Remote Sensing of Environment 114 (2010) 1277–1285

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Modelling tree diameter from airborne laser scanning derived variables: A comparison of spatial statistical models

Christian Salas [a,c,*], Liviu Ene [b], Timothy G. Gregoire [a], Erik Næsset [b], Terje Gobakken [b]

[a] School of Forestry and Environmental Studies, Yale University, New Haven, Connecticut 06511-2104, USA
[b] Department of Ecology and Natural Resource Management, Norwegian University of Life Sciences, Ås, Norway
[c] Departamento de Ciencias Forestales, Universidad de La Frontera, Temuco, Chile

ARTICLE INFO

Article history:
Received 17 July 2009
Received in revised form 13 January 2010
Accepted 16 January 2010

Keywords:
LiDAR
Spatial correlation
Mixed-effects models
Geographically weighted regression
Diameter–height models

ABSTRACT

Formerly, tree height has been more difficult to measure accurately in the field than tree diameter at breast height. As a consequence, models to predict height from diameter measurements have been widely developed in the forestry literature. Through the use of airborne laser scanning technology (e.g., LiDAR), tree variables such as height and crown diameter can be measured accurately, a development which has spawned the need for models to predict diameter from airborne laser-derived measurements. Although some work has been done for fitting such models, none have incorporated spatial information to improve the accuracy of the predicted diameters. Using a simple linear model for predicting tree diameter from laser-derived tree height and crown diameter measurements, we compared the performance of ordinary least squares (OLS), generalized least squares with a non-null correlation structure (GLS), linear mixed-effects model (LME), and geographically weighted regression (GWR). Our data were obtained from 36 sample plots established in Norway. This is the first study to examine the use of spatial statistical models for tree-level LiDAR data. Root mean square prediction errors in tree diameter with LME are 3.5%, with GWR are 10%, and with OLS and GLS are 17%. LME also exhibited low variability in predicting performance across all the validation classes (based on laser-derived height). Giving the difficulties of using parametric statistical inference (such as maximum likelihood-based indices) for GWR, we used permutation tests as a way for detecting statistical differences. LME was significantly better than the other models, as well as GWR was to OLS and GLS. Our results indicate that the LME model produced the best predictions of tree diameter from LiDAR-based variables to a degree that has previously not been possible.

1. Introduction

Airborne laser scanning (ALS) is a powerful remote sensing system for collecting topographic data having applications in a variety of disciplines. ALS systems have led a revolution in remote sensing technology during the last 10–14 years (Popescu & Wynne, 2004). As pointed out by Reutebuch et al. (2005), the basis of this revolution is the ability to measure directly the three-dimensional structure (i.e., terrain, vegetation, and infrastructure) of imaged areas and to separate biospatial data (measurements of aboveground vegetation) from geospatial data (measurements of the terrain surface) using active remote sensing technologies. In ALS, a scanner which distributes the transmitted pulses across the flight direction of the platform is attached to the laser, and such systems can measure the 3D position of points on the ground and in vegetation canopies with an accuracy of a few decimetres (Næsset et al., 2004). Among the laser scanner systems currently available, LiDAR (light detection and ranging) sensors offer impressive performance that challenge physical barriers in the optical and electronic domains by offering a high density of points at scanning frequencies of 50,000 pulses/second, multiple echoes per laser pulse, intensity measurements for the returning signal, and centimeter accuracy for horizontal and vertical positioning (Popescu & Wynne, 2004). These features, make LiDAR useful for directly assessing vegetation characteristics, and overall providing new tools for measuring and monitoring biospatial data across the landscape (Popescu & Wynne, 2004; Reutebuch et al., 2005). Further reviews of ALS and LiDAR studies in general and in forestry, can be found in Næsset et al. (2004) and references therein.

Height–diameter models are keys in conventional forest inventories. In conventional forest sampling, i.e., ground-based measurements of tree variables within sample plots, diameter at breast height (*dbh*) is measured for all the trees within plots, whereas height (*h*) is measured only for a subsample of trees within plots, because *h* is more difficult

* Corresponding author. School of Forestry and Environmental Studies, Yale University, 360 Prospect Street, New Haven, CT 06511-2104, USA. Tel.: +1 203 432 9398; fax: +1 203 432 3809.
E-mail address: christian.salas@yale.edu (C. Salas).

0034-4257/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2010.01.020

and expensive to measure. In this setting, models that predict *h* as a function of *dbh* (hereafter referred to as height–diameter models, or *h-dbh* models[1]) are fitted in order later to predict *h* of those trees based on tree *dbh* alone. Several studies have been conducted on *h-dbh* models (e.g. Thorey, 1932; Meyer, 1940; Curtis, 1967; Yuancai & Parresol, 2001), including model comparisons (e.g. Zhang, 1997; Peng et al., 2001), up to models incorporating stand-level variables (e.g. Staudhammer & LeMay, 2000; Sánchez et al., 2003) and random effects (e.g. Lappi & Bailey, 1988; Robinson & Wykoff, 2004; Lynch et al., 2005). Several model forms have been proposed and fitted, van Laar and Akça (p. 122, 2007) list a few.

Modelling efforts using data from laser scanning are mainly at the plot level. Forest canopy height model (or CHM) is a digital elevation map of the top-of-forest canopy (Nelson et al., 1998). A CHM is relatively easy to obtain through airborne laser scanning techniques, e.g., LiDAR, and methods for calibrating it have been advised. Several efforts for estimating plot-level tree heights derived from LiDAR have been conducted (e.g., Popescu et al., 2002; Popescu & Wynne, 2004). As pointed out by Næsset et al. (2004), who reviewed the Nordic experience in laser scanning, the main objective especially in Norway has been to develop methods that are directly suited for practical forest inventory at the stand level. In this setting, regression models are fitted using laser-derived variables and field plot variables, in order to predict stand variables (plot-wise). The stand variables of chief interest in this regard, according to Næsset et al. (2004) have been: mean tree height, basal area, and stand volume. Other variables have been also predicted such as biomass and carbon (e.g., Nelson et al., 2004), diameter and basal area distribution (e.g., Næsset & Gobakken, 2004), and number of trees (e.g., Hudak et al., 2006, 2008).

In a laser scanning-based forest inventory framework there is a shift from *h-dbh* to *dbh-h* models. When trees are measured on the ground, their positions can be linked with laser measurements (Persson et al., 2004), through a process known as segmentation, and the trees identified from the segmentation, are known as linked trees. Individual heights of the linked trees can be roughly estimated from laser scanning data by using the approximated ground level and the highest hit of the treetop (Rönnholm et al., 2004). Nevertheless, tree diameter is not available from LiDAR data, but it is a crucial variable in order to analyze the forest structure. Analogous to the need in ground-based forest inventory to develop *h-dbh* models to predict *h* of non-measured trees, in a laser-based forest inventory we need *dbh-h* models to predict *dbh*. More generally LiDAR-derived variables, not just *h*, can be used as predictor variables.

Spatial correlation especially affects the statistical inference of fitted models. When data are spatially collected, it is likely that points close together have more similar values (or different) of variables than points that are farther apart (Schabenberger & Gotway, 2005; Ives & Zhu, 2006). Due to this, residuals of models being fit using OLS would likely be autocorrelated. As pointed out by Kissling and Carl (2008), the presence of spatial autocorrelation is problematic for classical statistical tests because these methods assume independently distributed errors. Effective sample size decreases as the correlation between observations increases (Schabenberger & Gotway, 2005). Ignoring spatial correlation when selecting covariates for inclusion in models can lead to the exclusion of relevant covariates in the model (Hoeting et al., 2006). Models which incorporate spatial correlation allow for the direct incorporation of the sampling design, using model-based inference, in the modelling by accounting for the hierarchical dependences between members of a population (Cressie et al., 2009). Further references on applications mixing hierarchical modelling and sampling design can be found in Hoeting (2009).

[1] Height–diameter models are sometimes referred to as diameter–height models (e.g., Schröder & Álvarez, 2001; Zhang & Shi, 2004). We prefer to be consistent with mathematics, and reserve the left word in the couplet to identify the response variable, as in Henry & Aarssen (1999).

Models that explicitly use spatial information are not common in forest inventories. When a traditional regression model is fitted with ordinary least squares (OLS), beyond the implicit spatial arrangement of the data, spatial correlation plays no role in the modelling process (Brunsdon et al., 1998). The advantages of taking spatial information into account using ground-based measurements for modelling *h-dbh* (Zhang & Shi, 2004), tree basal area growth-*dbh* (Zhang et al., 2004), and tree crown area-*dbh* (Zhang & Gove, 2005; Zhang et al., 2005) have been examined. In a laser scanning context, the works on modelling *dbh* from LiDAR variables have not taken spatial information into consideration. For example, Hyyppä et al. (2001) built a model, using OLS that predict tree diameter *dbh* as a function of height (*lh*) and tree crown diameter derived from LiDAR (*lcd*). They only compared the forest inventory estimates of the mean height obtained using that function versus the same parameter estimates using ground-data, however, they did not report any statistics or comparison of the fitted model. The same model was later fit by Schardt et al. (2004) and García et al. (2007) with good results, but again not offering greater details on the performance of the model. Both Persson et al. (2004) and Heurich et al. (2004) proposed slightly different models using the same predictor variables. Persson et al. (2004) reported the fit statistics of the model, and obtained errors of around 10% in predicting *dbh*. In recent studies, Heurich and Thoma (2008) and Heurich (2008), both height and density-related laser variables (e.g., height percentiles and total penetration rate. Næsset, 2002), have been incorporated in models for predicting *dbh*.

Testing the specification of the spatial regression models, has not been a straightforward endeavor. Brunsdon et al. (1999b) and Leung et al. (2000) compared the use of OLS and geographically weighted regression (GWR) using a parametric test, a *F*-test based on the residual sum of squares of each model, and maximum likelihood-based tests have been later proposed (e.g., Fotheringham et al., 2002) too. However, as we will shall later argue, those tests are not valid because they are suitable for parametric models, but not for non-parametric ones such as GWR. Zhang and Shi (2004) and Zhang et al. (2004) compared the use of OLS and GWR using the *F*-test, as described in Leung et al. (2000). Later, comparisons of GWR were extended not only to OLS but also to other models (e.g., Zhang & Gove, 2005; Zhang et al., 2005, 2008) such as linear mixed-effects models (LME). Nevertheless, in the straight statistical sense of the model, their LME models are not actually mixed-effects models, because no random effects coefficients were included in their models. The statistical model that they were fitting was a generalized linear model with a non-null correlation structure. Brunsdon et al. (1999a) compared GWR and a model with random coefficients (in GWR, parameter coefficients are not assumed to be random, Brunsdon et al., 1999a), which they fitted in a Bayesian framework.

Reasoning that statistical models that use spatial information may offer both better prediction and a sounder statistical framework than using statistical models that do not use spatial information, we compare the performance of alternative models fitted using ordinary least squares (OLS), generalized least squares with spatial correlation structure (GLS), linear mixed-effects model (LME), and geographically weighted regression (GWR), for modelling the spatial variation in tree diameter at breast height as a function of laser-derived variables relationship in forest stands.

2. Methods

2.1. Data

2.1.1. Field data

The study area is located in the municipality of Aurskog, situated in the south-eastern Norway (59°80′ N, 11°55′ E, 172–388 m a.s.l.). The sample plots, 36 of 1000 m^2, were located in heterogeneous managed forest conditions regarding the tree species composition, site qualities,

and development classes. The main three species were Norway spruce [*Picea abies* (L.) Karst.], Scots pine (*Pinus silvestris* L.), and birch (*Betula* sp.). At the plot level, the species composition by volume ranged from 0–98% for spruce, 0–100% for pine, and 0–34% for birch. The terrain across the study area is gentle compared to average terrain conditions for productive forests in Norway. Nonetheless the local topography varied significantly among plots. The plot center coordinates (i.e., the x and y positions) were determined by differential dual-frequency Global Positioning System (GPS) and Global Navigation Satellite System (GLONASS) measurements, using two Topcon dual-frequency receivers. On each plot, all trees with $dbh \geq 5$ cm were callipered, and the stem locations were mapped using a Sokkia SET5F total station. The tree species and vegetation status were recorded for each measured tree. Total tree height (h) was measured using a Vertex III hypsometer on reference trees selected from across all plots.

2.1.2. LiDAR data

The LiDAR data were obtained under leaf-on canopy conditions on June, 2006. Five non-overlapping strips were flown from an average flying altitude of 800 m, using a PA31 Piper Navajo (LN-LAS) airplane. The strips were E-W oriented, with an N-S spacing of ca. 8.7 km. Data were acquired with an Optech ALTM 3100 system, operating with a pulse repetition frequency of 100 kHz, a scanning frequency of 70 kHz, and using a half-angle of 5°. Laser data were processed by the contractor Blom Geomatics A.S. After processing, the average values of laser footprint diameter, along track point spacing and across track point spacing were of ca. 0.2, 1.1 and 0.1 m, respectively. The GPS/INS processing was performed using Applanix POSPac software, and then planimetric coordinates (x and y) and ellipsoidal heights of the first and last return echoes were obtained from the navigation and laser data using Optech REALM tools. The 3D adjustment and classification were performed using TerraScan and TerraModel software from TerraSolid Ltd. The TIN model representing the terrain surface was subtracted from the height values of all the echoes recorded, to obtain the heights above the surface of the echoes, with an accuracy of ± 10 cm.

2.1.3. Individual tree detection

Only the first returns located inside buffers extending 10 m around the plot borders were retained for the canopy height models (CHM) creation. The average density of the first echoes within the buffers after data processing was 7.3 pulses/m^2. For each plot, a CHM was obtained by gridding the irregularly spaced laser returns into a 0.35 m regular grid. For single tree delineation, we followed the procedure described by Hyyppä et al. (2001). The CHM was smoothed using the low-pass Gaussian filter described by Hyyppä et al. (2001). Regional maxima were identified in the smoothed CHMs by searching in 8-connected neighborhoods for finding the pixels having higher values than their external boundary neighbors. A marker-based watershed algorithm was employed for individual tree crown delineation. Individual tree variables (tree height, crown width, and stem location) were obtained from the point clouds located inside the delineated segments. The tree height and stem location were obtained directly from the echo with the highest z-value within each segment, which was considered to represent the top of the delineated tree. The polygons representing the crown projections of the delineated trees on the ground were approximated using the convex hull of the (x, y)-coordinates of the returns located within each segment. Thus, the crown widths were estimated as the diameter of a circle with the same area as the crown polygons. Tree height and crown width are underestimated with 0.6 m (Standard deviation, SD, of 0.6 m) and 0.4 m (SD 1.1 m), respectively, while the stem location error is 0.35 m. In total, we have 1438 trees with both field-based and LiDAR-derived variables (Table 1). The minimum number of sample trees per plot was 22, which is a reasonable number of trees for fitting models to predict *dbh* with LiDAR-derived variables as covariates on a per plot basis.

Table 1
Descriptive statistics for 1438 trees. *dbh* is diameter at breast height, *h* is total height, *lh* and *lcd* are LiDAR derived height and crown diameter, respectively.

Statistic	*dbh* (cm)	*h* (m)	*lh* (m)	*lcd* (m)
Minimum	5.0	5.0	3.4	0.3
Maximum	49.5	35.8	29.6	4.2
Mean	19.2	15.2	15.1	1.5
Std. dev	7.8	4.7	4.8	0.7

2.2. Statistical models

Slight variations have been proposed for *dbh*-LiDAR derived variable models. Hyyppä et al. (2001), Schardt et al. (2004), and García et al. (2007) used the following mean function

$$\mathrm{E}[dbh] = \beta_0 + \beta_1 lh + \beta_2 lcd, \tag{1}$$

where β_0, β_1, and β_2 are the coefficients to be estimated, *lh* is the laser-derived height, and *lcd* is the laser-derived tree crown diameter. In the sequel, we shall use Eq. (1) as the mean function in various statistical models. Our aim is to compare a variety of ways to incorporate spatial information in statistical models for *dbh*. Rather than pool data from all plots together, we fit the various models to data from each plot separately, thereby avoiding having to account for variations in site quality, aspect, slope, and other topographical features among plots. We largely adopt the statistical notation of Schabenberger and Gotway (2005) and Pinheiro and Bates (2000), where a response and predictor variables y and xs have been recorded at spatial locations $\mathbf{s}_i$, ..., $\mathbf{s}_n$, i.e., specific geographic locations for each ith observation with spatial coordinates $\mathbf{s}_i = [x_i, y_i]'$, represented by $\mathbf{y}(\mathbf{s})$ and $\mathbf{X}(\mathbf{s})$, respectively. The additive error term is represented by $\mathbf{e}(\mathbf{s})$.

2.2.1. Model 1

The first statistical model is as follows

$$\begin{aligned} &dbh(\mathbf{s}_i) = \beta_0 + \beta_1 lh(\mathbf{s}_i) + \beta_2 lcd(\mathbf{s}_i) + e(\mathbf{s}_i), \\ &e(\mathbf{s}_i) \sim \mathcal{N}\left(0, \sigma^2\right), \\ &\sum \left[e(\mathbf{s}_i), e\left(\mathbf{s}_j\right)\right] = \sigma^2 I_n. \end{aligned} \tag{2}$$

where $dbh(\mathbf{s}_i)$, $lh(\mathbf{s}_i)$, and $lcd(\mathbf{s}_i)$ are *dbh*, *lh*, and *lcd*, recorded at spatial location $\mathbf{s}_i$, respectively. Errors of the model at location $\mathbf{s}_i$ are represented by $\boldsymbol{e}(\mathbf{s}_i)$, and they are assumed normally distributed with mean 0 and variance σ^2. β_0, β_1, and β_2 are coefficient parameters of the model. $\sum[e(\mathbf{s}_i), e(\mathbf{s}_j)]$ is the covariance matrix of the errors, and $\mathbf{I}_n$ is an identity matrix of size n (the sample size).

Eq. (2) can be more generally represented in matrix form as follows,

$$\mathbf{y}(\mathbf{s}) = \mathbf{X}(\mathbf{s})\beta + \mathbf{e}(\mathbf{s}) \tag{3}$$

where $\mathbf{y}(\mathbf{s})$ is the vector with the response variable recorded at location $\mathbf{s}$, $\mathbf{X}(\mathbf{s})$ is the matrix of predictor variables spatially recorded, β is a vector with the coefficient parameters of the model, and $\mathbf{e}(\mathbf{s})$ is the vector of random errors.

With ordinary least squares (OLS), the coefficient vector is estimated by

$$\hat{\beta}_{\mathrm{OLS}} = \{\mathbf{X}(\mathbf{s})'\mathbf{X}(\mathbf{s})\}^{-1}\mathbf{X}(\mathbf{s})'\mathbf{y}(\mathbf{s}). \tag{4}$$

Model 1 does not incorporate any kind of spatial information. As pointed out by Schabenberger and Gotway (2005, p. 316), model 1 is a spatial regression model since the response variable $\mathbf{y}(\mathbf{s})$, and the

predictor variables comprising $\mathbf{X}(\mathbf{s})$, are recorded at spatial locations. However, the spatial information serves only to link $\mathbf{y}(\mathbf{s})$ and $\mathbf{X}(\mathbf{s})$; there is nothing in the model that explicitly considers spatial pattern or spatial relationships.

2.2.2. Model 2

The second model differs from Eq. (2) only in its specification of the error covariance. For this model

$$\sum\left[\mathbf{e}(\mathbf{s}_i),\mathbf{e}\left(\mathbf{s}_j\right)\right]=\sigma^2\boldsymbol{\Omega}, \tag{5}$$

where Ω is the correlation matrix, the elements of which specify the spatial correlation between all pairs of trees on a plot. If the correlation parameters in Ω were known, then the generalized least squares (GLS) estimator

$$\hat{\beta}_{\mathrm{GLS}}=\left\{\mathbf{X}(\mathbf{s})'\boldsymbol{\Omega}^{-1}\mathbf{X}(\mathbf{s})\right\}^{-1}\mathbf{X}(\mathbf{s})'\boldsymbol{\Omega}^{-1}\mathbf{y}(\mathbf{s}) \tag{6}$$

would be feasible. Usually the parameters in Ω must be estimated from data or auxiliary information. Using $\hat{\Omega}$ as the estimated correlation matrix (see below), the estimated GLS estimator of β is

$$\hat{\beta}_{\mathrm{EGLS}}=\left\{\mathbf{X}(\mathbf{s})'\hat{\boldsymbol{\Omega}}^{-1}\mathbf{X}(\mathbf{s})\right\}^{-1}\mathbf{X}(\mathbf{s})'\hat{\boldsymbol{\Omega}}^{-1}\mathbf{y}(\mathbf{s}). \tag{7}$$

As pointed out by Cressie (1993) and exemplified by Pinheiro and Bates (2000), spatial correlation structures are generally represented by their semivariogram, instead of their correlation function. We derived $\hat{\Omega}$ for each plot after deciding the following: (a) the functional form of the semivariogram, (b) the distance metric to be used, and (c) whether the semivariogram would include a nugget effect. Regarding (a), we assessed the following semivariogram functions; exponential, Gaussian, linear, rational quadratic, and spherical. All of them are fully described in Pinheiro and Bates (2000, p. 230–233), as well as further details on them can be found in spatial statistics books (Waller & Gotway, 2004; Schabenberger & Gotway, 2005; Diggle & Ribeiro, 2007). Regarding (b), we assessed the following distance metrics between the x and y coordinates among trees: Euclidean, maximum, and Manhattan. Finally, regarding (c), we assessed the effect of adding a nugget to the semivariogram function. We fit model 2 and compared all the possible combinations of semivariogram function-distance metrics-nugget effect, and selected the best model using the Akaike Information Criterion (AIC, Akaike, 1973). A referee suggested that a Matern type of correlation function would better capture a nugget effect, however, we did not use this function in our study, but did add nugget effect.

2.2.3. Model 3

The third model adds random effects to the fixed effect coefficient parameters:

$$\begin{aligned}&dbh(\mathrm{s}_i)=(\beta_0+b_{0i})+(\beta_1+b_{1i})lh(\mathbf{s}_i)+(\beta_2+b_{2i})lcd(\mathbf{s}_i)+e(\mathbf{s}_i),\\&\mathbf{b_i}=(b_{0i},b_{1i},b_{2i})'\sim\mathcal{N}(0,\mathbf{D}(\phi)),\\&e(\mathbf{s}_i)\sim\mathcal{N}\left(0,\sigma^2\right),\\&\sum\left[e(\mathbf{s}_i),e\left(\mathbf{s}_j\right)\right]=\sigma^2\boldsymbol{\Omega},\end{aligned} \tag{8}$$

where $\mathbf{D}(\phi)$ is a 3×3 random effects covariance matrix parameterized by ϕ. The inclusion of a vector of random effects permits the model to be individualized to each tree, while also permitting a pooling of all the data when fitting the model. Importantly, the random individual effects induce an intra-individual correlation structure that accounts for the lack of independence among trees on the same plot. It is this feature that has a great appeal for modelers of correlated data, even though the induced correlation function may not be easily discernible (Schabenberger & Gregoire, 1996). Further details on mixed-effects models can be found in Pinheiro and Bates (2000) and Schabenberger and Pierce (2002), as well as applications in a spatial context in Schabenberger and Gotway (2005) and in Bivand et al. (2008).

In matrix notation, Eq. (8) may be written as

$$\mathbf{y}(\mathbf{s})=\mathbf{X}(\mathbf{s})\beta+\mathbf{Z}(\mathbf{s})\mathrm{b}+\mathbf{e}(\mathbf{s}). \tag{9}$$

The flexibility permitted by a mixed model of this sort comes at the price of added complexity, both statistical and computational. In practice it is often found that data do not support a random effect for each fixed effect parameter, and also that when the correlation structure is modelled by $\mathbf{D}(\phi)$, it is often sufficient to specify $\Omega=\mathbf{I}_n$.

Customarily β in Eq. (9) is estimated by maximum likelihood or restricted maximum likelihood. Followed by the best linear unbiased predictors ("BLUP", Robinson, 1991) of the random effects. We used the nlme package (Pinheiro et al., 2008) implemented in R for fitting LME models.

Having fitted a LME model, the mean response at x can be estimated by $x\hat{\beta}$, where x is the vector of covariate values. An individual observation can be predicted by adding the corresponding BLUP of the random effect.

In model 3 we are not directly incorporating spatial information. The random effects in Eq. (8), show that we are estimating an overall (i.e., fixed) mean, and allowing for local deviations from this mean. These deviations are random, in our case normally distributed.

2.2.4. Model 4

The fourth model allows the coefficients to vary by spatial location

$$\begin{aligned}&dbh(\mathbf{s}_i)=\beta_0(\mathbf{s}_i)+\beta_1(\mathbf{s}_i)lh(\mathbf{s}_i)+\beta_2(\mathbf{s}_i)lcd(\mathbf{s}_i)+e(\mathbf{s}_i)\\&e(\mathbf{s}_i)\sim\mathcal{N}(0,\sigma^2),\\&\sum\left[e(\mathbf{s}_i),e\left(\mathbf{s}_j\right)\right]=\sigma^2\mathbf{I}_n,\end{aligned} \tag{10}$$

where $\beta_0(\mathbf{s}_i)$, $\beta_1(\mathbf{s}_i)$, and $\beta_2(\mathbf{s}_i)$ represent the coefficient parameters for the ith observation, the other notation was previously defined. Model 4 can be fitted using geographically weighted regression (GWR), a non-parametric regression method that has been portrayed by Fotheringham et al. (2002), Brunsdon et al. (1996, 1998) and Fotheringham et al. (1998).

In GWR a vector of coefficient parameters is required for each observation, that is to say, each observation has different coefficient parameters, therefore, having a local feature. Sometimes, only a portion of the coefficients of a GWR model can be local and the other ones can remain global, which has been called "mixed-GWR" (Fotheringham et al., 2002, p. 65–73). The estimated parameter vector for each observation in GWR (Brunsdon et al., 1998) is obtained using weighted least squares regression, as follows

$$\hat{\beta}_{i\mathrm{GWR}}=\left\{\mathbf{X}(\mathbf{s})'\mathbf{W}(\mathbf{s}_i)^{-1}\mathbf{X}(\mathbf{s})\right\}^{-1}\mathbf{X}(\mathbf{s})'\mathbf{W}(\mathbf{s}_i)^{-1}\mathbf{y}(\mathbf{s}), \tag{11}$$

where $\mathbf{W}(\mathbf{s}_i)$ is the diagonal weight matrix for the ith observation (i.e., tree). This matrix in general is defined as $\mathbf{W}(\mathbf{s}_i)=\mathrm{diag}\{W(s_i,\ s_j)\}$, where the scalar weight $W(s_i, s_j)$ for the j-th observation depends on its proximity to the ith observation. A simple scheme of weighting would be to exclude from the model fitting observations beyond a distance b from the ith observation, which is termed the moving window approach (Fotheringham et al., 2002). With this approach the spatial weighting has the problem of discontinuity in the parameter estimates (Brunsdon et al., 1998; Schabenberger & Gotway, 2005). An alternative approach is to use a continuous spatial weighting function, as in general represented by

$$W\left(\mathbf{s}_i,\mathbf{s}_j\right)=f\left(\mathrm{dist}_{(\mathbf{s}_i,\mathbf{s}_j)},b\right), \tag{12}$$

where f is a kernel function with bandwidth b. With Eq. (12) both the kernel and the bandwidth need to be stipulated. We assessed a Gaussian and a near-Gaussian kernels as spatial weighting functions and selected the one that achieves a lower AIC[2] of the GWR model. The bandwidth is the searching radius around each observation and is a key in GWR. The bandwidth can greatly affect the properties of the $\hat{\beta}_{iGWR}$ (Brunsdon et al., 1998). It governs the extent to which the resulting local calibration results are smoothed (Fotheringham et al., 2002). We used a fixed bandwidth of 6 m for all the plots, determined based on the range of the empirical semivariogram (i.e., distance in which the semivariogram reach the sill) of *dbh* per plot, and is consistent with the studies by Zhang et al. (2004) and Pukkala (1989). We used the spgwr package (Bivand & Yu, 2009) implemented in R for fitting GWR models.

2.3. Comparing models

We first computed the studentized model residuals and plotted them in a spatial fashion. Zhang and Gove (2005) called this type of residuals "local *Z*-values". These values would allow us to compare the magnitudes of the residuals in space. Notice that we used the studentized residuals definition given by Schabenberger and Pierce (2002, Eq. (4.23)), and in order to compute this type of residual for GWR, we used the **L** matrix of Leung et al. (2000, Eq. (16)) as the hat matrix.

2.3.1. Goodness-of-fit

For each model (OLS, GLS, LME, and GWR), we computed the following statistics in order to assess the goodness-of-fit of the tested models: (a) the aggregated difference (*AD*) or mean residual, computed as follows

$$AD = \frac{1}{n}\sum_{i}^{n} r_i, \tag{13}$$

where $r_i = dbh(\mathbf{s}_i) - \widehat{dbh}(\mathbf{s}_i), dbh(\mathbf{s}_i)$, is the observed tree diameter at the ith location, $\widehat{dbh}(\mathbf{s}_i)$ is the predicted tree diameter at the $\mathbf{s}_i$ location, and n is the sample size in the field plot being analyzed. For the LME model, the prediction includes the BLUP of the random effects. (b) the root mean square differences (*RMSD*) as follows,

$$RMSD = \sqrt{\frac{1}{n}\sum_{i}^{n} r_i^2}, \tag{14}$$

and (c) the aggregated of the absolute value differences (*AAD*), as follows

$$AAD_j = \frac{1}{n}\sum_{i}^{n} |r_i|. \tag{15}$$

We fitted and validated the models using the same observations (on a per plot basis). We deliberately avoided the use of validation or cross-validation, because withholding observations alters the pattern of spatial correlation of the observations. We preferred to use the term *RMSD* rather than *RMSE* (root mean squared error) following Stage & Crookston (2007) and Hudak et al. (2008).

At this stage, we assess the models exclusively according to their goodness-of-fit. We follow the methodology for selecting models proposed by Salas (2002). The methodology is as follows. First, partition the data into validation classes. We use the percentiles of the variable *lh* in order to define five validation classes, with similar number of observations in each class. Second, compute the statistics detailed above (*AD*, *RMSD*, and *AAD*) within validation classes. Third, compute the mean and standard deviation of among-class values of these statistics, therefore having an among-class mean and an among-class standard deviation for *AD*, and so on). Fourth, assign ranking to the among-class mean and among-class standard deviation of each statistic. Ranking score 1 is assigned to the model with lowest among-class mean (or closer to 0 for *AD*) and lowest among-class standard deviation (which indicates more stable results across validation classes). Fifth, add all these ranking scores (six in total, two scores by three statistics) and sort the models according to this sum (no weighting of statistics was considered), the model with the lowest sum would be rank as the best. The methodology aims to rank best not only the models with overall better precision and accuracy but also with a consistent behavior across validation classes.

2.3.2. Statistical inference

Once we have maximized the log-likelihood function of a parametric model, it is straightforward to compare models using indices such as; Akaike Information Criterion (AIC, Akaike, 1973) or the corrected AIC of Hurvich et al. (1998), among others. However, this is possible only for parametric models, in our study; OLS, GLS, and LME[3]. Statistical inference for GWR offers several challenges. Although we could think of building a likelihood function for GWR (e.g., Eq. (42) in Fotheringham et al., 2002), the actual way in which the parameters are estimated does not involve any optimal solution for this likelihood function, because all the coefficients are fitted independently from one observation to another. The log-likelihood in Fotheringham et al. (2002) for GWR permits an infinity of maxima. Non-parametric statistical tests provide a more suitable framework for assessment of GWR. Permutation tests, also called randomization (Manly, 2006) or re-randomization (Welch, 1987) tests, offer a non-parametric simple and straightforward technique for conducting hypothesis testing among statistics of two groups. In order to conduct the permutation tests, we first defined the statistics that would allow us to compare the models, which we will denote as $\hat{\theta}$, as follows

$$\hat{\theta} = \frac{1}{m}\sum_{j=1}^{m} \left(RMSD_{aj} - RMSD_{bj}\right) \tag{16}$$

where m is the total number of sample plots, $RMSD_{aj}$ and $RMSD_{bj}$ are the root mean square differences (Eq. (14)) in % of models a and b, respectively, for the j-th plot. We set model a to be the base one and model b to be the new model, or the one with lower *RMSD*. Our hypothesis under test is H_0: $\mu_{RMSD_a} - \mu_{RMSD_b} = 0$, vs. H_a: $\mu_{RMSD_a} > \mu_{RMSD_b}$. The following step was to randomly assign (without replacement) m values of *RMSD*s to model a and the rest to model b, and based on this new sample (called "permutation sample" by Hesterberg et al., 2003), we computed $\hat{\theta}$ (Eq. (16)) for the permutation sample. The next step is to repeat this resampling many times, we used 10,000 permutations. The distribution of the statistic from these resamples forms the sampling distribution under the condition that H_0 is true. Finally, we computed the proportion of the distribution which exceeded the difference observed between the two models (i.e., $\hat{\theta}$ using the original sample), which is the estimate of the *P*-value for the one-sided test, given the null hypothesis is true. We conducted this simple two-sample permutation test procedure between every pair of models. We also used other variants of non-parametric tests, but with all of them we achieved the same conclusions.

[2] We used the AIC of GWR as define in Fotheringham et al. (2002).

[3] A referee raised the issue about the appropriate number of degrees of freedom for subject-specific inference in mixed models, which is still in debated among statisticians (Bates, 2006). The conventional literature, as embodied by Pinheiro and Bates (2000) is incorporated in LME, yet Hodges and Sargent (2001) and Vaida and Blanchard (2005) make a case for an alternative assessment. Further, Bates (2006) questions the appropriateness of the *F* distribution for inference altogether. The issues are beyond the scope of this paper to address, however, and while appreciative of having the matter brought to our attention by the referee, we have chosen to present our results following currently accepted practice in this regard.

3. Results

Most of our plots showed a random spatial pattern of trees, although a few exhibited a somewhat regular spatial pattern. We fit a smooth curve to the empirical semivariogram in order to depict a general trend of the autocorrelation of *dbh*, which shown spatial correlation but only at small distances. Overall, the forest was fairly homogenous with respect to the spatial distribution of trees at the size of our plots.

A rather simplified version of model 3 was sufficient for our data. Although our base formulation of model 3 (Eq. (8)) allowed both the intercept and the slopes coefficient parameters to have a random-effect, we found better fit, as measured by AIC, when having only the intercept random-effects. That is, the final model 3 used was $dbh(\mathbf{s}_i) = (\beta_0 + b_{0i}) + \beta_1 lh(\mathbf{s}_i) + \beta_2 lcd(\mathbf{s}_i) + e(\mathbf{s}_i)$. Furthermore, we did not find significant improvement by imposing a marginal correlation, therefore, we set $\Omega = \mathbf{I}_n$ in Eq. (8).

Compared to the OLS fit, the studentized residuals of the GWR model changed both in magnitude and spatial distribution. The other statistical models do not change the spatial pattern of the residuals, which is in agreement with the findings of Zhang and Shi (2004) and Zhang et al. (2004). However, LME produces much smaller studentized residuals than the other models (Fig. 1c).

Both goodness-of-fit and permutation tests portrayed LME as the best assessed model. There is a clear difference between the *RMSD*, computed based on all the trees of each plot, of LME and the rest of the models (Fig. 2). *AD* and *AAD* by validation classes point again to LME as the model with a more consistent predictive performance across all validation classes. In the same manner, the proportion of field plots (i.e. cases) where each model was assigned to be the best, according to the proposed methodology, portrait to LME as the best in almost all (95%) the field plots used in our study (Table 2). Permutation tests

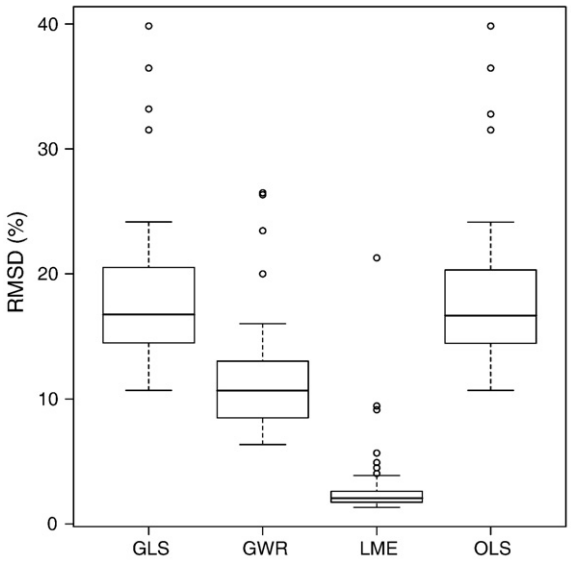

Fig. 2. Box plot of the root mean square differences (*RMSD*) for predicting tree diameter in a field plot (i.e., each data point represents a *RMSD* for a plot) from the assessed statistical models. GLS is generalized least squares with correlation structure, GWR is geographically weighted regression, LME is linear mixed-effects, and OLS is ordinary least squares.

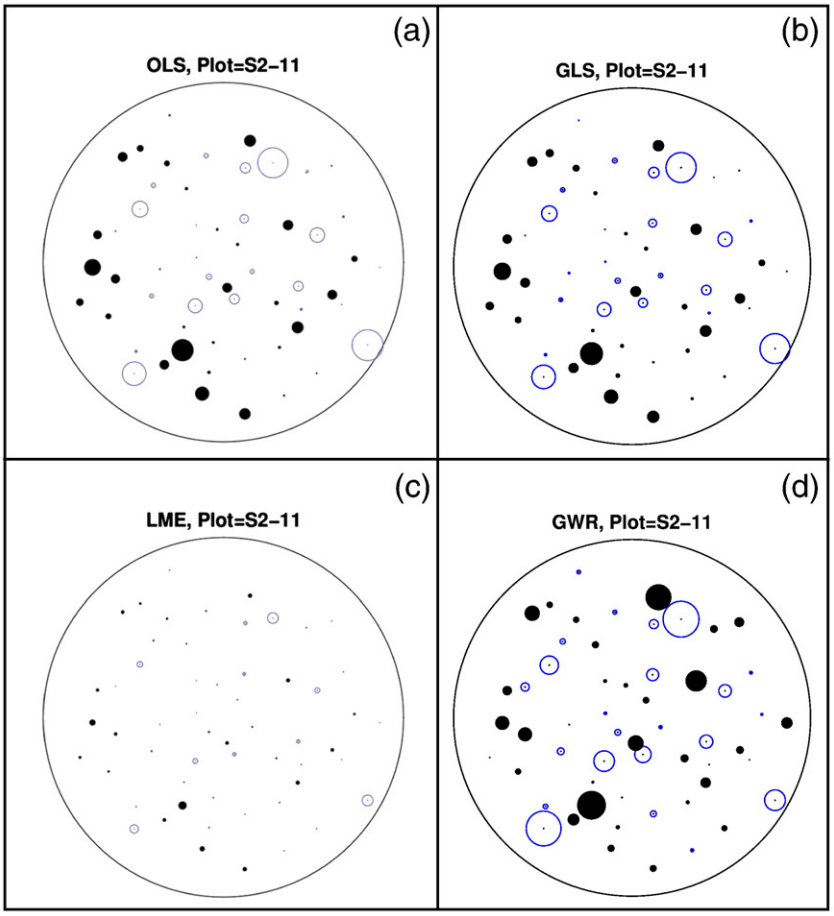

Fig. 1. Graph of studentized model residuals (local *Z*-values) from the four modelling techniques [OLS (a), GLS (b), LME (c), and GWR (d)] when predicting diameter in a sample plot. The black dots represent positive local *Z*-value, and the circles represent negative local *Z*-value. Both types of circles are proportional to the absolute value of the studentized residuals.

Table 2
Percentage of the plots (i.e., cases) where each model was chosen as the best according to its goodness-of-fit.

Model	Position			
	1st	2nd	3rd	4rd
GLS	0.0	2.8	44.4	52.8
GWR	5.6	86.1	5.6	2.8
LME	94.4	5.6	0.0	0.0
OLS	0.0	5.6	50.0	44.4

assessment of the *RMSD* between pairs of models showed significant differences. Statistical significance comparisons of the models did favor the ones with superior goodness-of-fit. There is indeed an ordering of the models, not only considering the pure *RMSD* but also its statistical significance, which is LME preferred to GWR preferred to OLS (Fig. 3). Nevertheless, we did not find strong evidence against the null hypothesis of no differences between the *RMSD*s of GLS and OLS. This is expected because GLS really affects (or improves) the estimates of the standard errors of the coefficient estimates, over the ones obtained using OLS.

4. Discussions

The linear mean function (Eq. (1)) seems to be appropriate for these data. We obtained a *RMSD* of 18% for our model fit by OLS. Among other studies with models having the same predictor variables: Persson et al. (2004) obtained a 10% error, computed over all their sample trees, but with a much smaller sample (562 trees) and a smaller diameter range too; and unfortunately neither Heurich et al. (2004) nor Schardt et al. (2004) and García et al. (2007) reported their *dbh* model errors. Heurich (2008), using models with more predictor

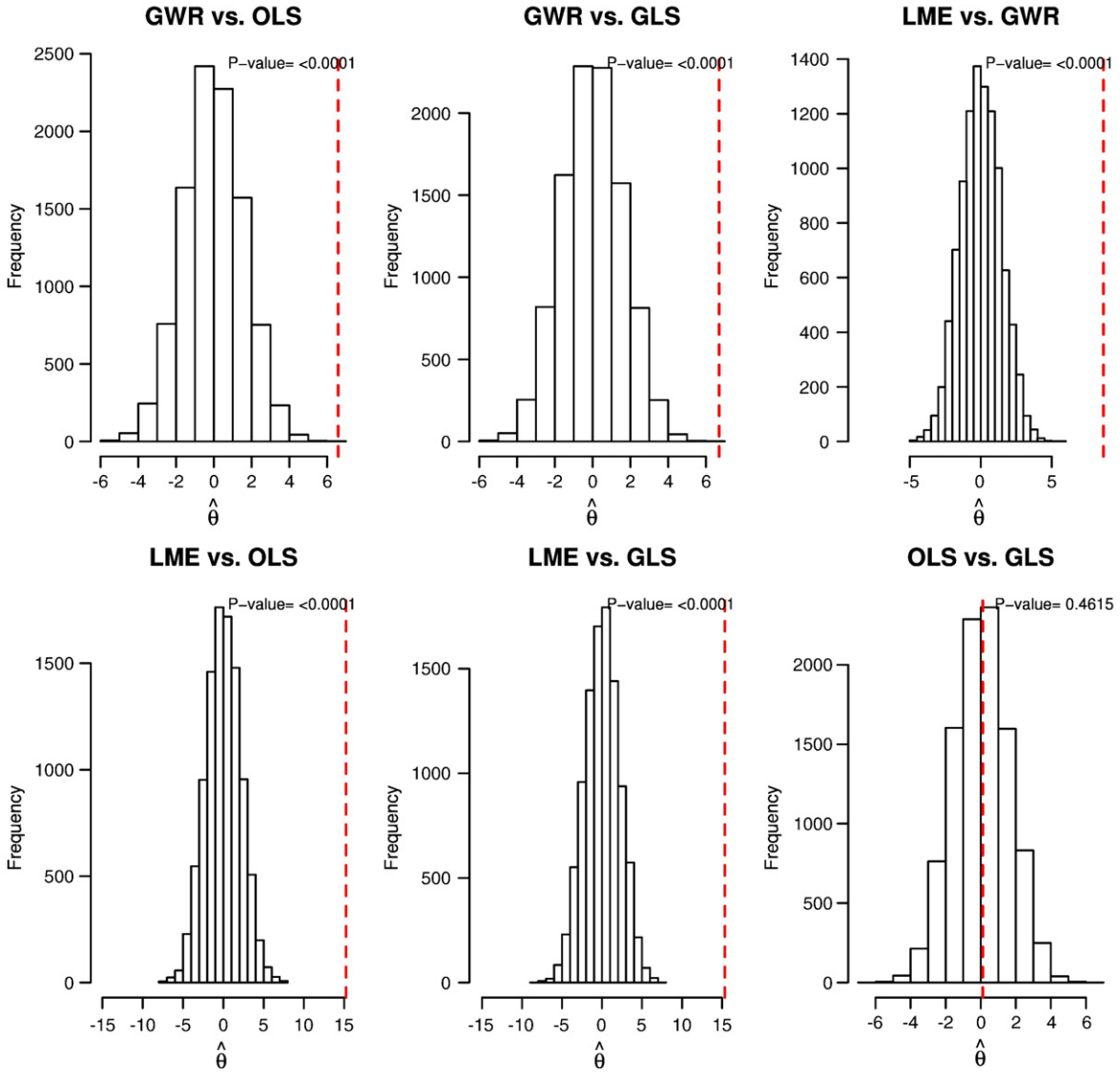

Fig. 3. Permutation distribution of $\hat{\theta}$ (i.e., the mean difference between *RMSD*s on predicting tree height from OLS, GLS, LME, and GWR models) using a simple two-sample permutation test. The vertical line shows the statistics for our dataset. 10,000 permutations were conducted. The empirical one-sided *P*-value for the significance of the statistic is also printed.

variables, obtained a *RMSD* of 19.9%[4] for his *dbh* models in 936 trees. Therefore we think our mean function performs reasonably well, and is appropriate as a baseline model.

We obtained the best precision and accuracy in predictions using LME. Prediction errors with LME are much lower (4%, Fig. 2) than what can be achieved having a rather overparametrized model such as GWR (10%), where a full model is fitted for each observation. Furthermore, the *RMSD*s of the LME model were significantly different from all other models considered. GLS models the spatial correlation of the residuals, its prediction errors are not significantly different from the one obtained from a simpler OLS model. However, standard error estimates under OLS are biased, which raises doubts about the correctness of statistical inference following OLS estimation. Finally, LME also offers us a statistical pathway to further explore the relationship being modelled.

The use of GWR in practical applications is unclear. Although several studies have portrayed GWR as performing well and having a series of advantages over other models (e.g., Zhang & Shi, 2004; Zhang et al., 2004; Zhang & Gove, 2005; Zhang et al., 2005), few studies have focused on how a GWR model could be used to predict *Y* for a new observation which was not used in fitting the model. One way for doing so involves to have the entire dataset used for fitting the model (which is complicated in practice), and this prediction is done as in a cross-validation setting (e.g., see the explanation of GWR prediction in Leung et al., 2000, p. 27). Indeed, alternatives such as interpolating in a map given the GWR predictions for known points or interpolating the parameter estimates (which is the final aim of GWR), is possible, but the statistical properties of interpolated predictions may be difficult to discern. Furthermore, in forestry, plots may just be too far separated (in space) to permit accurate interpolation of this sort.

5. Concluding remarks

The models we examined all shared a common mean function which included laser height and laser crown diameter in a linear fashion. The objective of this study was to examine alternative ways to include spatial information on tree location that would result in improved model fit. It was intended to be an exploratory initial investigation, rather than a comprehensive exercise to determine the optimal, in some sense, mean function or covariance structure for the data we had at our disposal. Of the models we considered, LME with random intercepts was judged to perform better than the GWR model, a model fitted by GLS with a spatial correlation error term, and an aspatial model fitted by OLS. To our knowledge, this is the first study assessing the utility of incorporating spatial relationships when modelling field-observed response variables and LiDAR extracted predictor variables. Our results show that a mixed-effects model offers the best goodness-of-fit for modelling *dbh* from LiDAR-based variables.

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov, & G. Czaki (Eds.), *Second international symposium on information theory* (pp. 267—281). Budapest, Hungary: Akademiai Kiadó.

Bates, D. M. (2006). *Statistical inferences in linear mixed-effects models.* Zürich, Zwitzerland: Presented at ETH.

Bivand, R., & Yu, D. (2009). spgwr: Geographically weighted regression. *R package version*, 5—7.

Bivand, R. S., Pebesma, E. J., & Gómez, V. (2008). *Applied spatial data analysis with R.* New York, NY, USA: Springer 374 p.

Brunsdon, C., Aitkin, M., Fotheringham, S., & Charlton, M. (1999). A comparison of random coefficient modelling and geographically weighted regression for spatially non-stationary regression problems. *Geographical & Environmental Modelling*, *3*(1), 47—62.

Brunsdon, C., Fotheringham, A. S., & Charlton, M. (1999). Some notes on parametric significance tests for geographically weighted regression. *Journal of Regional Science*, *39*(3), 497—524.

Brunsdon, C., Fotheringham, A. S., & Charlton, M. E. (1996). Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis*, *28*(4), 281—298.

Brunsdon, C., Fotheringham, S., & Charlton, M. (1998). Geographically weighted regression—Modelling spatial non-stationary. *Statistician*, *47*(3), 431—443.

Cressie, N. (1993). *Statistics for spatial data.* New York, USA: Wiley Interscience 928 p.

Cressie, N., Calder, C. A., Clark, J. S., Hoef, J. M. V., & Wikle, C. K. (2009). Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. *Ecological Applications*, *19*(3), 553—570.

Curtis, R. O. (1967). Height–diameter and height–diameter-age equations for second-growth Douglas-fir. *Forest Science*, *13*(4), 365—375.

Diggle, P. J., & Ribeiro, P. J., Jr. (2007). *Model-based geostatistics.* New York, USA: Springer 228 pp.

Fotheringham, A. S., Charlton, M. E., & Brunsdon, C. (1998). Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. *Environment and Planning A*, *30*, 1905—1927.

Fotheringham, S., Brunsdon, C., & Charlton, M. (2002). *Geographically weighted regression: The analysis of spatially varying relationships.* New York, USA: Wiley 269 p.

García, R., Súarez, J. C., & Patenaude, G. (2007). Delineation of individual tree crowns for LiDAR tree and stand parameter estimation in Scotish woodlands. In S. I. Fabrikant, & M. Wachowicz (Eds.), *The European information society: Leading the way with Geo-information* (pp. 55—87). New York, USA: Springer.

Henry, H. A., & Aarssen, L. W. (1999). The interpretation of stem diameter-height allometry in trees: Biomechanical constrainsts, neighbour effects, or biased regressions? *Ecology Letters*, *2*, 89—97.

Hesterberg, T., Monaghan, S., Moore, D. S., Clipson, A., & Epstein, R. (2003). *Bootstrap methods and permutation tests.* New York, USA: W. H. Freeman and Company 74 p.

Heurich, M. (2008). Automatic recognition and measurement of single trees based on data from airborne laser scanning over the richly structured natural forests of the bavarian forest national park. *Forest Ecology and Management*, *255*, 2416—2433.

Heurich, M., Persson, A., Holmgren, J., & Kennel, E. (2004). Detecting and measuring individual trees with laser scanning in mixed mountain forest of central Europe using an algorithm developed for Swedish boreal forest conditions. In M. Thiers, B. Kock, H. Spiecker, & H. Weinacker (Eds.), *Proceedings of the ISPRS working group part 8/2Int. Arch. Photogramm. Remote Sensing, Vol. 36*. (pp. 307—312) Freiburg, Germany. Institute for Forest Growth and Institute for Remote Sensing and Landscape Information Systems, Albert Ludwigs University. Freiburg, Germany.

Heurich, M., & Thoma, F. (2008). Estimation of forestry stand parameters using laser scanning data in temperate, structurally rich natural European beech (*Fagus sylvatica*) and Norway spruce (*Picea abies*) forests. *Forestry*, *81*(5), 645—661.

Hodges, J. S., & Sargent, D. J. (2001). Counting degrees of freedom in hierarchical and other richly parameterized models. *Biometrika*, *88*, 367—379.

Hoeting, J. A. (2009). The importance of accounting for spatial and temporal correlation in analyses of ecological data. *Ecological Applications*, *19*(3), 574—577.

Hoeting, J. A., Davis, R. A., Merton, A. A., & Thompson, S. E. (2006). Model selection for geostatistical models. *Ecological Applications*, *16*(1), 87—98.

Hudak, A., Crookston, N., Evans, J., Hall, D., & Falkowski, M. (2008). Nearest neighbor imputation of specie-level, plot-scale forest structure attributes from LiDAR data. *Remote Sensing of Environment*, *112*, 2232—2245 and corrigendum in 113:289–290.

Hudak, A. T., Crookston, N. L., Evans, J. S., Falkowski, M. J., Smith, A. M. S., Gessler, P., & Morgan, P. (2006). Regression modeling and mapping of coniferous forest basal area and tree density from discrete-return lidar and multispectral satellite data. *Canadian Journal of Remote Sensing*, *32*, 126—138.

Hurvich, C. M., Simonoff, J. S., & Tsai, C. -L. (1998). Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion. *Journal of the Royal Statistical Society Series B*, *60*(2), 271—293.

Hyyppä, J., Kelle, O., Lehikoinen, M., & Inkinen, M. (2001). A segmentation-based method to retrieve stem volume estimates from 3-d tree height models produced by laser scanners. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(5), 969—975.

Ives, A. R., & Zhu, J. (2006). Statistics for correlated data: Phylogenies, space, and time. *Ecological Applications*, *16*(1), 20—32.

Kissling, W. D., & Carl, G. (2008). Spatial autocorrelation and the selection of simultaneous autoregressive models. *Global Ecology and Biogeography*, *17*(1), 59—71.

Lappi, J., & Bailey, R. L. (1988). A height prediction model with random stand and tree parameters: An alternative to traditional site index methods. *Forest Science*, *34*, 907—927.

Leung, Y., Mei, C. L., & Zhang, W. X. (2000). Statistical tests for spatial nonstationarity based on the geographically weighted regression model. *Environment and Planning A*, *32*, 9—32.

Lynch, T. B., Holley, A. G., & Stevenson, D. J. (2005). A random-parameter height-*dbh* model for cherrybark oak. *Southern Journal of Applied Forestry*, *29*(1), 22—26.

Manly, B. F. J. (2006). *Randomization, bootstrap and Monte Carlo methods in biology*, 3rd Edition London, England: Chapman & Hall/CRC 455 p.

Meyer, H. A. (1940). A mathematical expression for height curves. *Journal of Forestry*, *38*(5), 415—420.

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*(1), 88—99.

Næsset, E., & Gobakken, T. (2004). Estimation of diameter and basal area distributions in coniferous forest by means of airborne laser scanner data. *Scandinavian Journal of Forest Research*, *19*, 529—542.

Næsset, E., Gobakken, T., Holmgren, J., Hyyppä, H., Hyyppä, J., Maltamo, M., Nilsson, M., Olsson, H., Persson, A., & Söderman, U. (2004). Laser scanning of forest resources: The Nordic experience. *Scandinavian Journal of Forest Research*, *19*, 482—499.

[4] This value was derived and computed proportional to the sample size of his forest types.

Nelson, R. E., Short, A., & Valenti, M. (2004). Measuring biomass and carbon in Delaware using airborne profiling LiDAR. *Scandinavian Journal of Forest Research*, *19*, 500−511.

Nelson, R. F., Gregoire, T. G., & Oderwald, R. G. (1998). The effects of fixed-area plot width on forest canopy height simulation. *Forest Science*, *44*(3), 438−444.

Peng, C., Zhang, L., & Liu, J. (2001). Developing and validating nonlinear height–diameter models for major tree species of Ontario's boreal forests. *Northern Journal of Applied Forestry*, *18*(3), 87−94.

Persson, Å., Holmgren, J., & Söderman, U. (2004). Detecting and measuring individual trees using an airbone laser scanner. *Photogrammetric Engineering and Remote Sensing*, *68*(9), 925−932.

Pinheiro, J., Bates, D., DebRoy, S., & Sarkar, D.The R Core Team. (2008). *nlme: Linear and nonlinear mixed effects models.* R package version 3.1–89.

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and Splus.* New York, USA: Springer-Verlag 528 p.

Popescu, S. C., & Wynne, R. H. (2004). Seeing the trees in the forest: using lidar and multispectral data fusion with local filtering and variable window size for estimating tree height. *Photogrammetric Engineering and Remote Sensing*, *70*(5), 589−604.

Popescu, S. C., Wynne, R. H., & Nelson, R. E. (2002). Estimating plot-level tree heights with lidar: local filtering with a canopy-height based variable window. *Computers and Electronics in Agriculture*, *37*(1–3), 71−95.

Pukkala, T. (1989). Predicting diameter growth in even-aged scots pine stands with a spatial and non-spatial model. *Silva Fennica*, *23*(2), 101−116.

Reutebuch, S. E., Andersen, H. -E., & McGaughey, R. J. (2005). Light detection and ranging (LIDAR): An emerging tool for multiple resource inventory. *Journal of Forestry*, *103*(6), 286−292.

Robinson, A. P., & Wykoff, W. R. (2004). Imputing missing height measures using a mixed-effects modeling strategy. *Canadian Journal of Forest Research*, *34*, 2492−2500.

Robinson, G. K. (1991). That BLUP is a good thing: The estimation of random effects. *Statistical Science*, *6*(6), 15−51.

Rönnholm, P., Hyyppä, J., Hyyppä, H., Haggrén, H., Yu, X., & Kaartinen, H. (2004). Calibration of laser-derived tree height estimates by means of photogrammetric techniques. *Scandinavian Journal of Forest Research*, *19*(6), 524−528.

Salas, C. (2002). Ajuste y validación de ecuaciones de volumen para un relicto del bosque de roble-laurel-lingue. *Bosque*, *23*(2), 81−92.

Sánchez, C. A. L., Varela, J. G., Dorado, F. C., Alboreca, A. R., Soalleiro, R. R., González, J. G. A., & Rodríguez, F. S. (2003). A height–diameter model for *Pinus radiata* D. Don in Galicia (Northwest Spain). *Annals of Forest Science*, *60*(3), 237−245.

Schabenberger, O., & Gotway, C. A. (2005). *Statistical methods for spatial data analysis.* Boca Raton, FL, USA: Chapman & Hall/CRC 512 p.

Schabenberger, O., & Gregoire, T. G. (1996). Population-avaeraged and subject-specific approaches for clustered categorical data. *Journal of Statistical Computation and Simulation*, *54*, 231−253.

Schabenberger, O., & Pierce, F. J. (2002). *Contemporary statistical models for the plant and soil sciences.* Boca Raton, FL, USA: CRC Press 738 p.

Schardt, M., Ziegler, M., Wimmer, A., Wack, R., & Hyyppä, J. (2004). Assessment of forest parameters by means of laser scanning. In M. Thiers, B. Kock, H. Spiecker, & H. Weinacker (Eds.), *Proceedings of the ISPRS (International Society of Photogrammetry and Remote Sensing) working group part 8/2Int. Arch. Photogramm. Remote Sensing, Vol. 36*. (pp. 272−276) Freiburg, Germany.

Schröder, J., & Álvarez, J. G. (2001). Comparing the performance of generalized diameter–height equations for maritime pine in Northwestern Spain. *Forstwissenschaftliches Centralblatt*, *120*, 18−23.

Stage, A. R., & Crookston, N. L. (2007). Partitioning error components for accuracy-assessment of near-neighbor methods of imputation. *Forest Science*, *53*, 62−72.

Staudhammer, C., & LeMay, V. M. (2000). Height prediction equations using diameter and stand density measures. *Forestry Chronicle*, *76*(2), 303−309.

Thorey, L. G. (1932). A mathematical method for the construction of diameter height curves based on site. *Forestry Chronicle*, *8*(2), 121−132.

Vaida, F., & Blanchard, S. (2005). Conditional Akaike information for mixed-effects models. *Biometrika*, *92*(2), 351−370.

van Laar, A., & Akça, A. (2007). *Forest Mensuration.*Dordrecht, The Netherlands: Springer 383 p.

Waller, L. A., & Gotway, C. A. (2004). *Applied spatial statistics for public health data.* Hoboken, N.J., USA: John Wiley & Sons 494 p.

Welch, W. J. (1987). Rerandomizing the median in matched-pairs designs. *Biometrika*, *74*(3), 609−614.

Yuancai, L., & Parresol, B. R. (2001). Remarks on height–diameter modeling. *USDA For. Serv. Res. Not. SE-10, USA* (pp. 8).

Zhang, L. (1997). Cross-validation of non-linear growth functions for modelling tree height–diameter relationships. *Annals of Botany*, *79*(3), 251−257.

Zhang, L., Bi, H., Cheng, P., & Davis, C. J. (2004). Modeling spatial variation in tree diameter–height relationships. *Forest Ecology and Management*, *189*(2), 317−329.

Zhang, L., & Gove, J. H. (2005). Spatial assessment of model errors from four regression techniques. *Forest Science*, *51*(4), 334−346.

Zhang, L., Gove, J. H., & Heath, L. (2005). Spatial residual analysis of six modeling techniques. *Ecological Modelling*, *186*, 154−177.

Zhang, L., Ma, Z., & Guo, L. (2008). Spatially assessing model errors of four regression techniques for three types of forest stands. *Forestry*, *81*(2), 209−225.

Zhang, L., & Shi, H. (2004). Local modeling of tree growth by geographically weighted regression. *Forest Science*, *50*(2), 225−244.