

Evaluation of structure specification in linear mixed models for modeling the spatial effects in tree height-diameter relationships

J. Lu, L. Zhang

Lu J., Zhang L., 2013. Evaluation of structure specification in linear mixed models for modeling the spatial effects in tree height-diameter relationships. *Ann. For. Res.* 56(1): 137-148, 2013.

Abstract. In recent years, linear mixed models (LMM) have become more popular to deal with spatial effects in forestry and ecological data. In this study, different structure specifications of linear mixed model were applied to model tree height-diameter relationships, including LMM with random blocks only (LMM-block), LMM with spatial covariance only (LMM-covariance), and the combination of the last two (LMM-block-covariance). Further, the between-group heterogeneous variances were incorporated into LMM-covariance and LMM-block-covariance. The results indicated that, in general, LMM-covariance significantly reduced spatial autocorrelation in model residuals, while LMM-block was effective in dealing with spatial heterogeneity. LMM-block treated the blocks as random effects and avoided the estimation of parameters of the variogram model. Thus, it produced better model predictions than LMM-covariance. LMM-block-covariance took both block effects and spatial covariance into account, and significantly improve model fitting. However, it did not produce better model predictions due to the increase of model complexity and estimation of the local variogram within each block.

Keywords linear mixed model, spatial heterogeneity, spatial autocorrelation, height-diameter equation.

Authors. Junfeng Lu - Lerfob, INRA, Nancy 54000, France; Lianjun Zhang (lizang@syr.edu) - Department of Forest and Natural Resources Management, State University of New York, College of Environmental Science and Forestry, One Forestry Drive, Syracuse, NY 13210, USA.

Manuscript received July 23, 2012; revised October 29, 2012; accepted November 09, 2012; online first November 16, 2012.

Introduction

The development of geographic information systems (GIS) in recent years has made

geographically referenced information easily available for researchers and made more comprehensive data exploration, analysis and modeling possible. As one type of spatial analysis, spatial regression is used to capture the spatial

dependency in the relationships between variables in order to overcome statistical problems such as unstable parameter estimation and unreliable significance testing. However, the use of geo-spatial information leads to a number of features that merit special attention, because spatial regression models are different from traditional (aspatial) regression models or time series models in terms of model errors. The two main characteristics of the model errors in spatial regression are spatial dependence and heterogeneity (Anselin 1999, Fotheringham et al. 2002). Spatial dependence is mostly due to the existence of spatial spillovers because of the miss-match between the scale of the spatial unit of observations and the phenomenon of interest, while spatial heterogeneity is due to the structural differences between locations and led to different error distributions (Anselin 1999). These distinguished characteristics have important implications for description, explanation, and prediction in spatial analysis and modeling.

For a linear regression model, the spatial dependence can be incorporated in two ways: (1) using a spatially lagged dependent variable as an additional regressor or predictor in the model, and (2) specifying a particular variance-covariance structure for the model errors. The former is referred to as a spatial lag model and the latter is called a spatial error model (Anselin 1988). In recent years, geographically weighted regression (GWR) has become popular to explore and model spatial heterogeneity (Brunsdon et al. 1996, Fotheringham et al. 2002, Zhang et al. 2004). GWR attempts to capture spatial variations by calibrating a regression model at different locations in space that allows the different relationships between variables of interests (Zhang et al. 2004). In forestry and ecological applications, Fox et al. (2007) used moving average autoregressive models for individual tree growth models. Zhang & Shi (2004) and Shi et al. (2006) tried GWR as a local modeling technique for tree growth.

Linear mixed models (LMM) have become popular for dealing with spatial dependence and heterogeneity over the last 15 years. Researchers have been applying different techniques for this purpose under various circumstances. One way is to use blocking approach to diminish the effect of variations among experimental or sampling units. The block effects are usually considered random because the blocks in an experiment are only a subset of all available blocks on which the statistical inference about treatment means is to be made (Littell et al. 2006). Thus, one can separate the entire study area or data observations into “blocks” such as forest plots or stands, or tree species, or even an individual tree with repeated measures. By taking these “blocks” as the random effect, LMM can incorporate spatial heterogeneity and/or spatial dependence into the modeling process. Another way is to specify a spatial variance-covariance structure or function to represent the spatial dependence or autocorrelation in the model errors, which is incorporated into the process of parameters estimation (Littell et al. 2006). It is similar to the way of dealing with repeated measurements (Wolfinger 1996). Further, the spatial prediction can be implemented by a Kriging procedure based on the spatial variance-covariance function estimated by LMM (Littell et al. 2006, Minasny & McBratney 2007).

Researchers also utilize other forms of LMM, which have limited practical applications, but may have beneficial impacts on the outcomes of spatial data analysis and modeling. For example, random coefficient models are built on the analysis of covariance models and can be used to treat the regression coefficients for one or more covariates as a random sample from a population of possible coefficients (Littell et al. 2006). Heteroscedasticity (i.e., heterogeneous variances) can be incorporated into the covariance structures of LMM to overcome the violation of the homogeneous variance assumption for the model errors (Dutilleul & Legendre 1993). There are within-group and

between-group heterogeneous variances. A particular and useful term of LMM with within-group variances is known as the weighted linear mixed model (WLMM), where the variance is modeled as a function of exploratory variable(s) (Bates 2008). Therefore, there are multiple features in LMM that researchers can use to explore and model spatial dependence and heterogeneity.

LMM has drawn attentions from forestry and ecological researchers. Calama and Montero (2004), Meng et al. (2007), and Vazquez & Pereira (2005) applied LMM with the random effects specified for geographical regions, plots or individual trees. Calegario et al. (2005), Budhathoki et al. (2008), Nothdurft et al. (2006), and Gregoire & Schabenberger (1996) used nonlinear mixed models for tree growth. Zhang et al. (2008) incorporated the spatial dependence into the modeling process using LMM with a spatial covariance structure.

In this study, we attempted to apply several structures of LMM to model the tree height-diameter relationships. There are basically two ways to capture the spatial effects: blocking for spatial heterogeneity and spatial covariance for spatial dependence. For convenience, we call the former LMM-block and the latter LMM-covariance. Further, a LMM model can incorporate both blocking and covariance, namely LMM-block-covariance. The between-group heterogeneity of variance can be specified in either LMM-covariance or LMM-block-covariance in which blocks are present. Consequently, we specified five LMM models with the ordinary least squares (OLS) model as a benchmark. These 5 models were compared in terms of model fitting, residuals, and prediction. Our main objectives were: (1) to investigate which model structure or specification was more appropriate for describing the spatial effects on the tree height-diameter relationships, and (2) to choose a best model specification for the difference between blocks, if spatial heterogeneity was proved to be the main spatial effect.

Materials and methods

Theory and methods

LMM-block

Linear mixed models (LMM) extend the generalized linear model by allowing a more flexible specification of the variance-covariance matrix of model errors. In other words, LMM allows for both spatial autocorrelation and heterogeneous variances although it still assumes normality for the model errors (Littell et al. 2006). LMM can be expressed as:

$$y = X\beta + Zu + \varepsilon \quad (1)$$

where y is a vector of the observed response variable, X is a known matrix of explanatory variables for the fixed effects, Z is a known design matrix of the random effects, β is a vector of unknown fixed-effects parameters, u is a vector of unknown random-effects parameters, and ε is a vector of unobserved random errors (Littell et al. 2006).

A key assumption for LMM is that u and ε are normally distributed such that

$$E \begin{bmatrix} u \\ \varepsilon \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (2)$$

$$\text{VAR} \begin{bmatrix} u \\ \varepsilon \end{bmatrix} = V = \begin{bmatrix} G & 0 \\ 0 & R \end{bmatrix} \quad (3)$$

where G is the variance-covariance matrix of u and R is the variance-covariance matrix of ε . The variance of the response variable 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 . If R is specified as a diagonal matrix, LMM only considers the random effects such as blocks which can be reflected by G . Thus, we call the LMM with the diagonal R and a specific G as the LMM-block model.

For the purpose of model prediction, it is required not only to estimate the parameters of the fixed effects and covariance, but also to estimate the random effects. The solutions for β and u in LMM are called the best linear unbiased predictor (BLUP), and the prediction of y can be obtained by:

$$\hat{y} = X\hat{\beta} + ZGZV^{-1}(y - X\hat{\beta}) \tag{4}$$

LMM-covariance

By integrating spatial variability into the model, the LMM with spatial covariance is derived. The function describing the spatial covariance actually comes from the models of spatial variogram. The R matrix can be derived by:

$$R = I\sigma_1^2 + \sigma^2 F \tag{5}$$

where I is diagonal matrix, σ_1^2 is the nugget of the variogram model, σ^2 is the sill of the variogram model, and F is a matrix whose ij^{th} element is $f(d_{ij})$ as a function of distance (d_{ij}) between observations i and j . The three most commonly used functions are (Littell et al. 2006):

Exponential $f(d_{ij}) = \exp\left(-\frac{d_{ij}}{\rho}\right)$ (6)

Gaussian $f(d_{ij}) = \exp\left(-\frac{d_{ij}^2}{\rho^2}\right)$ (7)

Spherical

$$f(d_{ij}) = \left[1 - 1.5\left(\frac{d_{ij}}{\rho}\right) + 0.5\left(\frac{d_{ij}}{\rho}\right)^3\right] \quad (d_{ij} < \rho) \tag{8}$$

where ρ is the range. The spatial covariance structure is the same across the geographical region. The between-group heterogeneity can also be reflected by a different spatial covariance structure from different groups (i.e., blocks). In other words, each block has its own

parameters, including sill, nugget and range, in its spatial covariance function.

The prediction for LMM-covariance is different from LMM-block, in which Kriging is needed for the prediction based on the estimated spatial covariance parameters. Residual maximum likelihood-empirical best linear unbiased predictor (REML-EBLUP) was proposed and applied for the spatial prediction of soil properties by Minasny & McBratney (2007). This method includes two steps: (1) estimating the parameters of spatial covariance by LMM using REML; and (2) Kriging with external drift (i.e., the exploratory variables) using these parameters for prediction. When we predict y at an unsampled location v_o , the solution is

$$\hat{y}(v_o) = x(v_o)' \hat{\beta} + k' K^{-1} (Y - X\hat{\beta}) \tag{9}$$

$$\hat{\beta} = (X'K^{-1}X)^{-1} X'K^{-1}Y \tag{10}$$

where $Y = [y(v_1), y(v_2), \dots, y(v_n)]$ and $X = [x(v_1), x(v_2), \dots, x(v_n)]$ are the vectors of observations at the sampled locations v_1, v_2, \dots, v_n , v_1, v_2, \dots, v_n , and $K = \text{cov}(Y, Y')$, and $k = \text{cov}(Y, y(v_i)v_o)$ are the covariance functions.

Kriging is designed to predict the response variable at unsampled locations in the plot using the information from the sampled locations. In contrast, the common way for assessing the models is to use the model residuals (i.e., residual = observation - predictions at the same location). Therefore, for every observation we make the estimation of fixed effects using the LMM generated from the whole data, but use the whole data minus this observation to predict the random effects of spatial dependence. To compare LMM-covariance and LMM-block, a cross-validation process is conducted to assess the effects of Kriging.

LMM-block-covariance

It is also possible to build the LMM models with both random effects of block and spatial

covariance simultaneously. We call this model as LMM-block-covariance, which has the potential to improve the model fitting. However, it is also more complex in model specification and structure. Thus, it is necessary to statistically test the significance of improvement in model fitting.

Model assessment

Akaike's information criterion (AIC) and Bayesian information criterion (BIC) are commonly used for model selection and comparison (Hoeting et al. 2006):

$$AIC = -2 \log L + 2(p + k + 1) \quad (11)$$

where L is the likelihood (restricted likelihood in this study), p is the number of fixed effect terms, and k is the number of random effect terms. The candidate model with the lowest AIC is selected as the best model. AIC can also be used for assessing possible variance-covariance structures.

For models selection, it is necessary to test the differences among the candidate models with different model structures and specifications. Commonly, a LMM model with more complex variance-covariance structure fits the data better. For example, the more terms of random effects specified, the more complex the LMM model is. However, the complexity may cause the over-parameterization of the variance-covariance structure, which leads to inefficient estimation and poor assessment of standard errors for estimating the mean response profiles (fixed effects). On the other hand, an overly restricted specification of the variance-covariance structure invalidates the inferences on the mean response profile when the assumed covariance structure does not hold (Altham 1984).

Likelihood ratio test (LRT) has been used extensively as a tool for testing the significance of random effects in LMM. To test the significance of one random effect, it assumes

this effect has zero variance in the null hypothesis. Thus, the test statistic of LRT is defined as (Verbeke & Molenberghs 1997):

$$-2 \ln \lambda_N = -2 \ln \left[\frac{L_{REML}(\hat{\theta}_{REML,0})}{L_{REML}(\hat{\theta}_{REML,1})} \right] \quad (12)$$

where $\hat{\theta}_{REML,0}$ and $\hat{\theta}_{REML,1}$ are the restricted maximum likelihood estimates under the null hypothesis and under the alternative hypothesis, respectively. The statistic, $-2 \ln \lambda_N - 2 \ln \lambda_N$ follows a χ^2 distribution with the degrees of freedom equal to the difference of the number of parameters of random effects in the model under the null and alternative hypotheses.

Note that AIC does not enable the evaluation of model prediction for the response variable. More importantly, the prediction of LMM-covariance involves Kriging. Therefore, the root mean squared deviation (RMSD) and standardized squared deviation can be used for comparing the predictions from different models/methods (Minasny & McBratney 2007). Because RMSD actually provides the same information as the residual sum of squares (RSS) commonly used in nonparametric regression, we adopted RSS as an index for the comparison between prediction methods:

$$RSS = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (13)$$

Residual diagnostics

Moran's I can be used examine the spatial autocorrelation in model residuals from different models (Anselin 1995, Tiefelsdorf 2000, Boots 2002, Zhang & Gove 2005, Zhang et al. 2005) as follows:

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n c_{ij}(\theta) (e_i - \bar{e})(e_j - \bar{e})}{\sum_{i=1}^n \sum_{j=1}^n c_{ij}(\theta) \sum_{i=1}^n (e_i - \bar{e})^2} \quad (14)$$

where e_i and e_j denote the model errors at the

location i and j , respectively, \bar{e} is the mean of e_i over the n locations, and $c_{ij}(\theta)$ is the spatial weight within a given bandwidth θ . If the location j is a neighbor of the subject location i , $c_{ij}(\theta) = 1$. Otherwise, $c_{ij}(\theta) = 0$. The expected mean of the Moran's I is $-1 / (n - 1)$. The Moran's I is positive when the observed values of locations within the bandwidth tend to be similar, negative when they tend to be dissimilar, and approximately 0 when the observed values are arranged randomly and independently over space (Lee and Wong 2001, Zhang and Gove 2005). Further, spatial correlograms with a 5-m lag increment were obtained to investigate the changes of Moran's I in model residuals across different lag distances (Zhang et al. 2008).

Spatial heterogeneity in model residuals can be evaluated by the relative structured variability (RSV) which is defined as (partial sill / sill) and used to represent the proportion of the autocorrelated spatial heterogeneity (Schabenberger & Gotway 2005). Others also used the intra-block spatial variation of model residuals at a range of block sizes (spatial scales) to quantify the local spatial variability (Garrigues et al. 2006, Zhang et al. 2009).

Implementation of models

Because the scatterplot of the tree height (HT) against diameter (DBH) was quadratic in shape, we chose the following model to fit the height-diameter relationship:

$$\ln(HT) = \beta_0 + \beta_1 \cdot \ln(DBH) + \varepsilon \quad (15)$$

where \ln is natural logarithm, β_0 and β_1 are regression coefficients to be estimated, and ε is the model error term. Model residuals were defined as the difference between the observed and predicted $\ln(HT)$.

To analyze the spatial effects on the height-diameter relationship, LMM-block and LMM-covariance were fitted to the example data.

LMM-covariance can directly use the spatial coordinates of observations. For LMM-block, blocking in the region needs to be implemented. However, there were no universal criteria for blocking. Because the example plot size was 100×100 m and the number of trees was 659, we decided to use 20×20 m blocks to compromise the number of blocks and the number of trees within each block. The between-group heterogeneous variances can be added into either LMM-covariance or LMM-block-covariance. Finally, we proposed five LMM models to compare as the combinations of different model structures and specifications as follows: (i) LMM-block, (ii) LMM-covariance, (iii) LMM-block-covariance, (iv) LMM-covariance (between-group heterogeneity), (v) LMM-block-covariance (between-group heterogeneity).

The comparison among the model structures provides us the information on which one describes the spatial effects best, e.g., blocking or spatial covariance or the combination of blocking and spatial covariance. In addition, the use of between-group heterogeneous variances can explore the difference between blocks, because the LMM model with the between-group heterogeneous variances allows each block a unique within-block spatial variogram.

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, and uneven-aged. An example plot of 100×100 m in size with 659 trees was used. The average tree diameter at breast height (DBH) was 17.9 cm (ranging from 10.2 to 74.2 cm), and the average total height (HT) was 13.1 m (ranging from 6.4 to 32.9 m). The position of every tree was recorded in spatial coordinates. The major species in the example plot were balsam fir (*Abies balsamea* (L.) Mill.) (58.1% in number of trees) and black spruce (*Picea mariana* (Mill.) BSP) (36.6%). Other minor species included white birch (*Betula papyrifera* Marsh.) (0.3%), white spruce (*Picea*

glauca (Moench) Voss) (2.6%), and white pine (*Pinus strobus* L.) (2.4%).

Results

Model fitting and estimation of variance

Table 1 indicated that the LMM-block model (AIC = -819.8) had the significant improvement over OLS (AIC = -745.4) in terms of model fitting. LRT showed a strong evidence (p -value < 0.0001) to statistically reject the null hypothesis of no block effects. The estimated variance of random errors, σ_1^2 (0.01536) of LMM-block was smaller than that of OLS

($\sigma_1^2 \sigma_1^2 = 0.01852$). The estimated variance of the block effects, σ_b^2 , was 0.00324. From the perspective of LMM-block, some variability of observations can be attributed to the spatial effects across the example plot.

There are several ways to specify the LMM-covariance model such as exponential, Gaussian, or spherical variogram models, which can also be with or without a nugget effect. The covariance among the observations can be assumed existing either within blocks or throughout the entire study area. First, the exponential model with a nugget effect was tried within blocks and across the example plot, resulting in that AIC was -754.4 for the LMM-covariance within blocks, while AIC was -842.7 for the LMM-covariance across the example plot. Thus, it seemed that it was more reasonable and beneficial to apply the LMM-

covariance across the entire study region rather than within blocks. In addition, the exponential model was compared with or without a nugget effect. The model AICs were -845.1 for the exponential model with a nugget and -820.9 for the exponential model without a nugget. LRT also rejected the null hypothesis of no nugget effect (p -value < 0.0001). The exponential model was compared with other commonly used variogram models. The result indicated that the exponential variogram model (AIC = -845.1) was a better choice than either Gaussian (AIC = -836.7) or spherical (AIC = -825.7) for our data. Finally, the model parameters of the LMM-covariance with exponential variogram and nugget were sill = 0.0106, range = 4.96, and nugget = 0.0083.

The combination of blocking and covariance generated the LMM-block-covariance model, which had AIC equaled to -848.1 for model fitting (Table 1). The LRT tests indicated that the LMM-block-covariance model was significantly improved in model fitting over either LMM-block or LMM-covariance, respectively (p -value < 0.0001 for both LRT tests). The estimated model parameters were $\sigma_b^2 = 0.0019$

$\sigma_{\beta_0}^2 = 0.0019$, sill = 0.0099, range = 2.92, and nugget = 0.0069, which were significantly smaller than the estimate of block variance of LMM-block and the estimates of three variogram parameters of LMM-covariance.

To consider different covariance structures in different blocks, the LMM-covariance with between-group heterogeneity was used to fit the data and the resultant model AIC was -880.7 (Table 1). However, allowing each block

Table 1 Model fitting statistics

Models	RSS	RSS-fixed	AIC
OLS	12.1676	-	-745.4
LMM-block	9.7849	-	-819.8
LMM-covariance	10.5703	12.2252	-845.1
LMM-block-covariance	10.0103	10.3115	-848.1
LMM-covariance (between-group heterogeneity)	9.6079	12.2426	-880.7
LMM-block-covariance (between-group heterogeneity)	8.9662	10.2300	-919.2

has different covariance structures added many parameters into the model because each block had its own variogram parameters. However, LRT still showed the significant improvement of LMM-covariance with between-group heterogeneity over LMM-covariance (p -value < 0.0001). Similarly, the addition of between-group heterogeneity to the LMM-block-covariance model reduced the model AIC from -848.1 to -919.2 (Table 1) and LRT indicated a significant improvement for incorporating the between-group heterogeneity into the LMM-block-covariance model (p -value < 0.0001).

Model coefficients of fixed effects

The estimates of coefficients of fixed effects and the standard error (S.E.) of estimates of all the models were presented in Table 2. LMM-block and LMM-covariance had similar coefficients of fixed effects, β_0 and β_1 , to those of the OLS model, but their β_0 was lower than that of OLS and β_1 higher than that of OLS. In contrast, the two coefficients of LMM-block-covariance were just between LMM-block and LMM-covariance. The incorporation of between-group heterogeneity into LMM-covariance made the model coefficients more different from those of OLS, while this addition of the between-group heterogeneity to LMM-block-covariance produced the model coefficients closer to those of OLS (Table 2).

Both LMM-block and LMM-covariance had smaller S.E. of β_0 and β_1 than those of OLS. LMM-block yielded smaller S.E. for β_0 , but larger S.E. for β_1 than those of LMM-covari-

ance. The corporation of between-group heterogeneity produced smaller S.E. than those of both LMM-covariance and LMM-block-covariance (Table 2).

Model prediction

Good model fitting in LMM does not necessarily guarantee good performance of prediction, which is measured by the residual sum of squares (RSS, eq. [24]). However, the model RSS for LMM-covariance is actually from the combination of prediction by fixed effects and Kriging by cross-validation. We separated the RSS of model prediction by the fixed effects alone and denoted it as RSS-fixed. The RSS, RSS-fixed (only for models with covariance structures) and AIC (model fitting) of different models were listed in Table 1.

LMM-block performed much better (RSS = 9.7849) than OLS (RSS = 12.1676). LMM-covariance had smaller RSS (10.5703) than OLS, but larger RSS than LMM-block. An interesting point was if LMM-covariance was used for prediction without using Kriging its prediction was similar to OLS because the RSS-fixed of LMM-covariance was 12.2252 (slightly larger than the RSS of OLS). It indicated that although LMM-covariance fitted the data better (smaller AIC) than OLS, it relied on the Kriging process to achieve better model prediction. Incorporating block or random effects into LMM-covariance not only improved the model fitting, but also helped with model prediction (smaller RSS) (Table 1).

Table 2 The estimates of model coefficients and standard errors (S.E.) of the fixed effects

Models	β_0	S.E.(β_0)	β_1	S.E.(β_1)
OLS	0.972	0.0470	0.559	0.0170
LMM-block	0.920	0.0458	0.576	0.0156
LMM-covariance	0.938	0.0461	0.568	0.0155
LMM-block-covariance	0.936	0.0459	0.569	0.0155
LMM-covariance (between-group heterogeneity)	0.927	0.0377	0.572	0.0132
LMM-block-covariance (between-group heterogeneity)	0.957	0.0374	0.566	0.0125

Residual analysis

The model residuals were analyzed using Moran's *I* (Fig. 1). It is clear that all LMM models produced much smaller Moran's *I* in the model residuals than OLS. In general, Moran's *I* are positive across a range of spatial lags and approach to zero after the spatial lag 15 m for all LMM models. LMM-block has higher Moran's *I* than that of LMM-covariance. LMM-block-covariance also has higher Moran's *I* than that of LMM-covariance models. It seems that the addition of between-group heterogeneity leads to higher Moran's *I* in LMM-covariance.

The intra-block variances (Fig. 2) show that LMM-block has smaller spatial heterogeneity in the model residuals than LMM-covariance. In fact, LMM-covariance is very close to OLS in the intra-block variances across the spatial lags. The addition of between-block heterogeneity does not change the intra-block variances much.

Discussion

It was evident that LMM models with better model fitting (lower AIC) did not necessarily produce better model prediction. For example the AIC of LMM- covariance was smaller than that of LMM-block, but LMM-block produced smaller RSS for model prediction than LMM-covariance. The reason was that LMM-covariance required a Kriging process which was based on the estimated parameters of a variogram model. The errors of prediction by the Kriging process were added to the total errors of prediction by the LMM-covariance model. In addition, if the study area was spatially heterogeneous such as the example plot in this study (a mixed-species stand), the stationarity assumption needed for the Kriging was violated, consequently, affected the prediction by the LMM-covariance model. On the other hand, LMM-covariance was designed to incorporate spatial autocorrelation in the data into the modeling process. It yielded much smaller Moran's *I* in the model residuals than the mod-

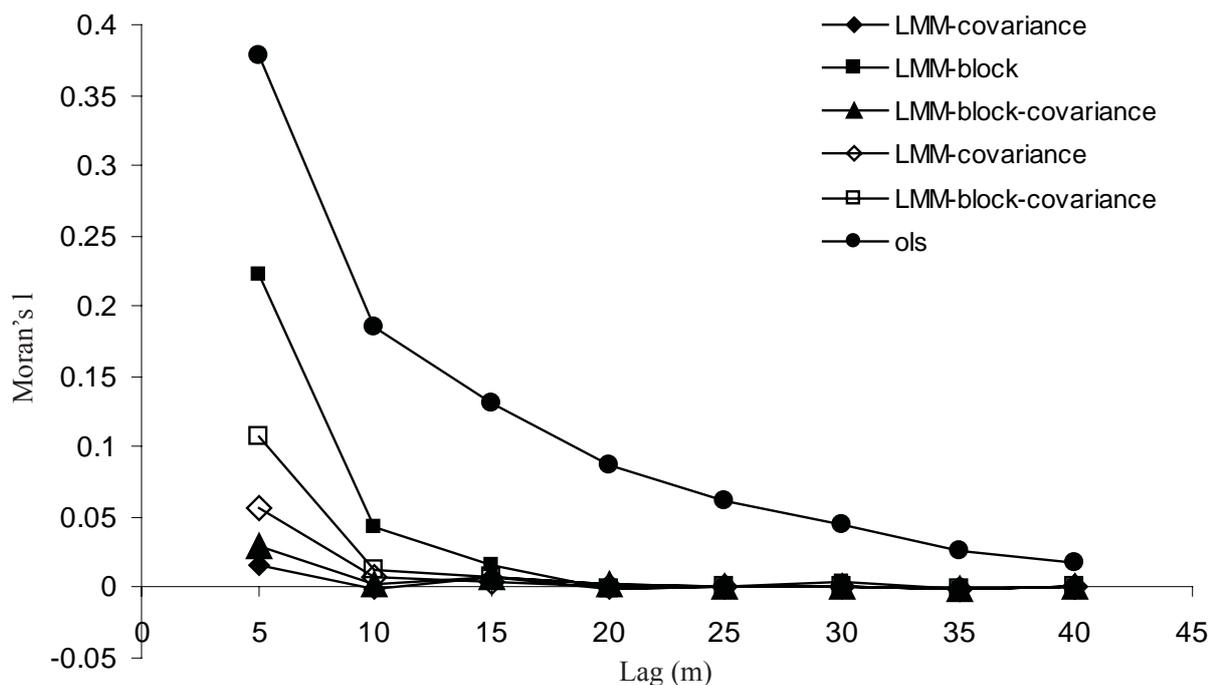


Figure 1 Moran's *I* of model residuals across a range of spatial lags

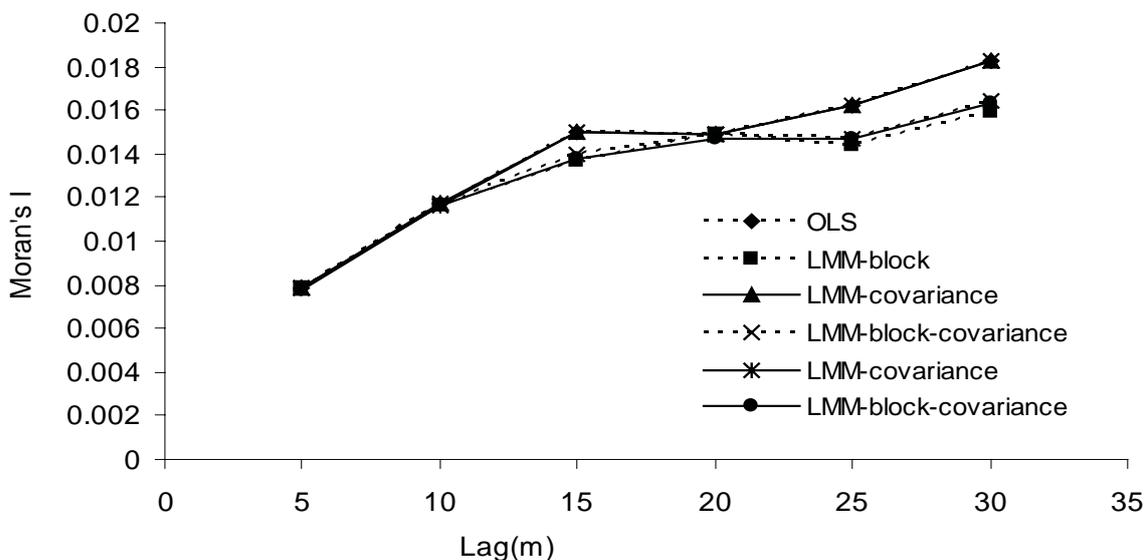


Figure 2 Intra-block variances of model residuals across a range of spatial lags

els without using spatial covariance.

In contrast, LMM-block was designed to deal with the spatial heterogeneity in the data. Thus, it produced more reliable estimates of fixed effects and, better prediction, and smaller intra-block variances of residuals (indicating the reduction of spatial heterogeneity in the model residuals). However, the selection of block size was ad hoc and relatively arbitrary for LMM-block, if the segmentation of the study area was used as the blocks. A good balance between number of blocks and number of trees within a block needed to be kept. Otherwise, some blocks may end up without sufficient observations for fitting the model, especially in a spatially clustered stand (e.g. the example plot).

LMM-block-covariance, which considered both blocking and spatial covariance simultaneously, had better AIC than both LMM-block and LMM-covariance, but its RSS was between LMM-block and LMM-covariance. Although LMM-block-covariance took the advantage of incorporating both blocking and spatial covariance into the modeling process, it did not yield a better model prediction as one

would expect. If improving model fitting is intended to obtain accurate estimations on model coefficients, LMM-block-covariance should be considered. If the model prediction is the primary objective of the study, LMM-block is a better choice.

Incorporating between-group heterogeneity into model structure allowed the local variogram to be used in LMM and was proved to enable the Kriging to make better predictions (Walter et al. 2001, Corstanje et al. 2008). It not only improved model fitting but also produced better model prediction. However, its drawback was that it significantly increases the number of model parameters to be estimated, i.e. the complexity of the model structure. Nevertheless, our results provided some insights on using local variogram in spatial analyses and modeling, which may be compared with other local spatial modeling techniques such as geographically weight regression (Fotheringham et al. 2002, Zhang et al. 2004).

Conclusions

Previous studies showed that linear mixed models with spatial covariance can effectively reduce spatial autocorrelation and heterogeneity in model residuals (e.g., Zhang et al. 2005, Zhang et al. 2008). In this study we compared the LMM with different specifications of model structure. The results indicated that, in general, LMM-covariance significantly reduced spatial autocorrelation in model residuals, while LMM-block was effective in dealing with spatial heterogeneity. The tradeoff in model selection is always a challenge, and the selection of an appropriate model structure may depend on the specific situation of data.

LMM-block treated the blocks as random effects and avoided the estimation of parameters of the variogram model. Thus, it produced better model predictions than LMM-covariance. It showed a different way of incorporating spatial heterogeneity into a “global model”, unlike local models such as GWR.

LMM-block-covariance took both block effects and spatial covariance into account, and significantly improve model fitting. However, it did not produce better model predictions due to the increase of model complexity and estimation of the local variogram within each block. Therefore, we would recommend one considering LMM-block for the prediction purpose and LMM-block-covariance for the analytic purpose of spatial effects.

Linear mixed models with different model structures provide prosperous perspectives on spatial analysis and modeling. Future works may focus on developing more effective algorithms for estimating parameters of local variograms, and combining between-group heterogeneity in spatial covariance and within-group heterogeneity in error variance together.

References

- Altham P.M.E., 1984. Improving the precision of estimation by fitting a model. *Journal of the Royal Statistical Society Series B (Methodological)* 46: 118-119.
- Anselin L., 1999. Spatial econometrics. Available online at: http://www.csiss.org/learning_resources/content/papers/baltchap.pdf. [Last accessed Dec. 20, 2009].
- Anselin L., 1995. Local indicator of spatial association - LISA. *Geographical Analysis* 27: 93-115.
- Anselin L., Griffith D.A., 1988. Do spatial effects really matter in regression analysis. *Papers of Regional Science Association* 65: 11-34.
- Bates D.M., 2008. Penalized least squares versus generalized least squares representations of linear mixed models. Available online at cran.r-project.org/web/packages/lme4/vignettes/PLSvGLS.pdf [last accessed Nov. 10 2009].
- Boots B., 2002. Local measures of spatial association. *Ecoscience* 9: 168-176.
- Brunsdon C., Fotheringham A.S., Charlton M., 1998. Spatial nonstationarity and autoregressive models. *Environment and Planning A* 30: 957-973.
- Budhathoki C.B., Lynch T.B., Guldin J.M., 2008. Nonlinear mixed modeling of basal area growth for shortleaf pine. *Forest Ecology and Management* 255: 3440-3446.
- Calama R., Montero G., 2005. Multilevel linear mixed model for tree diameter increment in stone pine (*Pinus pinea*): a calibrating approach. *Silva Fennica* 39: 37-54.
- Calegario N., Daniels, R.F., Maestri R., Neiva R., 2005. Modeling dominant height growth based on nonlinear mixed-effects model: a colonial Eucalyptus plantation case study. *Forest Ecology and Management* 204:11-20.
- Corstanje R., Grunwald S., Lark R.M., 2008. Inferences from fluctuations in the local variogram about the assumption of stationarity in the variance. *Geoderma* 143: 123-132.
- Dutilleul P., Legendre P., 1993. Spatial heterogeneity against heteroscedasticity: An ecological paradigm versus a statistical concept. *Oikos* 66: 152-171.
- 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., Brunsdon C.A., Charlton M.E., 2002. Geographically weighted regression: the analysis of spatially varying relationships. John Wiley & Sons, New York, 269 p.
- Fox J.C., Bi H., Ades P.K., 2007. Spatial dependence and individual-tree growth models: I. Characterizing spatial dependence. *Forest Ecology and Management* 245: 10-19.
- Garrigues S., Allard D., Baret F., Weiss M., 2006. Quantifying spatial heterogeneity at the landscape scale using variogram models. *Remote Sensing of Environment* 103: 81-96.
- Gregoire T.G., Schabenberger O., 1996. Nonlinear mixed-effects modeling of cumulative bole volume with spatially correlated within-tree data. *Journal of Agricultural, Biological, and Environmental Statistics* 1: 107-119.

- Hoeting J.A., Davis R.A., Merton A.A., Thompson S.E., 2006. Model selection for geostatistical models. *Ecological Applications* 16: 87-98.
- Lee J., Wong D.W.S., 2001. *Statistical analysis with ArcView GIS*. John Wiley & Sons, New York, 192 p.
- Littell R.C., Milliken G.A., Stroup W.W., Wolfinger R.D., Schabenberger A., 2006. *SAS for mixed models*. SAS Press, Cary, USA, 441 p.
- Meng Q., Cieszewska C., Madden M., 2008. Large area forest inventory using Landsat ETM+: A geostatistical approach. *Journal of Photogrammetry and Remote Sensing* 64: 27-36.
- Minasny B., McBratney A.B., 2007. Spatial prediction of soil properties using EBLUP with the Matérn covariance function. *Geoderma* 140: 324-336.
- Nothdurft A., Kublin E., Lappi J., 2006. A non-linear hierarchical mixed model to describe tree height growth. *European Journal of Forest Research* 125: 281-289.
- Schabenberger O., Gotway C.A., 2005. *Statistical Methods for Spatial Data Analysis*. Chapman & Hall/CRC. Boca Raton, Fla.
- Shi H., Zhang L., Liu L. 2006. A new spatial-attribute weighting function for geographically weighted regression. *Canadian Journal of Forestry Research* 36: 996-1005.
- Tiefelsdorf M., 2000. Modeling spatial processes: the identification and analysis of spatial relationships in regression residuals by means of Moran's I. P.167 in *Lecture Notes in Earth Sciences*. Springer, Berlin, Germany.
- Verbeke G., Molenberghs G., 1997. *Linear mixed models in practice*. Springer-Verlag, New York, USA, pp. 63-66.
- Walter C., McBratney A.B., Douaoui A., Minasny B., 2001. Spatial prediction of topsoil salinity in the Chelif Valley, Algeria, using local ordinary kriging with local variograms versus whole-area variogram. *Australian Journal of Soil Research* 39: 259-272.
- Wolfinger R.D., 1996. Heterogeneous variance-covariance structures for repeated measures. *Journal of Agricultural, Biological, and Environmental Statistics* 1:205-230.
- Zhang L., Shi H., 2004. Local modeling of tree growth by geographically weighted regression. *Forest Science* 50: 225-244.
- Zhang L., Gove J.H., 2005. Spatial assessment of model errors from four regression techniques. *Forest Science* 51: 334-346.
- Zhang L., Gove J.H., Heath L., 2005. Spatial residual analysis of six modeling techniques. *Ecological Modeling* 186: 154-177.
- Zhang L., Ma Z., Guo L., 2008. Spatially assessing model errors of four regression techniques for three types of forest stands. *Forestry* 81: 209-225.