A new spatial-attribute weighting function for geographically weighted regression

Haijin Shi, Lianjun Zhang, and Jianguo Liu

Abstract: In recent years, geographically weighted regression (GWR) has become popular for modeling spatial heterogeneity in a regression context. However, the current weighting function used in GWR only considers the geographical distances of trees in a stand, while the attributes (e.g., tree diameter) of the neighboring trees are totally ignored. In this study, we proposed a new weighting function that combines the "geographical space" and "attribute space" between the subject tree and its neighbors, such that (1) neighbors with greater geographical distances from the subject tree are assigned smaller weights, and (2) at a given geographical distance, neighboring trees with sizes that are similar to that of the subject tree are assigned larger weights. The results indicate that the GWR model with the new spatial-attribute weighting function performs better than the one with the spatial weighting function in terms of model residuals and predictions for different spatial patterns of tree locations.

Résumé : Dans les dernières années, la régression géographiquement pondérée (RGP) a souvent été utilisée pour modéliser l'hétérogénéité spatiale dans un contexte de régression. Toutefois, les fonctions de pondération actuellement disponibles pour la RGP considèrent seulement la distance entre les arbres dans un peuplement, alors que sont totalement ignorés les attributs des arbres voisins comme le diamètre. Dans la présente étude, nous proposons une nouvelle fonction de pondération qui combine l'information spatiale à celle d'attributs de l'arbre étudié et de ses voisins, de telle sorte (1) que les voisins les plus éloignés reçoivent une pondération plus faible et (2) qu'à une distance donnée, les voisins de taille similaire à l'arbre étudié reçoivent une pondération plus élevée. Les résultats montrent qu'une RGP utilisée avec la nouvelle fonction de pondération proposée a une meilleure performance que celle qui utilise un fonction de pondération uniquement spatiale, que ce soit en termes de résidus ou de valeurs prédites pour différents patrons de distribution spatiale des arbres.

[Traduit par la Rédaction]

Introduction

Existence of spatial patterns in forest stands has been widely recognized. The complex historical and environmental mosaic imposed by initial establishment patterns, microenvironmental conditions, climate factors, and competing vegetation may result in various spatial compositions and structures in different forest stands (Moeur 1993; Rouvinen and Kuuluvainen 1997). The spatial distribution of trees strongly affects tree size, growth, crown structure, and mortality (Miller and Weiner 1989; Kenkel et al. 1989; Weiner 1990; Newton and Jolliffe 1998; Dovciak et al. 2001). On the contrary, differences in tree sizes, crown structures, and other tree characteristics can result in different spatial patterns of trees over time. Therefore, the attributes (e.g., tree height), competition (e.g., neighboring tree size), and location (e.g., spatial coordinate) of trees are all of great importance (Liu and Ashton 1999; Lee and Wong 2001). All these factors lead to spatial heterogeneity across space. The traditional way of modeling forest growth and yield is to use ordinary least-square (OLS) regression. However, OLS cannot deal with spatial heterogeneity in forestry data (Zhang and Shi 2004). It is necessary to understand the influence of spatial heterogeneity on tree competition and growth, and to improve the performance of the traditional forest growth and yield models by incorporating spatial information in the model systems.

Various spatial modeling methods have been applied to explore the effect of spatial heterogeneity in forest and ecological studies (Haining 1978; Anselin 1988; Fox et al. 2001). For example, the spatial expansion model takes the parameters of a linear regression model as a function of spatial coordinates. Hence, the resultant parameter estimates drift over space (Casetti 1972, 1997; Jones and Casetti 1992; Fotheringham and Brunsdon 1999; Páez et al. 2002). The spatially adaptive filtering model is based on a "predictor–corrector" approach, which works iteratively to adjust the parameter estimates in terms of adjacent neighbors (Gorr and Olligschlaeger 1994; Fotheringham et al. 2002). Thus, the model coefficients vary locally across space. An alternative method is using a random coefficient model in which

Received 3 November 2004. Resubmitted 5 September 2005. Accepted 29 November 2005. Published on the NRC Research Press Web site at http://cjfr.nrc.ca on 4 April 2006.

H. Shi[1] and J. Liu. Center for Systems Integration and Sustainability, Department of Fisheries and Wildlife, 13 Natural Resources Building, Michigan State University, East Lansing, MI 48824, USA.
L. 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, USA.

[1]Corresponding author (e-mail: hshi@msu.edu).

 doi:10.1139/X05-295

the regression coefficients are assumed to vary from case to case and follow a probability distribution such as the normal distribution (Anselin 1988).

A new approach, geographically weighted regression (GWR), for depicting the spatial heterogeneity in a regression context has been developed and has become popular in recent years (Fotheringham et al. 1996, 2002; Brunsdon et al. 1996, 1998; Zhang and Shi 2004; Zhang et al. 2004; Shi et al. 2006). In GWR, any spatial nonstationarity in the relationship of interest is accounted for by the local estimation of model coefficients through a spatial weighting function. This spatial weighting function is a decreasing function of distance (geographical space) from the focal point (x_0), so that the impact of the neighbors (x_i, i = 1, …, k, where k is the number of neighbors) nearby is stronger than that of neighbors farther away. However, the use of distance (geographical space) only for determining the weights in the GWR model may not be realistic and reasonable because the attribute effects of the focal point and its neighbors are totally ignored.

In fact, the development of GWR follows the general principle of local smoothing and locally weighted regression (Leung et al. 2000; Páez et al. 2002), in which the weights are determined according to the size of the residuals (Cleveland 1979; Cleveland and Devlin 1988; Casetti 1982; Casetti and Can 1999). For a given focal point x_0 in the locally weighted regression, if the size of its neighbor x_i is similar to the size of x_0, the "distance of attribute space" between x_0 and x_i is small. Thus, the neighbor is assigned a large weight by the weighting function. In contrast, a neighbor x_i with a size that is dissimilar to the size of x_0 is assigned a small weight, since it is far away from x_0 in the distance of attribute space. In other words, the weights are determined by the "attribute space" instead of the "geographical space" (Leung et al. 2000). This approach pays more attention to the fitting of the dependent variable rather than to spatially varying parameters. Clearly, the "attribute space" approach does not consider geographical locations of the neighbors and the relative distance (geographical space) between x_0 and x_i in spatial data.

In this study, we propose an approach that incorporates the tree attribute into the spatial weighting function used in GWR. The new weighting function will combine the geographical space and attribute space between the subject tree (x_0) and its neighboring trees (x_i), such that (1) the neighbors (x_i) with large geographical distances from x_0 will be assigned small weights, and vice versa, and (2) at a given geographical distance, the neighbors (x_i) with attributes that are similar to those of x_0 will be assigned large weights, and vice versa. Therefore, the spatial-attribute weighting function takes into account both geographical distance and size of the subject tree and of its neighbors. Biologically, this function implies that competition is a reciprocal process. Large trees have an influence on small trees, while small trees also compete for resources with large trees. The properties of this spatial-attribute weighting function will be tested with regard to spatial continuity and statistical and biological interpretations.

The objectives of this study were (1) to model the relationship between tree size and growth using OLS and GWR with different weighting functions (i.e., spatial weighting and spatial-attribute weighting functions), (2) to compare model fitting for the OLS and the two GWR models using a goodness-of-fit test and through mapping parameter estimates used to interpret individual tree growth, and (3) to evaluate the performance of the two GWR models with different weighting functions.

Data

The data used in this study were collected from 48 plots of mixed-species, second-growth northern hardwoods in the Bartlett Experimental Forest at Bartlett, New Hampshire, which were established as a density study in 1963 (Leak and Solomon 1975; Solomon 1977). The dominant species in these plots were beech (*Fagus grandifolia* Ehrh.), red maple (*Acer rubrum* L.), and paper birch (*Betula papyrifera* Marsh). Some other species included yellow birch (*Betula alleghaniensis* Britt.), sugar maple (*Acer saccharum* Marsh.), white ash (*Fraxinus americana* L.), and miscellaneous softwoods. Each plot was a 36.58 m × 36.58 m square and was surrounded by a 15.24 m isolation strip. All trees on the plots were marked, tagged, and recorded by species. Tree diameter at breast height (DBH) was measured for trees with DBH equal to or greater than 11.43 cm at plot establishment and remeasured several times afterward. Plot maps were also constructed in 1991 following the field procedure developed by Reed et al. (1989) for measuring tree coordinates. Thus, the coordinates of each tree and the distances between trees were available. The last remeasurements were taken in summer 2000. The tree basal area growth between 1991 and 2000 was used in the regression model in this study.

Methods

GWR model

Suppose we have a set of n observations $\{X_{ij}\}$ with the spatial coordinates $\{(u_i, v_i)\}$, i = 1, 2, …, n, on p independent or predictor variables, j = 1, 2, …, p, and a set of n observations on a dependent or response variable $\{y_i\}$. The underlying model for GWR is

$$[1] \qquad y_i = \beta_0(u_i, v_i) + \sum_{j=1}^{p} X_{ij}\, \beta_j(u_i, v_i) + \varepsilon_i$$

where $\{\beta_0(u_i, v_i), \beta_1(u_i, v_i), \ldots, \beta_p(u_i, v_i)\}$ are p + 1 continuous functions of the location (u_i, v_i) in the study area. The ε_i is the random error term with a distribution $N(0, \sigma^2 I)$.

The parameter estimation is a moving window process. A region or window was drawn around a location i, and all the data points within this region or window were then used to estimate the parameters in eq. 1. The estimator of β_i is given at each location i by a weighted least-squares approach:

$$[2] \qquad \hat{\beta}_i = [X^T \mathbf{W}_i(u_i, v_i) X]^{-1} X^T \mathbf{W}_i(u_i, v_i) y$$

where $\mathbf{W}_i(u_i, v_i)$ is an n by n matrix:

$$[3] \qquad \mathbf{W}_i(u_i, v_i) = \begin{pmatrix} w_{i1} & 0 & \ldots & 0 \\ 0 & w_{i2} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & w_{in} \end{pmatrix}$$

In locally weighted regression models, the values of $\mathbf{W}_i(u_i, v_i)$ are constant. In the GWR model, on the other hand, $\mathbf{W}_i(u_i, v_i)$ varies with the location i depending on the distance between location i and its neighboring locations (see eq. 4). The above process was repeated for each observation in the data, and consequently, a set of parameter estimates was obtained for each location.

The weights (w_{ij}) in the weight matrix $\mathbf{W}_i(u_i, v_i)$ is a decreasing function of distance d_{ij} between the subject i and its neighboring location j. In general, the spatial weighting function is taken as the exponential distance-decay form:

$$[4] \qquad w_{ij} = \mathrm{e}^{\left(-\frac{d_{ij}^2}{h^2}\right)}$$

where h is the kernel bandwidth. 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. However, the weights are nonzero for all data points, no matter how far they are from the center i (Fotheringham et al. 2002).

Kernel bandwidth plays an important role in model fitting. There are three common approaches available for choosing the "optimal" kernel bandwidth: (1) a predefined bandwidth, (2) using a cross-validation procedure, and (3) using a minimum Akaike information criterion (AIC) (Brunsdon et al. 1998; Fotheringham et al. 2000, 2002). Using the AIC and cross-validation methods to process large samples is computationally intense. Pàáez et al. (2002) also found that the cross-validation procedure sometimes might not result in a reasonable kernel bandwidth. In contrast, the use of the predefined bandwidth is quite simple. It generally depends on existing knowledge and the researchers' experience.

Spatial-attribute weighting function

The spatial weighting function (eq. 4) takes only the geographical distance into account and ignores the influence of the trees' attributes. Since both tree size and location have a strong impact on competition among trees, growth, crown structure, and mortality (e.g., Miller and Weiner 1989; Moeur 1993; Newton and Jollife 1998), we propose to modify eq. 4 as follows:

$$[5] \qquad w_{ij} = \mathrm{e}^{\left(-\frac{d_{ij}^2}{h^2}\times f(\tau)\right)}$$

where $f(\tau)$ is a function that changes the weight w_{ij} according to the difference (τ) between the size of the subject tree and the size of its neighbor. According to the idea of the weighting function in locally weighted regression techniques (Cleveland 1979; Castti 1982; Cleveland and Devlin 1988), the weight should decrease as the difference between the focal point and its neighbors increases. The symmetric weight is one of the important properties of the weighting function, because it reduces bias (Cleveland and Devlin 1988). The $f(\tau)$ function can be a bisquare, a "tribcube", or an exponential function. In this study we propose the following format for the $f(\tau)$ function:

$$[6] \qquad f(\tau) = \exp\left(\left|1 - \frac{\mathrm{DBH}_{ij}}{\mathrm{DBH}_{ii}}\right|\right)$$

where DBH_{ii} is the DBH of the subject tree i, and DBH_{ij} is the DBH of the neighboring tree j. According to eqs. 5 and 6, large weights are assigned to the neighboring trees with DBHs that are similar to that of the subject tree, and the small weights are assigned to the neighboring trees if their DBHs are different from that of the subject tree. When the size of a neighboring tree is the same as that of the subject tree (i.e., $f(\tau) = 1$), the weight w_{ij} for that tree is determined by the spatial distance only.

Regression model

We selected the following model for the relationship between tree growth and size, which has been used successfully in similar situations (Vanclay 1994; Zhang and Shi 2004).

$$[7] \qquad \log(\mathrm{BAG}+1) = \beta_0(u,v) + \beta_1(u,v)\log(\mathrm{DBH}) + \beta_2(u,v)\,\mathrm{DBH}^2 + \varepsilon$$

where BAG is the tree basal area growth between 1991 and 2000, DBH is the tree DBH in 1991, log is a 10-based logarithm, $\beta_0(u, v)$~ $\beta_2(u, v)$ are regression coefficients to be estimated, and ε is the model random error. If the spatial coordinates are removed from the above model, eq. 7 becomes the derivative model of the Bertalanffy growth function. This model has been used as a basic function in several forest growth and yield models because of its simplicity and robust predictions (e.g., Wykoff 1990; Hann and Larsen 1991; Vanclay 1994; Monserud and Sterba 1996).

Goodness-of-fit test

It is important to test whether the GWR models offer a statistically significant improvement over the OLS model. In GWR, the parameter estimates change as the spatial coordinates vary. By incorporating the local spatial information in the model estimation, GWR always provides a better model fit in terms of the residual sum of squares (Brunsdon et al. 1998; Fotheringham et al. 2002). In this study, we performed the goodness-of-fit test using the approach proposed by Leung et al. (2000) to test whether GWR is an improvement over OLS (Zhang and Shi 2004).

Evaluation of the spatial and spatial-attribute weighting functions

Bootstrapping, jack-knifing, and cross-validation are common methods based on "resampling" for evaluating regression models when new data are not available for the same purpose (Efron and Gong 1983; Shao and Tu 1995). In a recent study, Kozak and Kozak (2003) found that cross-validation or data splitting provides little additional information in the process of evaluating regression models. Jack-knifing or predicted residual error sum of squares (PRESS) in regression analysis may not be an appropriate method for evaluating the GWR models because the sample size for any subject tree within the kernel bandwidth is usually small. Omitting one or more observation(s) will make the sample size even smaller. Thus, we decided to use bootstrapping for the evaluation of the GWR model with our proposed spatial-attribute weighting function against the model with the spatial weighting function (Efron and Gong 1983; Mooney and Duval 1993; Fox 1997). During the bootstrapping process, 100 random samples with replacement were drawn from all

Fig. 1. Maps of tree locations for the three example plots: (*a*) regularity, (*b*) randomness, and (*c*) clustering. The circle is proportional to the tree DBH.

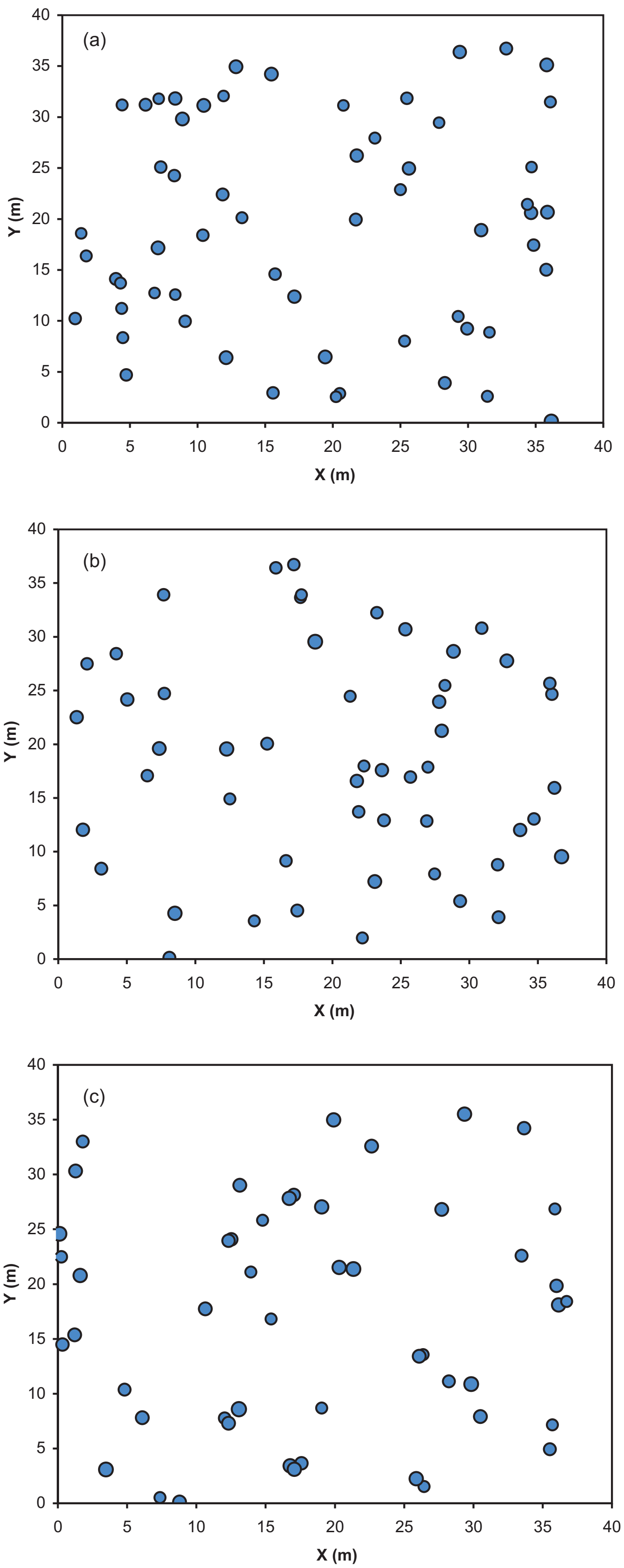
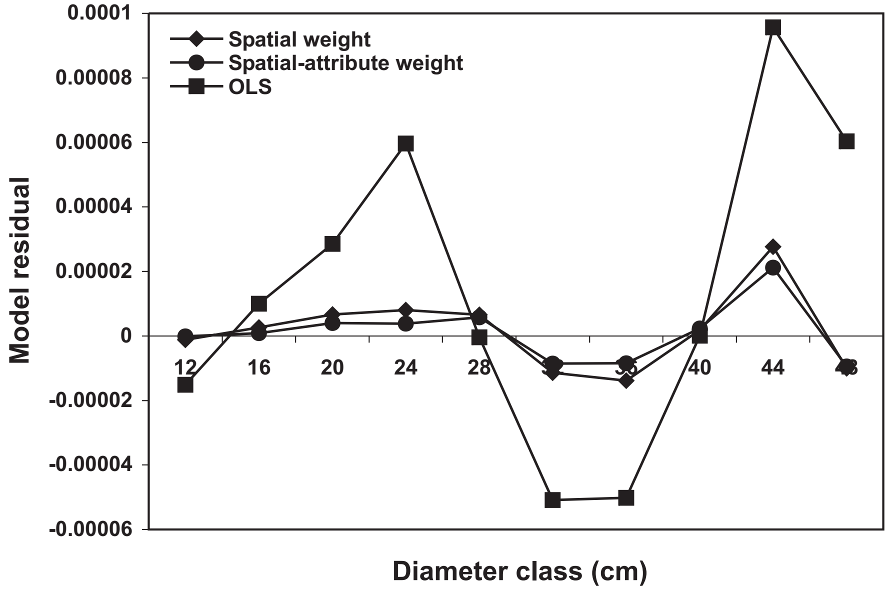


Table 1. Goodness-of-fit test for the improvement of the geographically weighted regression (GWR) models over the ordinary least-square (OLS) model for the 48 plots.

	Spatial weighting GWR		Spatial-attribute weighting GWR	
Plot	F^a	P^a	F	P
01	2.5223	0.0017	2.6110	0.0025
02	2.0390	0.0130	1.8809	0.0333
03	5.2687	<0.0001	5.7664	0.0001
04	6.1597	<0.0001	6.6956	<0.0001
05	3.9897	0.0324	6.2960	0.0246
06	2.4673	0.0031	2.5129	0.0045
07	2.2733	0.0014	2.9239	0.0001
08	2.3188	0.0056	2.3097	0.0086
09	1.8045	0.0485	1.7131	0.0810
10	1.9815	0.0482	2.4301	0.1635
11	2.0855	0.0414	2.2536	0.0369
12	1.6308	0.0474	1.6758	0.0642
13	2.2761	0.0073	2.2218	0.0136
15	1.7089	0.0376	1.6955	0.0449
16	1.5265	0.0766	1.9351	0.0172
17	2.7232	0.0031	2.6660	0.0073
18	3.1648	0.0046	3.9078	0.0041
19	4.6006	<0.0001	4.8624	0.0000
20	1.2741	0.2406	1.4967	0.1381
21	1.9180	0.0122	2.1911	0.0050
22	4.2975	0.0001	4.6588	0.0002
25	2.9870	0.0005	3.4297	0.0003
26	1.7848	0.0447	1.7015	0.1861
29	3.7249	0.0028	4.0095	0.0038
30	2.0904	0.0462	2.6593	0.0340
31	2.0691	0.0077	2.3319	0.0043
32	3.9460	0.0136	4.5785	0.0222
33	3.4593	0.0041	3.3886	0.0085
34	2.7529	0.0333	3.0080	0.0346
35	1.0696	0.4150	0.9851	0.5288
36	2.5141	0.0063	2.9397	0.0046
37	5.3968	0.0063	5.2163	0.0156
38	3.0901	<0.0001	3.0259	<0.0001
39	2.0412	0.0353	2.1114	0.0440
40	2.2591	0.0173	2.2903	0.0209
41	4.3077	0.0107	5.2538	0.0386
42	3.0153	0.0227	3.9896	0.0315
43	5.7198	<0.0001	5.9300	<0.0001
45	1.1837	0.3279	3.7878	0.0008
46	2.3584	0.0087	2.2756	0.0191
47	3.6691	0.0003	4.3857	0.0002
49	2.3901	0.0041	2.1904	0.0108
50	2.3224	0.0014	2.6537	0.0005
51	3.8221	<0.0001	3.8169	<0.0001
53	3.3976	0.0001	4.0204	<0.0001
54	1.8322	0.0457	1.8244	0.1710
56	3.1504	0.0004	4.3216	0.0001
59	4.8137	0.0010	6.0975	0.0021

[a]See Zhang and Shi (2004).

Fig. 2. Model residuals across 4 cm diameter classes over 48 plots.



the trees (including the subject tree) surrounding the subject tree within the kernel bandwidth for each location. The following equation was used to compute the bootstrapping root mean square error (RMSE):

$$[8]\qquad \mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}\left[\left(\sum_{j=1}^{100}\mathrm{MRES}_{ij}^{2}\right)/100\right]}{n}}$$

where MRES_{ij} is the model residual, n is the number of observations, i is the ith tree, j is the index of bootstrapping replication. RMSE provides an overall measure of predictive performance of the two GWR models with different weighting functions.

Example plots

Using Ripley's K function (Ripley 1977; Diggle 1983; Haase 1995), we grouped the 48 plots into three categories: regularity (12 plots), randomness (30 plots), and clustering (6 plots). Three example plots (plots 17, 39, and 59) were selected from the three categories to further assess the spatial variation of parameter estimates and differences between the two GWR models with different weighting functions. We selected these three example plots because their tree locations represented typical regular, random, and clustered spatial patterns, respectively, based on the analysis of the Ripley's K function. In addition, the number of trees in these three example plots was similar (n = 59 in plot 17, n = 52 in plot 39, and n = 50 in plot 59). We evaluated the impact of spatial patterns on tree growth through the contour maps of the parameter estimates for each example plot. The initial spatial locations of each tree in the three example plots are shown in Fig. 1. The circles in the figure are proportional to the sizes of tree DBH.

Results

Determination of kernel bandwidth

In this study, we decided to predefine the bandwidth for the 48 plots based on the variogram of the OLS model residuals. Our analysis indicated that the variogram curve tended to flatten out after 5.2 m (i.e., range = 5.2 m), meaning that there was no spatial autocorrelation between trees beyond the distance (Isaaks and Srivastava 1989; Kohl and Gertner 1997). Therefore, we chose 5.2 m as the predefined kernel bandwidth. The reasons for using predefined kernel bandwidth were (1) the plot size is relatively small (0.135 ha), and therefore unreasonable kernel bandwidths could be obtained using the AIC and cross-validation methods, and (2) different kernel bandwidths had to be chosen for these 48 plots. To make comparisons among these 48 plots, we need to make a compromise by using a predefined kernel bandwidth. In addition, similar distances (or kernel bandwidth) have been used in other studies of distance-dependent competition indices (e.g., Pukkala 1989; Kenkel et al. 1989; Rouvinen and Kuuluvainen 1997; Shi and Zhang 2003; Zhang and Shi 2004).

Improvement of the GWR models over the OLS model

All trees in each plot were used to fit eq. 7 by OLS, GWR with the spatial weighting function, and GWR with the spatial-attribute weighting function. The null hypothesis of no improvement of the GWR models over the OLS model was tested using the approximate F test proposed by Leung et al. (2000). The results indicated that both GWR models performed better than the OLS model (Table 1). For the GWR model with the spatial weighting function, the model fitting was significantly improved over the OLS model (α = 0.05) for 44 (92%) of the 48 plots. In the case of the GWR model with the spatial-attribute weighting function, 42 (87.5%) plots showed an improvement in model fitting. The goodness-of-fit test indicated that the model parameter estimates of eq. 7 were better modeled as a spatially variable parameter from location to location within each plot. In other words, the relationship between BAG and DBH was not constant across each plot.

The two GWR models generally produce smaller model residuals than the OLS model across 4 cm diameter classes

Table 2. The comparison of root mean square error (RMSE) obtained from the geographically weighted regression (GWR) models with the spatial and spatial-attribute weighting functions using bootstrapping.

	RMSE	
Plot	Spatial weighting GWR	Spatial-attribute weighting GWR
01	0.00035	0.00025
02[a]	0.00048	0.00049
03	0.00066	0.00036
04	0.00037	0.00021
05	0.00060	0.00041
06	0.00040	0.00029
07	0.00041	0.00023
08	0.00035	0.00026
09	0.00046	0.00034
10[a]	0.00097	0.00204
11	0.00049	0.00046
12	0.00053	0.00038
13	0.00037	0.00028
15	0.00038	0.00028
16	0.00041	0.00025
17	0.00062	0.00049
18	0.00078	0.00053
19	0.00078	0.00072
20	0.00054	0.00042
21	0.00046	0.00029
22	0.00045	0.00026
25	0.00049	0.00024
26	0.00080	0.00067
29[a]	0.00128	0.00137
30[a]	0.00406	0.00468
31	0.00036	0.00022
32	0.00071	0.00041
33	0.00063	0.00050
34	0.00057	0.00029
35	0.00075	0.00059
36	0.00045	0.00037
37	0.00108	0.00091
38	0.00031	0.00023
39[a]	0.00064	0.00129
40	0.00059	0.00055
41	0.00066	0.00036
42	0.00076	0.00130
43	0.00040	0.00024
45	0.00041	0.00030
46	0.00039	0.00033
47	0.00047	0.00030
49[a]	0.00051	0.00052
50[a]	0.00044	0.00063
51	0.00044	0.00021
53	0.00043	0.00027
54[a]	0.00045	0.00046
56	0.00081	0.00071
59	0.00064	0.00046

[a]The RMSE obtained from the GWR model with the spatial-attribute weighting function is larger than that from the GWR model with the spatial weighting function.

(Fig. 2). The OLS model appears to produces much larger negative biases (overestimation) for intermediate-sized trees (28–40 cm in diameter) and much larger positive biases (underestimation) for both small trees (16–24 cm in diameter) and large trees (>40 cm in diameter) than the two GWR models. However, the two GWR models had similar patterns of model residuals across the diameter classes.

The equivalence test using paired *t* test was performed to compare the average model residuals and absolute residuals between the two GWR models across the 48 plots. We followed the testing procedure proposed by Robinson and Froese (2004). We chose three criteria (i.e., ε = 10%, 25%, and 50%) that were relative to the standard deviation of the difference of the two model residuals. The cutoff ($C_{\alpha;n-1}(\varepsilon)$) was obtained from the noncentral *F* distribution. According to the equivalence test, if the *t* value was greater than the cutoff value, then the null hypothesis of dissimilarity would not be rejected (Robinson and Froese 2004). Our equivalence test indicated that the GWR model with the spatial-attribute weighting function produced significantly smaller model residuals (t = 2.43, $C_{0.01;47}(10\%)$ = 0.02, $C_{0.01;47}(25\%)$ = 0.06, and $C_{0.01;47}(50\%)$ = 1.13), and absolute residuals (t = 18.13) than did the GWR model with the spatial weighting function. Therefore, our proposed spatial-attribute weighting function performed better than the spatial weighting function currently used in GWR in terms of model fitting.

Evaluation of the spatial and spatial-attribute weighting functions

The comparison of the bootstrapping RMSE in Table 2 indicated that the GWR model with the spatial-attribute weighting function had smaller RMSE in 40 (83.33%) of the 48 plots than the GWR model with the spatial weighting function. Similarly, we used the equivalence test to test whether there was a significant difference between these two RMSEs. The results indicated that the GWR model with the spatial-attribute weighting function produced a significantly smaller bootstrapping RMSE than did the GWR model with the spatial weighting function (t = 1.51, the cutoff values were the same as above), implying that although the former model is more complicated because of the incorporation of the tree attribute information in the weighting function, it does provide better predictions for the response variable.

Further comparisons between the two GWR models using the three example plots

To further investigate the difference of model residuals between the two GWR models, we studied the model residual and absolute model residuals for the three example plots in detail. In general, the GWR model with the spatial-attribute weighting function produced smaller model residuals than did the GWR model with the spatial weighting function across tree diameter classes (Fig. 3). The difference between the two models was smaller for the regular plot (Fig. 3*a*) than for the clustered plot (Fig. 3*c*).

The difference between the two GWR models was clearer for the absolute model residuals (Fig. 4). The absolute model residuals obtained from the GWR model with the spatial-attribute weighting function were smaller than those of the

Fig. 3. Model residuals across the diameter classes for the three example plots: (*a*) regularity, (*b*) randomness, and (*c*) clustering.

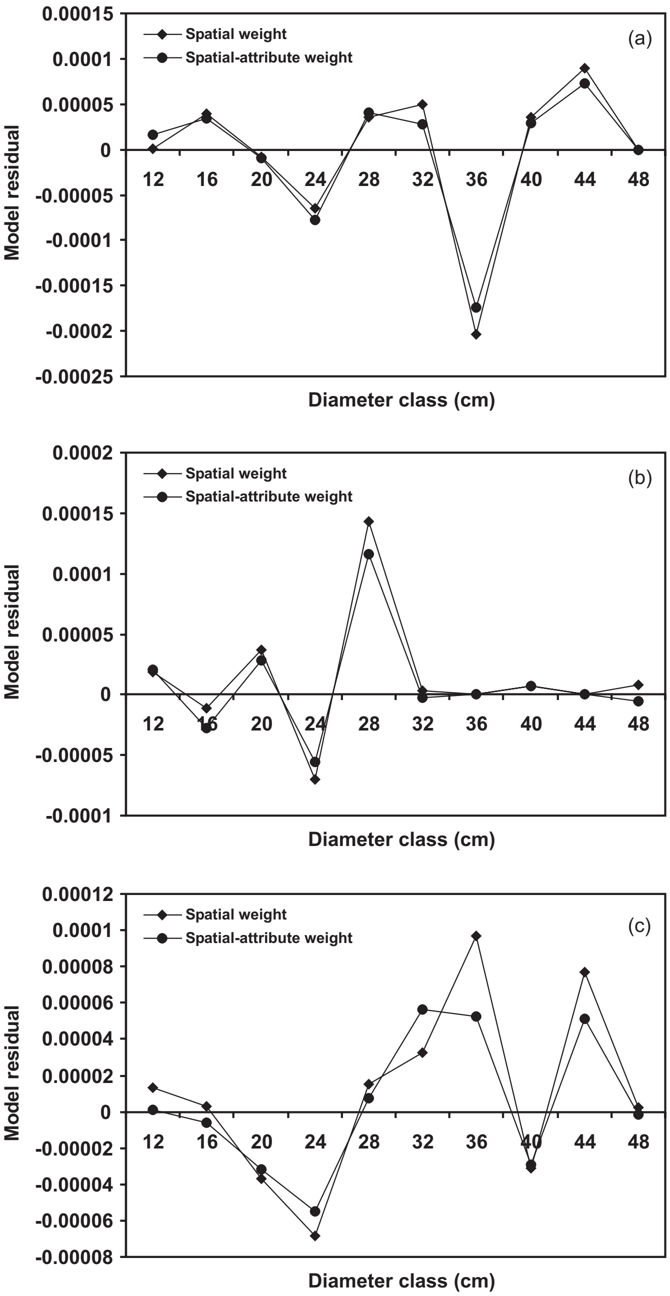
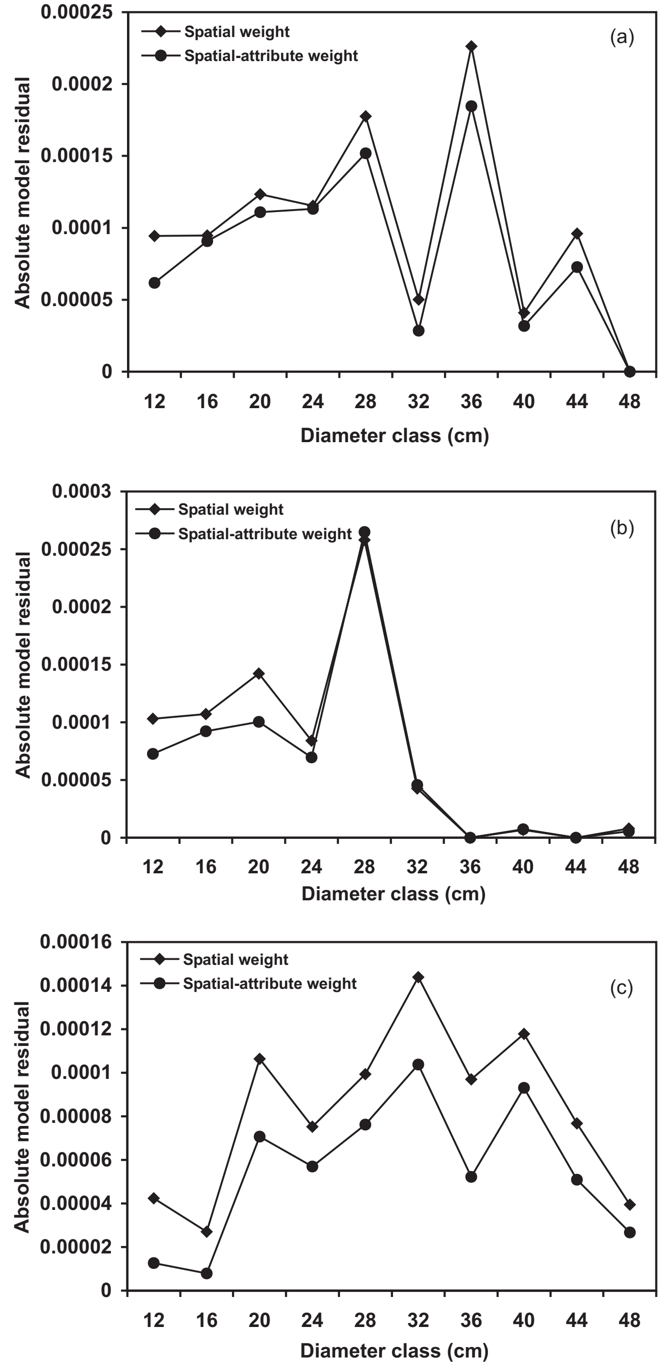


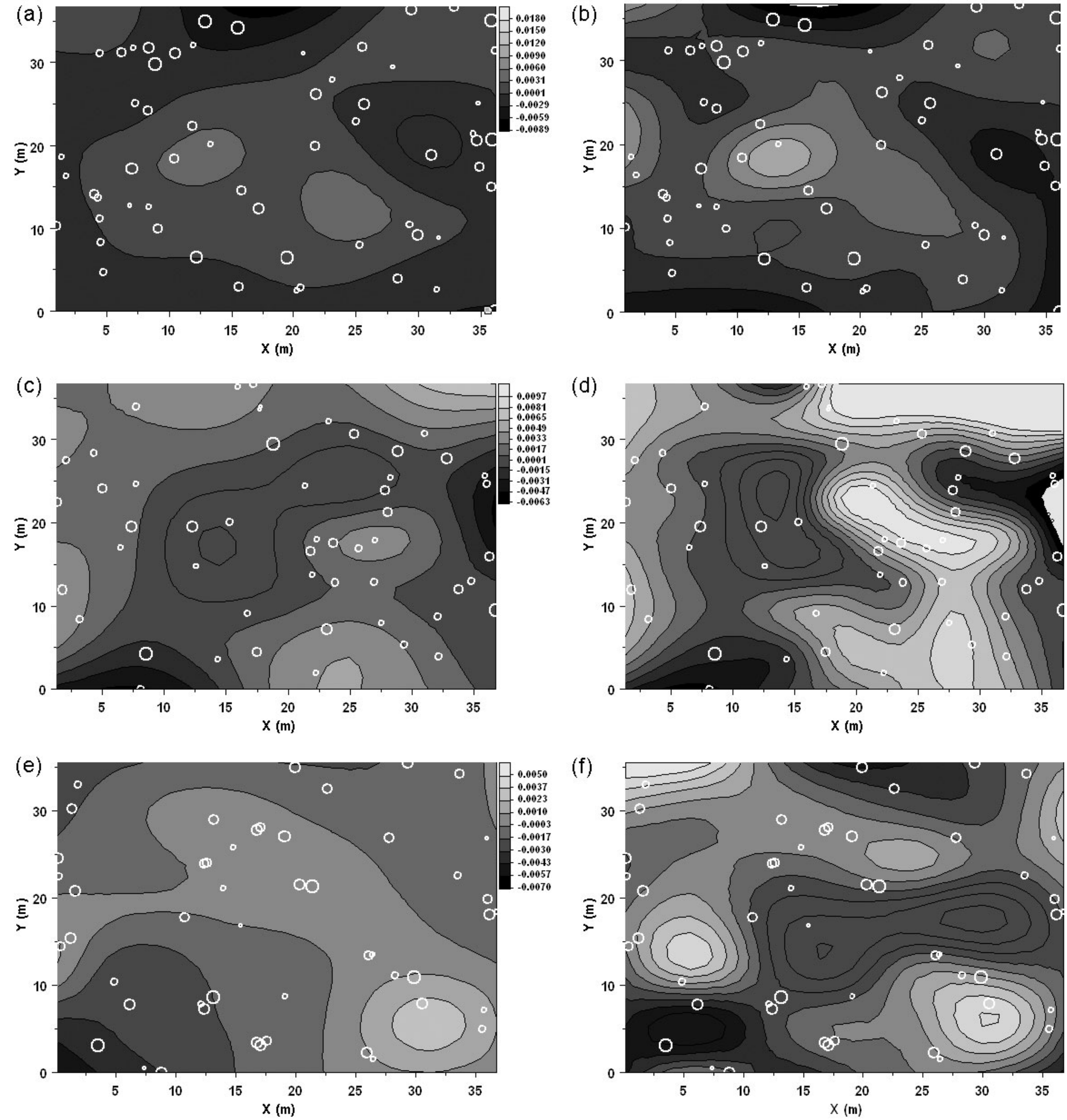
Fig. 4. Absolute model residuals across the diameter classes for the three example plots: (*a*) regularity, (*b*) randomness, and (*c*) clustering.



GWR model with the spatial weighting function across tree diameter classes, especially for the clustered plot (Fig. 4*c*).

In general, the range of the parameter estimates obtained from the GWR model with the spatial-attribute weighting function was wider than that of GWR model with the spatial weighting function. Figure 5 shows the contour plots of the parameter estimates (β_1) of eq. 7 for the three example plots. To make the comparison among these contour maps, the same range of parameter estimates was used for each example plot. The wide range of parameter estimates locally highlighted the large variation of parameter estimates in each example plot. For the regular spatial pattern, the large variation of the parameter estimates (β_1) was close to the center of this plot. The range of the parameter estimates (β_1) was from –0.0063 to 0.0041 for the GWR model with the spatial weighting function (Fig. 5*a*), while it varied from –0.0089 to 0.018 for the GWR model with the spatial-attribute weighting method (Fig. 5*b*). For the random and clustered spatial patterns, similar results can be seen (Figs. 5*c*, 5*d*, 5*e*, and 5*f*). The GWR model with the new spatial-attribute weighting function took the local spatial variation of tree density, growth conditions, and competition (e.g., tree size) into ac-

Fig. 5. Contour plots of the parameter estimates of β_1 in eq. 7: (*a*) regular plot, spatial weighting function; (*b*) regular plot, spatial-attribute weighting function; (*c*) random plot, spatial weighting function; (*d*) random plot, spatial-attribute weighting function; (*e*) clustered plot, spatial weighting function; and (*f*) clustered plot, spatial-attribute weighting function.



count, and therefore the wider range of parameter estimates might represent the local conditions better than the smaller range. It would improve the precision and accuracy of model predictions.

Discussion

The choice of spatial kernel is very important in GWR. Two types of spatial kernels are available: fixed and adaptive spatial kernels. In this study, we employed a fixed spatial kernel for each example plot. In general, the adaptive spatial kernel is suitable for sparse data. In other words, the kernel would be different at different locations. In this case, if the fixed spatial kernel was used, there might have been few data points in the kernel bandwidth for the sparse data, which can cause large standard errors for parameter estimates and "undersmoothed" surfaces (Fotheringham et al. 2002). In this study, the cross-validation and AIC methods cannot be used to obtain the adaptive kernels because of the size limitation of our example plots. Although it would be useful to test the model fitting using the spatial adaptive kernel, it is almost unrealistic to determine the spatial adaptive kernels for each data point in our example plots according to subjective judgment. If the plot size is large enough, the adaptive kernel can be obtained using the cross-validation or AIC method. We are aware of the size limitation of our example plots. Our proposed spatial-attribute weighting function needs to be tested using large field plots in the future.

The weights matrices obtained from the weighting functions are key elements in most regression models (Getis and Aldstadt 2004). The spatial weights matrix in GWR is defined as the expression of spatial dependence between observations (Fotheringham et al. 2002). However, according to our study, constructing a weights matrix should take into account dependences not only from "geographical space" but also from "attribute space" in local statistical analysis and spatial modeling. We proposed the spatial-attribute weighting function (eq. 5) to obtain the weights matrix, and therefore the local variation being measured in attribute space and geographical space can be coped with at the same time. It can be further used to investigate local conditions influencing individual tree growth.

Although the contour plots indicate the existence of spatial heterogeneity for each example plot, it is difficult to determine the impact of attribute space in the local parameter estimates. The impact of attribute space is the combined effect of the neighbors of the subject tree. Therefore, it is a multidimensional space and cannot be visually investigated. The effect of attribute space should be cautiously interpreted across space. However, from the comparison of the contour plots between the two GWR models, we can see that the tree attributes do have an influence on the regression coefficients.

Conclusions

The GWR method is a useful tool for modeling spatial heterogeneity according to the analysis of the model fitting with our example plots (i.e., regular, random, and clustered plots). Not only the spatial information but also the attributes of trees can be incorporated into the weighting function of the GWR model. Our results indicate that no matter what spatial patterns existed in the stand, the performance of the GWR model with the spatial-attribute weighting function would be better than that of the GWR model with the spatial weighting function, especially for the plots where trees are clustered. The visualization of the parameter estimates highlights the spatial distribution of the multivariate relationship under study, such as the relationship between tree growth and spatial locations. With the rapid development of GIS, the GWR model can be easily incorporated into GIS for forest growth simulations.

Acknowledgments

The financial support for this study was provided by the Michigan Department of Natural Resources and a USDA National Research Initiative grant.

References

Anselin, L. 1988. Spatial econometrics: methods and models. Kluwer Academic Publishers, Dordrecht, Netherlands

Brunsdon, C.A., Fotheringham, A.S., and Charlton, M.E. 1996. Geographically weighted regression: a method for exploring spatial nonstationarity. Geogr. Anal. **28**: 281–298.

Brunsdon, C.A., Fotheringham, A.S., and Charlton, M.E. 1998. Geographically weighted regression — modeling spatial nonstationary. Statistician, **47**(3): 431–443.

Casetti, E. 1972. Generating models by the expansion method: application to geographic research. Geogr. Anal. **4**: 81–91.

Casetti, E. 1982. Drift analysis of regression parameters: an application to the investigation of fertility development relations. Model Simul. **13**: 961–966.

Casetti, E. 1997. The expansion method, mathematical modeling, and spatial econometrics. Int. Reg. Sci. Rev. **20**: 9–32.

Casetti, E., and Can, A. 1999. The economic estimation and testing of DARP models. J. Geogr. Syst. **1**: 91–106.

Cleveland, W. 1979. Robust locally weighted regression and smoothing scatterplots. J. Am. Stat. Assoc. **74**: 829–836.

Cleveland, W.S., and Devlin, S.J. 1988. Locally weighted regression: an approach to regression analysis by local fitting. J. Am. Stat. Assoc. **83**: 596–610.

Diggle, P.J. 1983. Statistical analysis of spatial point patterns. Academic Press, London.

Dovciak, M., Frelich, L.E., and Reich, P.B. 2001. Discordance in spatial patterns of white pine (*Pinus strobus*) size-classes in a patchy near-boreal forest. J. Ecol. **89**: 280–291.

Efron, B., and Gong, G. 1983. A leisurely look at the bootstrap, the jackknife, and cross-validation. Am. Stat. **37**: 36–48.

Fotheringham, A.S., and Brunsdon, C. 1999. Local forms of spatial analysis. Geogr. Anal. **31**: 340–358.

Fotheringham, A.S., Charlton, M.E., and Brunsdon, C. 1996. The geography of parameter space: an investigation into spatial nonstationarity. Int. J. Geogr. Inf. Syst. **10**: 605–627.

Fotheringham, A.S., Brunsdon, C., and Charlton, M.E. 2000. Quantitative geography: perspectives on spatial data analysis. SAGE Publications, Newbury Park, Calif.

Fotheringham, A.S., Brunsdon, C., and Charlton, M.E. 2002. Geographically weighted regression: the analysis of spatially varying relationships. John Wiley & Sons, Inc., Chichester, UK.

Fox. J. 1997. Applied regression analysis, linear models, and related methods. SAGE Publications, Newbury Park, Calif.

Fox, J.C., Ades, P.K., and Bi, H. 2001. Stochastic structure and individual-tree growth models. For. Ecol. Manage. **154**: 261–276.

Getis, A., and Aldstadt, J. 2004. Constructing the spatial weights matrix using a local statistic. Geogr. Anal. **36**: 90–104.

Gorr, W.L., and Olligschlaeger, A.M. 1994. Weighted spatial adaptive filtering: Monte Carlo studies and application to Illicit Drub Market Modeling. Geogr. Anal. **26**: 67–87.

Haase, P. 1995. Spatial pattern analysis in ecology based on Ripley's K-function: introduction and methods of edge correction. J. Veg. Sci. **6**: 575–582.

Haining, R. 1978. Estimating spatial interaction models. Environ. Plann. A, **10**: 305–320.

Hann, D.W., and Larsen, D.R. 1991. Diameter growth equations for fourteen tree species in Southwest Oregon. Forest Research Laboratory, Oregon State University, Corvallis, Ore. Res. Bull. 69.

Isaaks, E.H., and Srivastava, R.M. 1989. An introduction to applied geostatistics. Oxford University Press, New York.

Jones, J., and Casetti, E. 1992. Applications of the expansion method. Routledge, London.

Kenkel, N.C., Hoskins, J.A., and Hoskins, W.D. 1989. Local competition in a naturally established jack pine stand. Can. J. Bot. **67**: 2630–2635.

Kohl, M., and Gertner, G. 1997. Geostatistics in evaluating forest damage surveys: considerations on methods for describing spatial distributions. For. Ecol. Manage. **95**: 131–140.

Kozak, A., and Kozak, R. 2003. Does cross validation provide additional information in the evaluation of regression models. Can. J. For. Res. **33**: 976–987.

Leak, W.B., and Solomon, D.S. 1975. Influence of residual stand density on regeneration of northern hardwoods. USDA For. Ser. Res. Pap. NE-310.

Lee, J., and Wong, D.W.S. 2001. Statistical analysis with ArcView GIS. John Wiley and Sons, Inc., New York.

Leung, Y., Mei, C.L., and Zhang, W.X. 2000. Statistical tests for spatial nonstationarity based on the geographically weighted regression model. Environ. Plann. A, **32**: 9–32.

Liu, J., and Ashton, P.S. 1999. Simulating effects of landscape contexts and timber harvest on tree species diversity. Ecol. Appl. **9**: 186–201.

Miller, T.E., and Weiner, J. 1989. Local density variation may mimic effects of asymmetric competition on plant size variability. Ecology, **70**: 1188–1191.

Moeur, M. 1993. Characterizing spatial patterns of trees using stem-mapped data. For. Sci. **39**: 756–775.

Monserud, R.A., and Sterba, H. 1996. A basal area increment model for individual trees growing in even- and uneven-aged forest stands in Austria. For. Ecol. Manage. **80**: 57–80.

Mooney, C.Z., and Duval, R.D. 1993. Bootstrapping: a nonparametric approach to statistical inference. SAGE Publications, Newbury Park, Calif. Series 07-095.

Newton, P.F., and Jolliffe, P.A. 1998. Assessing processes of intraspecific competition with spatially heterogeneous black spruce stands. Can. J. For. Res. **28**: 259–275.

Pàáez, A., Uchida, T., and Miyamoto, K. 2002. A general framework for estimation and inference of geographically weighted regression models: 1. Location-specific kernel bandwidths and a test for locational heterogeneity. Environ. Plann. A, **34**: 733–754.

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.

Reed, D.D., Liechty, H.O., and Burton, A.J. 1989. A simple procedure for mapping tree location in forest stands. For. Sci. **35**: 657–662.

Ripley, B.D. 1977. Modelling spatial patterns (with discussion). J. R. Stat. Soc. B, **39**: 172–212.

Robinson, A.P., and Froese, R.E. 2004. Model validation using equivalence tests. Ecol. Model. **176**: 349–358.

Rouvinen, S., and Kuuluvainen, T. 1997. Structure and asymmetry of tree crowns in relation to local competition in a natural mature Scots pine forest. Can. J. For. Res. **27**: 890–902.

Shao, J., and Tu, D. 1995. The jackknife and bootstrap. Springer-Verlag, New York.

Shi, H., and Zhang, L. 2003. Local analysis of tree competition and growth. For. Sci. **49**: 938–955.

Shi, H., Laurent, E.J., LeBouton, J., Racevskis, L., Hall, K.R., Donovan, M., Doepker, R.V., Walters, M.B., Lupi, F., and Liu, J. 2006. Local spatial modeling of white-tailed deer distribution. Ecol. Model. **190**: 171–189.

Solomon, D.S. 1977. The influence of stand density and structure on growth of northern hardwoods in New England. USDA For. Serv. Res. Pap. NE-362.

Vanclay, J.K. 1994. Modeling forest growth and yield: application to mixed tropical forests. CAB International, Oxon, UK.

Weiner, J. 1990. Asymmetric competition in plant populations. Trends Ecol. Evol. **5**: 360–364.

Wykoff, W.R. 1990. A basal area increment model for individual conifers in the northern Rocky Mountains. For. Sci. **36**: 1077–1104.

Zhang, L., and Shi, H. 2004. Local modeling of tree growth by geographically weighted regression. For. Sci. **50**: 225–244.

Zhang, L., Bi, H., Cheng, P., and Davis, C.J. 2004. Modeling spatial variations in tree diameter–height relationships. For. Ecol. Manage. **189**: 317–329.