

# Spatial Assessment of Model Errors from Four Regression Techniques

Lianjun Zhang and Jeffrey H. Gove

**Abstract:** Forest modelers have attempted to account for the spatial autocorrelations among trees in growth and yield models by applying alternative regression techniques such as linear mixed models (LMM), generalized additive models (GAM), and geographically weighted regression (GWR). However, the model errors are commonly assessed using average errors across the entire study area and across tree size classes. Little attention has been paid to the spatial heterogeneity of model performance. In this study, we used local Moran coefficients to investigate the spatial distributions of the model errors from the four regression models. The results indicated that GAM improved model-fitting to the data and provided better predictions for the response variable. However, it is nonspatial in nature and, consequently, generated spatial distributions for the model errors similar to the ones from ordinary least-squares (OLS). Furthermore, OLS and GAM yielded more clusters of similar (either positive or negative) model errors, indicating that trees in some subareas were either all underestimated or all overestimated for the response variable. In contrast, LMM was able to model the spatial covariance structures in the data and obtain more accurate predictions by accounting for the effects of spatial autocorrelations through the empirical best linear unbiased predictors. GWR is a varying-coefficient modeling technique. It estimated the model coefficients locally at each tree in the example plot and resulted in more accurate predictions for the response variable. Moreover, the spatial distributions of the model errors from LMM and GWR were more desirable, with fewer clusters of dissimilar model errors than the ones derived from OLS and GAM. *FOR. SCI.* 51(4):334–346.

**Key Words:** Spatial autocorrelation and heterogeneity, local indicator of spatial autocorrelation, ordinary least-squares, linear mixed model, generalized additive model, geographically weighted regression.

FOREST MODELERS HAVE LONG REALIZED that using ordinary least-squares (OLS) regression to estimate the parameters of growth and yield models may violate (at least) one of the OLS assumptions (i.e., independence of observations, due to temporal or/and spatial autocorrelations in forest data; Liu and Burkhart 1994, Fox et al. 2001). Violating the assumption leads to a biased estimation of the standard errors of model parameters and reduces estimation efficiency, while regression coefficients remain unbiased. Thus, significance tests and measures of model fit may be misleading (West et al. 1984, Gregoire 1987, Anselin and Griffith 1988, Fox et al. 2001). Temporal autocorrelation received attention in forest modeling practice beginning in the 1960s (Buckman 1962, Curtis 1967, West et al. 1984, Gregoire 1987). A number of statistical methods have been applied to deal with the temporal dependency such as the first-order autoregressive model (e.g., Yang 1984, Monserud 1986), the first-order mixed autoregressive moving average model (e.g., Stage and Wykoff 1993), and linear and nonlinear mixed models (e.g., Lappi 1991, Gregoire et al. 1995, Penner et al. 1995, Schabenberger and Gregoire 1995, Gregoire and Schabenberger 1996, Tasissa and Burkhart 1998).

The violation of the OLS assumption due to spatial dependency in tree data has drawn attention from forest

modelers only in recent years. Commonly, distance-dependent competition indices were developed and incorporated into individual-tree growth and yield models to account indirectly for the presence of the spatial dependency (e.g., Pukkala and Kolström 1987, Pukkala 1989, Biging and Dobbertin 1995). Other spatial analyses (e.g., Moran coefficient, variogram, kriging, Ripley's K-function, Gibbs process, nonparametric kernel, and nearest-neighbor method) have been used to detect, analyze, and model spatial patterns and autocorrelations of many forest variables, including tree diameter, diameter distribution, total height, growth, mortality, and regeneration (e.g., Reed and Burkhart 1985, Schoonderwoerd and Mohren 1988, Kenkel et al. 1989, Moer 1993, Biondi et al. 1994, Liu and Burkhart 1994, Frohlich and Quednau 1995, Kangas and Korhonen 1996, Wells and Getis 1999, Fox et al. 2001, Goreaud et al. 2002, Nanos and Montero 2002). Modern statistical techniques have lately become popular to deal with spatial effects (i.e., autocorrelation and heterogeneity) in predicting forest composition and attributes, species distribution, biodiversity, forest type and class, and insect attack. Techniques used include linear mixed models (LMM), generalized additive models (GAM), classification and regression tree (CART), multivariate adaptive regression splines (MARS), artificial neural networks (ANN), and

Lianjun Zhang, Faculty of Forest and Natural Resources Management, State University of New York, College of Environmental Science and Forestry, One Forestry Drive, Syracuse, NY 13210—Phone: (315) 470-6558; Fax: (315) 470-6535; lizhang@esf.edu. Jeffrey H. Gove, Research Forester, USDA Forest Service Northeastern Research Station, PO Box 640, Durham, NH 03824—Phone: (603) 868-7667; jgove@fs.fed.us.

**Acknowledgments:** This research was supported by funds provided by the USDA Forest Service, Northeastern Research Station RWU 4104. The authors thank the associate editor and two anonymous reviewers for their valuable suggestions and helpful comments on the manuscript.

Manuscript received April 19, 2004, accepted February 21, 2005

Copyright © 2005 by the Society of American Foresters

geographically weighted regression (GWR) (e.g., Austin and Meyers 1996, Preisler et al. 1997, Lehmann 1998, Moisen and Edwards 1999, Moisen et al. 1999, Frescino et al. 2001, Guisan et al. 2002, Moisen and Frescino 2002, Lehmann et al. 2003, Zhang and Shi 2004, Zhang et al. 2004). Many of these modeling techniques are considered more robust for spatially correlated data, less restricted in model assumptions, and flexible in modeling nonlinearity and nonconstant variance structure (Guisan et al. 2002, Moisen and Frescino 2002).

Performance assessment of these modeling techniques has often been focused on the overall model errors (Laffan, S.W. 1999. Spatially assessing model error using geographically weighted regression. GeoComputation at [http://www.geocomputation.org/1999/086/gc\\_086.htm](http://www.geocomputation.org/1999/086/gc_086.htm). March 22, 2005, Moisen and Frescino 2002), assuming the models predict the response variable equally well across the entire study area. Although previous studies showed the improvement on model-fitting and parameter standard errors using global accuracy assessments, there is little information available on how well these models perform spatially. Areas where predictions are not acceptable may indicate the need for other variables in the models or a different approach. One way to locally assess the model errors is to map them across the study area (Rathert et al. 1999, Fotheringham et al. 2002). Visualization, however, does not enable us to identify significant clusters of positive or negative model errors at multiple scales. An alternative is to compute local spatial autocorrelation indices for the model errors (Tiefelsdorf and Boots 1997, Laffan 1999 web site above, Tiefelsdorf 2000).

A set of indices known collectively as local indicator of spatial association (LISA) has been developed to study local spatial nonstationarity (e.g., Anselin 1995, Fotheringham and Brunsdon 1999, Boots 2002). The strength of LISA is its ability to identify spatial associations in a variable in one or more portions of the study area even though there is no discernible pattern as summarized by a global autocorrelation statistic (Getis and Ord 1996). LISA has been used to successfully identify clusters in biological data sets (Sokal et al. 1998a, 1998b). Shi and Zhang (2003) explored the relationships among LISA, tree growth, and traditional individual-tree competition indices. They found that LISA was useful to identify “hot spots” (positive autocorrelation, or similarity) and “cold spots” (negative autocorrelation, or dissimilarity) of tree growth and competition.

The objectives of this study were (1) to apply four alternative regression techniques (i.e., OLS, LMM, GAM, and GWR) to model tree crown area using tree diameter, (2) to evaluate the performance of the four models using overall (global) model errors and errors across tree size classes, and (3) to assess the performance of the four models in terms of spatial distributions and clustering of positive/negative model errors using spatial maps and local Moran coefficients.

## Description of Modeling Methods

Four regression techniques were used in this study.

1. Ordinary least-squares (OLS). Given a set of  $n$  ( $k = 1, 2, \dots, n$ ) observations on  $p$  ( $g = 1, 2, \dots, p$ ) independent or predictor variables  $X$ , and a dependent or response variable  $Y$ , the relationship between  $Y$  and  $X$  can be regressed using OLS as

$$Y = X\beta + \varepsilon, \quad (1)$$

where  $Y$  is a vector of the observed response variable,  $X$  is a known matrix including a column of 1s (for intercept) and  $p$  independent variables,  $\beta$  is a vector of unknown fixed-effects parameters, and  $\varepsilon$  is a random error term with assumed distribution  $N(0, \sigma^2 I)$ , where  $I$  denotes an identity matrix and  $\sigma^2$  represents the common error variance. The OLS estimate of  $\beta$  is obtained by

$$\hat{\beta} = (X^T X)^{-1} X^T Y, \quad (2)$$

where superscript T denotes the transpose of a matrix. The relationship represented by Equation 1 is assumed to be universal or constant across the geographic area.

2. Linear mixed model (LMM). The linear mixed model is a special case of generalized linear models, and can be expressed as

$$Y = X\beta + Z\gamma + \varepsilon, \quad (3)$$

where  $Y$ ,  $X$ , and  $\beta$  are as defined in Equation 1,  $Z$  is a known design matrix,  $\gamma$  is a vector of unknown random-effects parameters, and  $\varepsilon$  is a vector of unobserved random errors. It is assumed: (1)  $E(\gamma) = 0$  and  $\text{var}(\gamma) = G$ , (2)  $E(\varepsilon) = 0$  and  $\text{var}(\varepsilon) = R$ , (3)  $\text{cov}(\gamma, \varepsilon) = 0$ , and (4) both  $\gamma$  and  $\varepsilon$  are normally distributed. The variance of  $Y$  is  $V = ZGZ^T + R$ , and can be estimated by setting up the random-effects design matrix  $Z$  and by specifying covariance structures for  $G$  and  $R$  (Littell et al. 1996).

LMM can be used to (1) characterize or model the spatial covariance structure in the data and (2) remove the effects of spatial autocorrelation to obtain more accurate estimates for the response variable or treatment means (Littell et al. 1996). In principle, spatial autocorrelation can be reflected in  $G$  or  $R$  or both. For this study, no random effects were considered. Thus,  $Z = 0$ ,  $V = R$ , and Equation 3 is reduced to

$$Y = X\beta + \varepsilon. \quad (4)$$

Spatial autocorrelations among observations are modeled through  $R = \text{var}(\varepsilon)$  such that

$$\text{Var}(\varepsilon_i) = \sigma^2 + \sigma_1^2, \quad (5)$$

$$\text{Cov}(\varepsilon_i, \varepsilon_j) = \sigma^2 f(d_{ij}, \theta), \quad (6)$$

where  $d_{ij}$  is the distance between locations  $i$  and  $j$ . This is an LMM model with a nugget effect in which  $\sigma_1^2$ ,  $\sigma^2 + \sigma_1^2$ , and  $\theta$  correspond to the geostatistics parameters nugget, sill, and range, respectively. Different covariance models  $f(d_{ij}, \theta)$  are available including spherical,

exponential, Gaussian, and power. For example, the exponential covariance model has the form of

$$f(d_{ij}, \theta) = e^{(-d_{ij}/\theta)}. \quad (7)$$

A likelihood ratio test can be used to determine whether it is necessary to model the spatial covariance structure of the data (Littell et al. 1996, SAS Institute, Inc. 2002).

In general, OLS is no longer the best approach to estimating LMM (Littell et al. 1996). Likelihood-based methods (e.g., maximum likelihood (ML) and restricted/residual maximum likelihood (REML) methods) are usually used to solve for  $\beta$ . Numerical algorithms, such as the Newton–Raphson algorithm, can be used to obtain

$$\hat{\beta} = (X^T \hat{V}^{-1} X)^{-1} X^T \hat{V}^{-1} Y. \quad (8)$$

Furthermore, the empirical best linear unbiased predictions (EBLUP) should be used to take spatial autocorrelations into account for predicting the response variable (Schabenberger and Pierce 2002).

3. Generalized additive model (GAM). GAM is a nonparametric extension of generalized linear models (GLM) (Hastie and Tibshirani 1990, Guisan et al. 2002, SAS Institute, Inc. 2002). Whereas GLM emphasizes estimation and inference for model parameters, GAM focuses on exploring data nonparametrically. The strength of GAM is its ability to deal with highly nonlinear and nonmonotonic relationships between the response variable and a set of explanatory variables. Thus, GAM is sometimes referred to as data-driven rather than model-driven (Guisan et al. 2002). In general, GAM can be expressed as

$$Y = S_0 + \sum_{g=1}^p S_g(X_g) + \varepsilon, \quad (9)$$

where  $S_0$  is the intercept and  $S_g(X_g)$  is a nonparametric smoothing function for the  $g$ th independent variable  $X$ . The only underlying assumption is that the functions in GAM are additive and that the component functions are smooth. The restriction on additivity allows us to interpret the GAM model in much the same way as a conventional regression analysis. Many methods have been proposed to approach the formulation and estimation of GAM. One of the algorithms is known as backfitting, in which we seek the nonparametric functions  $S_g(X_g)$  to minimize

$$\varepsilon^2 = \left[ Y - \left( S_0 + \sum_{g=1}^p S_g(X_g) \right) \right]^2. \quad (10)$$

A wide range of nonparametric functions is available for GAM, and combinations of these functions are also possible. Several smoothers are commonly used. Cubic smoothing spline finds the function  $S_g(X_g)$  that mini-

mizes the penalized least-squares (Hastie and Tibshirani 1990),

$$\sum_{k=1}^n \left( Y - \left( S_0 + \sum_{g=1}^p S_g(X_g) \right) \right)^2 + \sum_{g=1}^p \lambda_g \int S_g''(X_g)^2 dX_g, \quad (11)$$

where  $\lambda_g$  is a nonnegative smoothing parameter for the  $g$ th variable  $X_g$ , and  $S_g''$  is the second derivative of the spline function. The first term in Equation 11 measures closeness to the data, whereas the second term penalizes curvature in the function. The parameter  $\lambda_g$  governs the “wiggleness” of the function  $S_g(X_g)$ ; when  $\lambda_g = 0$ ,  $S_g(X_g)$  produces an interpolating curve, whereas larger values of  $\lambda_g$  force  $S_g(X_g)$  to be smoother. The  $S_g(X_g)$  is a straight line when  $\lambda_g \rightarrow \infty$ . The selection of the best  $\lambda_g$  value can be achieved using the concept of effective degrees of freedom. Other spline smoothers include B-spline, polynomial spline, thin-plate spline, etc. Locally weighted scatterplot smoothing (LOESS) is also used to fit a regression surface to the data points within a chosen neighborhood of a given point, in which a weighted least-squares algorithm is used to fit linear and quadratic functions of the predictors at the centers of neighborhoods (Venables and Ripley 1997, SAS Institute, Inc. 2002).

4. Geographically weighted regression (GWR). Given a set of location coordinates  $(u_i, v_i)$  for each observation, the underlying model for GWR is

$$Y = \beta_0(u_i, v_i) + \sum_{g=1}^p \beta_g(u_i, v_i) X_g + \varepsilon, \quad (12)$$

where  $\{\beta_0(u_i, v_i), \beta_1(u_i, v_i), \dots, \beta_p(u_i, v_i)\}$  are  $(P + 1)$  continuous functions of the location  $(u_i, v_i)$  in the study area. Again,  $\varepsilon$  is the random error term with a distribution  $N(0, \sigma^2 I)$ . The aim of GWR is to obtain the estimates of these functions for each independent variable  $X$  and each geographic location  $i$  (i.e., each tree in this study) using data near the location. The estimation procedure of GWR is as follows: (1) draw a circle of a given radius around one particular location  $i$  (the center), (2) compute a weight ( $w_{ij}$ ) for each neighboring observation  $k$  ( $k = 1, 2, \dots, n$ ) according to the distance ( $d_{ij}$ ) between the location  $j$  and the center  $i$  at  $(u_i, v_i)$ , and (3) estimate the model coefficients using weighted least-squares regression such that

$$\hat{\beta}_i = (X^T W_i X)^{-1} X^T W_i y,$$

where the weight matrix,  $W_i$ , is

$$W_i = \begin{pmatrix} w_{i1} & 0 & \cdots & 0 \\ 0 & w_{i2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & w_{in} \end{pmatrix} \quad (14)$$

If  $W_i = I$  (identity matrix), i.e., each observation has a weight of unity, the GWR model is equivalent to the OLS model and the coefficients are the same for every location. Equation 13 is not a single equation but a matrix of equations. Once  $w_{ij}$  has been calculated, the coefficient matrix  $\beta$  can be computed row by row by repeated application of Equation 13. A set of estimates of spatially varying parameters is obtained without specifying a functional form for the spatial variation. The standard error for each coefficient can also be estimated. Essentially, GWR lets the data speak for themselves when estimating each regression coefficient for each independent variable and each geographic location (Brunsdon et al. 1998).

Different kernel functions can be used to compute the weight matrix for each location (Fotheringham et al. 2002). In this study we used a Gaussian distance-decay-based kernel function,

$$w_{ij} = e^{-(d_{ij}/h)^2}, \quad (15)$$

where  $h$  is referred to as the bandwidth. This kernel function assumes that the bandwidth at each center  $i$  is a constant across the study area (i.e., a fixed kernel). If the locations  $i$  and  $j$  coincide (i.e.,  $d_{ij} = 0$ ),  $w_{ij}$  equals one while  $w_{ij}$  decreases according to a Gaussian curve as the distance  $d_{ij}$  increases. For this weighting function, the weights are nonzero for all data points no matter how far they are from the center  $i$ . However, this is not necessarily the case when other weighting functions are used (Fotheringham et al. 2002).

## Data and Methods

### Data

The data used in this study were a part of the stem map data of a softwood stand located near Sault Ste. Marie, Ontario, Canada (Ek 1969). The stand was mature, second growth, uneven-aged, and 8.1 hectares (20 acres) or  $402 \times 201$  m ( $1,320 \times 660$  ft) in size. Major tree species included balsam fir (*Abies balsamea* [L.] Mill.), black spruce (*Picea mariana* [Mill.] BSP.), white spruce (*Picea glauca* [Moench] Voss), and white pine (*Pinus strobus* L.). A total of 6,881 trees were measured in 1967. The information included tree location coordinates, diameters at breast height (dbh), heights, and crown area (crown) for trees  $>8.9$  cm (3.5 in.). Pielou's index of nonrandomness (Pielou 1959) was 1.35, indicating a clustered spatial distribution of trees. The mapping and measurement methods were detailed in Ek (1969).

An example plot was selected from the stand to reduce

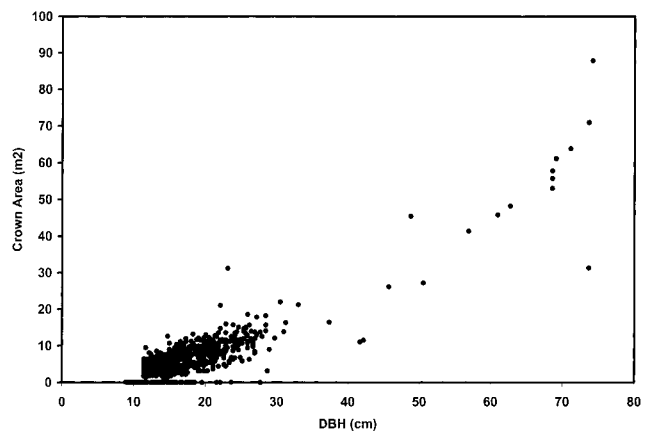


Figure 1. Scatterplot of tree crown area versus diameter ( $n = 941$  trees).

computation time for the LMM, GAM, and GWR methods. The example plot had 941 trees and was located in the right, lower corner of the stand ( $100 \times 100$  m in size). It was a close representation of the whole stand in terms of species composition, tree attributes, and the spatial distributions of trees. Balsam fir was the dominant tree species (59.1% in number of trees), followed by black spruce (36.8%). Minor species included white spruce, white pine, and white birch (*Betula papyrifera* Marsh.). The average tree diameter was 15.9 cm (ranging from 8.9 to 74.2 cm), and the average crown area was  $5.1 \text{ m}^2$  (ranging from 0 to  $87.9 \text{ m}^2$ ).

### Regression Model

Based on the data for this study (Figure 1), we chose the following linear regression model to fit the relationship between tree crown and dbh:

$$\text{crown} = \beta_0 + \beta_1 \text{dbh} + \beta_2 \text{dbh}^2 + \varepsilon, \quad (16)$$

where crown is the tree crown area ( $\text{m}^2$ ), dbh is the tree dbh (cm),  $\beta_0$ – $\beta_2$  are regression coefficients to be estimated, and  $\varepsilon$  is the model random error. This parabola model may result in estimating smaller crown areas for larger tree sizes. However, our objective was not necessarily to develop a predictive model for the tree crown area. Rather, we attempted to investigate the spatial heterogeneity of the model errors for fitting the relationship by the four regression methods.

### Model-Fitting and Evaluation

Equation 16 was fitted to the example plot by OLS, LMM, GAM, and GWR methods, respectively. Statistical Analysis System 9.0 (SAS) (SAS Institute, Inc. 2002) was used to fit the first three models. The GWR model was fitted using a computer software program, GWR 2.0 (Fotheringham et al. 2002). Detailed information on the software is available from A.S. Fotheringham, National Centre for Geocomputation, National University of Ireland <http://www.nuim.ie/ncg/GWR/software.htm> Mar. 22, 2005.

For the LMM model, different spatial covariance models



were tried to account for the spatial autocorrelations among trees, including Gaussian, exponential, power, and spherical functions (Littell et al. 1996, SAS Institute, Inc. 2002). The exponential covariance structure was selected according to model-fitting statistics such as Akaike's information criterion (AIC). Note that the PRED from the OUP = option under MODEL statement in PROC MIXED is based on EBLUP in two cases: (1) the LMM model contains a RANDOM statement, and (2) the response value for an observation is missing (SAS Institute, Inc. 2002). However, we used only  $R$  to model the spatial autocorrelations among the trees (i.e., no random effects). The following procedure was used to obtain the appropriate predictions for crown based on EBLUP. The trees in the example plot were duplicated, resulting in  $941 + 941 = 1,882$  trees/observations. The observed values of crown in the second half were then set to missing. When Equation 16 was fit to the 1,882 trees, only the first half of the data (941 trees) with observed crown was used to estimate the LMM model coefficients. However, the predictions for these trees were not based on EBLUP because of the absence of a RANDOM statement. The second half of the data (i.e., the duplication of the same 941 trees) with missing values of crown was not used in model-fitting, but the predictions of crown for these trees were based on EBLUP, which were used to compute the LMM model errors. One can easily detect whether the predictions of the response variable are based on EBLUP or not. The EBLUP predictions at two locations with identical dbh values should differ depending on where the trees are located, whereas the regular predictions (not based on EBLUP) would be the same for a given value of dbh regardless of the locations of trees (Schabenberger and Pierce 2002).

The GAM model was fitted with different nonparametric smoothers for both dbh and  $\text{dbh}^2$  in Equation 16, respectively (SAS Institute, Inc. 2002). The GAM model with the cubic smoothing spline was selected because it had a smaller model deviance than the one with LOESS. For the GWR model, the Gaussian kernel function (Equation 15) was used to estimate the weights with the optimal band-

width  $h$  determined as 7.55 m by AIC minimization (Fotheringham et al. 2002).

The model error or residual was defined as the difference between the observed and appropriately predicted crown (e.g., EBLUP predictions in LMM). The absolute model error was calculated by taking the absolute value of the model error. To examine the model errors across tree sizes, all trees in the example plot were grouped into diameter classes, and average model error and absolute error were calculated for each diameter class.

The spatial distributions of the model errors from the four regression methods were investigated using the global and local Moran coefficients (Anselin 1995, Tiefelsdorf 2000, Boots 2002). In this study we used the range of the variogram for the OLS model errors (Figure 2) as the predefined bandwidth ( $h = 7.27$  m) for calculating the Moran coefficients, because the range of a variogram indicates there is no spatial autocorrelation among the OLS model errors beyond this distance (Isaaks and Srivastava 1989).

The global Moran coefficient (MC) is defined by

$$MC = \frac{n \sum_{k=1}^n \sum_{l=1}^n c_{ij}(h)(e_i - \bar{e})(e_j - \bar{e})}{\sum_{k=1}^n \sum_{l=1}^n c_{ij}(h) \sum_{k=1}^n (e_i - \bar{e})^2}, \quad (17)$$

where  $e_i$  and  $e_j$  denote the model errors at locations  $i$  and  $j$ , respectively,  $\bar{e}$  is the mean of  $e_i$  over  $n$  locations, and  $c_{ij}(h)$  is the spatial weight measure within the bandwidth. If location  $j$  is a neighbor of the subject location  $i$ , then  $c_{ij}(h) = 1$ ; otherwise  $c_{ij}(h) = 0$ . The expected mean of the MC is  $-1/(n - 1)$ . A positive global MC indicates that the observed values of locations tend to be similar, whereas a negative MC indicates that the locations tend to be dissimilar in the observed values. The global MC is approximately zero when the observed values are arranged randomly and independently over space. The expected value and variance of the MC for the sample size  $n$  can be calculated using two sampling assumptions: normality or randomization (Cliff and Ord 1981, Lee and Wong 2001). A Z-test for the null hypothesis of no spatial autocorrelation between observed values over the  $n$  locations can be conducted based on the standardized MC.

Anselin (1995) showed that global MC of spatial autocorrelation can be decomposed into local values as follows:

$$MC_i = (e_i - \bar{e}) \sum_{k=1}^n c_{ij}(h)(e_j - \bar{e}). \quad (18)$$

The first component  $(e_i - \bar{e})$  is the difference between the model error  $e_i$  at the reference location  $i$  and the mean, whereas the second component,

$$\sum_{k=1}^n c_{ij}(h)(e_j - \bar{e}),$$

is the sum of differences between the neighboring model errors  $e_j$  and the mean. A positive local  $MC_i$  indicates a

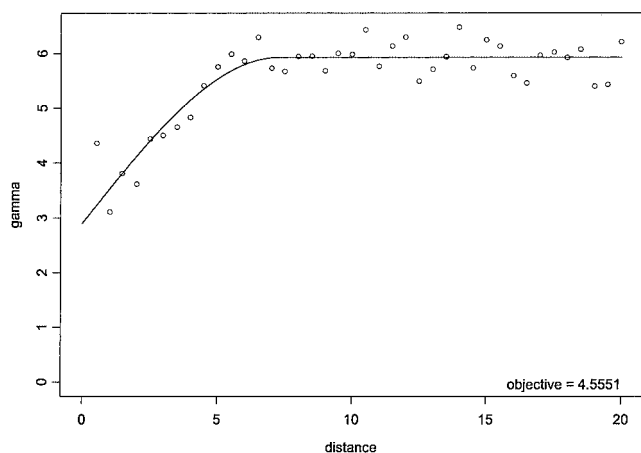


Figure 2. Variogram of the OLS model residuals (range = 7.27 m).

cluster of error values around  $i$  with values similar to those at  $i$ , and deviates strongly (either positively or negatively) from  $\bar{e}$ . A negative local  $MC_i$  describes a cluster of error values around  $i$  with values dissimilar to those at  $i$ . If either  $e_i$  or the values of  $e_j$  in the neighborhood of  $i$  are close to  $\bar{e}$ , the local  $MC_i$  will indicate no spatial autocorrelation (Tiefelsdorf and Boots 1997, Boots 2002). When the local  $MC_i$  is standardized by division by the variance,

$$\left( \sum_{k=1}^n (e_j - \bar{e})^2 / n \right),$$

a pseudosignificant level of  $MC_i$  can be obtained by a conditional randomization or permutation approach (Lee and Wong 2001). A significant positive  $Z$ -test indicates that location  $i$  is associated with the similar values of the surrounding locations, whereas a significant negative  $Z$ -test indicates that location  $i$  is associated with the dissimilar values of the surrounding locations.

Most forest modelers are also interested in the performance of a growth-and-yield model when it applies to an independent data set. To validate the performance of the four regression models, the example plot ( $n = 941$  trees) was split into two sets. A total of 800 trees (about 85%) were randomly selected from the example plot as the model development data. The remaining 141 trees (15%) of the data, which were not used in model-fitting, served as the model validation data. The two data sets had similar averages, dispersions, and ranges (see Table 6). Then the model errors and absolute errors were computed for the model validation data in the same way as mentioned above.

## Results

### Model-Fitting

The OLS model fitted the example plot reasonably well with an error sum of squares (SSE) = 8,671.15, and AIC = 4,785.1 (Table 1). The model  $R^2$  was 0.83 and the two coefficients for dbh ( $\hat{\beta}_1 = 0.4714$ ) and  $\text{dbh}^2$  ( $\hat{\beta}_2 = 0.0059$ ) were significantly different from zero (both  $P$ -values < 0.0001). The LMM model, using the exponential covariance structure, had much smaller SSE (5,748.80) and AIC (4,770.8) than the OLS model. The null model likelihood ratio test was statistically significant ( $P$ -value < 0.0001), indicating that the exponential covariance structure was preferred to the simple diagonal covariance structure of the OLS model (SAS Institute, Inc. 2002).

Similarly, both GAM and GWR models fitted the exam-

ple plot better than the OLS model (Table 1). The GAM model's SSE (6,557.49) was smaller than that of the OLS model. The  $F$ -test ( $F = 10.10$ ) for testing the improvement of GAM over OLS (Hastie and Tibshirani 1990, Venables and Ripley 1997) was highly significant ( $P$ -value < 0.0001). The GWR model's SSE was 4,799.40, and AIC was 4,523.4. The approximate goodness-of-fit likelihood ratio test was used to investigate the improvement of the GWR model over the OLS model (Fotheringham et al. 2002). The results indicated that the GWR model improved model-fitting significantly ( $F = 4.90$  and  $P$ -value < 0.0001) over the OLS model. Evidently, the relationships between crown and dbh are not constant across the example plot. For instance, the GWR slope coefficient  $\hat{\beta}_1$  for dbh ranged from  $-0.659$  to  $3.03$ , and the slope coefficient  $\hat{\beta}_2$  for  $\text{dbh}^2$  ranged from  $-0.072$  to  $0.029$ , indicating the spatial heterogeneity or nonstationarity of the relationship between crown and dbh. In general, spatial heterogeneity is related to locations in space, missing variables, and functional misspecification (Anselin and Griffith 1988, Zhang and Shi 2004).

### Global Analysis of Model Errors

Conventionally, the assessment of regression models is conducted through the residual analysis such as the examination of overall (global) model errors and errors across tree size classes (e.g., Moisen and Frescino 2002, Zhang et al. 2004). The OLS model errors have larger range and are more negatively skewed, the errors of LMM and GAM have similar ranges, skewness, kurtosis, and quantiles, and the GWR model produces much smaller range for the model errors than the above three models (Figure 3). The four models appear to produce similar model errors for trees up to 40 cm in diameter (Figure 4). However, the OLS and LMM models produce much larger negative residuals for large-sized trees (40–60 cm in diameter). In general, LMM, GAM, and GWR produce smaller absolute errors than OLS across the diameter classes (Figure 5). The GWR model consistently yields smaller absolute errors for larger trees, except the GAM model is the best for trees in the diameter class of 40–60 cm.

### Spatial Assessment of Model Errors

The global MC was computed for the model errors from the four regression models. Table 2 shows that the global MC for the OLS and GAM model errors were significantly positive ( $Z$ -values > 1.96), indicating that the above two

**Table 1. Model-fitting statistics for the four regression techniques.**

Model	SSE	AIC	Test <sup>a</sup>	$P$ -value
OLS	8671.15	4785.1	N.A.	N.A.
LMM	5748.80	4770.8	$\chi^2 = 567.8$	<0.0001
GAM	6557.49		$F = 10.10$	<0.0001
GWR	4799.40	4523.4	$F = 4.90$	<0.0001

$n = 941$  trees.

<sup>a</sup>The hypothesis test for the improvement of alternative modeling methods over the OLS model.

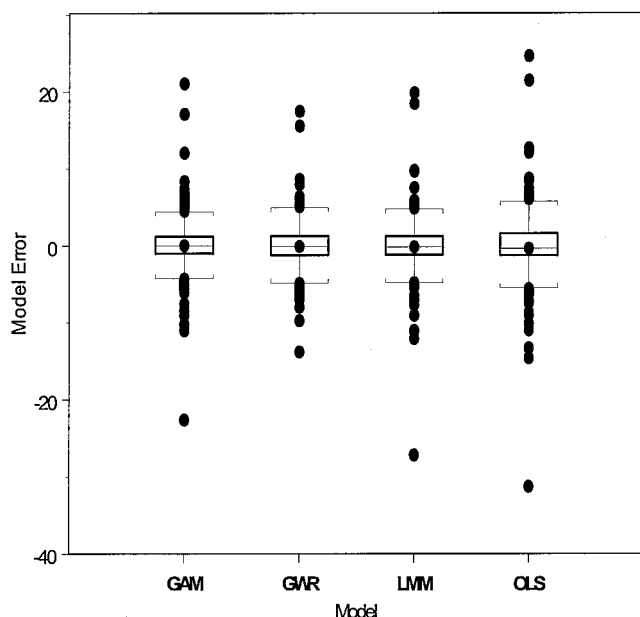


Figure 3. Box plot of the model errors from the four regression models.

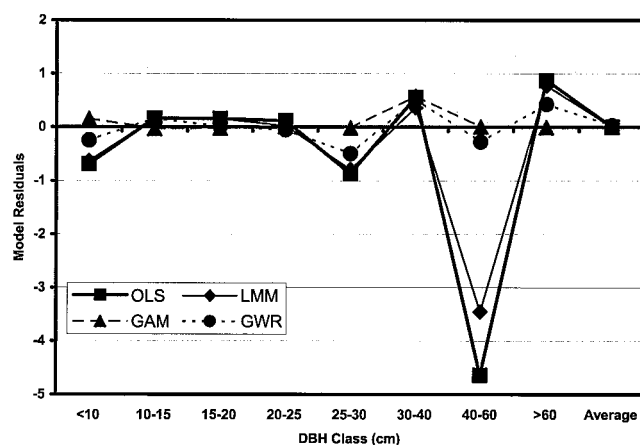


Figure 4. Model errors across tree diameter classes.

models produced model errors in clusters of similar values (i.e., either positive or negative values) across the example plot. Although the global MC for the LMM and GWR model errors were significantly negative ( $Z\text{-value} < -1.96$ ), the magnitudes of the global MC were much smaller than those of OLS and GAM. It was evident that both LMM and GWR did account for the spatial autocorrelations among the trees and, consequently, reduced the autocorrelations among the model errors. The negative global MC meant that the model errors from the LMM and GWR models were clustered in dissimilar values across the example plot.

Figure 6a and c illustrate that the spatial distributions of the model errors derived from OLS and GAM have very similar patterns. There are several clusters of positive or negative errors. Figure 6b shows that the LMM model errors are smaller in magnitude and have fewer numbers of clusters of positive or negative errors across the example plot.

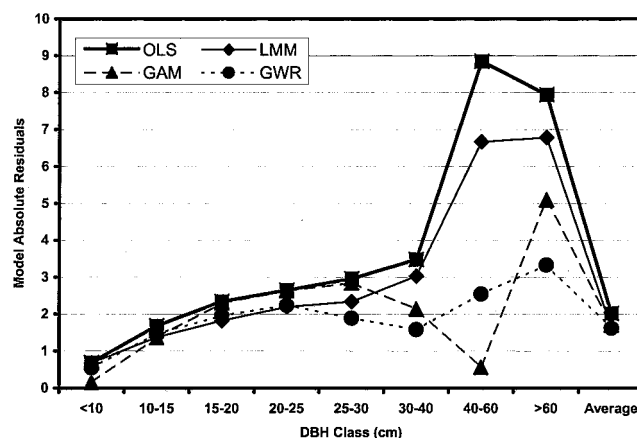


Figure 5. Model absolute errors across tree diameter classes.

Table 2. Global Moran coefficients (MC) of the model errors from the four regression techniques.

Model	Global MC	Z-value <sup>a</sup>	Z-value <sup>b</sup>
OLS	0.0458	4.03	4.07
LMM	-0.0264	-2.17	-2.20
GAM	0.0645	5.64	5.67
GWR	-0.0259	-2.14	-2.15

<sup>a</sup>Standard normal test based on the normality assumption.

<sup>b</sup>Standard normal test based on the randomization assumption.

Figure 6d demonstrates that the GWR model errors are much smaller and have different spatial patterns than the last three models.

LISA has proven to be a useful tool to identify “hot spots” (positive autocorrelation, or similarity) and “cold spots” (negative autocorrelation, or dissimilarity) of values (Boots 2002, Shi and Zhang 2003). Local Moran coefficient ( $MC_i$ ) was computed for each model error from each of the four models, and Z-value was also computed for each corresponding local  $MC_i$  (Tables 3 and 4; Figure 7).

The local  $MC_i$  for the model errors from the OLS and GAM models had similar averages, standard deviations, and percentiles (Table 3). The two models seemed to produce more and larger positive local  $MC_i$  values. There are several “hot spots” of the model errors existing in the example plot, meaning the clusters of either positive or negative errors (Figure 7a and c). In contrast, the LMM and GWR models had similar averages, standard deviations, and percentiles, and more negative local  $MC_i$  or more “cold spots” of dissimilar model errors (Table 3; Figure 7c and d). The Z-values of the local  $MC_i$  had patterns similar to the local  $MC_i$  values for the four regression models (Table 4). Again, the OLS and GAM models had more positive Z-values, whereas the LMM and GWR models had more negative Z-values. The spatial distributions of the local Z-values are very similar to those of local  $MC_i$  for the four models (not shown).

### Comparison of Four Regression Models

The local  $MC_i$  is computed for each location in all the data. However, the values of the statistic are correlated for

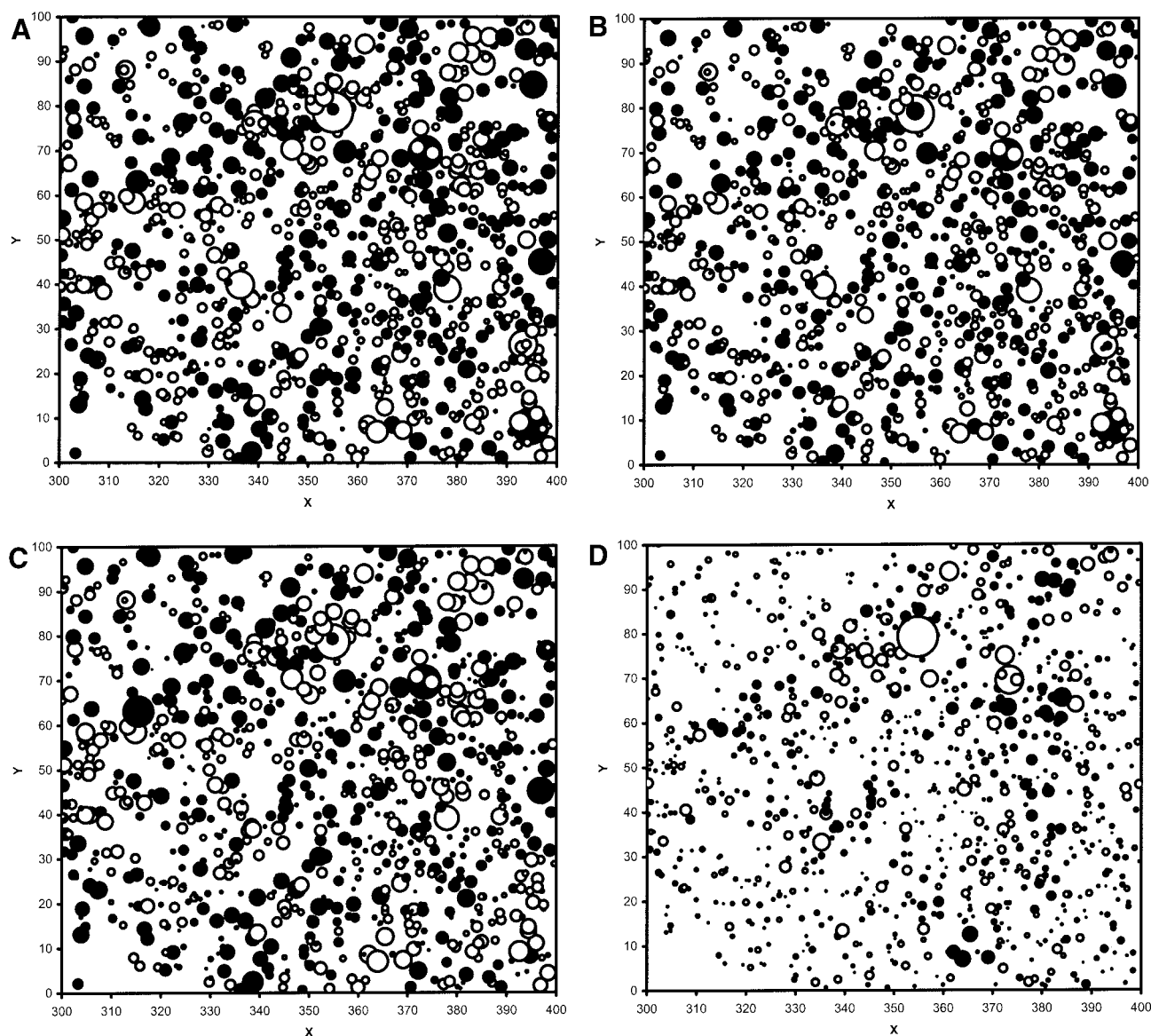


Figure 6. Plot of model errors: (a) OLS, (b) LMM, (c) GAM, and (d) GWR. The size of the symbols (black dot and circle) is proportional to the model errors/residuals. The black dots represent positive residuals, and the circles represent negative residuals.

Table 3. Local Moran coefficients of the model errors from the four regression techniques.

Model	Mean	Std	Minimum	5% Q	25% Q	Median	75% Q	95% Q	Maximum
OLS	0.71	5.97	-78.00	-4.55	-0.73	0.10	1.53	6.83	95.60
LMM	-0.41	4.15	-84.76	-4.36	-1.04	-0.06	0.51	2.91	15.76
GAM	0.99	5.86	-34.48	-3.96	-0.42	0.11	1.58	7.91	93.27
GWR	-0.40	4.63	-97.72	-4.87	-0.99	-0.06	0.65	3.67	18.10

Respectively, 5% Q, 25% Q, 75% Q, and 95% Q are 5%, 25%, 75%, and 95% quantiles.

neighboring locations. The nature of the correlation is dependent on how the neighbors are defined and how the locations are arranged in the study area. Due to the problems of multiple comparisons for the local  $MC_i$  for all the data, the significance levels should be adjusted when testing the significance of the local  $MC_i$  for each location. One possibility is to apply Bonferroni adjustment in which the significance level for each individual location is  $\alpha/n$ , where  $n$  is the sample size. However, this adjusted local significance

level is too conservative for a large sample size (i.e.,  $n = 941$  in this study), and may not be appropriate for testing local LISA (Anselin 1995, Boots 2002). Therefore, the local Z-values for the local  $MC_i$  were evaluated for the significance levels of 0.05 ( $Z_{\alpha/2} = 1.96$ ), 0.01 ( $Z_{\alpha/2} = 2.58$ ), and 0.001 ( $Z_{\alpha/2} = 3.27$ ) (Table 5).

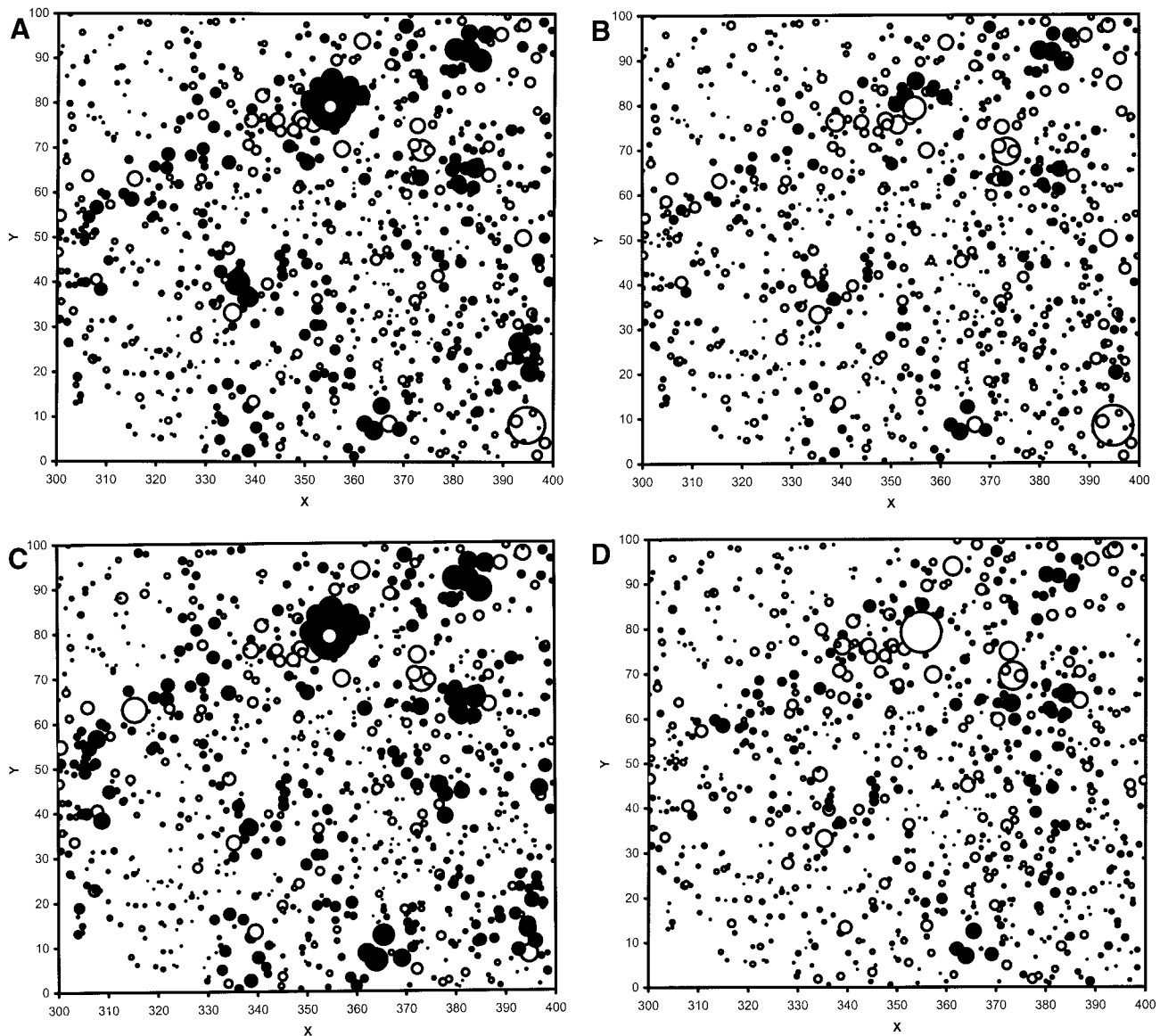
For  $\alpha = 0.05$ , the OLS and GAM models produced more than 7% significant local Z-values of the 941 values, whereas the LMM and GWR models yielded about 4%



**Table 4.** Z-value of the local Moran coefficient of the model errors from the four regression techniques.

Model	Mean	Std	Minimum	5% Q	25% Q	Median	75% Q	95% Q	Maximum
OLS	0.18	1.58	-22.92	-1.21	-0.20	0.03	0.43	1.74	24.38
LMM	-0.11	1.16	-24.95	-1.23	-0.26	-0.014	0.152	0.84	4.46
GAM	0.27	1.54	-10.09	-1.13	-0.10	0.03	0.46	2.32	23.69
GWR	-0.10	1.15	-23.36	-1.40	-0.25	-0.01	0.17	0.96	4.46

Respectively, 5% Q, 25% Q, 75% Q, and 95% Q are 5%, 25%, 75%, and 95% quantiles.



**Figure 7.** Plot of the local Moran coefficient: (a) OLS, (b) LMM, (c) GAM, and (d) GWR. The size of the symbols (black dot and circle) is proportional to the local Moran coefficient (MC) of the model errors/residuals. The black dots represent positive local MC, and the circles represent negative local MC.

significant Z-values (Table 5). Among the significant Z-values there were about 60% positive Z-values and 40% negative Z-values for the OLS and GAM models. This implied that OLS and GAM tended to generate more clusters of either positive or negative model errors in some subareas of the example plot. Trees in those subareas were either all underestimated (positive errors) or all overestimated (negative errors) for the response variable. However,

there were about 40% positive Z-values and 60% negative Z-values among the significant Z-values for the LMM and GWR models (Table 5). If there are clusters of the model errors or residuals existing, a large error tends to be surrounded by smaller neighboring errors and a small error tends to be surrounded by larger neighboring errors. Similar trends can be seen for the other two significance levels in Table 5. The LMM and GWR models produced relatively

**Table 5. Comparison of the significant Z-values for the local Moran coefficients.**

Model	Number of $ Z  > 1.96$	Among the significant Z-values	
		$Z \leq -1.96$ (%)	$Z > 1.96$ (%)
OLS	67 (7.12%)	27 (40.3%)	40 (59.7%)
LMM	38 (4.04%)	22 (57.9%)	16 (42.1%)
GAM	78 (8.29%)	23 (29.5%)	55 (70.5%)
GWR	38 (4.04%)	24 (63.2%)	14 (36.8%)
Number of $ Z  > 2.58$			
OLS	44 (4.68%)	15 (34.1%)	29 (65.9%)
LMM	23 (2.44%)	16 (69.6%)	7 (30.4%)
GAM	56 (5.95%)	15 (26.8%)	41 (73.2%)
GWR	20 (2.13%)	12 (60.0%)	8 (40.0%)
Number of $ Z  > 3.27$			
OLS	29 (3.08%)	9 (31.0%)	20 (69.0%)
LMM	12 (1.28%)	8 (66.7%)	4 (33.3%)
GAM	38 (4.04%)	11 (28.9%)	27 (71.1%)
GWR	12 (1.28%)	9 (75.0%)	3 (25.0%)

Numbers in parentheses are percentages;  $n = 941$  trees.

**Table 6. Descriptive statistics of the tree variables for the model development data (800 trees) and model validation data (141 trees).**

Data	Variable	Mean	Std	Minimum	Maximum
Model	dbh (cm)	15.88	7.59	8.89	73.66
Development	crown (m <sup>2</sup> )	5.03	6.96	0.00	70.98
Model	dbh (cm)	16.13	8.71	8.89	74.17
Validation	crown (m <sup>2</sup> )	5.35	9.13	0.00	87.88

fewer numbers of significant local  $MC_i$  for smaller  $\alpha$  levels than the OLS and GAM models. The split between positive and negative Z-values was about 30% (positive) and 70% (negative) for LMM and GWR, and vice versa for OLS and GAM at  $\alpha = 0.01$  and  $\alpha = 0.001$ .

### Model Validation

To validate model performance, we randomly split the example plot ( $n = 941$  trees) into the model development data set (85%) and validation data set (15%) (Table 6). Because the spatial continuity among the validation trees was lost because of the random selection, we only computed the descriptive statistics for the model errors and absolute errors for the four regression techniques based on the validation data. The results showed that, on average, GWR and GAM produced smaller model errors than OLS and LMM. GAM yielded the smallest absolute model errors among the four regression techniques. The performance of OLS and

LMM was similar for the model validation data in terms of averages, dispersions, and ranges of the model errors and absolute errors (Table 7).

### Discussion

In this study, we investigated the spatial heterogeneity of model performance for four regression methods, using a case study of one small forest population. An exhaustive comparison of estimation and diagnostic methods for larger areas or simulated populations that encompass a wide range of spatial variability would be difficult because of the computer-intensive nature of the LMM, GAM, and GWR methods. Still, the results discovered here on a clustered plot of trees may hold for other similar forested areas, because spatial modeling techniques such as LMM and GWR are able to account for the irregularity or heterogeneity in the population.

**Table 7. Descriptive statistics of the model residuals and absolute residuals based on model validation data (141 trees).**

Model	Variable	Mean	Std	Minimum	Maximum
OLS	Residual	0.08	3.58	-14.25	27.70
	Absolute residual	2.11	2.89	0.03	27.70
LMM	Residual	0.09	3.53	-13.99	27.55
	Absolute residual	2.06	2.86	0.01	27.55
GAM	Residual	-0.03	2.77	-18.17	6.53
	Absolute residual	1.65	2.23	0.01	18.17
GWR	Residual	-0.02	3.18	-10.26	21.33
	Absolute residual	2.11	2.37	0.03	21.33

GWR, a local modeling technique, may hold some promise in forest growth modeling under an individual tree distance-dependent paradigm. Each tree would have an associated set of locally calibrated coefficients for model components that would be stored for subsequent prediction. The complexity of the underlying growth model might dictate several such individual tree equations. One unresolved question in such an approach would be whether these local relationships stand the test of time; that is, given several prediction cycles, do local coefficient estimates still represent a given locality as well as smoothed global estimates would? One could conjecture that, due to factors such as tree senescence, mortality, canopy dynamics, and recruitment, locally calibrated models might not continue to fit a given tree or region over time and may require updating or adjustment. Global models, however, having drawn from the pool of larger variation in the overall dynamics and states of the forest, might tend to represent these changes with more robustness over time. Also, local models based on techniques such as GWR are population-specific. One could not estimate local models for trees on a forest and hope to export them to another population with different spatial-size interrelationships. Therefore, these methods are of limited use in developing regional models.

One possible way to make local models, such as those fitted under GWR, more robust to the changes in the neighborhood might be to adopt a different strategy for weighting observations in the parameter estimation stage. In this study, the weight function used in GWR estimation took the form of Gaussian decay. Although this weight function has certain desirable properties with regard to spatial continuity in general, it probably has little biological justification, nor do its competitors (Fotheringham et al. 2002). It would undoubtedly make more sense to use some form of competition index within a local neighborhood, either alone or combined in concert with the Gaussian decay function, as the weighting function in GWR for forest trees. In the latter case, the competition index might be used as an adjustment to the Gaussian weights, so that the competition potential of nearby trees upweights their values.

## Conclusion

Forest modelers have realized that the misspecification of covariance structure for spatially correlated data will produce biased standard error estimators, consequently affecting hypothesis tests and confidence intervals of the model. Generalized additive models (GAM) do not make any assumptions on model errors, and fit the data nonparametrically. Although it improves the model-fitting and produces better prediction due to its robustness and flexibility, GAM is nonspatial in nature because it focuses on multidimensional space of predictor variables. However, a linear mixed model is able to characterize the spatial covariance structures in the data with different geostatistics models. More accurate predictions for the response variable can be obtained by accounting for the effects of spatial autocorrelation through the empirical best linear unbiased predictors

(Littell et al. 1996, Schabenberger and Pierce 2002). GWR is clearly a spatial model. Its kernel function is located in two-dimensional geographical space, and takes spatial locations explicitly into account for estimating the model coefficients at each tree in the example plot (Brunsdon et al. 1999). Therefore, GWR produces a different spatial distribution for model errors than the ones from other regression techniques.

Our results showed that the OLS and GAM models yielded significant clusters of positive or negative residuals, indicating that trees in some subareas of the study area were either all underestimated or all overestimated for the response variable. Since LMM and GWR were able to adjust the estimation of the model coefficients according to the local spatial autocorrelations, they produced more accurate predictions for the response variable. The LMM and GWR model residuals also had more desirable spatial distributions (fewer clusters or clusters of dissimilar model errors) than the ones from the OLS and GAM models.

## Literature Cited

- ANSELIN, L. 1995. Local indicator of spatial association—LISA. *Geog. Anal.* 27:93–115.
- ANSELIN, L., AND D.A. GRIFFITH. 1988. Do spatial effects really matter in regression analysis. *Papers of Reg. Sci. Assoc.* 65:11–34.
- AUSTIN, M.P., AND J.A. MEYERS. 1996. Current approaches to modeling the environmental niche of eucalypts: Implication for management of forest biodiversity. *For. Ecol. Manage.* 85:95–106.
- BIGING, G.S., AND M. DOBBERTIN. 1995. Evaluation of competition indices in individual tree growth models. *For. Sci.* 41:360–377.
- BIONDI, F., D.E. MYERS, AND C.C. AVERY. 1994. Geostatistically modeling stem size and increment in an old-growth forest. *Can. J. For. Res.* 24:1354–1368.
- BOOTS, B. 2002. Local measures of spatial association. *EcoScience* 9:168–176.
- BRUNSDON, C., M. AITKIN, A.S. FOTHERINGHAM, AND M.E. CHARLTON. 1999. A comparison of random coefficient modeling and geographically weighted regression for spatially non-stationary regression problems. *Geo. Envir. Model.* 3:47–62.
- BRUNSDON, C.A., A.S. FOTHERINGHAM, AND M.E. CHARLTON. 1998. Geographically weighted regression—modeling spatial non-stationary. *The Statistician* 47:431–443.
- BUCKMAN, R.R. 1962. Growth and yield of red pine in Minnesota. *USDA For. Serv. Tech. Bull.* 1272. 50 p.
- CLIFF, A.D., AND J.K. ORD. 1981. *Spatial processes: Models and applications.* Pion, London. 266 p.
- CURTIS, R.O. 1967. A method of estimation of gross yield of Douglas-fir. *For. Sci. Monogr.* 13. 24 p.
- EK, A.R. 1969. Stem map data for three forest stands in northern Ontario. *For. Res. Lab., Sault Ste. Marie, Ontario. Information Report O-X-113.* 23 p.

- FOTHERINGHAM, A.S., AND C. BRUNSDON. 1999. Local forms of spatial analysis. *Geog. Anal.* 31:340–358.
- FOTHERINGHAM, A.S., C. BRUNSDON, AND M. CHARLTON. 2002. Geographically weighted regression: The analysis of spatially varying relationships. John Wiley & Sons, New York. 269 p.
- FOX, J.C., P.K. ADES, AND H. BI. 2001. Stochastic structure and individual-tree growth models. *For. Ecol. Manage.* 154:261–276.
- FRESCINO, T.S., T.C. EDWARDS JR., AND G.G. MOISEN. 2001. Modeling spatially explicit forest structure attributes using generalized additive models. *J. Veg. Sci.* 12:15–26.
- FROHLICH, M., AND H.D. QUEDNAU. 1995. Statistical analysis of the distribution pattern of natural regeneration in forests. *For. Ecol. Manage.* 73:45–57.
- GETIS, A., AND J.K. ORD. 1996. Local spatial statistics: An overview. P. 261–277 in *Spatial analysis: Modeling in a GIS environment*. P. Longley and M. Batty (eds.). John Wiley and Sons, New York.
- GOREAUD, F., M. LOREAU, AND C. MILLIER. 2002. Spatial structure and the survival of an inferior competitor: A theoretical model of neighborhood competition in plants. *Ecol. Model.* 158:1–19.
- GREGOIRE, T.G. 1987. Generalized error structure for forestry yield models. *For. Sci.* 33:423–444.
- GREGOIRE, T.G., AND O. SCHABENBERGER. 1996. A non-linear mixed-effects model to predict cumulative bole volume of standing trees. *J. Appl. Stat.* 23:257–271.
- GREGOIRE, T.G., O. SCHABENBERGER, AND J.P. BARRETT. 1995. Linear modeling of irregular spaced, unbalanced, longitudinal data from permanent-plot measurements. *Can. J. For. Res.* 25:137–156.
- GUISAN, A., T.C. EDWARDS JR., AND T. HASTIE. 2002. Generalized linear and generalized additive models in studies of species distributions: Setting the scene. *Ecol. Model.* 157:89–100.
- HASTIE, T.J., AND R.J. TIBSHIRANI. 1990. Generalized additive models. Chapman & Hall, New York. 335 p.
- ISAAKS, E.H., AND R.M. SRIVASTAVA. 1989. An introduction to applied geostatistics. Oxford University Press, New York. 561 p.
- KANGAS, A., AND K.T. KORHONEN. 1996. Application of non-parametric kernel regression and nearest-neighbor regression for generalizing sample tree information. P. 631–638. in *Proc. of the spatial accuracy assessment in natural resources and environmental sciences symposium*, May 21–23, 1996, Fort Collins, Colorado. Mowrer, H.T., R.L. Czaplewski, and R.H. Hamre (eds.). USDA For. Serv. Gen. Tech. Rep. RM-GTR-277.
- KENKEL, N.C., J.A. HOSKINS, AND W.D. HOSKINS. 1989. Local competition in a naturally established jack pine stand. *Can. J. Bot.* 67:2630–2635.
- LAPPI, J. 1991. Calibration of height and volume equations with random parameters. *For. Sci.* 37:781–801.
- LEE, J., AND D.W.S. WONG. 2001. Statistical analysis with Arc-View GIS. John Wiley and Sons, Inc., New York. 192 p.
- LEHMANN, A. 1998. GIS modeling of submerged macrophyte distribution using generalized additive models. *Plant Ecol.* 139:113–124.
- LEHMANN, A., J. MCC. OVERTON, AND J.R. LEATHWICK. 2003. GRASP: Generalized regression analysis and spatial prediction. *Ecol. Model.* 160:165–183.
- LITTELL, R.C., G.A. MILLIKEN, W.W. STROUP, AND R.D. WOLFIN-GER. 1996. SAS system for mixed models. SAS Institute, Inc., Cary, NC. 633 p.
- LIU, J., AND H.E. BURKHART. 1994. Spatial autocorrelation of diameter and height increment predictions from two stand simulators for loblolly pine. *For. Sci.* 40:349–356.
- MOEUR, M. 1993. Characterizing spatial patterns of trees using stem-mapped data. *For. Sci.* 39:756–775.
- MOISEN, G.G., D.R. CULTER, AND T.C. EDWARDS JR. 1999. Generalized linear mixed models for analyzing error in a satellite-based vegetation map of Utah. P. 37–44 in *Quantifying spatial uncertainty in natural resources. Theory and application for GIS and remote sensing*, Mowrer, H.T., and R.G. Congalton (eds.). Ann Arbor Press, Chelsea, MI.
- MOISEN, G.G., AND T.C. EDWARDS JR. 1999. Use of generalized linear models and digital data in a forest inventory of northern Utah. *J. Agr. Biol. Envir. Stat.* 4:372–390.
- MOISEN, G.G., AND T.S. FRESCINO. 2002. Comparing five modeling techniques for predicting forest characteristics. *Ecol. Model.* 157:209–225.
- MONSERUD, R.A. 1986. Time-series analysis of tree-ring chronologies. *For. Sci.* 32:349–372.
- NANOS, N., AND G. MONTERO. 2002. Spatial prediction of diameter distribution models. *For. Ecol. Manage.* 161:147–158.
- PENNER, M., T. PENTTILÄ, AND H. HÖKKÄ. 1995. A method for using random parameters in analyzing permanent sample plots. *Silva Fenn.* 29:287–296.
- PIELOU, E.C. 1959. The use of point-to-point distances in the study of the pattern of plant populations. *J. Ecol.* 47:607–613.
- PREISLER, H.K., N.G. RAPPAPORT, AND D.L. WOOD. 1997. Regression methods for spatially correlated data: An example using beetle attacks in a seed orchard. *For. Sci.* 43:71–77.
- PUKKALA, T. 1989. Prediction of tree diameter and height in a Scots pine stand as a function of the spatial pattern of trees. *Silva Fenn.* 23:83–99.
- PUKKALA, T., AND T. KOLSTRÖM. 1987. Competition indices and the prediction of radial growth in Scots pine. *Silva Fenn.* 21:55–67.
- RATHERT, D., D. WHITE, J.C. SIFNEOS, AND R.M. HUGHES. 1999. Environmental correlates of species richness for native freshwater fish in Oregon, USA. *J. Biogeog.* 26:1–17.
- REED, D.D., AND H.E. BURKHART. 1985. Spatial autocorrelation of individual tree characteristics in loblolly pine stands. *For. Sci.* 31:575–587.
- SAS INSTITUTE, INC. 2002. SAS/STAT Users' guide, version 9.0. SAS Institute, Inc., Cary, NC.
- SCHABENBERGER, O., AND T.G. GREGOIRE. 1995. A conspectus on



- estimating function theory and its application to recurrent modeling issues in forest biometry. *Silva Fenn.* 29:49–70.
- SCHABENBERGER, O., AND F.J. PIERCE. 2002. Contemporary statistical models for the plant and soil sciences. CRC Press, Boca Raton, FL. 738 p.
- SCHOONDERWOERD, H., AND G.M.J. MOHREN. 1988. Autocorrelation and competition in even-aged stands of Douglas-fir in the Netherlands. P. 619–626 *in* Forest growth modeling and prediction. USDA For. Serv. Gen. Tech. Rep. NC-120.
- SHI, H., AND L. ZHANG. 2003. Local analysis of tree competition and growth. *For. Sci.* 49:938–955.
- SOKAL, R.R., N.L. ODEN, AND B.A. THOMSON. 1998a. Local spatial autocorrelation in a biological model. *Geog. Anal.* 30:331–354.
- SOKAL, R.R., N.L. ODEN, AND B.A. THOMSON. 1998b. Local spatial autocorrelation in biological variables. *Biol. J. Linnean Soc.* 65:41–62.
- STAGE, A.R., AND W.R. WYKOFF. 1993. Calibrating a model of stochastic effects on diameter increment for individual-tree simulation of stand dynamics. *For. Sci.* 39:692–705.
- TASSISA, G., AND H.E. BURKHART. 1998. An application of mixed effects analysis to modeling thinning effects on stem profile of loblolly pine. *For. Ecol. Manage.* 103:87–101.
- TIEFELSDORF, M. 2000. Modeling spatial processes: The identification and analysis of spatial relationships in regression residuals by means of Moran's I. *Lecture notes in earth science* 87. Springer, New York. 167 p.
- TIEFELSDORF, M., AND B. BOOTS. 1997. A note on the extremities of local Moran's  $I_i$ 's and their impact on global Moran's I. *Geog. Anal.* 29:248–257.
- VENABLES, W.N., AND B.D. RIPLEY. 1997. Modern applied statistics with S-Plus, 2nd ed. Springer, New York. 548 p.
- YANG, Y. 1984. Autoregression analysis of forest growth. *Quart. J. Chin. For.* 17:1–11.
- WELLS, M.L., AND A. GETIS. 1999. The spatial characteristics of stand structure in *Pinus torreyana*. *Plant Ecol.* 143:153–170.
- WEST, P.W., D.A. RATKOWSKY, AND A.W. DAVIS. 1984. Problems of hypothesis testing of regressions with multiple measurements from individual sampling units. *For. Ecol. Manage.* 7:207–224.
- ZHANG, L., H. BI, P. CHENG, AND C.J. DAVIS. 2004. Modeling spatial variations in tree diameter-height relationships. *For. Ecol. Manage.* 189:317–329.
- ZHANG, L., AND H. SHI. 2004. Local modeling of tree growth by geographically weighted regression. *For. Sci.* 50:225–244.