing display the effect of management intensity on forest structure. *Ecological Indicators, 60,* 574–585.

Valbuena, R., Maltamo, M. & Packalen, P. (2016b). Classification of multi-layered forest development classes from low-density national airborne lidar datasets. *Forestry, 89*, 392–401.

van Aardt, J.A.N., R.H. Wynne, & Oderwald, R.G. (2006). Forest volume and biomass estimation using small-footprint LiDAR-distributional parameters on a per-segment basis. *Forest Science*, *52* (6), 636-649.

Wang, Q. J. (1996). Direct sample estimators of L moments. *Water Resources Research*, 32 (12), 3617−3619.

Weiner, J. (1990) Asymmetric competition in plant populations. *Trends in Ecology and Evolution, 5,* 360–364.

Westfall, J. A., Patterson, P. L., & Coulston, J. W. (2011). Post-stratified estimation: within-strata and total sample size recommendations. *Canadian Journal of Forest Research, 41* (5), 1130–1139.

White, J.C.; Wulder, M.A.; Varhola, A.; Vastaranta, M.; Coops, N.C.; Cook, B.D.; Pitt, D.; Woods, M. (2013). *A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach*. Natural Resources Canada, Information Report FI-X-010. Canadian Forest Service, Victoria, BC.

Zenner, E.K. (2005). Development of tree size distributions in Douglas-fir forests under differing disturbance regimes. *Ecological Applications*, *15*, 701-714

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

Appendix A. L-moments and their relationship to Gini Coefficient

A.1. L-moments for describing a distribution

Let an order statistic $X_{k:r}$ be the k-th smallest observation in a sample of size r of the random variable X (e.g. ALS return heights), and let $E(X_{k:r})$ be its expected value. For example, consider $E(X_{1:2})$ in the following population of size 3: $\{12,16,14\}$. There are three possible samples of size $r = 2$, with sample minima ($k = 1$): $\{12,12,14\}$. The expected value is the mean over these, i.e., $E(X_{1:2}) = 12.67$. In the analysis of this paper, the population is the unknown infinite set of all possible ALS returns over the primary calculation unit (sample plot or grid cell). The expected value is estimated using the observed sample of returns.

L-moments describe the distribution of a scalar random variable X through weighted sums of $E(X_{k:r})$. Hosking (1990) defined the L-moments as:

(A1) $$Lr = r^{-1} \sum_{k=0}^{r-1} (-1)^k \cdot \binom{r-1}{k} \cdot E(X_{r-k:r}).$$

The first L-moment ($L1$) is obtained by substituting $r = 1$ in equation (A1) to get:

(A2) $$L1 = E(X_{1:1}) = E(X),$$

which is thus equivalent to the first product-moment (expectation) of X. Hence, $L1$ is the L-measure for the location or central tendency of the distribution. If observations of X are available, $L1$ can be estimated as the arithmetic mean:

(A3) $$\widehat{L1} = \bar{X}.$$

The second L-moment ($L2$), follows the case for $r = 2$:

(A4) $$L2 = \tfrac{1}{2} E(X_{2:2}) - \tfrac{1}{2} E(X_{1:2}) = \tfrac{1}{2} E[X_{2:2} - X_{1:2}],$$

which is the expected value of half difference between minimum ($X_{1:2}$) and maximum ($X_{2:2}$) in a sample of size two. It therefore provides the mean of half differences, and thus it is the L-measure for the dispersion of the distribution.

Following a similar logic for the third L-moment ($L3$), substituting $r = 3$ in (A1) yields:

(A5) $$L3 = \tfrac{1}{3} E(X_{3:3}) - \tfrac{2}{3} E(X_{2:3}) + \tfrac{1}{3} E(X_{1:3}),$$

which is a weighted sum of minimum, ($X_{1:3}$), median ($X_{2:3}$), and maximum ($X_{3:3}$) of a sample with size three. It can further be written as:

(A6) $$L3 = \tfrac{1}{3} E[(X_{3:3} - X_{2:3}) - (X_{2:3} - X_{1:3})],$$

to show that $L3$ expresses the expected difference between the maximum-median and median-minimum differences in a sample of size three, which provides a L-measure for the asymmetry of the distribution of X. Hence, $L3 = 0$ corresponds to a symmetric distribution, $L3 > 0$ describes positive asymmetry (left-skewed distribution) and $L3 < 0$ describes negative asymmetry (right-skewed distribution).

A.2. L-moment ratios

Hosking (1990) also defined the ratios for L-moments. They have the advantage of being bounded by finite intervals (Hosking 1989), yielding comparable relative descriptions for the distribution of X.

The second L-moment ratio is obtained as the ratio of the second to the first L-moments. It is called the L-coefficient of variation (Lcv) for its comparison to conventional moments. From equations (A2) and (A4) it can be observed that Lcv equals:

(A7) $$Lcv = \frac{L2}{L1} = \frac{E(X_{2:2}) - E(X_{1:2})}{2E(X)}.$$

For positive random variables, the values for the second L-moment ratio are bounded by the [0, 1] range (Hosking, 1989). Just like the coefficient of variation of conventional moments, Lcv is a descriptor of dispersion relative to central tendency; that is to say, concentration. This brings the advantage that concentration measures are comparable among distributions differing in their location or central tendency ($L1$), and also independently of the units of measure. It is worthwhile to note that Hosking never defined a second L-moment ratio, as their generalized definition stands only for $r = 3, 4$ ... (Hosking 1990: 108), and the L-coefficient of variation was simply presented alongside. It was only later that many authors have regarded Lcv to be the second L-moment ratio.

The third L-moment ratio is obtained by division between the third and the second L-moments. It is called the L-skewness ($Lskew$), as it has been found to be a robust descriptor for the asymmetry of the distribution of X. From equations (A4) and (A6), and using the equivalence $E(X_{3:3} - X_{1:3}) = \frac{3}{2} E(X_{2:2} - X_{1:2})$ (Robbins, 1944: Eq. 22; David & Nagaraja, 2003: 44, 56) it yields:

(A8) $$Lskew = \frac{L3}{L2} = \frac{E(X_{3:3}) - 2E(X_{2:3}) + E(X_{1:3})}{E(X_{3:3}) - E(X_{1:3})}.$$

As explained for $L3$, $Lskew = 0$ corresponds to a symmetric distribution, while positive or negative values denote the type of asymmetry for the distribution. Additionally, $Lskew$ has the advantage of presenting theoretical bounds within the [-1, 1] interval (Hosking 1989). Consequently, $Lskew$ is a descriptor of asymmetry relative to dispersion, and therefore independent of the units of measure and the dispersion of the distribution of X.

A.3. Equivalence between the Gini coefficient and the L-coefficient of variation

The Gini coefficient of a scalar random variable X (GC) is the ratio of the area comprised between the Lorenz curve and the diagonal line of equality (Gini, 1921):

(A9) $$GC = 1 - 2\int_0^1 L(X)dX,$$

Where $L(X)$ is the Lorenz curve: the relative cumulative distribution of a variable against the cumulative frequency distribution of the proportion of individuals in the population. From Eq. (A9), Kleiber (2005: Eq. 6) showed that:

(A10) $$GC = 1 - \frac{E(X_{1:2})}{E(X)}.$$

On the other hand, the Lcv gives also the GC. From Eq. (A7) it derives:

(A11a) $$Lcv = \frac{E(X_{2:2}) - E(X_{1:2})}{2E(X)}$$

(A11b) $$= \frac{E(X_{2:2} - X_{1:2}) + 2E(X_{1:2}) - 2E(X_{1:2})}{2E(X)}$$

(A11c) $$= \frac{E(X_{2:2} + X_{1:2}) - 2E(X_{1:2})}{2E(X)}$$

(A11d) $$= \frac{2E(\mathrm{X}) - 2E(X_{1:2})}{2E(X)}$$

(A11e) $$= 1 - \frac{E(X_{1:2})}{E(X)}$$

Equation (A11d) results from (A11c) because $X_{1:2} + X_{2:2}$ is the sum of two independent and identically distributed samples, and it is therefore equivalent to $X_1 + X_2$. Consequently, (A10) and (A11e) demonstrate:

(A12) $$GC = Lcv$$

The result in Eq. (A12) is essentially a special case of a 140-years-old result (Helmert, 1876; as cited in David and Nagaraja, 2003: 249) presented in equation 9.4.2 of David and Nagaraja (2003), which might even provide interesting extensions using expectations of order statistics in sample sizes larger than $r = 1, 2, 3$.

Appendix B. Criteria for determining forest development classes

Silvicultural development classes are used in Finland to classify forest stands and assist in decision-making for forest management planning. It was possible to apply stratified sampling using the stand register dataset employed by the Finnish Forest Centre (SMK, Suomen Metsäkeskus) for their operational management planning, since a development class has been explicitly assigned to each stand from previous inventories. The development class to which each sample plot belonged to was nevertheless ultimately corroborated in the field, being the criteria used *in-situ* prevalent over the stand register data. Minor differences in per-stratum sample sizes were simply caused by such type of discrepancies found in few plots. The criteria that segregated forest areas into different forest classes were:

- *Seedling*: stands with average tree height lower than 1.3 m, and absence of mature trees (overstorey).
- *Sapling*: stands with average tree height greater than 1.3 m, and average diameter at breast height (DBH) smaller than 8 cm, and absence of mature trees (overstorey).
- *Young*: stands with average DBH ranging 8-16 cm and average tree height ranging 7-9 m high.
- *Advanced*: stands with average DBH greater than 16 cm.
- *Mature*: stands reaching a quadratic mean DBH (QMD) greater than 18 cm.
- *Shelterwood*: stands including a dense overstorey of mature trees (DBH > 16 cm) which reaches at least 100-300 stems·ha^{-1}, and also a dense understorey of seedlings (height < 1.3 m) of shade-tolerant species, usually Norway spruce (1500-1800 stems·ha^{-1}).
- *Seed-tree*: stands including a sparse overstorey of mature trees (DBH > 16 cm) of only 50-100 stems·ha^{-1}, and also a dense understorey of seedlings (height < 1.3 m) of shade-intolerant species, usually Scots pine (1500-2200 stems·ha^{-1}) or Birch species (1100-1600 stems·ha^{-1}).
- *Multi-storied*: stands including a dense understorey (above-mentioned densities) of seedlings (height < 1.3 m) and saplings (height > 1.3 m, DBH < 8 cm) of any species, usually deciduous but also Scots pine or Norway spruce. The size of trees in the overstorey is not a determinant criterion, but trees in the understory must reach their sapling stage.