

Matching remotely sensed and field-measured tree size distributions

Jari Vauhkonen and Lauri Mehtätalo

Abstract: Undetected trees and inaccuracies in the predicted allometric relationships of tree stem attributes seriously constrain single-tree remote sensing of seminatural forests. A new approach to compensate for these error sources was developed by applying a histogram matching technique to map the transformation between the cumulative distribution functions of crown radii extracted from airborne laser scanning (ALS) data and field-measured stem diameters (dbh, outside bark measured at 1.3 m aboveground). The ALS-based crown data were corrected for the censoring effect caused by overlapping tree crowns, assuming that the forest is an outcome of a homogeneous, marked Poisson process with independent marks of the crown radii. The transformation between the cumulative distribution functions was described by a polynomial regression model. The approach was tested for the prediction of plot-level stem number (N), quadratic mean diameter (DQM), and basal area (G) in a managed boreal forest. Of the 40 plots studied, a total of 18 plots met the assumptions of the Poisson process and independent marks. In these plots, the predicted N , DQM, and G had best-case root mean squared errors of 299 stems·ha⁻¹ (27.6%), 2.1 cm (13.1%), and 2.9 m²·ha⁻¹ (13.0%), respectively, and the null hypothesis that the mean difference between the measured and predicted values was 0 was not rejected ($p > 0.05$). Considerably less accurate results were obtained for the plots that did not meet the assumptions. However, the goodness-of-fit of the predicted diameter distribution was especially improved compared with the single-tree remote sensing prediction, and observations realistically obtainable with ALS data showed potential to further localize the predictions. Remarkably, predictions of N showing no evidence against zero bias were derived solely based on the ALS data for the plots meeting the assumptions made, and limited training data are proposed to be adequate for predicting the stem diameter distribution, DQM, and G . Although this study was based on ALS data, we discuss the possibility of using other remotely sensed data as well. Taken together with the low requirements for field reference data, the presented approach provides interesting practical possibilities that are not typically proposed in the forest remote sensing literature.

Key words: forest inventory, airborne laser scanning, light detection and ranging (LiDAR), Boolean model, marked point pattern, histogram matching, mixed-effects modeling, diameter distribution, forest structure.

Résumé : Les arbres non détectés et les imprécisions dans la prédiction des relations allométriques d'attributs de la tige des arbres restreignent l'usage de la télédétection à l'arbre près (TAP) dans les forêts seminaturelles. Une nouvelle approche de compensation de ces sources d'erreur a été développée en appliquant une technique d'appariement d'histogrammes. La démarche avait pour but de cartographier la transformation entre la fonction cumulative de distribution (FCD) des rayons des cimes, extraits par balayage laser aéroporté (BLA), et celle du diamètre des tiges mesuré sur le terrain (diamètre sur écorce mesuré à 1,3 m du sol). Les mesures de cime obtenues par BLA ont été corrigées pour tenir compte de l'effet de troncature causé par la superposition de la cime des arbres, en supposant que la forêt est le résultat d'un processus ponctuel marqué de Poisson homogène et que les rayons des cimes correspondent à des marques indépendantes. La transformation entre les FCD a été décrite au moyen d'un modèle de régression polynomiale. L'approche a été testée pour la prédiction du nombre de tiges par placette (N), du diamètre quadratique moyen (DQM) et de la surface terrière (G) dans une forêt boréale aménagée. Sur les 40 placettes étudiées, 18 au total respectaient les hypothèses du processus de Poisson et des marques indépendantes. Dans ces placettes, les valeurs prédites de N , de DQM et de G avaient des erreurs quadratiques moyennes optimales de respectivement 299 tiges·ha⁻¹ (27,6 %), 2,1 cm (13,1 %) et 2,9 m²·ha⁻¹ (13,0 %) et l'hypothèse nulle impliquant que la différence moyenne entre les valeurs mesurées et prédites est égale à 0 n'a pas été rejetée ($p > 0,05$). Des résultats considérablement moins précis ont été obtenus pour les placettes qui ne respectaient pas les hypothèses. Cependant, la qualité de l'ajustement de la distribution diamétrale prédite a été particulièrement améliorée par rapport à la prédiction de la TAP. Les observations réalistes obtenues avec les données de BLA ont démontré qu'il est ainsi possible d'obtenir des prédictions à l'échelle locale. Il est remarquable que des prédictions de N , pour lesquelles aucun biais ne pouvait être démontré, aient été obtenues en se basant uniquement sur les données de BLA des placettes respectant les hypothèses retenues. Ainsi, nous croyons qu'une quantité limitée de données d'étalonnage serait suffisante pour prédire la distribution diamétrale des tiges, le DQM et la G . Bien que l'étude soit fondée sur des données de BLA, nous discutons de la possibilité d'utiliser également d'autres données de télédétection. Tout en exigeant peu de données terrain de référence, l'approche présentée ici offre des possibilités pratiques intéressantes qui ne sont généralement pas proposées dans la littérature traitant de télédétection des forêts. [Traduit par la Rédaction]

Mots-clés : inventaire forestier, balayage laser aéroporté, détection et télémétrie par ondes lumineuses (LiDAR), modèle booléen, motif ponctuel marqué, appariement d'histogrammes, modélisation avec effets mixtes, distribution diamétrale, structure de la forêt.

Received 17 June 2014. Accepted 19 November 2014.

J. Vauhkonen. University of Eastern Finland, School of Forest Sciences, P.O. Box 111, FI-80101 Joensuu, Finland; University of Helsinki, Department of Forest Sciences, P.O. Box 27, FI-00014 Helsinki, Finland.

L. Mehtätalo. University of Eastern Finland, School of Computing, P.O. Box 111, FI-80101 Joensuu, Finland.

Corresponding author: Jari Vauhkonen (e-mail: jari.vauhkonen@uef.fi).

1. Introduction

The diameter (dbh, outside bark measured at 1.3 m aboveground) distribution of trees within a sample plot is a highly important attribute, characterizing both the economic and ecological values of a forest. In field inventories, a diameter distribution may be obtained by measuring diameters of trees within a certain area or by estimating the parameters of a theoretical distribution model using plot- or stand-level variables as predictors. In contrast, remote sensing provides distributions of auxiliary variables from which stem diameters must be predicted (e.g., Maltamo and Gobakken 2014). In this respect, detecting and delineating single trees from remotely sensed (RS) data is a particularly interesting alternative, because the observations of individual tree heights or crown dimensions readily describe tree size and, in simplified terms, correspond to allometric transformations of the stem diameters (see, for example, Korpela 2007).

Forest inventories based on detecting single trees from RS products such as aerial and satellite images and airborne laser scanning (ALS) data have indeed been a popular research topic in the 2000s (Hyypä et al. 2008; Ke and Quackenbush 2011). The most common tree detection methods (Hyypä et al. 2008) and the later processing steps (cf. Korpela and Tokola 2006; Breidenbach and Astrup 2014) are fundamentally similar for different RS data sources. In the following text, we use “single-tree remote sensing” (STRS; Korpela and Tokola 2006) as an umbrella term for the full sequence of steps, including tree detection, feature extraction, and estimation of tree attributes. By “RS crown radii”, we refer to observations of crown radii derived from RS data by applying a tree detection and delineation method such as the one described in section 3.1. The RS crown radii are thus subject to the parameterization of the method but are likely affected more by the prevailing forest structure than by an algorithm or data source (Larsen et al. 2011; Vauhkonen et al. 2012).

Multiple studies have focused on the individual steps of the STRS, but only a few have reported plot- or stand-level estimates obtained as a result of the full detection and estimation procedure (Korpela et al. 2007; Peuhkurinen et al. 2007, 2011; Vauhkonen et al. 2014). The accuracy of the aggregated estimates has been found to depend on multiple factors, including both tree detection and tree-level estimation errors. Undetected trees and errors in allometric model predictions degrade the accuracies of plot- and stand-level estimates and seriously constrain STRS of seminatural forests (Korpela et al. 2007). As a result, area-based estimation approaches are preferred in practical forest inventories (e.g., McRoberts et al. 2010; Maltamo et al. 2011).

Despite the limited ability of delineated tree crowns to depict the entire tree stock, crown segments produce information on physical tree dimensions (e.g., crown radii) that is not provided by plot- or stand-level analyses. Therefore, it is intuitive to develop approaches that combine the detected trees with an attempt to minimize the systematic errors in tree detection and estimation. Several approaches to compensate for the undetected trees have been developed, typically relying on either theoretical (Maltamo et al. 2004; Mehtätalo 2006) or empirical (Flewelling 2008; Lindberg et al. 2010; Breidenbach et al. 2010) models to predict the actual number of trees within the segmented tree crowns. The empirical approaches, in particular, require field-measured tree coordinates to link the trees with the segments. However, accurate tree locations are rarely determined in practical forest inventories (e.g., Maltamo et al. 2011), and the practical applications of the proposed approaches may thus be currently limited.

Assuming certain conditions that prevent the trees from being detected, it may be possible to derive a detection probability, i.e., the detectability of a tree. Although the use of detection probabilities to improve estimates of wildlife populations is common, it has been rare in forest-inventory literature. Mehtätalo (2006) proposed a model-based approach to estimate the detectability of a tree accord-

ing to its relative size in the stand. The approach assumes that the tree locations are generated by the spatial marked Poisson point process, with density λ (trees·m⁻²) and crown radii as independent random marks. This model is called the Boolean model in the literature of stochastic geometry (Stoyan et al. 1995). It is then assumed that a tree is detectable if the crown center does not fall within a larger tree crown. The approach provides an estimator for stand density based on the observed canopy closure and the estimated mean crown area of the stand. When tested on data simulated to meet all of the stated assumptions, the model-based approach successfully compensated for the undetected trees and provided accurate estimates of stand density. However, the assumptions made are unrealistic for all seminatural forest types, and the sensitivity of failing the assumptions and the resulting effects on the obtained result are unknown.

Once the undetected trees are compensated for, the fundamental problem with STRS is the transformation from the RS crown-size distribution to the stem-diameter distribution. Due to error propagation resulting from the steps of conventional STRS (e.g., Korpela et al. 2007) and complications to the overall inventory system due to the requirement for mapped trees for model fitting or calibration, it could be beneficial if the entire transformation was carried out at the aggregate (e.g., plot) level, omitting the need to match RS trees with field-measured ones. Histogram matching, or more universally distribution matching (DM), is a well-known technique established in the field of digital-image processing for this purpose (Gonzalez and Woods 2008). Earlier forest remote-sensing applications of DM are especially related to the calibration of various data sources such as satellite data (e.g., Olsson 1993; Gilichinsky et al. 2011) or radar data (Pantze et al. 2014), multitemporal ALS data (Nyström et al. 2013), multiplatform laser scanning data (Hopkinson et al. 2013), and predicted forest attribute distributions (Baffetta et al. 2012; Gilichinsky et al. 2012; Xu et al. 2014). It is proposed here that by predicting an accurate transformation function for the considered forest structure type (section 3.2; Fig. 1), the DM technique could be used to map the transformation from the RS distributions of crown radii to those of stem diameters. Due to inherently operating with distributions, the approach could simplify conventional STRS (cf. Korpela et al. 2007) and potentially improve its operational premises, particularly with respect to data collection.

The purpose of this study was to test the DM technique for modeling the transformation from the RS tree crown radii to the field-measured stem diameter. The Boolean model was used as a starting point to correct the observed distribution of crown radii for the censoring effect caused by overlapping tree crowns and to estimate the stand density. The results were evaluated against the assumptions of the Boolean model. The proposed approach was compared with a conventional STRS prediction. The simulations were carried out using ALS data; however, the reader is advised to also pay attention to the applicability of the technique to tree size distributions produced by other remote sensing materials suitable for STRS.

2. Material

The data were based on a typical, managed boreal forest in eastern Finland (62°31'N, 30°10'E). The ALS data for the area were collected on 26 June 2009 using an Optech ALTM Gemini laser scanning system from approximately 720 m above ground level with a field of view of 26°. Pulse repetition frequency was set to 125 kHz, and when the instrument was operated in a multipulse mode, the nominal sampling density was 11.9·m⁻².

Field measurements of a total of 79 field plots were carried out in May–June 2010 (for further details, see Packalén et al. (2013) and Vauhkonen et al. (2014)). A subset of these data, 40 plots with >95% of basal area consisting of scots pine (*Pinus sylvestris* L.), was considered here to focus the study on one tree species and simple stand structure, similar to Maltamo et al. (2012). In addition, the plot size was standardized to 400 m², rather than varying from 400 to 900 m² as in

Fig. 1. The principle of the DM technique applied (schematic). The matching of two cumulative distribution functions (top figure) produces a transformation function (middle), the variation of which among all the plots studied is shown (bottom). DBH, diameter at breast height.

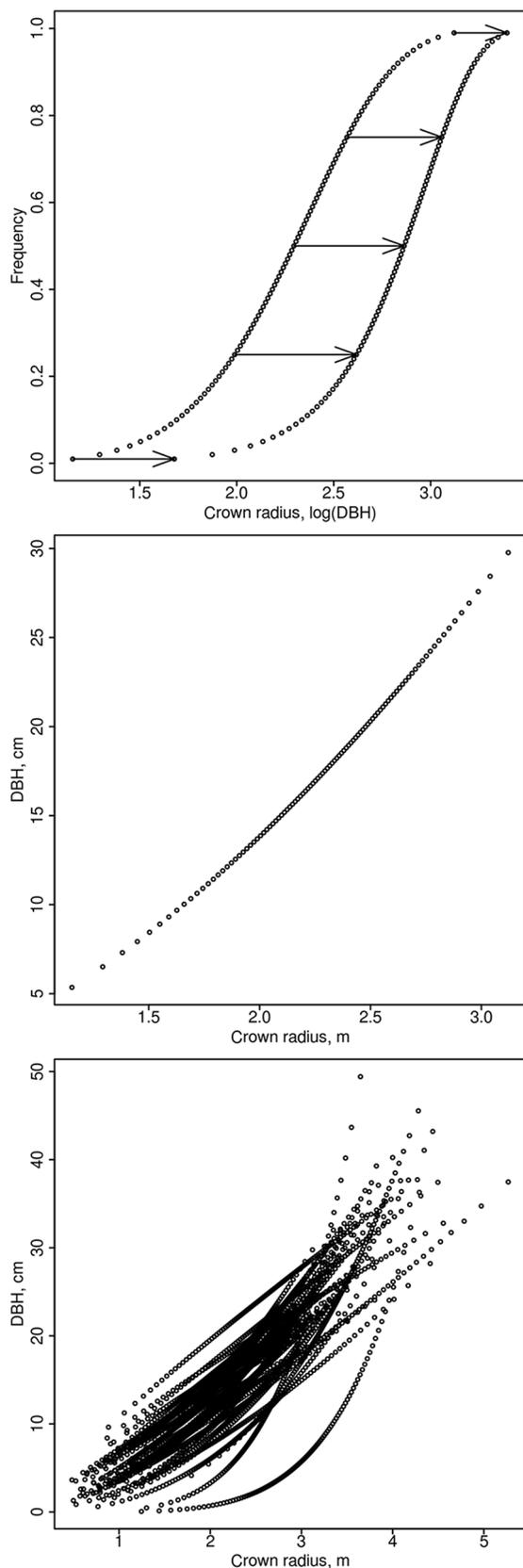


Table 1. Mean, standard deviation (SD), minimum, and maximum of the forest attributes in the plots studied are presented.

Attribute	Data	Mean	SD	Minimum	Maximum
N, stems·ha ⁻¹	All	1224	564	400	2250
	OK	1085	486	400	2025
DQM, cm	All	15.3	3.8	8.7	24.5
	OK	15.8	3.7	10.1	22.2
G, m ² ·ha ⁻¹	All	23.4	5.9	12.2	35.1
	OK	22.1	5.1	12.2	32.1

Note: Column “Data” refers to the studied data set, being either all plots ($n = 40$) or a subset of OK plots ($n = 18$) as defined in section 3.5. N, stem number; DQM, quadratic mean diameter; G, basal area.

Maltamo et al. (2012). Standardization was achieved by creating rectangular windows of 20 m × 20 m, with the center and orientation corresponding to the original plot. The trees located within this window based on the measured tree coordinates were extracted for the analyses. Table 1 presents the central characteristics of the field data.

3. Methods

DM is a technique to map the transformation between cumulative distribution functions (CDFs) $F_1(x_1)$ and $F_2(x_2)$ of variables x_1 and x_2 by looking for a monotonic transformation $g(x_1)$ that produces as accurate a match, $F_1(g(x_1)) = F_2(x_2)$, as possible. In our case, $F_1(x_1)$ and $F_2(x_2)$ were the CDFs of the RS crown radii and field-measured stem diameters, respectively. STRS produces information on the crowns of detected trees only. Therefore, the observed distribution of crown radii is censored from the left tail, i.e., small trees are underrepresented. In addition, scaling the observed distribution to a per-plot basis requires information on stand density, which is underestimated when based only on the detected trees. A satisfactory solution to this problem is needed before the DM can be conducted. In the following text, section 3.1 presents our algorithm for producing the crown radii data based on ALS data and correcting the observed data for the censoring effect, section 3.2 presents the subsequent DM, and section 3.3 describes an additional analysis in which the matching function of the DM was augmented with a plot-specific predictor to account for the varying forest structure.

3.1. Generating crown-radius distributions

The observations of crown radii were based on tree crown segments extracted from a canopy height model (CHM) raster with a resolution of 0.5 m. The CHM was first low-pass filtered using Gaussian kernels, with the size of the smoothing window increasing as a stepwise function of the heights of the CHM (Pitkänen et al. 2004). The segments were created around the local maxima using watershed segmentation with a drainage-direction-following algorithm (Pitkänen 2005). Pixels lower than 2 m were masked out from the crown segments, and small segments, at most three pixels in size, were combined with one of the neighbor segments based on the smallest average gradient on the segment boundary between two segments. The method and the parameters applied are described in more detail by Packalén et al. (2013).

The resulting crown segments are censored by the (unknown) size-dependent detectability, $p(r)$, which expresses the probability of detecting a tree of certain crown radius, r , from the air. The detectability can also be interpreted as the inclusion probability of the tree in the sample (see chapter 16 in Thompson (2012)). We start by assuming that the tree-crown radii on the plot (including the trees that are not detectable) are a sample from a continuous distribution, with a CDF of the following Weibull form:

$$(1) \quad F(r) = 1 - \exp\left[-\left(\frac{r}{\beta}\right)^\alpha\right]$$

where α and β are unknown parameters.

We assume that the tree locations are generated by the marked Poisson point process, where the marks are circular crown segments with random radii centered at the tree location. Under this Boolean model, we further assume that each individual tree is detectable in the ALS data if the tree location does not fall within the crown of a larger tree. Under these assumptions, the stand density has the following relationship with the canopy closure (cc) and expected squared crown radius $E(r^2)$ (Mehtätalo 2006):

$$(2) \quad \lambda = -\frac{\ln(1 - \text{cc})}{\pi E(r^2)}$$

Furthermore, the detectability of a tree with crown radius r is given by

$$(3) \quad p(r; \lambda) = (1 - \text{cc})^{-\lambda} \frac{1}{E(r^2)} \int_0^r t^2 f(t) dt$$

where λ is the stand density, $f(r) = \frac{d}{dr} F(r)$ is the probability density function (PDF) of crown radius, and $E(r^2) = \int_0^\infty r^2 f(r) dr$ is the mean squared crown radius. The PDF of crown radii for the detectable trees is of the form of weighted distribution:

$$(4) \quad f_D(r) = \frac{p(r; \lambda) f(r)}{E(p(r; \lambda))}$$

where $E(p(r; \lambda)) = \int_0^\infty p(r; \lambda) f(r) dr$ is the normalizing constant of the weighted PDF.

To apply these equations, we first estimated the canopy closure as the area of the segmented CHM within the plot in proportion to the total plot area. This allowed λ to be eliminated from $f_D(r)$ (eq. 4) by using eq. 2 and the observed canopy closure, so that the PDF of detectable trees could be parameterized using the Weibull parameters α and β only. These parameters were estimated by fitting the PDF of the detectable trees to the detected trees. We used the standard method of maximum likelihood. Using the resulting parameter estimates in the Weibull distribution (eq. 1) provided the estimated distribution of crown radii for all trees, which was uncensored for the effect of detectability. This distribution further allowed computing the expected value of crown radii to estimate the stand density (eq. 2), which was used in scaling the estimated distribution to the plot or per-hectare level.

3.2. Distribution matching

After estimating the distribution of crown radii, a crucial question is how to predict the diameter distribution based on it. A traditional solution is to match the detected tree crowns with the field-measured stem diameters based on locations and to fit a regression model based on these data. This approach requires that tree locations are determined in the field, which is highly labor intensive. DM adopts another approach by matching the corresponding percentiles of the estimated distribution of crown radii (eq. 4) with the field-measured distribution of stem diameters at the plot level. This kind of matching is justified if the size order of trees is (at least approximately) the same with respect to both stem diameter and crown diameter, i.e., the relationship of crown radius to stem diameter is monotonic. From a practical point of view, an important benefit to this matching is that field-measured tree locations are not needed.

The transformation between the tree size distributions was modeled by means of polynomial regression analysis. A second-order polynomial was found to adequately describe the transformation without introducing an extra number of parameters to be estimated. To consider the between-plot differences of the transformations (Fig. 1), a mixed-effects model was used (Lindstrom and

Bates 1990; Pinheiro and Bates 2000) in which the fixed-effects part of the model was augmented with random parameters that were estimated for each plot. To produce smooth transformation functions as in Fig. 1, the CDF of the stem diameters was modeled by the Weibull function similar to the crown radii (section 3.1). Let d_{si} be the i th percentile ($i = 1, 2, \dots, 99$) of the Weibull distribution of diameters of plot s and r_{si} be the corresponding percentile of the Weibull distribution for crown radii (eq. 4), which has been corrected for the censoring effect as described in section 3.1. The model to transform crown radii to stem diameter was specified as

$$(5) \quad d_{si} = \beta_1 r_{si} + \beta_2 r_{si}^2 + z_{1s} r_{si} + z_{2s} r_{si}^2 + \varepsilon_{si}$$

where β_1 and β_2 are fixed parameters, z_{1s} and z_{2s} are random plot effects, which allow the transformation to be plot specific, and ε_{si} is the residual error for plot s and percentile i . The second-order polynomial of eq. 5 does not mathematically restrict the transformation to be monotonic; however, this did not lead to any problems in the empirical part of this study. This model was tested in the DM in two ways: (i) as a population-averaged prediction, where the random effects were given their expected values $z_{1s} = 0$ and $z_{2s} = 0$, and (ii) as a plot-specific prediction, where the predicted plot-specific random effects were utilized.

3.3. Augmenting the distribution matching with a plot-specific predictor

Obtaining plot-specific predictions with eq. 5 would have required at minimum a pair of d_{si} and r_{si} observed from the target forest, which is unrealistic from an operational inventory point of view. In practice, only population-averaged estimates could thus be obtained for a nonsampled plot, i.e., using the fixed part of eq. 5. To improve this, an additional analysis was carried out to augment the model with fixed effects to predict the random effects. First, eq. 5 was reduced to include the first random-effects term and the residual error only. Then, the random parameters by plot were predicted using a single ALS-based variable, computed from the plot-level distribution of height values with an attempt to characterize the forest structure. Specifically, the mean and standard deviation of the height values, as well as the 5th, 10th, 20th, ..., 90th, and 95th percentiles and the corresponding proportional densities of the ALS-based canopy height distribution, calculated according to Korhonen et al. (2008, pp. 502–503), were all considered as a variable x to predict the random plot-level effects. The variable having the highest correlation with the random plot-level effects was included in the model in the following form:

$$(6) \quad d_{si} = \beta_1 r_{si} + \beta_2 r_{si}^2 + z_{1s} r_{si} + \varepsilon_{si}$$

where d_{si} , r_{si} , β_1 , β_2 , z_{1s} , and ε_{si} correspond to eq. 5. The random effect z_{1s} was predicted by a submodel

$$(7) \quad z_{1s} = \alpha + \delta x_s + \varepsilon_s$$

where α and δ are the fixed parameters for plot variable x of plot s , fit together with eq. 6. When predicting the percentiles of the stem diameter distribution, eq. 6 was applied by replacing z_{1s} with the value predicted by eq. 7. The model fitting was done in a single step by the nlme package of R (Pinheiro et al. 2013).

3.4. Benchmark method

A benchmark diameter distribution was derived according to a conventional STRS (cf. Korpela et al. 2007), i.e., by assigning the delineated tree crowns with the largest stem diameter observed

Table 2. A summary of the RMSE and bias values obtained for each attribute.

Attribute	Equation(s)	Distribution	<i>n</i>	RMSE	RMSE%	Bias	Bias%
<i>N</i> , stems·ha ⁻¹	—	C	40	629	51.4	500	40.9*
	—	C	18	538	49.6	442	40.7*
	3	U	40	1030	84.2	-284	-23.2
	3	U	18	299	27.6	83	7.6
DQM, cm	8	C	40	4.6	29.9	-3.5	-22.8*
	8	C	18	4.2	26.5	-3.6	-23.1*
	5, f	U	40	3.1	20.6	0.3	2.1
	5, f	U	18	2.1	13.1	-0.2	-1.4
	6, 7	U	40	2.3	15.0	-0.1	-0.4
<i>G</i> , m ² ·ha ⁻¹	5, f	U	40	4.7	20.3	-0.9	-3.8
	5, f	U	18	2.9	13.0	1.0	4.4
	5	U	40	15.8	67.7	-5.0	-21.3
	5	U	18	4.9	22.0	0.8	3.5
	6, 7	U	40	10.8	46.3	-3.8	-16.2

Note: “Distribution” indicates whether the prediction (obtained from an equation) was based on censored (C) or uncensored (U) crown-radius distribution. *n* is the number of validation plots, being either 40 (all plots) or 18 (OK plots, i.e., plots where all assumptions were met). The letter “f” indicates that only the fixed part of eq. 5 was used. An asterisk (*) indicates statistical significance at the 95% confidence level according to a paired *t* test. *N*, stem number; DQM, quadratic mean diameter; *G*, basal area.

within the segment and fitting a tree-level model to these censored tree-crown data. The model used tree height and crown radius as the predictor variables. The fitted model was

$$(8) \quad d_{sj} = -0.57 + 0.28h_{sj}^{1.32} + 2.04r_{sj}^{0.92} + \varepsilon_{sj}$$

where d_{sj} is the stem diameter (cm) of the matched tree *j* on plot *s*, and h_{sj} and r_{sj} are the corresponding ALS-based tree height (m) and crown radius (m) estimates, respectively. The coefficient of determination (R^2) and the root mean squared error (RMSE; see section 3.5) for this model were 0.79 cm and 2.6 cm (14.6%), respectively. The plot-specific diameter distribution was composed by applying eq. 8 to each tree crown observed within the plot.

3.5. Evaluation and performance measures

The full field reference data of 40 plots were employed for both the model fitting and evaluation. The forest attributes estimated and evaluated for their accuracy and precision were the stem number (*N*), quadratic mean diameter (DQM), and basal area (*G*). Attributes *N* and DQM resulted from the analyses of sections 3.1 and 3.2–3.3, respectively, whereas *G* was derived as $G = 1/40\,000 \times N \times \pi \times DQM^2$ based on the estimated diameter distribution of the plot.

The accuracy of the plot-level aggregated statistics (i.e., DQM, *N*, and *G*) was evaluated using graphs and two goodness-of-fit measures, RMSE and average bias, calculated as

$$(9) \quad RMSE = \sqrt{\frac{\sum_{s=1}^n (y_s - \hat{y}_s)^2}{n}}$$

and

$$(10) \quad bias = \frac{\sum_{s=1}^n (y_s - \hat{y}_s)}{n}$$

where *n* is the number of plots, and y_s and \hat{y}_s are the measured and predicted attributes, respectively. The relative RMSE and bias were calculated by dividing the RMSE and bias by the mean of the reference attribute. A paired *t* test ($\alpha = 0.05$) was used to test the null hypothesis that the mean difference between measured and estimated values was 0.

The goodness-of-fit between the measured and predicted diameter distribution of each plot was assessed by the well-known Kolmogorov–Smirnov (KS) test statistic ($\alpha = 0.05$). Furthermore, the correspondence of these distributions was quantified by means of the error index (EI) proposed by Reynolds et al. (1988):

$$(11) \quad EI = \sum_{i=1}^k |f_i - \hat{f}_i|$$

where f_i and \hat{f}_i are the true and predicted stem number, respectively, within a diameter class *i*, and *k* is the number of classes with width = 2 cm. The absolute values were scaled to the observed total number of stems and presented additionally in terms of relative frequencies, the latter ones scaled between 0 (perfect fit) and 1 (no overlap at all) (Packalén and Maltamo 2008).

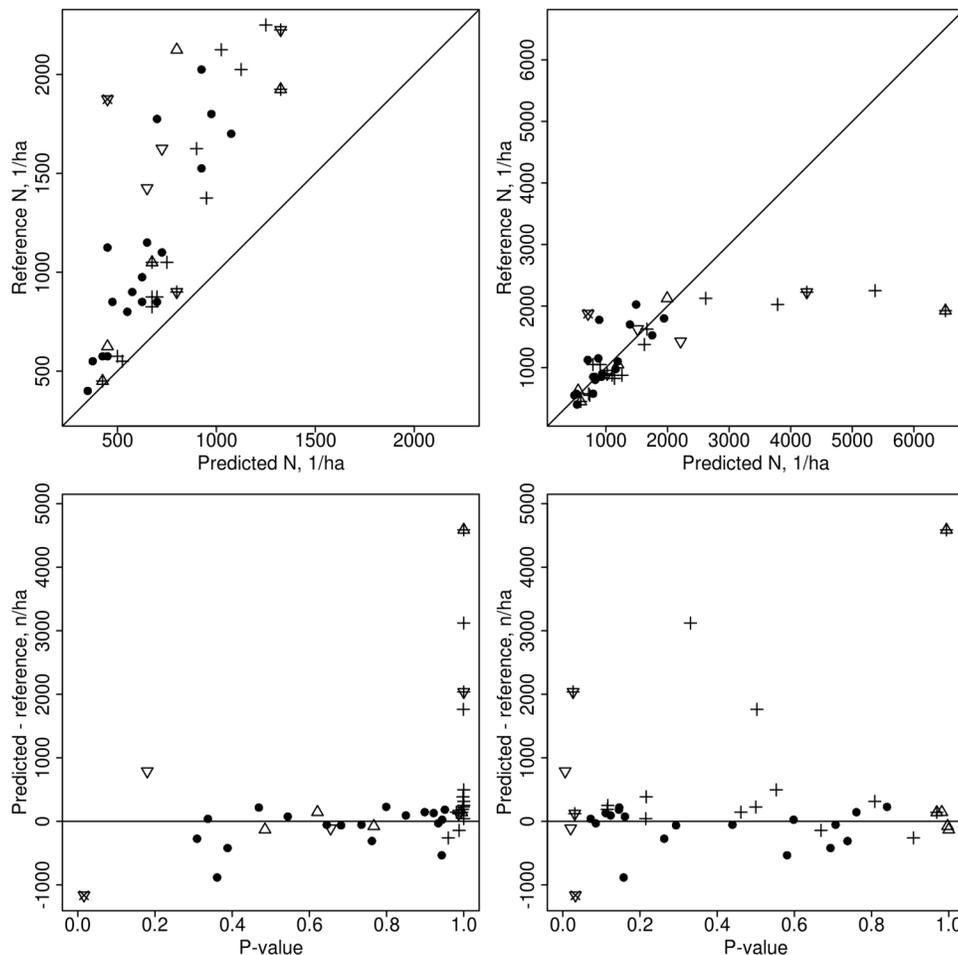
The underlying assumptions on the spatial pattern and autocorrelation (see section 3.1) were assessed by comparing the Clark Evans index (CEI; Clark and Evans 1954) of the aggregation of a point pattern and Moran’s *I* autocorrelation coefficient (Gittleman and Kot 1990), respectively, against those statistics under the Boolean model (complete spatial randomness and independence of marks; see also Bailey and Gatrell 1995). The tests were based on the field data, i.e., tree coordinates and diameters, respectively. Due to the small plot size, an edge-effect correction proposed by Donnelly (1978) was applied to the derived CEIs. The tests were carried out using the spatstat package of R (Baddeley and Turner 2005). In the following text, the plots with significantly regular (REG, *n* = 16) or clustered (CLUS, *n* = 1) spatial pattern or positive (POS, *n* = 5) or negative (NEG, *n* = 5) autocorrelation are treated separately from the plots in which all assumptions were met (OK, *n* = 18).

4. Results

The censored crown-radius distribution emphasized the largest trees and produced a significant underestimate of *N* (Table 2). Accounting for the detectability of the trees by means of uncensoring the distribution compensated for the bias but produced a higher RMSE for *N*. However, this increase was caused by a few plots with large errors in predicted stand density (Fig. 2, upper right), in particular, by four plots with a seriously overestimated *N*.

The goodness-of-fit tests on the assumed Boolean model explained the errors obtained in the uncensored result, in that the

Fig. 2. Upper panels: N ($1\cdot\text{ha}^{-1}$) estimated from the censored (left) and uncensored (right) distributions of crown radii vs. the reference attributes. Lower panels: the p values of the goodness-of-fit tests against alternative hypotheses “CEI < 1 (clustered spatial pattern)” (left) and “observed Moran’s I < expected Moran’s I (negative autocorrelation)” (right) vs. errors in N estimates ($n\cdot\text{ha}^{-1}$) produced from the uncensored crown radius distributions. The plus (+), times (x), triangle (Δ), upside-down triangle (∇), and solid circle (\bullet) symbols depict the REG, CLUS, POS, NEG, and OK plots, respectively.



plots not meeting the assumptions at the 95% confidence level produced the highest errors in N (Fig. 2). More precisely, the overestimates most frequently originated from the regular spatial patterns and the underestimates from the clustered spatial patterns. The errors resulting from lack of independence of crown radii were marginal compared with the errors caused by violating the assumption of a completely random spatial pattern of tree locations. The estimates derived from the uncensored distributions were considerably more accurate on the plots in which these assumptions were met (Table 2). No similar tendencies on the accuracy of the estimates were observed with the censored distribution (Table 2; Fig. 2).

In total, 25 of the 40 predicted diameter distributions differed significantly from the reference ones when the prediction was based on conventional STRS (the censored crown radii and eq. 8). The corresponding figures when using the DM technique were 19 (eq. 5, population-averaged prediction) and 1 (eq. 5, plot-specific prediction). Also based on the error indices (Table 3), both the population-averaged and plot-specific predictions clearly improved the correspondence between the predicted and reference diameter distributions. Graphical assessment showed that, in most cases, the DM technique was able to reproduce density functions closely corresponding to the fitted Weibull functions, especially when applying the plot-specific prediction (Fig. 3). The second-order polynomial (eq. 5) was not able to reproduce the extreme shapes of the CDF; however, in such cases, the initial Weibull function had only a mod-

erate fit with the reference distribution (see Fig. 3, lower panels). Although the DQM extracted from the censored distribution was an overestimate and emphasized the largest trees (Table 2), the null hypothesis on equal observed and estimated DQMs could not be rejected when the population-averaged prediction was used. When employing the plot-specific prediction, the DM technique produced a nearly 1:1 relationship between the reference and estimated DQMs. The errors in N thus provide a dominating source of errors in the further results.

Several ALS-based plot variables showed moderate to high correlations with the random plot effects of eq. 5. The proportional density at the 10th percentile height resulted in the highest R^2 value ($R^2 = 0.52$) and was thus used to predict the random plot parameter in eqs. 6 and 7. The use of eqs. 6 and 7 typically produced density functions that were closer to those observed (Fig. 3), particularly in plots that deviated greatly from the population-averaged estimates, and slightly improved the accuracy of DQM compared with using the population-averaged estimates (Table 2).

The predictions of G based on uncensored crown radii and by fitting eq. 5 to the full reference data are shown in Fig. 4. In terms of RMSE and bias values, the population-averaged prediction provided more accurate estimates than the plot-specific prediction of eq. 5 (Table 2). However, the fixed part clearly averaged the form of the transformation function, whereas the full mixed-effects model was able to describe the properties of the transformation

Table 3. Mean, standard deviation (SD), minimum, and maximum of the absolute (and relative, in parenthesis) values of Reynolds' error indices for the diameter distributions predicted using the given crown-radius distribution and equation.

Distribution	Equation(s)	Mean	SD	Minimum	Maximum
C	8	943.8 (0.38)	582.9 (0.12)	250.6 (0.20)	3063.8 (0.82)
U	5, f	840.1 (0.36)	543.5 (0.13)	327.3 (0.15)	2836.4 (0.76)
U	5	619.9 (0.28)	329.2 (0.11)	223.4 (0.12)	2200.0 (0.59)
U	6, 7	723.5 (0.31)	418.8 (0.12)	234.5 (0.16)	2443.1 (0.65)

Note: The letter "F" indicates that only the fixed part of eq. 5 was used. C, censored; U, uncensored.

functions, as well as the errors due to the unmet assumptions, in more detail. In particular, the errors in N were more pronounced in the G predictions based on mixed-effect prediction, and the absolute bias values of these estimates were considerably high and close to the critical value ($p = 0.11$ and 0.07 , respectively). In general, higher accuracies were obtained for the 18 OK plots for which the absolute values of the biases were reasonable and the null hypothesis on equal predicted and reference values could not be rejected (Table 2).

The violations against the assumed independence of crown radii caused the highest errors in G when the population-averaged prediction of eq. 5 was used, and those against the spatial pattern caused the highest errors in G when the plot-specific prediction of eq. 5 was used (Fig. 4). However, these plots also had a high canopy cover, which may have an effect in terms of the function form of eq. 3 (see section 5). In these data, the plots with the highest errors could be restricted by limiting the considerations to the plots with an estimated canopy cover <90%, for example.

5. Discussion

The DM approach proposed here constitutes a previously unstudied means of combining tree-level and area-based analyses. The approach is adapted from digital image processing (cf. Gonzalez and Woods 2008), but when applied in the present form ("predictive DM"), our approach differs considerably from typical DM applications, which assume that both the reference and the target-image histograms are simultaneously available. The approach is thus innovative and could also have further potential applications when taken back to the image analysis domain.

The approach presented here is based on single-tree crowns detected by means of ordinary STRS, but the estimation does not directly rely on these observations. Our approach compensates for the (systematic) errors caused by the forest structure (Larsen et al. 2011; Vauhkonen et al. 2012), which are proposed to affect the success of both tree detection (Persson et al. 2002; Korpela 2004; Falkowski et al. 2008) and segmentation (Vauhkonen et al. 2010, 2014; Peuhkurinen et al. 2011). To the best of our knowledge, to date, the only result showing no evidence against zero bias based on a pure STRS procedure (i.e., without local calibration of the systematic errors) has been presented by Vauhkonen et al. (2011), who utilized detailed forest plantation records for model localization in a similar mixed-effects modeling approach to the one considered here.

The idea to combine tree-level and area-based approaches somewhat resembles corresponding attempts by Maltamo et al. (2004) and Lindberg et al. (2010), who augmented or adjusted the individually detected tree distribution using theoretical or area-based distributions. On the other hand, to compensate for the errors in tree detection and estimation, Breidenbach et al. (2010) developed a "semi-individual" tree-detection approach in which the tree-crown segments derived from the ALS data were imputed with data from 0 to n trees depending on the nearest neighbor segments considered in the estimation. However, the aforementioned approaches require mapped tree data, which do not exist in operational forest inventories (e.g., Maltamo et al. 2011). Therefore, our approach, based on plot-level tree distributions, is able to overcome this limitation.

The approach proposed here also resembles area-based ALS inventories with respect to the use of a two-stage modeling and prediction procedure (cf. Næsset 2002) but with the exception that the estimation is based on DM instead of linear regression or another modeling technique. Conventional area-based methods such as parameter prediction, recovery, or nearest neighbor based approaches (see Maltamo and Gobakken 2014) could also be seen as alternative means to predicting diameter distribution. However, the requirements by the area-based approaches for the amount of field reference plots (see Maltamo et al. 2011) exclude such a comparison in these data. Estimation based on the detected tree-crown radii could be expected to reduce the requirements for field data because it utilizes information on physical tree dimensions (cf. Hyypä and Inkinen 1999), and conventional STRS (cf. Korpela et al. 2007) therefore provided the closest benchmark. In fact, due to underestimating N and overestimating DQM, the pure STRS approach already provided a fairly accurate G , but it was based on unrealistic stem diameters (cf. Fig. 3). Earlier comparisons proposing that the STRS and area-based approaches yielded approximately similar results may thus have suffered from comparing the averaged results only; however, considerably more attention should be focused on assessing the shape of diameter distribution as in Peuhkurinen et al. (2011).

Compared with conventional STRS, the goodness-of-fit of the diameter distribution was particularly improved by applying the DM technique proposed. When interpreting our results, the reader should note that the predictions obtained by the use of full eq. 5 represent a theoretical potential obtainable by plot-specific observations of the percentiles of both the stem diameter and crown radius, being therefore unrealistic in practice. However, either the fixed part of eq. 5 or a combination of eqs. 6 and 7 could be realistically applied with auxiliary observations obtainable from ALS data. In addition to the tree detection based on the CHMs and augmenting the obtained model with a plot-level predictor (eqs. 6 and 7), our approach did not explicitly use tree height for any other purpose, which suggests that predictions based on eq. 5 and crown radii extracted from aerial images, for instance, could potentially be carried out in a similar manner. The latter would reduce the costs of data acquisition significantly, therefore constituting an extremely interesting option from a practical application point of view.

An accurate estimate of the total stem number was crucial for unbiased predictions. Our tree detection approach was thus augmented with estimated tree detectability based on the Boolean model (Mehtätalo 2006; Stoyan et al. 1995). According to the goodness-of-fit tests made, the underlying assumptions of the Boolean model were met only in 45% of the plots studied. The assumption of a homogeneous Poisson process for tree locations, which appeared to be more critical than the independence of marks, was met in 58% of the plots. Departures from the direction of regular pattern caused especially large overestimates of N , which further affected the estimates of G . Therefore, the model-based estimation of detectability and stand density should be developed specifically for regular patterns. Work in this direction, motivated by remote sensing of forests, has been reported in Bondesson and Fahlen (2003). Another starting point could be the quermass model proposed by Kendall et al. (1999). In situations with a low degree of fit between the empirical and Weibull distributions, the latter could be replaced by, for example, a

Fig. 3. The probability density functions and cumulative distribution functions ($F(x)$) of three example plots showing different goodness-of-fit between the observed and predicted diameter distributions. Eq. 5, f indicates that only the fixed part of eq. 5 was used. DBH, diameter at breast height.

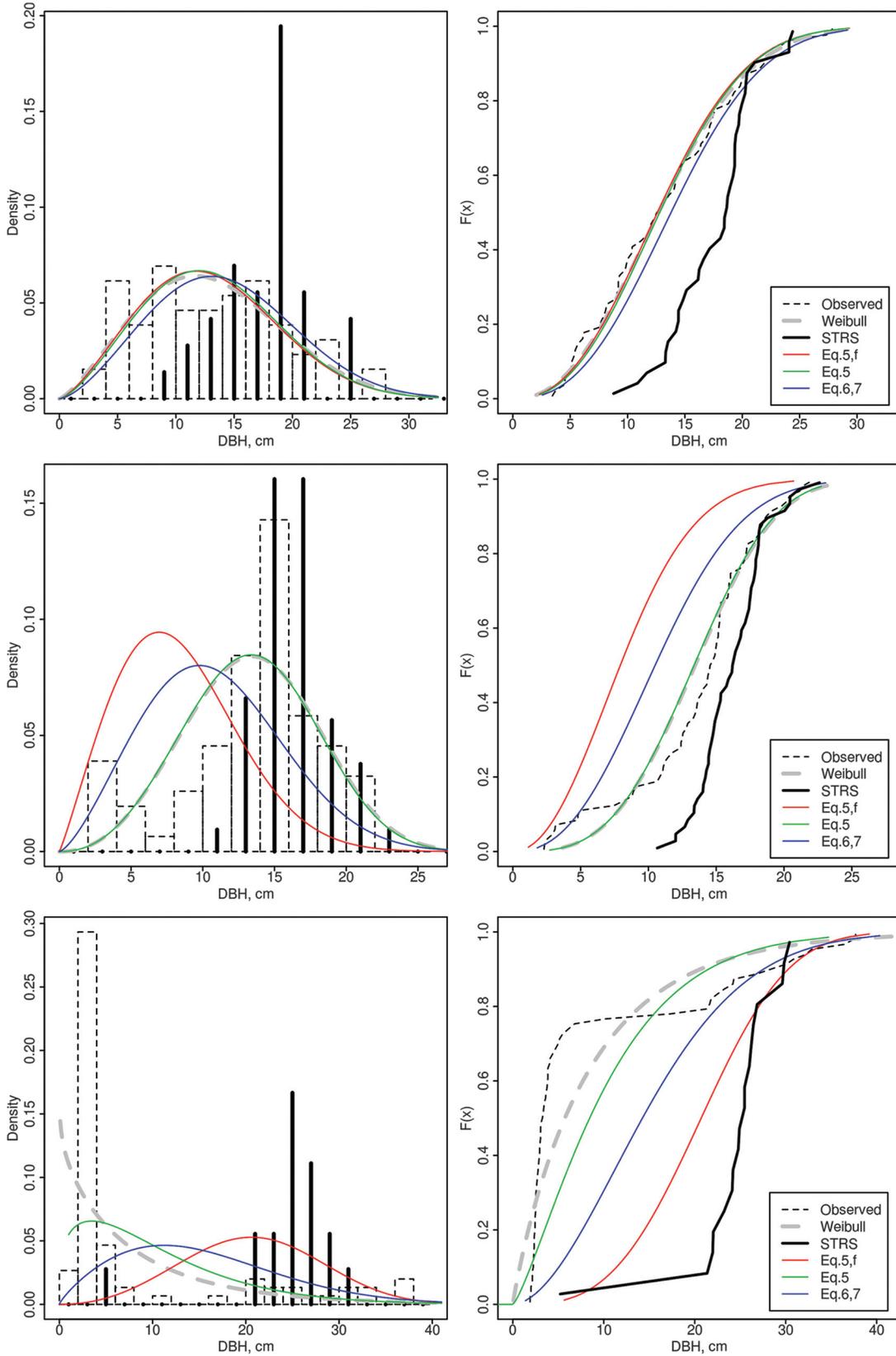
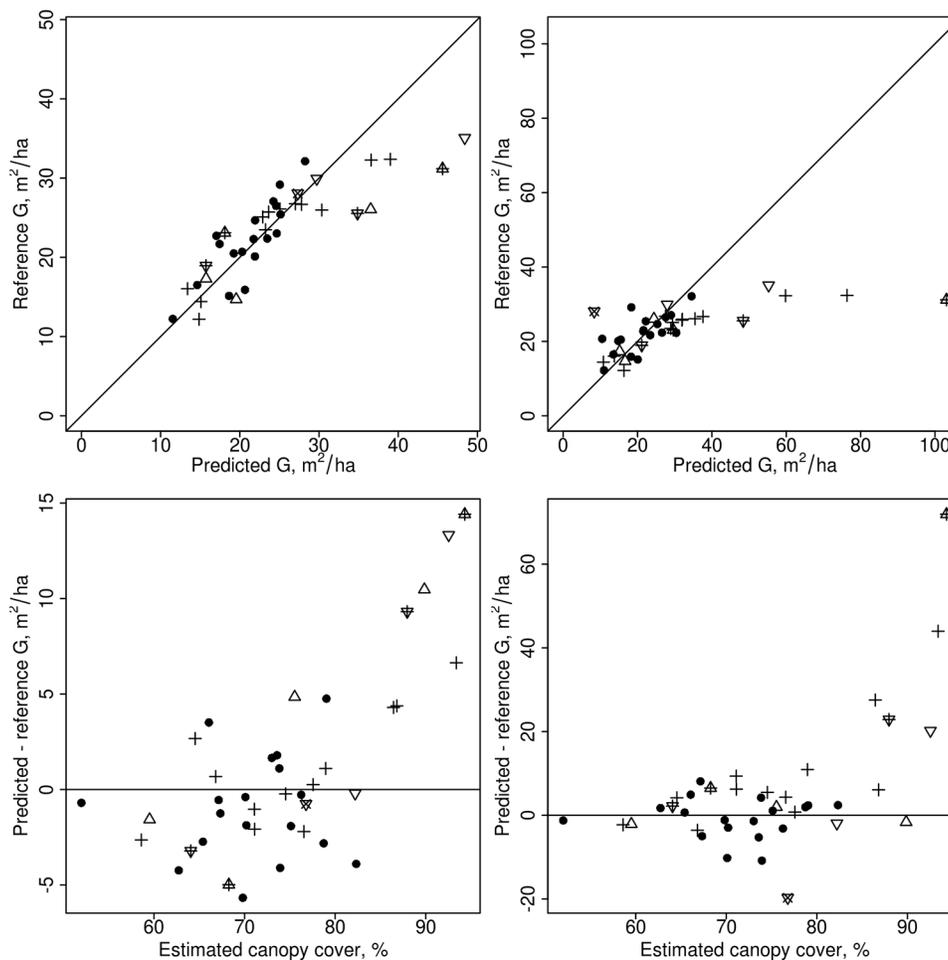


Fig. 4. Upper panels: estimated vs. reference G ($\text{m}^2 \cdot \text{ha}^{-1}$) based on the uncensored crown radii and predicting with the fixed part (left) or full eq. 5 (right). Lower panels: the corresponding residuals plotted against estimated canopy cover (%). The plus (+), times (\times), triangle (Δ), upside-down triangle (∇), and solid circle (\bullet) symbols depict the REG, CLUS, POS, NEG, and OK plots, respectively.



mixture of two Weibull distributions (Mehtätalo 2006). Furthermore, an empirical estimation of the detectability could also be tested, even though such an approach would increase the data collection costs.

Defining the spatial pattern of trees is ambiguous: for example, our test results are in 70% agreement with Packalén et al. (2013), who analyzed the same plots but used subsampling and wider confidence intervals. Furthermore, the plot size was different for some plots (cf. section 2). Nevertheless, the application of our approach is restricted due to these reasons, and detecting the underlying spatial pattern is difficult by means of remote sensing (e.g., Packalén et al. 2013). On the other hand, the used critical value ($p = 0.05$) is just one choice. It is clearly visible in Fig. 2 that only the most severe departures from the random pattern led to the largest overestimates. Therefore, allowing different levels of significance, our approach could probably be applied with a larger subset of the plots than the 18 considered. The largest errors were found in plots with a high estimated canopy cover ($>90\%$; Fig. 4). Limiting the estimation based on the canopy cover would be justified because the estimator of stand density (eq. 3) is strongly based on it. As a function of cc, eq. 3 gets very steep, particularly for canopy covers above 90%. Therefore, the largest estimation errors occur on plots in which the high canopy cover is actually a result of regular pattern instead of an exceptionally high stand density. This is quite logical, because for a certain model for tree crown radii, a regular pattern of tree locations naturally results in higher canopy closure than a random pattern. The canopy cover

could thus be used as an accuracy indicator in practical predictions when the actual spatial pattern is unknown.

When generating the crown radii distributions, several alternative measures of crown dimensions could be considered instead of crown radii used. We also ran tests with the distributions of crown area and crown radius extracted from the circle area equation, assuming the crown areas to be circularly distributed around the tree top. These variations mainly affected the magnitude of the observations but not their mutual order relative to Fig. 3, for example. Furthermore, although some alternative ALS point-based measures are found to be better in line with the field-measured canopy closure (Korhonen et al. 2011), a CHM-based estimate was reasoned to be more consistent with the crown radii extracted from the same CHMs. In our study, the ultimate predictions were obtained by means of ALS-based crown radii and further height distribution variables to account for the forest structure (eqs. 6 and 7). However, auxiliary variables extracted from other remote sensing data sources such as aerial photographs could be studied for the same purpose. Notably, coarse predictions based on eq. 5 and crown radii and canopy closure extracted from aerial photographs could readily be studied following the description given in sections 3.2 and 3.3.

Our approach inherently predicts the stem diameter distribution. Distribution modeling based on two-parameter Weibull functions was selected to produce smooth transformation functions; however, a more complex approach and more field calibration data will likely be required to produce realistic diameter

distributions (cf. Maltamo and Gobakken 2014). If it would be more substantial to produce a detailed and flexible description of the prevailing diameter distribution rather than a smooth transformation, the use of the observed percentiles of the stem size distributions (cf. Mehtätalo 2006) or techniques based on mixture or nonparametric modeling (see Maltamo and Gobakken 2014) could be studied, especially in more complex forest structural types. However, it is noted that several error sources in the transformation modeling may contribute to the description of the distributions and, therefore, average the details obtained by a more complex modeling technique.

It is also acknowledged that the forest area selected for these initial simulations was rather simple, i.e., we had to consider only one species occurring in a managed forest. Because the crowns of other Scandinavian coniferous species present similar properties regarding stem diameter modeling (cf. Kalliovirta and Tokola 2005), the proposed approach could be assumed to work with these species and their mixtures, whereas deciduous tree species could cause problems due to more frequently erroneous crown delineations (Koch et al. 2006; Vauhkonen et al. 2012). Besides species composition, further problems could be caused by complex, multilayered forest structures by preventing the detection of lower canopy layers and thus hindering the estimation of the transformation functions (Fig. 1). On the other hand, the overall approach could possibly perform better in an unmanaged forest, provided a better hold of the assumptions on spatial pattern and autocorrelation. Overall, future tests of the approach should include other species and their mixtures, more complex diameter distributions and their models, and data from different remote sensing systems.

6. Conclusions

A new approach to relate RS and field-measured tree size distributions was developed based on the well-known histogram matching technique to map the transformation between their CDFs. The observed distribution of crown radii was corrected for the censoring effect caused by overlapping tree crowns, assuming the forest as an outcome of a homogeneous, marked Poisson process with independent marks of the crown radii. Of the 40 plots studied, 18 met the assumptions of the Poisson process and independent marks. In these plots, the approach successfully compensated for the undetected trees, but departures from random patterns caused especially large overestimates of stand density, which further affected the estimates of the basal area.

The transformation from the CDF of crown radius to the CDF of stem diameter particularly improved the goodness-of-fit of the predicted diameter distribution compared with conventional STRS, and observations realistically obtainable with RS data showed potential to localize the predictions. Ultimately, the violations against the assumed independence of crown radii and the spatial pattern caused large errors in basal area predictions. However, the plots with the highest errors could be identified by high canopy cover, and thus they could be restricted by limiting the considerations to the plots with an estimated canopy cover <90%, for example. The model-based estimation of the detectability and stand density was the main error source, which should be further developed for regular spatial patterns.

The proposed approach provides the following interesting possibilities not typical to the forest inventory literature based on the use of ALS, which should be further examined:

- A low requirement for field reference data. It is proposed that only a few field plots could be adequate for producing averaged predictions, yet an increased number is likely required for detailed diameter distribution modeling.
- No need for tree mapping in the reference data. The field measurements required for model fitting could be based on angle count sampling, for example, provided that the description

obtainable by the use of a theoretical distribution model is considered adequate.

- No explicit requirement for tree height data. Although height values based on ALS data were implicitly used for tree detection and model localization, it is proposed that whether high-resolution aerial or satellite images provide consistent crown radius distributions and canopy cover estimates, such observations could be used in the averaged manner presented here to produce coarse predictions of stem diameter distributions.

Acknowledgements

We would like to thank Matti Maltamo, Petteri Packalen, and Juho Pitkänen for kindly allowing the use of the data prepared earlier for the Spearhead project “Multi-scale Geospatial Analysis of Forest Ecosystems” of the University of Eastern Finland for the present purpose. We also acknowledge the anonymous reviewers for their dedication to improve the quality of the presentation. The preparation of the manuscript was financially supported by the Research Funds of the University of Helsinki.

References

- Baddeley, A., and Turner, R. 2005. Spatstat: an R package for analyzing spatial point patterns. *J. Stat. Softw.* **12**: 1-42. Available from www.jstatsoft.org/v12/i06/.
- Baffetta, F., Corona, P., and Fattorini, L. 2012. A matching procedure to improve k-NN estimation of forest attribute maps. *For. Ecol. Manage.* **272**: 35–50. doi:10.1016/j.foreco.2011.06.037.
- Bailey, T.C., and Gatrell, A.C. 1995. *Interactive spatial data analysis*. Addison Wesley Longman Limited, Essex, England.
- Bondesson, L., and Fahlen, J. 2003. Mean and variance of vacancy for hard-core disc processes and applications. *Scandinavian Journal of Statistics*, **30**(4): 797–816. doi:10.1111/1467-9469.00365.
- Breidenbach, J., and Astrup, R. 2014. The semi-individual tree crown approach. *In* *Forestry applications of airborne laser scanning — concepts and case studies*. Edited by M. Maltamo, E. Næsset, and J. Vauhkonen. *Managing Forest Ecosystems* 27. Springer, the Netherlands. pp. 113–133.
- Breidenbach, J., Næsset, E., Lien, V., Gobakken, T., and Solberg, S. 2010. Prediction of species specific forest inventory attributes using a nonparametric semi-individual tree crown approach based on fused airborne laser scanning and multispectral data. *Remote Sens. Environ.* **114**: 911–924. doi:10.1016/j.rse.2009.12.004.
- Clark, P.J., and Evans, F.C. 1954. Distance to nearest neighbour as a measure of spatial relationships in populations. *Ecology*, **35**: 445–453. doi:10.2307/1931034.
- Donnelly, K. 1978. Simulations to determine the variance and edge-effect of total nearest neighbour distance. *In* *Simulation studies in archaeology*. Edited by I. Hodder. Cambridge University Press, New York. pp. 91–95.
- Falkowski, M.J., Smith, A.M.S., Gessler, P.E., Hudak, A.T., Vierling, L.A., and Evans, J.S. 2008. The influence of conifer forest canopy cover on the accuracy of two individual tree measurement algorithms using lidar data. *Can. J. Remote Sens.* **34**(Suppl. 2): S338–S350. doi:10.5589/m08-055.
- Flewelling, J.W. 2008. Probability models for individually segmented tree crown images in a sampling context. *In* *Proceedings of SilviLaser 2008, 8th International Conference on LIDAR Applications in Forest Assessment and Inventory*, Heriot-Watt University, Edinburgh, UK, 17–19 September 2008. Edited by R. Hill, J. Rosette, and J. Suárez. pp. 284–294.
- Gilichinsky, M., Sandström, P., Reese, H., Kivinen, S., Moen, J., and Nilsson, M. 2011. Mapping ground lichens using forest inventory and optical satellite data. *Int. J. Remote Sens.* **32**: 455–472. doi:10.1080/01431160903474962.
- Gilichinsky, M., Heiskanen, J., Barth, A., Wallerman, J., Egberth, M., and Nilsson, M. 2012. Histogram matching for the calibration of kNN stem volume estimates. *Int. J. Remote Sens.* **33**: 7117–7131. doi:10.1080/01431161.2012.700134.
- Gittleman, J.L., and Kot, M. 1990. Adaptation: statistics and a null model for estimating phylogenetic effects. *Syst. Zool.* **39**: 227–241. doi:10.2307/2992183.
- Gonzalez, R., and Woods, R.E. 2008. *Digital image processing*. 3rd edition. Prentice Hall, U.S.A.
- Hopkinson, C., Lovell, J., Chasmer, L., Jupp, D., Kljun, N., and van Gersel, E. 2013. Integrating terrestrial and airborne lidar to calibrate a 3D canopy model of effective leaf area index. *Remote Sens. Environ.* **136**: 301–314. doi:10.1016/j.rse.2013.05.012.
- Hyypä, J., and Inkinen, M. 1999. Detecting and estimating attributes for single trees using laser scanner. *The Photogrammetric Journal of Finland*, **16**: 27–42.
- Hyypä, J., Hyypä, H., Leckie, D., Gougeon, F., Yu, X., and Maltamo, M. 2008. Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests. *Int. J. Remote Sens.* **29**: 1339–1366. doi:10.1080/01431160701736489.
- Kalliovirta, J., and Tokola, T. 2005. Functions for estimating stem diameter and

- tree age using tree height, crown width and existing stand database information. *Silva Fenn.* **39**: 227–248. doi:10.14214/sf.386.
- Ke, Y., and Quackenbush, L.J. 2011. A review of methods for automatic individual tree-crown detection and delineation from passive remote sensing. *Int. J. Remote Sens.* **32**: 4725–4747. doi:10.1080/01431161.2010.494184.
- Kendall, W.S., van Lieshout, M.N.M., and Baddeley, A.J. 1999. Quermass-interaction processes: conditions for stability. *Adv. Appl. Probab.* **31**(2): 315–342. doi:10.1239/aap/1029955137.
- Koch, B., Heyder, U., and Weinacker, H. 2006. Detection of individual tree crowns in airborne LIDAR data. *Photogramm. Eng. Remote Sens.* **72**: 357–363. doi:10.14358/PERS.72.4.357.
- Korhonen, L., Peuhkurinen, J., Malinen, J., Suvanto, A., Maltamo, M., Packalén, P., and Kangas, J. 2008. The use of airborne laser scanning to estimate sawlog volumes. *Forestry*, **81**: 499–510. doi:10.1093/forestry/cpn018.
- Korhonen, L., Korpela, I., Heiskanen, J., and Maltamo, M. 2011. Airborne discrete-return LIDAR data in the estimation of vertical canopy cover, angular canopy closure and leaf area index. *Remote Sens. Environ.* **115**: 1065–1080. doi:10.1016/j.rse.2010.12.011.
- Korpela, I. 2004. Individual tree measurements by means of digital aerial photogrammetry. Doctoral thesis, University of Helsinki, Department of Forest Resource Management, Helsinki, Finland. *Silva Fennica Monographs* 3.
- Korpela, I. 2007. Incorporation of allometry in single-tree remote sensing with LIDAR and multiple images. In *Proceedings of the ISPRS Hannover Workshop 2007 — High Resolution Earth Imaging for Geospatial Information*, 29 May – 1 June 2007, Hannover, Germany. Edited by C. Heipke, K. Jacobsen, and M. Gerke. Vol. XXXVI, Part I/W51.
- Korpela, I., and Tokola, T. 2006. Potential of aerial image-based monoscopic and multiview single-tree forest inventory — a simulation approach. *For. Sci.* **52**: 136–147.
- Korpela, I., Dahlin, B., Schäfer, H., Bruun, E., Haapaniemi, F., Honkasalo, J., Ilvesniemi, S., Kuutti, V., Linkosalmi, M., Mustonen, J., Salo, M., Suomi, O., and Virtanen, H. 2007. Single-tree forest inventory using LiDAR and aerial images for 3D tree-top positioning, species recognition, height and crown width estimation. In *Proceedings of ISPRS Workshop on Laser Scanning 2007 and SilviLaser 2007*, 12–14 September 2007, Espoo, Finland. Edited by P. Rönnholm, H. Hyyppä, and J. Hyyppä. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVI, Part 3/W52. pp. 227–233. Available from www.isprs.org/proceedings/XXXVI/3-W52/final_papers/Korpela_2007.pdf.
- Larsen, M., Eriksson, M., Descombes, X., Perrin, G., Brandtberg, T., and Gougeon, F. 2011. Comparison of six individual tree crown detection algorithms evaluated under varying forest conditions. *Int. J. Remote Sens.* **32**: 5827–5852. doi:10.1080/01431161.2010.507790.
- Lindberg, E., Holmgren, J., Olofsson, K., Wallerman, J., and Olsson, H. 2010. Estimation of tree lists from airborne laser scanning by combining single-tree and area-based methods. *Int. J. Remote Sens.* **31**: 1175–1192. doi:10.1080/01431160903380649.
- Lindstrom, M.J., and Bates, D.M. 1990. Nonlinear mixed-effects models for repeated measures data. *Biometrics*, **46**: 673–687. doi:10.2307/2532087.
- Maltamo, M., and Gobakken, T. 2014. Predicting tree diameter distributions. In *Forestry applications of airborne laser scanning — concepts and case studies*. Edited by M. Maltamo, E. Næsset, and J. Vauhkonen. *Managing Forest Ecosystems* 27, Springer, the Netherlands. pp. 177–191.
- Maltamo, M., Eerikäinen, K., Pitkänen, J., Hyyppä, J., and Vehmas, M. 2004. Estimation of timber volume and stem density based on scanning laser altimetry and expected tree size distribution functions. *Remote Sens. Environ.* **90**: 319–330. doi:10.1016/j.rse.2004.01.006.
- Maltamo, M., Packalén, P., Kallio, E., Kangas, J., Uuttera, J., and Heikkilä, J. 2011. Airborne laser scanning based stand level management inventory in Finland. In *Proceedings of SilviLaser 2011 — 11th International Conference on LiDAR Applications for Assessing Forest Ecosystems*, 16–20 October 2011, Hobart, Australia. Available from www.iufro.org/download/file/8239/5065/40205-silvilaser2011.pdf.
- Maltamo, M., Mehtätalo, L., Vauhkonen, J., and Packalén, P. 2012. Predicting and calibrating tree attributes by means of airborne laser scanning and field measurements. *Can. J. For. Res.* **42**(11): 1896–1907. doi:10.1139/x2012-134.
- McRoberts, R.E., Tomppo, E.O., and Næsset, E. 2010. Advances and emerging issues in national forest inventories. *Scand. J. For. Res.* **25**: 368–381. doi:10.1080/02827581.2010.496739.
- Mehtätalo, L. 2006. Eliminating the effect of overlapping crowns from aerial inventory estimates. *Can. J. For. Res.* **36**(7): 1649–1660. doi:10.1139/x06-066.
- Næsset, E. 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sens. Environ.* **80**: 88–99. doi:10.1016/S0034-4257(01)00290-5.
- Nyström, M., Holmgren, J., and Olsson, H. 2013. Change detection of mountain birch using multi-temporal ALS point clouds. *Remote Sens. Lett.* **4**: 190–199. doi:10.1080/2150704X.2012.714087.
- Olsson, H. 1993. Regression functions for multitemporal relative calibration of Thematic Mapper data over boreal forest. *Remote Sens. Environ.* **46**: 89–102. doi:10.1016/0034-4257(93)90034-U.
- Packalén, P., and Maltamo, M. 2008. Estimation of species-specific diameter distributions using airborne laser scanning and aerial photographs. *Can. J. For. Res.* **38**(7): 1750–1760. doi:10.1139/X08-037.
- Packalén, P., Vauhkonen, J., Kallio, E., Peuhkurinen, J., Pitkänen, J., Pippuri, I., Strunk, J., and Maltamo, M. 2013. Predicting the spatial pattern of trees with airborne laser scanning. *Int. J. Remote Sens.* **34**: 5154–5165. doi:10.1080/01431161.2013.787501.
- Pantze, A., Santoro, M., and Fransson, J.E.S. 2014. Change detection of boreal forest using bi-temporal ALOS PALSAR backscatter data. *Remote Sens. Environ.* **155**: 120–128. doi:10.1016/j.rse.2013.08.050.
- Persson, Å., Holmgren, J., and Söderman, U. 2002. Detecting and measuring individual trees using an airborne laser scanner. *Photogramm. Eng. Remote Sens.* **68**: 925–932.
- Peuhkurinen, J., Maltamo, M., Malinen, J., Pitkänen, J., and Packalén, P. 2007. Preharvest measurement of marked stands using airborne laser scanning. *For. Sci.* **53**: 653–661.
- Peuhkurinen, J., Mehtätalo, L., and Maltamo, M. 2011. Comparing individual tree detection and the area-based statistical approach for the retrieval of forest stand characteristics using airborne laser scanning in Scots pine stands. *Can. J. For. Res.* **41**(3): 583–598. doi:10.1139/X10-223.
- Pinheiro, J.C., and Bates, D.M. 2000. *Mixed-effects models in S and S-PLUS*. Springer-Verlag, New York.
- Pinheiro, J., Bates, D., DebRoy, S., and Sarkar, D., and R Core Team. 2013. nlme: linear and nonlinear mixed effects models. R package version 3.1-111. Available from cran.r-project.org/web/packages/nlme.
- Pitkänen, J. 2005. A multi-scale method for segmentation of trees in aerial images. In *Proceedings of the SNS Meeting at Sjusjøen — Forest Inventory and Planning in Nordic Countries*, Norway, 6–8 September 2004. Edited by K. Hobbelstad. NIJOS Report 09/05, Norwegian Institute of Land Inventory, Oslo, Norway.
- Pitkänen, J., Maltamo, M., Hyyppä, J., and Yu, X. 2004. Adaptive methods for individual tree detection on airborne laser based canopy height model. In *Proceedings of ISPRS working group VIII/2: Laser-Scanners for Forest and Landscape Assessment*. Edited by M. Theis, B. Koch, H. Spiecker, and H. Weinacker. University of Freiburg, Freiburg, Germany. pp. 187–191.
- Reynolds, M.R., Burk, T.E., and Huang, W.C. 1988. Goodness-of-fit tests and model selection procedures for diameter distribution models. *For. Sci.* **34**: 373–399.
- Stoyan, D., Kendall, W.S., and Mecke, J. 1995. *Stochastic geometry and its applications*. 2nd edition. Wiley, Chichester, UK.
- Thompson, S.K. 2012. *Sampling*. 3rd edition. John Wiley & Sons, Hoboken, New Jersey.
- Vauhkonen, J., Korpela, I., Maltamo, M., and Tokola, T. 2010. Imputation of single-tree attributes using airborne laser scanning-based height, intensity, and alpha shape metrics. *Remote Sens. Environ.* **114**: 1263–1276. doi:10.1016/j.rse.2010.01.016.
- Vauhkonen, J., Mehtätalo, L., and Packalén, P. 2011. Combining tree height samples produced by airborne laser scanning and stand management records to estimate plot volume in *Eucalyptus* plantations. *Can. J. For. Res.* **41**(8): 1649–1658. doi:10.1139/x11-083.
- Vauhkonen, J., Ene, L., Gupta, S., Heinzl, J., Holmgren, J., Pitkänen, J., Solberg, S., Wang, Y., Weinacker, H., Hauglin, K.M., Lien, V., Packalén, P., Gobakken, T., Koch, B., Næsset, E., Tokola, T., and Maltamo, M. 2012. Comparative testing of single-tree detection algorithms under different types of forest. *Forestry*, **85**: 27–40. doi:10.1093/forestry/cpr051.
- Vauhkonen, J., Packalén, P., Malinen, J., Pitkänen, J., and Maltamo, M. 2014. Airborne laser scanning based decision support for wood procurement planning. *Scand. J. For. Res.* **29**(Suppl. 1): 132–143. doi:10.1080/02827581.2013.813063.
- Xu, Q., Hou, Z., Maltamo, M., and Tokola, T. 2014. Calibration of area based diameter distribution with individual tree based diameter estimates using airborne laser scanning. *ISPRS J. Photogramm. Remote Sens.* **93**: 65–75. doi:10.1016/j.isprsjprs.2014.03.005.