APPLICATION OF SPATIAL REGRESSION MODELS FOR FOREST BIOMASS ESTIMATION IN GUIZHOU PROVINCE, SOUTHWEST CHINA

QI, Y. J.[1] – ZHANG, Y. C.[1] – WANG, K.[1] – HE, S. Q.[2] – TAN, W.[1*]

[1]*College of Forestry, Guizhou University, 550025 Guiyang, PR China*

[2]*Longli Forest Farm, 551200 Longli, PR China*

**Corresponding author*
e-mail: xiaoweitan@163.com; phone: +86-135-1198-5162

(Received 25th Apr 2020; accepted 11th Aug 2020)

Abstract. At the regional scale, many studies have been devoted to the construction of biomass models to improve the accuracy of regional biomass estimation, while only a few studies were carried out to investigate spatial influence. Therefore, the current research examined the spatial autocorrelations as well as variations between forest biomass and forest variables using 419 forest biomass plots sampled in the Guizhou Province in the year 2010. Besides, 4 global models, including the ordinary least squares model (OLS), linear mixed model (LMM), spatial lag model (SLM), spatial error models (SEM), together with geographically weighted regression model (GWR, the local model), were fitted to the associations of forest biomass with basal area, height and age of the stand. As suggested by our findings, distinct spatial autocorrelations as well as variations exist between forest biomass and these variables. OLS is not appropriate for modeling. SLM and SEM efficiently accounted for the spatial autocorrelations within model residual; however, they were unable to manage spatial heterogeneities. However, LMM and GWR, which had combined spatial variations as well as dependence during the modeling process, performed well in data fitting and response variable predicting. Of them, GWR reduced spatial heterogeneity to a greater extent than LMM.
Keywords: *linear mixed model, spatial lag model, spatial error model, geographically weighted regression*

Introduction

It is of crucial importance to estimate the forest biomass to understand source dynamics and atmospheric carbon sinks (Brown and Lugo, 1984). To be specific, forest biomass, particularly that is located in the subtropics, accounts for a leading global carbon emission source. In order to evaluate the ecological quality and the effectiveness of forestry construction and to provide decision-making for promoting rational use of energy and forest management, the large-scale study of forest biomass and its spatial distribution is necessary (Rodríguez-Veiga et al., 2016).

Forest biomass has been frequently predicted based on the timber volume data extracted via the forest inventories, for which, field plots were used for statistical sample collection to directly measure the forest parameters (such as breast height diameter, height, as well as tree species). Forest biomass estimation methods are continuously improving with the advancement of research. Thereafter, timber volume is converted into the above-ground biomass through the application of the biomass expansion factor (BEF) where timber volume information is obtained (Brown and Lugo, 1984). Fang et al. (2001) improved the BEF method for forest biomass estimation overcoming the shortcoming of static conversion parameters of the BEF method improving the biomass-volume model and providing a more accurate calculation of the parameters for various dominant tree species. Remote sensing methods, often used to

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1805_72157232

estimate large-scale forest parameters (Patenaude et al., 2005; Ploton et al., 2017), can hardly accurately estimate spatial distributions due to the limited sensitivity of the selected satellite sensors to forest parameters (Benson et al., 2016; Ghosh et al., 2018). Different technical means and auxiliary data were employed to improve the estimation precision. Remote sensing methods also use process models, which have a perfect theoretical foundation and clear physical meaning; however, they require more complex data sources as input parameters (Chiesi et al., 2011).

As we all know, forest biomass information is usually extracted across the huge geographic areas, which leads to distinct differences of topographic characteristic, species abundance and compositions, as well as vegetation coverage; as a result, the carbon storage and forest biomass among the different locations are also different (Liu et al., 2014). The neighbor locations are similar to each other, but the remote locations are not similar. This can be explained by two aspects of spatial effects, spatial heterogeneity (non-stationarity of space), together with spatial autocorrelation (namely, spatial dependence) (Anselin and Griffith, 1988). Of them, spatial heterogeneity represents the instability of structure that manifests as the systemically altering model variables or the diverse response functions. Meanwhile, spatial autocorrelations are those correlations of the random variable value in a region with identical variable values in adjacent regions. Ignoring spatial heterogeneity and spatial autocorrelation lead to false significance test results, inferior predictions (Anselin and Griffith, 1988), as well as damaging influence on modeling and data analytic results (Páez and Scott, 2005). In recent years, the indexes of global and local spatial autocorrelations, such as Geary's C, Moran's I, Getis'G*, and Getis'G, are extensively utilized for measuring spatial autocorrelation degrees across different regions (Boots, 2002; Páez and Scott, 2005; Fu et al., 2014). Several statistical regression approaches have been adopted in incorporating the spatial autocorrelations to model the associations across variables, including the spatial error model (Lichstein et al., 2002), spatial lag model (Anselin, 1993), spatial Durbin model (Overmars et al., 2003), spatial filter model (Borcard and Legendre, 2002), and linear mixed model (Zhang et al., 2009), and others (Dormann et al., 2010). Spatial heterogeneity represents the function of spatial scales that are referred to as the measuring units. A variety of regression approaches are developed for modeling the local alterations in the complicated associations of spatial random variables, including random coefficient model (Fotheringham and Brunsdon, 1999), spatial expansion approach (Anselin, 1992), multilevel modeling approach (Duncan, 1997; Jones1997), and spatial adaptive filtering approach (Gorr and Olligschlaeger, 1994). Recently, the geographically weighted regression (GWR) is increasingly used for exploring spatial heterogeneities (Fotheringham et al., 2002; Zhang et al., 2004; Liu et al., 2014).

The GWR is the local regression model that can effectively solve spatial nonstationarity. The parameters of each sample are estimated by using the locally weighted least square method and the spatial location is taken into account in the estimation process. In other words, GWR offers diverse regression modeling parameters of a localized model in every region to incorporate the locoregional spatial variation into the modeling processes. Therefore, it is becoming an important method to detect the non-stationarity of space and has been showing better performance than ordinary models when applied to many fields such as social economics (Öcal and Yildirim, 2010; Wang et al., 2019), geography (Erdoğan, 2010), soil (Song et al., 2019; Li et al., 2020), environment (Qin et al., 2019) and forestry (Propastin, 2012; Lin et al., 2018; Monjarás-Vega et al., 2020). Consequently, in the forestry fields, the influences of vegetation

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1805_72157232

competition, micro-environment, and growth potential; besides, effects of managing activities on carbon storage and forest biomass are assessed, analyzed, simulated and visualized (Liu et al., 2014; Zhen et al., 2013), while the study about spatial influence in modeling forest parameters is still rare, especially in regional scale (Liu et al., 2014).

The current research analyzed forest biomass for its spatial distribution in the Guizhou Province of southwest China. Typically, this study aimed to achieve the following objectives: (1) to use five models, including the ordinary least squares (OLS) model, the spatial lag model (SLM), the spatial error model (SEM), the linear mixed model (LMM), together with the geographically weighted regression model (GWR) fitting forest biomass information; (2) to analyze and compare these five models for their performances; and (3) to evaluate model residues for spatial autocorrelation as well as spatial heterogeneity.

Materials and methods

Study area

Guizhou Province (103°36′E to 109°35′E and 24°37′N to 29°13′N) is located in southwest China (*Fig. 1*) with a total land area of 176128 km^2 and 57% forest coverage rate. 92.5% of the whole province is mountainous and hilly, with an average altitude of 1100 m. This area is of subtropical monsoon climate. Affected by mountainous terrain, the weather conditions are variable. The average annual temperature is 10~18 °C, the annual precipitation is 1,000~1,500 mm, the relative humidity is above 70%, the annual sunshine hours are 1300 h, and the frost-free period is about 270 d (Editorial Committee of Guizhou Forest, 1991). Affected by climate, soil and mountainous terrain, the vegetation types in the province are diverse. The central and northern parts of the country are dominated by mid-subtropical evergreen broad-leaved forests, while the southern part is a south-subtropical evergreen broad-leaved forest. The middle-eastern part is humid forest and the western part is semi-humid forest. Cold and warm sub alpine coniferous forests are distributed in high altitude areas, while intrazonal karst evergreen deciduous broad-leaved mixed forest and secondary deciduous broad-leaved forest are distributed in limestone and dolomite mountains (Editorial Committee of Guizhou Forest, 1991).

According to the forest management plan of Guizhou province (2016-2050) and our study aims, the study area was divided into six ecological geographical sub-regions. The details are presented in *Table 1* and *Figure 1*.

Table 1. *The six ecological geographical sub-regions of Guizhou province*

Code of sub-regions	Name of sub-regions
A	The low-middle mountain sub-region of *Pinus massoniana*, bamboo and broad-leaved mixed forest for soil and water conservation
B	The low-middle mountain sub-region of *Pinus massoniana* and broad-leaved forest for dual-purpose of precious large-diameter timber
C	The high-middle mountain sub-region of *Pinus yunnanensis, Pinus armandii, Cryptomeria fortune* and broad-leaved mixed forest for soil and water conservation
D	The middle mountain sub-region of *Pinus massoniana*, broad-leaved mixed forest for special purpose
E	The low mountain sub-region of *Cryptomeria fortune* forest for fast-growing and high-yielding timber
F	The middle mountain sub-region of *Cryptomeria fortune, Pinus yunnanensis* and broad-leaved forest for dual-purpose of precious large-diameter timber

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1805_72157232

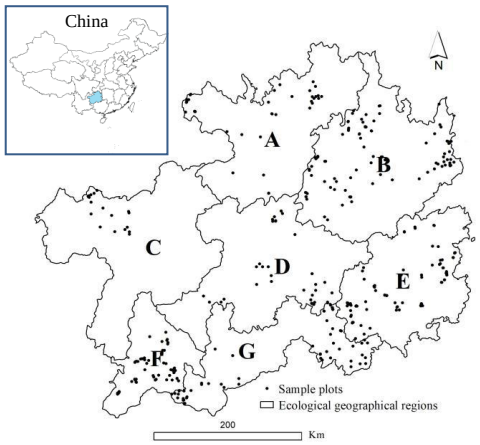

Figure 1. *Geographical location of the six regions with the number of sample plots of the study area in Guizhou, China. A, B, C, D, E and F are the code of sub-regions corresponding to name of sub-regions respectively described in Table 1*

Data

For the current research, the utilized stand as well as tree data were extracted based on 6 locations in the year 2010 from 419 permanent sample plots, including 283 Arbor plots (666.67 m^2 per plot), 51 bamboo plots (100 m^2 per plot), and 85 shrub plots (16 m^2 per plot), observed via the Chinese National Forest Inventory (CNFI), together with the Guizhou Provincial Forestry Department. The topographical descriptors and geographical locations were also collected for every plot. In addition, the stand as well as tree parameters were determined and analyzed for describing the slope (°), elevation (m), numbers of tree species and trees of every species, as well as diameter at breast height in cm (DBH, > 2 cm for bamboo and shrub plots, > 5 cm for Arbor plots), and living tree height in m (HT). In addition, additional stands as well as tree variables were subsequently calculated based on unit hectare, which included the tree number per hectare (TPH) (trees/hm^2), the living tree volume per hectare (m^3/hm^2), the mean living tree age (year), the basal area of trees per hectare (BA, the accumulated area in cross-section determined based on the breast height (1.3 m) for each tree within one stand) (m^3/hm^2), the average height of living trees (H), and vegetation coverage (%). The 1-year meteorological data were extracted based on 77 weather stations Guizhou Province in the year 2010 and Kriging interpolation was conducted with the data for obtaining the precipitation and temperature information for each sample plot. In each arbor plot, the biomass models based on the Guizhou province universal biomass equation were used (Zeng et al., 2011). The tree species, without a clear corresponding model, were referred to as the approximate dominant tree species (group) parameters. In each Bamboo plot, the biomass models were calculated according to Tian (2011) and that of the shrub plot was calculated according to Liu et al. (2009).

Theoretical background

Ordinary least squares (OLS)

Assume there are diverse *n* observations regarding the response variable y, together with *p* predictor variables x. The association of y with x is regressed based on OLS (*Eq. 1*):

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1805_72157232

$$y = x\beta + \varepsilon \quad \text{(Eq.1)}$$

where, β stands for the unclear fixed-effects parameter vector, whereas ε indicates the model error term following a normal distribution N (0, σ^2). Thus, the OLS estimator is acquired through (*Eq. 2;* Littell et al., 2006)

$$\hat{\beta} = (X'X)^{-1}X'Y \quad \text{(Eq.2)}$$

It is supposed that, the association given based on *Equation 1* is constant or universal within the geographical area.

Linear mixed model (LMM)

The LMM represents a special case of generalized linear models, which is presented in the form of *Equation 3:*

$$y = X\beta + Z\gamma + \varepsilon \quad \text{(Eq.3)}$$

in the formula, y stands for the response variable vector, x represents the fixed-effects predictor matrix, β indicates the unclear fixed-effects model coefficient vector, whereas Z suggests the given random-effects design matrix, γ represents the unclear random-effects parameter vector, while ε stands for a random error term. The following assumptions are given: (1) E (γ) = 0 and Var (γ) = G represents the random-effects covariance matrix; (2) E (ε) = 0 and Var (ε) = R suggests the model residual covariance matrix; (3) Cov (γ, ε) = 0; meanwhile, (4) γ and ε show normal distribution. Besides, variance y is calculated based on V = ZGZ′ + R, which is predicted through establishing the design matrix Z of random-effects and indicating the G and R covariance structures. Generally speaking, OLS has not been deemed as the optimal parameter prediction approach, whereas the maximum likelihood approaches are frequently adopted for obtaining γ and β (Littell et al., 2006).

Spatial lag model (SLM)

The SLM represents the formal spatial diffusion procedure, which obtains great spatial data dependences (Anselin, 1993, 2001). Generally, SLM is appropriate for assessing the spatial dependence presence and strength. To be specific, SLM is completed through incorporating one spatial lag term for dependent variable y to that OLS model (*Eq. 1*) mentioned above, as shown below (*Eq. 4*):

$$y = X\beta + \rho Wy + \varepsilon = (1 - \rho W)^{-1}X\beta + (1 - \rho W)^{-1}\varepsilon \quad \text{(Eq.4)}$$

where *W* represents the row-sum weight matrix after standardization processing, *W*y stands for the response variable that is lagged spatially, ρ indicates the spatial autocorrelation variable with normal distribution N (0, $\sigma^2 I$), and *I* denotes the identity matrix. For response y, its value in every region depends on the x in a specific region and in adjacent regions based on spatial multiplier $(1-\rho W)^{-1}$. It should be noted that, according to *Equation 1*, ε becomes correlated with predictor variable (namely, the spatially lagged *W*y). As OLS is not an appropriate parameter prediction method any

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1805_72157232

more, since it will produce ineffective and biased predictions, while the maximum likelihood approach is adopted.

Spatial error model (SEM)

It is assumed in SEM, that spatial autoregressive procedure takes place within error term only, rather than within response or predictor variables (Anselin, 1993, 2001). This is because that, the variables with spatial correlation or spatial region boundaries that do not coincide with the practical behavior units are eliminated (Graaff et al., 2001). SEM is recognized to be the special regression using one non-spherical error term, and this approach is suitable to examine the possible spatial autocorrelation effect induced by using the spatial data, regardless of the model spatiality. SEM combines OLS regression model with the spatial autoregression model within an error term ε, as shown below (*Eq. 5*):

$$y = X\beta + \varepsilon = X\beta + \lambda W\varepsilon + \xi = X\beta + (1 - \lambda W)^{-1}\xi \quad \text{(Eq.5)}$$

In the formula, W stands for the row-sum weight matrix after standardization, $W\varepsilon$ represents the error term with spatial lag, λ indicates the spatial autocorrelation variable, whereas ξ suggests the ordered error term with normal distribution N (0, $\sigma^2 I$)**,** and I denotes the identity matrix. For y, its value in every region is subjected to influences of error in every region based on spatial multiplier $(1-\lambda W)^{-1}$. Nonetheless, different from β, λ has been identified to be the nuisance variable, and this factor itself is not the focus, yet it is of necessity for correcting spatial dependencies. For y, its average value is independent from error spatial dependence, and that maximum likelihood approach has been used to estimate the parameters.

Geographically weighted regression (GWR)

The GWR is extended based on the conventional regression that allows variations, in this way, the regression coefficient is location-specific, and does not represent a universal estimate (Fotheringham and Brunsdon, 1999). The possible GWR model is shown below (*Eq. 6*):

$$y = \beta_0(u_i, v_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i) X_k + \varepsilon \quad \text{(Eq.6)}$$

In the formula, y stands for a response variable, whereas X_k represents diverse p predictor variables (k = 1, 2, …, p), and β_0 (ui, vi), β_1 (ui, vi), …, β_p (ui, vi) suggests those regression coefficients for the kth predictor variable and ith location. εi stands for a random error term with normal distribution N (0,$\sigma^2 I$), and I denotes the identity matrix. Generally, GWR aims to estimate the above-mentioned coefficients for every independent variable x in every geographical region i using neighbors within a given bandwidth and weighted least-squares regression. Those GWR model parameters in every region i in the matrix are predicted according to the following formula (*Eq. 7*):

$$\hat{\beta} = (X'W_iX)^{-1}X'W_iY \quad \text{(Eq.7)}$$

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1805_72157232

where W_i stands for the (m × m) diagonal matrix of spatial weights; X indicates the [m × (n + 1)] matrix of independent parameters, in which n is the explanatory variable number; whereas Y represents the (m × 1) matrix of dependent variables.

The weighted function is used to determine the weighted approach, and it is adaptive or fixed (Mcmillen, 2004). With regard to the size of adaptive kernel, every point weight is determined based on the Gaussian function (*Eq. 8*):

$$W_{ij} = \exp(-\frac{1}{2}(\frac{d_{ij}}{r})^2) \quad \text{(Eq.8)}$$

In the formula, d_{ij} stands for the Euclidean distance of an estimated site i compared with a sampling site j, r stands for the parameter of bandwidth. Notably, the selection of bandwidth within the as-mentioned GWR model represents a critical factor that affects the results of regression analysis. At present, both the corrected Akaike information criterion (AICc) and cross-validation approaches have been extensively utilized for determining bandwidth.

Model specification

The current research aimed was to examine forest biomass for its spatial distributions as well as patterns using the regression models. In every plot, forest biomass was used as the dependent response parameter (ton/hm^2). The stand as well as tree parameter number was analyzed and screened through gradual regression to be model predictors, including the tree basal area per hectare (BA), stand age (Year) and stand height (m). *Table 2* shows the baseline variable data.

Table 2. *Baseline variable data adopted in the current research*

Variable	Plot number	Average	Maximum	Minimum	Std
Biomass (t/hm^2)	419	51.25	261.57	0.01	47.25
Basal area (BA, m^2/hm^2)	419	161962.70	772947.11	477.59	126107.37
Age (Year)	419	20.83	98.00	0.00	15.06
Height (m)	419	8.79	25.00	0.24	5.44

The multiple linear model shown below was adopted for regressing the forest biomass relative to 3 estimators (Basal area (BA)), stand age and stand height) using OLS, SLM, SEM, LMM as well as GWR models:

$$\text{Biomass} = \beta_0 + \beta_1 \cdot \text{BA} + \beta_2 \cdot \text{Age} + \beta_3 \cdot \text{Height} + \varepsilon \quad \text{(Eq.9)}$$

where β_0, β_1, β_2, and β_3 stand for the regression coefficients predicted based on statistics, whereas ε is model residual that represents the heterogeneity of observed forest biomass compared with the predicted one. OLS model was adopted to be the model comparison benchmark in the current research. In LMM model, seven regions were used to be fixed effects, and covariance matrix R was used to model those spatial autocorrelations across various sample plots in those seven regions. For both SLM and SEM models, distance bandwidths performed better than k-nearest neighbors as spatial weights. For the GWR model, the best bandwidth size was 30 km determined according to the golden section

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1805_72157232

search (automated) and adaptive bisque kernel function, because it has a higher coefficient of determination (R^2) and lower residuals than other methods. Meanwhile, for those five models, model residuals were determined based on global Moran's Z- and I-values, with the bandwidths range of 5-45 km at an interval of 5 km, which described the pooled spatial autocorrelations within those model residuals at a variety of spatial scales. Specifically, local Moran's I-value, an index for local spatial autocorrelation (LISA), was calculated according to the best bandwidth (30 km) in every model residual for OLS, LMM, as well as GWR (Boots, 2002). Meanwhile, Z-value is used as the standard deviation (SD). At a significant level of $\alpha = 0.05$, $Z > 1.96$ and $Z < -1.96$ indicate that the model residual has significant spatial autocorrelation, and $-1.96 < Z < 1.96$ indicates otherwise. Intra-block spatial variances, illustrating the local spatial variability, were computed for the above-mentioned five model residuals, and the block size was 5-30 km at an interval of 5 km (Zhang et al., 2009). It is assessed based on the fact that 1/2 GWR coefficients should fall in the range of Q1 (25% quartile) to Q3 (75% quartile), whereas approximately 68% normally distributed LMM or OLS coefficients must fall in the range of 1 SD. The studied association might be not fixed spatially when inter-quartile range was > 1 SD of an equal global variable for the GWR local coefficients (Fotheringham et al., 2002; Zhang et al., 2004).

In this research, we aimed to examine model errors for their spatial heterogeneity and autocorrelation using the above 5 regression methods to fit the associations of biomass with relevant variables, but not to develop a predictive regional forest biomass model. Then, Akaike's Information Criterion (AIC), R^2, and mean squared error (MSE) summation were adopted to evaluate the pooled model fitting effect. Moreover, the LMM and OLS models were fitted using SAS 9.3 (SAS Institute Inc., 2011). In addition, SLM and SEM were fitted using GeoDa software (Anselin et al., 2006; GeoDa Center for Geospatial Analysis and Computing, 2009), while GWR 4.0 software (Nakaya et al., 2009; Department of Geography, Ritsumeikan University, Kyoto, Japan) was used to fit GWR model.

Results

Model fitting

The model fitting data of those five regression models adopted in this study are displayed in *Table 3*. As observed, OLS model had good data fitting effect (R^2 was 0.91) when not considering the independence assumption violation. With regard to SLM model, it adopted those spatially lagged dependent variables to be the predicting variables, so as to calculate spatial autocorrelation; meanwhile, SEM model managed spatial autocorrelation within a model error term (residuals). Besides, SLM and SEM performed better in data fitting (larger R^2, smaller MSE, smaller AIC, and lower global Moran's I (Z)) than the OLS models. For SME model, those tested values [Lagrange multiplier (LM) (error), as well as robust LM (error)] were relatively high relative to those for SLM model, indicating that the SEM model was better (*Table 4*).

According to *Table 3*, LMM performed well in data fitting (higher R^2, lower MSE, lower AIC, and lower global Moran's I (Z)) than OLS, SLM, and SEM. However, the best fit was observed for the GWR model that had greater ^{R2}, smaller MSE, lower AIC, and lower global Moran's I (Z) and SH%. As suggested by Z-values, for SLM and OLS models, their model residuals displayed distinct spatial autocorrelations ($Z > 1.96$, $\alpha = 0.05$). Nonetheless, for SEM, LMM and GWR models, their model residuals

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1805_72157232

displayed no distinct autocorrelations ($-1.96 < Z < 1.96$, at $\alpha = 0.05$) indicating a reduction in the spatial autocorrelation.

Table 5 lists the model coefficient estimates, standard errors (SE), as well as p-values for three predicting factors. Each model coefficient was deemed to be of statistical significance upon $\alpha = 0.05$ in all the five models. For local models, their means for 2 regression coefficients were similar to those for global counterparts (*Table 5*). In addition, the inter-quartile range for GWR model, local intercepts, BA coefficients, age coefficients and height coefficients did not fall within the scope of ± 1 SD of those other four models, indicating the non-stationary and/or different association of forest biomass with those 5 predicting factors. However, the GWR model generated the geographically different model coefficients (a model coefficient set was generated in every region). *Figure 2* illustrates the spatial variations for GWR model coefficients, which shows that the association of forest biomass with those three predicting factors are different at different study locations. Generally speaking, local intercepts were negative over the study area, Age coefficients were positive across our study locations, whereas age and height coefficients were either positive or negative based on different regions. Local intercepts were smaller within the eastern study locations while larger across the northern and southern study locations. BA coefficient was opposite to that of local intercepts, larger in the east, but smaller in north and west. The values of age coefficients were smaller in west and south, and larger in north and east, while age coefficients were smaller in north and south, and larger in the central region. As demonstrated from *Figure 2*, those magnitudes and signs for such regression coefficients depended on the geographical location, suggesting the region-specific influences of those three tree predictors on the forest biomass. The coefficient shows a positive value indicating a positive impact on forest aboveground biomass, otherwise it is a negative contribution. The trend between coefficients distribution was similar indicating their impacts on estimating forest aboveground biomass have regional similarity, and vice versa.

Spatial autocorrelation and heterogeneity in model residuals

Compared with OLS model, the rest four models generated markedly lower global Moran's *I*- values (*Table 5*). Generally, GWR model generated negative values, while SLM, LMM and SEM generated positive values. *Figure 3* shows spatial correlation diagram for model residuals of those five models, with the lag distance ranging from 5 to 45 km. For OLS model, its residuals displayed great spatial autocorrelation compared with those for other models at every tested lag distance. For SLM model, its residuals behaved similar to that of the OLS model but had smaller values across the lag distances (*Fig. 3*). The LMM and SEM models generated similar Moran's *I* values (<0.05) among various lag distances. For GWR model, its model residual showed no significant spatial autocorrelation (Z-value < 1.96) across the range of lag distances and approached zero especially at larger spatial scales (15–45 km).

Figure 3 also shows the intra-block within model residuals representing mean local spatial variation of a known block size. As observed, OLS, SLM, as well as SEM models had remarkably greater and relatively the same intra-block sizes. The GWR had the smallest intra-block local spatial variance among the various block sizes, and then LMM model ranked the second place. Nonetheless, the different intra-block variability across those five models were low at the block size interval of 5 km. It became greater and approached one stable value with the increase in block size.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1805_72157232

Table 3. *Model fitting statistics for the five regression models and measures of spatial autocorrelation and heterogeneity for the model errors*

Models	AIC	MSE	R^2	Moran's I (Z)
OLS	3244.32	55486.60	0.91	0.21(5.33278)
SLM	3235.55	54063.50	0.94	0.15(3.768887)
SEM	3215.54	50833.67	0.95	0.037(1.018091)
LMM	3169.11	43422.45	0.95	0.064(1.720469)
GWR	3103.43	28780.41	0.96	-0.076(-1.908456)

Table 4. *Spatial dependence diagnostics for the OLS model residuals*

Test statistics	Value	P
Moran's I	0.106588	0
Lagrange Multiplier (lag)	1	0.0005686
Robust LM (lag)	1	0.0215412
Lagrange Multiplier (error)	1	0
Robust LM (error)	1	0

Table 5. *Model coefficient estimates, standard error (SE), and p-values of the five models*

Model	Statistics	$\widehat{\beta}_0$	$\widehat{\beta}_1$	$\widehat{\beta}_2$	$\widehat{\beta}_3$	Spatial parameter
OLS	Estimate (SE)	-15.8505(1.1568)	$3.24\times10^{-4}(5.7\times10^{-6})$	0.3102(0.0439)	0.9348(0.1486)	
	β−1·S.D~β + 1·S.D	-16.6371~-15.0639	3.20×10^{-4}~3.28×10^{-4}	0.2803~0.3400	0.8338~1.0358	
SLM	Estimate (SE)	-17.9163(1.3020)	$3.15\times10^{-4}(6.2\times10^{-6})$	0.3320(0.0440)	0.8063(0.1529)	$\hat{\rho}$ = 0.0813
	β−1·S.D~β + 1·S.D	-18.8017~-17.0309	3.11×10^{-4}~3.19×10^{-4}	0.3020~0.3619	0.7023~0.9103	
SEM	Estimate (SE)	-14.7527(1.6506)	$3.17\times10^{-4}(5.9\times10^{-6})$	0.2290(0.0480)	1.1458(0.1593)	$\hat{\lambda}$ = 0.5569
	β−1·S.D~β + 1·S.D	-15.8751~-13.6303	3.13×10^{-4}~3.21×10^{-4}	0.1964~0.2616	1.0375~1.2541	
LMM	Estimate (SE)	-14.0511(2.0003)	$3.10\times10^{-4}(1.27\times10^{-5})$	0.2865(0.0656)	0.9334(0.1769)	
	β−1·S.D~β + 1·S.D	-15.4113~-12.6909	3.01×10^{-4}~3.19×10^{-4}	0.2419~0.3311	0.8131~1.0537	
GWR	Estimate (SE)	-14.3394(-29.6306~-5.1777)	$3.01\times10^{-4}(1.91\times10^{-4}$~$4.23\times10^{-4})$	0.2928(0.0054~0.8564)	1.003(-0.4457~2.6462)	
	Q1~Q3	-19.6553~-8.5991	2.66×10^{-4}~3.33×10^{-4}	0.1541~0.3795	0.4845~1.344	

Numbers in parentheses are the standard errors of the regression coefficients for OLS, SLM, SEM, and LMM and the ranges of the coefficients for GWR models. $\widehat{\beta}_0$ for Intercept, $\widehat{\beta}_1$ for Basel area, $\widehat{\beta}_2$ for Average age, $\widehat{\beta}_3$ for Average height, the same as follow. Q1 and Q3 represents 25% quartile and 75% quartile of GWR coefficients, respectively

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1805_72157232

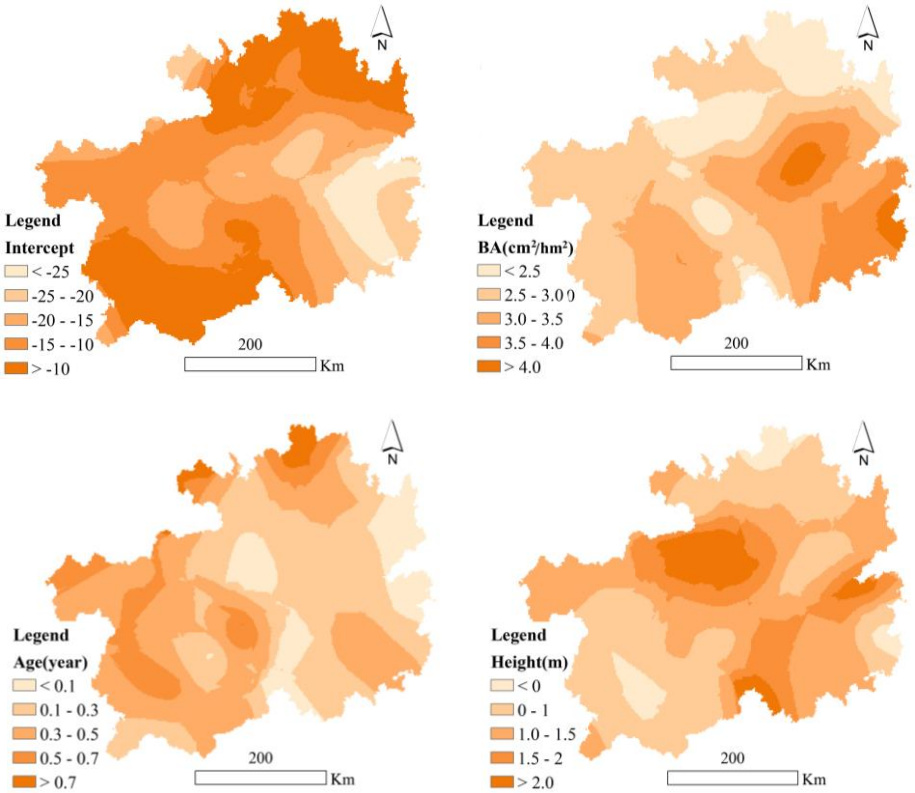

Figure 2. *Contour maps of the model coefficient estimates from the GWR model at the bandwidth 30 km.* $\hat{\beta}_0$ *for Intercept,* $\hat{\beta}_1$ *for BA,* $\hat{\beta}_2$ *for Age,* $\hat{\beta}_3$ *for height*

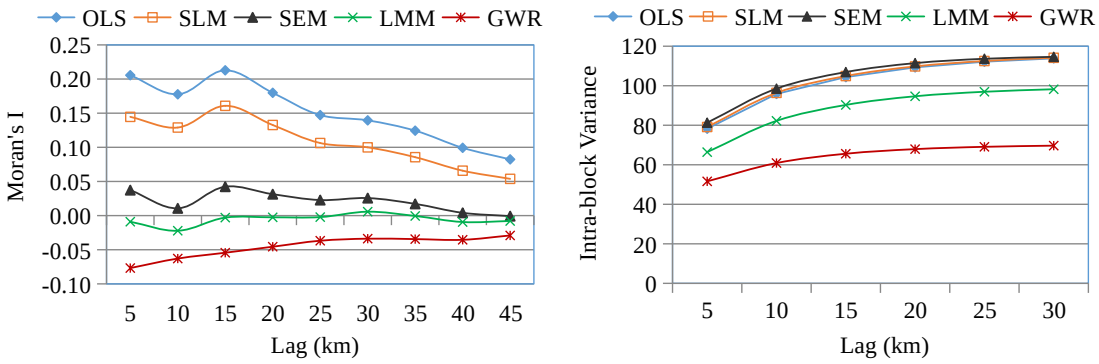

Figure 3. *Correlogram and intrablock variance of the five model residuals*

To illustrate the spatial details of residual autocorrelations, distributions of the Local Moran's *I* at 30 km spatial scales were computed for residuals from each of the five models (*Fig. 4*). It was observed that the local Moran's *I* values of OLS model residual were greater, which were positive (black circles) among the study region, which indicated that OLS model under-predicted or over-predicted the clusters (hotspots). SLM had similar spatial patterns for those 3 plots, and the local Moran's *I* values were greater and more positive (black dots), which indicated that the negative or positive model residuals were clustered. SEM produced few hot spots compared with those of SLM and OLS models. In contrast, the LMM and GWR models produced lower values,

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1805_72157232

which were more negative (white circles), which indicated the opposite signs of model residuals were clustered (cold spots).

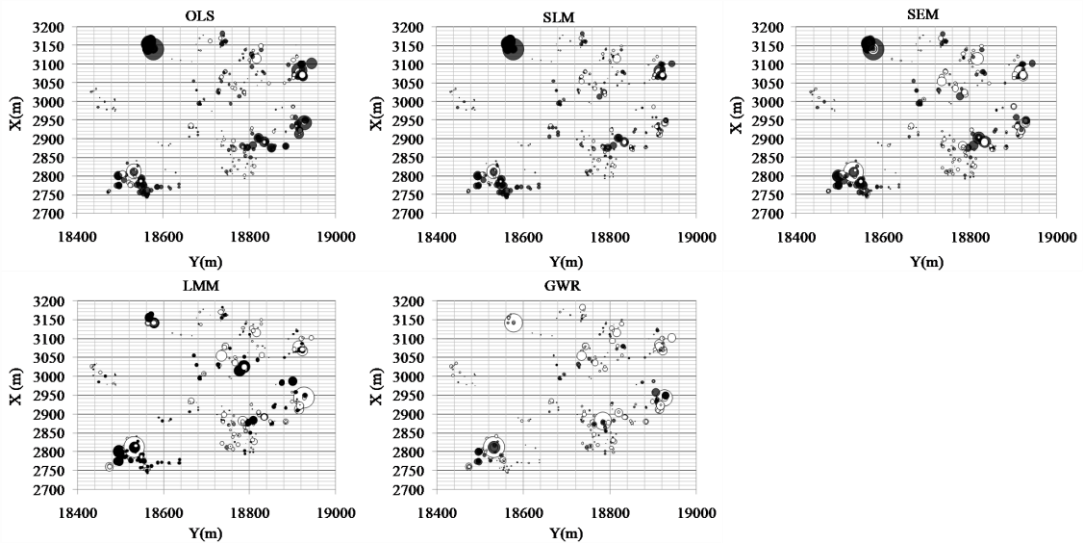

Figure 4. *Local Moran's I value distributions at 30 km spatial scales of the five model residuals. Black circles indicate positive values of local Moran's I with cluster "hot spots", while white circles indicate negative values of the local Moran's I with cluster "cold spots".*

Discussion

Model variables and coefficients

In our study, three variables, basal area, stand age, stand height were selected to estimate forest biomass. They all had statistically significant effects on the amount and distribution of forest biomass. In both global and local models, the three variables were the most important stand variables for our study area. The model coefficients of the three variables were all positive (*Table 4*), indicating that larger basal area, stand age and stand height would sequester more biomass in the forests, which is consistent with most studies (Mani and Parthasarthy, 2007; Cannell, 1984; Fang, 2001).

Other factors, including the temperature and precipitation data, were also considered as variables, but their influence was not significant. Our study also showed that the relationship between basal area and forest biomass was better than the relationship between DBH (diameter at breast height) and forest biomass. That is basal area explained 85% of the model, age explained 3%, height explained 3% (data not shown). Brown (2002) found that DBH alone explained more than 95% of the variation in above ground carbon content in tropic forests. Liu et al. (2014) also selected DBH as the main variable when estimating forest carbon stock of north forests in Heilongjiang province. However, Basal area has been used more frequently as a surrogate for biomass and carbon in tropical forests (Mani and Parthasarthy, 2007; Sagar and Singh, 2006), and often has been found to be the best predictor of biomass in combination with the mean tree height at the stand level (Cannell, 1984; Masera et al., 1997). Fang (2001) improved BEF for forest biomass estimation, which using only a simple linear relationship between biomass and volume when taking into account the influence of forest age, because other factors, site conditions, climate factors, etc., are already included in the volume. Basal area, stand age, stand height used in our study is the

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1805_72157232

explanation of this information including density, forest age, and site conditions. In many studies, shrubbery and bamboo plots often not be included or be estimated using other methods instead of BEF when forest biomass is estimated (Guo et al., 2013). Owing to the investigation standards of shrubbery and bamboo plots are different from those of arbor forests, for their starting DBH usually lower. And forest types and tree species in Guizhou province are diverse. At present, the volume equation of over 30 tree species including *Pinus massoniana*, *Cunninghamia lanceolata*, *Pinus yunnanensis* and *Pinus armandii* and so on have been established, while there is no corresponding volume equation for other tree species whose volume equation have not been established. And the volume equation also varies with the region. To avoid the impact of volume calculation errors on biomass modeling, this study established the relationship between three variables, namely basal area, stand age and stand height, with forest biomass, which not only avoided the trouble of establishing biomass according to different tree species, but these three variables also reflected the factors of volume, environment and other factors of the sample plot. In addition, the purpose of this study is to attach more importance to reveal the influence of spatial correlation and heterogeneity of biomass distribution on model fitting than the biomass model itself. Facts have proved that the three variables, basal area, stand age, and stand height are feasible for the estimation of regional forest biomass, and the model has a high explanatory degree with R^2 of the least square model reaches 91%.

OLS, SLM, SEM, and LMM global models remained unchanged among the studied regions, however, their coefficients were insufficient for the accurate description of three predicting factor effects on the forest biomass across various local regions. To be specific, GWR model offered the local model coefficients, and they were observed by the GIS technique to observe more details about on the association of forest biomass with the predicting factors (*Fig. 2*). Clearly, those predicting factors had different influences on forest biomass, depending on the different locations. For GWR model, its local model coefficients offered more details about the effects of micro-site variation and managing measures on forest biomass and tree growth. It would help to plan the management and make a decision (Lu and Zhang, 2012; Zhen et al., 2013). In view of the above analysis, GWR can be combined with remote sensing data to improve the accuracy of traditional models for estimating forest parameters, and in addition, the forest parameters can be further predicted and validated through each Reflectivity Pixels from remote sensing model and remote sensing data can be used to predict the spatial distribution of and a relatively more accurate spatial distribution map can be draw (Propastin, 2012; Chen, 2012). However, the distance between sampling points, the number of samples could influence the estimation of GWR model, and even the influence of outliers, weak data problem and even lack of independence weaken and limit its application (Zhang et al., 2004; Zhang and Shi, 2004; Shi et al., 2006).

Spatial autocorrelations as well as heterogeneities within the model residuals

Our results indicated that when considering spatial autocorrelations within model residuals, SLM, SEM, LMM, and GWR obtained a significant improvement over the traditional OLS model, which resulted in a biased hypothesis tested according to model coefficient. Particularly in that figure of local Moran's *I*, those spatial autocorrelations in residuals showed obvious characters. Across the entire study area, for OLS model, the local Moran's *I* values for its residuals were mostly positive (presented as "hot spots") while that of other models, especially the GWR model residuals were mostly negative

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1805_72157232

(presented as "cold spots"). Zhang et al. (2009) obtained similar conclusions while studying the relationship between DBH and tree height using these spatial regression models. As suggested by Moran's *I* and LM tests (*Table 4*) for the model residuals (*Fig. 3*), SEM model showed higher suitability to manage spatial autocorrelations than SLM. Therefore, spatial autocorrelations are the troublesome things due to model misspecification rather than the dependent variable being influenced by the values of the neibouring dependent variables (Luo et al., 2016). Though SLM and SEM models allowed for direct and effective correction for spatial autocorrelations among data corrected spatially, they were unable to manage spatial heterogeneity, as shown by those intra-block variance patterns among various block sizes like those for OLS model (*Table 3*; *Fig. 4*). Data fitted by LMM model were superior to those fitted by SLM and SEM, which offered a larger number of spatial autocorrelations in the expected model residuals (*Table 4*; *Fig. 3*), like in other reported studies (Zhang et al., 2009). LMM model, one of the global models, is also used to manage spatial correlation through 2 manners, including adjustment and characterization. Of them, the adjustment manner obtains the rate estimates for response variables using EBLUP, the global kriging in Geostatistics, whereas the characterization manner estimates the spatial covariance variables, such as semivariogram range, partial still, and nugget (Ozdenerol, 2006). LMM used the exponential spatial covariance structure, like geographical weight function in GWR model, to calculate those spatial weights for adjacent forest plots. As a result, LMM model emphasized the "local" data determined based on those semivariogram variables, so as to provide a large number of expected spatial autocorrelations and heterogeneities in model residuals relative to other global models.

According to other research (e.g., Zhang et al., 2009; Kupfer and Farris, 2007), GWR model accommodates the spatial heterogeneity in the meantime of markedly reducing the spatial autocorrelation within model residuals (*Fig. 3*). GWR, the local spatial model, adopts the moving window across an observation set distributed spatially to produce a model coefficient set based on data subsamples surrounding certain points spatially (Páez and Scott, 2005). Although GWR does not merge the spatial autocorrelation during the process of modeling, since it assumes a normal distribution N $(0, \sigma^2 I)$ in model error term. Besides, the GMR model definitely considers the spatial locations and emphasizes the local variations regarding the associations among variables. It is precisely due to the above characteristics of GWR, it represents an efficient approach in the mountainous regions with complicated terrains (Wang et al., 2020). The GWR model can be improved according to the need of the research. For example, Propastin (2012) extended the GWR model to develop a geographically altitudinal weighted regression (GAWR) model for managing altitudinal (vertical) as well as spatial (horizontal) instabilities in estimating the aboveground biomass for a rainforest region in tropics. Therefore, considering the diversity and complexity of forest ecosystems, GWR offers the highly precise visual data for afforestation planning and management measures with cost- and labor-effectiveness as long as it is properly applied (Zhang et al., 2009; Liu et al., 2014).

Conclusion

It is concluded that basal area, stand age, stand height were closely associated forest biomass in our study. They were the explanation of density, forestage, and site conditions. The distinct spatial autocorrelations as well as variations exist between

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1805_72157232

forest biomass and the variables, thus the OLS was not appropriate for modeling. SLM and SEM efficiently accounted for the spatial autocorrelations within model residual, but insufficient to deal with the problem of spatial heterogeneity. In contrast, the LMM and GWR incorporated the spatial dependence and variation into modeling processes, and consequently, fitted the data better and predicted the response variable more accurately. Therefore, in the presence of obvious spatial variations and autocorrelations between dependent and independent variables, and spatial regression models such as LMM, especially GWR should be considered.

Acknowledgements. The authors thank for the data from Chinese National Forest Inventory (CNFI) and Guizhou Provincial Forestry Department. This work was supported by the Science and Technology Planning Project of Guizhou Province of China (QKHJC[2017]1042; QKHZC[2017]2520-1); the National Natural Science Foundation of China (31700385); the National Key Research and Development Program of China (2017YFD0600302); and the First-class Discipline Construction Project of Guizhou Province of China (GNYL[2017]007).

REFERENCES

[1] Anselin, L. (1992): Spatial Dependence and Spatial Heterogeneity: Model Specification Issues in the Spatial Expansion Paradigm. – In: Jounes, J. P., Casetti, E. (eds.) Applications of the Expansion Method. Routledge, London.
[2] Anselin, L. (1993): Discrete Space Autoregressive Models. – In: Goodchild, M. F., Parks, B. O., Steyaert, L. T. (eds.) Environmental Modeling with GIS. Oxford University, New York.
[3] Anselin, L. (2001): Spatial Econometrics. – In: Baltagi, B. (ed.) A Panion to Theoretical Econometrics. Blackwell, Oxford, UK.
[4] Anselin, L., Griffith, D. A. (1988): Do spatial effects really matter in regression analysis? – Papers in Regional Science 65: 11-34.
[5] Anselin, L., Syabri, I., Kho, Y. (2006): GeoDa: an introduction to spatial data analysis. – Geographical Analysis 38: 5-22.
[6] Benson, M., Pierce, L., Sarabandi, K. (2016): Estimating boreal forest canopy height and above ground biomass using multi-modal remote sensing; a database driven approach. – IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2016, Beijing, China, July 10-15.
[7] Boots, B. (2002): Local measures of spatial association. – Écoscience 9: 168-176.
[8] Borcard, D., Legendre, P. (2002): All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. – Ecological Modelling 153: 51-68.
[9] Brown, S. (2002): Measuring carbon in forests: current status and future challenges. – Environmental Pollution 116: 363-372.
[10] Brown, S., Lugo, A. E. (1984): Biomass of tropical forests: a new estimate based on forest volumes. – Science 223: 1290-1293.
[11] Cannell, M. G. R. (1984): Woody biomass of forest stands. – Forest Ecology and Management 8: 299-312.
[12] Chen, G., Zhao, K., Mcdermid, G. J., Hay, G. J. (2012): The influence of sampling density on geographically weighted regression: a case study using forest canopy height and optical data. – International Journal of Remote Sensing 33: 2909-2924.
[13] Chiesi, M., Fibbi, L., Genesio, L., Gioli, B., Magno, R., Maselli, F., Moriondo, M., Vaccari, F. P. (2011): Integration of ground and satellite data to model Mediterranean forest processes. – International Journal of Applied Earth Observations and Geoinformation 13: 504-515.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1805_72157232

[14] Dormann, C. F., Mcpherson, J. M., Araújo, M. B., Bivand, R., Bolliger, J., Carl, G., Davies, R. G., Hirzel, A., Jetz, W., Kissling, D. W., Kühn, I., Ohlemüller, R., Peres-Neto, P. R., Reineking, B., Schröder, B., Schurr, F. M., Wilson, R. (2010): Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. – Ecography 30: 609-628.
[15] Duncan, C. (1997): Applying Mixed Multivariate Multilevel Models in Geographical Research. – In: Wstert, P., Verhoeff, R. N. (eds.) Place and People: Multilevel Modeling in Geographical Research. Nederlandse Geografische Studies 227, Univ. of Utrecht, Utrecht.
[16] Editorial Committee of Guizhou Forest (1991): Guizhou Forest. – Guizhou Science and Technology, Guiyang (in Chinese).
[17] Erdoğan, S. (2010): Modelling the spatial distribution of dem error with geographically weighted regression: an experimental study. – Computers & Geosciences 36: 34-43.
[18] Fang, J., Chen, A., Peng, C., Zhao, S., Ci, L. (2001): Changes in forest biomass carbon storage in China between 1949 and 1998. – Science 292: 2320-2322.
[19] Fotheringham, A. S., Brunsdon, C. (1999): Local Forms of Spatial Analysis. – Geographical Analysis 31(4): 340-358.
[20] Fotheringham, A. S., Brunsdon, C., Charlton, M. (2002): Geographically Weighted Regression: The Analysis of Spatially Varying Relationships. – John Wiley & Sons, New York.
[21] Fu, W. J., Jiang, P. K., Zhou, G. M., Zhao, K. L. (2014): Using Moran's I and GIS to study the spatial pattern of forest litter carbon density in a subtropical region of southeastern China. – Biogeosciences 11: 2401-2409.
[22] Ghosh, S. M., Behera, M. D. (2018): Aboveground biomass estimation using multi-sensor data synergy and machine learning algorithms in a dense tropical forest. – Applied Geography 96: 29-40.
[23] Gorr, W. L., Olligschlaeger, A. M. (1994): Weighted spatial adaptive filtering: Monte Carlo studies and application to illicit drug market modeling. – Geographical Analysis 26: 67-87.
[24] Graaff, T. D., Florax, R. J. C. M., Nijkamp, P., Reggiani, A. (2001): A general misspecification test for spatial regression models: dependence, heterogeneity, and nonlinearity. – Journal of Regional Science 41: 255-276.
[25] Guo, Z. D., Hu, H. F., Li, P., Li, N., Fang, J. (2013): Spatio-temporal changes in biomass carbon sinks in China's forests during 1977-2008. – Science China Life Sciences 56: 661-671.
[26] Jones, J. (1997): Multilevel Approaches to Modeling Contextuality. – In: Dale, A. (ed.) Nederlandse Geografische Studies. Univ. of Utrecht, Utrecht, pp. 19-40.
[27] Kupfer, J. A., Farris, C. A. (2007): Incorporating spatial non-stationarity of regression coefficients into predictive vegetation models. – Landscape Ecology 22: 837-852.
[28] Li, R., Cui, L., Fu, H., Meng, Y., Li, J., Guo, J. (2020): Estimating high-resolution PM1 concentration from Himawari-8 combining extreme gradient boosting-geographically and temporally weighted regression (XGBoost-GTWR). – Atmospheric Environment 229: 117434.
[29] Li, X., Shang, B., Wang, D., Wang, Z., Wen, X., Kang, Y. (2020): Mapping soil organic carbon and total nitrogen in croplands of the corn belt of northeast China based on geographically weighted regression kriging model. – Computers & Geosciences 135: 104392.
[30] Lichstein, J. W., Simons, T. R., Franzreb, S. K. E. (2002): Spatial autocorrelation and autoregressive models in ecology. – Ecological Monographs 72: 445-463.
[31] Lin, Z., Chao, L., Wu, C., Hong, W., Hong, T., Hu, X. (2018): Spatial analysis of carbon storage density of mid-subtropical forests using geostatistics: a case study in Jiangle County, southeast China. – Acta Geochimica 37: 92-103.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1805_72157232

[32] Littell, R. C., Milliken, W. W., Stroup, R. D., Wolfinger, A. (2006): SAS for Mixed Models. 2nd Ed. – SAS Press, Cary, USA.
[33] Liu, C. C., Wei, Y. F., Liu, Y. G., Guo, K. (2009): Biomass of canopy and shrub layers of karst forests in Puding, Guizhou, China. – Chinese Journal of Plant Ecology 33: 74-81 (in Chinese).
[34] Liu, C., Zhang, L., Li, F., Jin, X. (2014): Spatial modeling of the carbon stock of forest trees in Heilongjiang Province, China. – Journal of Forest Research 25: 269-280.
[35] Lou, M., Zhang, H., Lei, X., Li, C. (2016): Spatial autoregressive models for stand top and stand mean height relationship in mixed Quercus mongolica broadleaved natural stands of northeast China. – Forests 7: 43.
[36] Lu, J., Zhang, L. (2012): Geographically local linear mixed models for tree height-diameter relationship. – Forest Science 58: 75-84.
[37] Mani, S., Parthasarthy, N. (2007): Above-ground biomass estimation in ten tropical dry evergreen forest sites of peninsular India. – Biomass Bioenergy 31: 284-290.
[38] Masera, O. R., Ordóñez, M. J., Dirzo, R. (1997): Carbon emissions from Mexican forests: current situation and long term scenarios. – Climate Change 35: 265-295.
[39] McMillen, D. P. (2004): Geographically weighted regression: the analysis of spatially varying relationships. – American Journal of Agricultural Economics 86: 554-556.
[40] Monjarás-Vega, N. A., Briones-Herrera, C. I., Vega-Nieva, D. J., Calleros-Flores, E., Corral-Rivas, J. J., López-Serrano, P. M., Pompa-García, M., Rodríguez-Trejo, D. A., Carrillo-Parra, A., González-Cabán, A., Alvarado-Celestino, E., Jolly, W. M. (2020): Predicting forest fire kernel density at multiple scales with geographically weighted regression in Mexico. – Science of the Total Environment 718: 137313.
[41] Nakaya, T., Fotheringham, A., Charlton, M., Brusdon, C. (2009): Semiparametric geographically weighted generalized linear modelling in GWR 4.0. – 10th International Conference on GeoComputation, 30th November-2nd December 2009, UNSW, Sydney.
[42] Öcal, N., Yildirim, J. (2010): Regional effects of terrorism on economic growth in turkey: a geographically weighted regression approach. – Journal of Peace Research 47: 477-489.
[43] Overmars, K. P., Koning, G. H. J. D., Veldkamp, A. (2003): Spatial autocorrelation in multi-scale land use models. – Ecological Modelling 164: 257-270.
[44] Ozdenerol, E. (2006): Statistical methods for spatial data analysis. – Publications of the American Statistical Association 101: 389-340.
[45] Páez, A., Scott, D. M. (2005): Spatial statistics for urban analysis: a review of techniques with examples. – GeoJournal 61: 53-67.
[46] Patenaude, G., Milne, R., Dawson, T. P. (2005): Synthesis of remote sensing approaches for forest carbon estimation: reporting to the Kyoto Protocol. – Environmental Science & Policy 8: 161-178.
[47] Ploton, P., Barbier, N., Couteron, P., Antin, C. M., Ayyappan, N., Balachandran, N., Barathan, N., Bastin, J. F., Chuyong, G., Dauby, G., Droissart, V., Gastellu-Etchegorry, J. P., Kamdem, N. G., Kenfack, D., Libalah, M., Mofack II, G., Momo, S. T., Pargal, S., Petronelli, P., Réjou-Méchain, M., Sonké, B., Texier, N., Thomas, D., Verley, P., Dongmo, Z. D., Berger, U., Pélissier, R. (2017): Toward a general tropical forest biomass prediction model from very high resolution optical satellite images. – Remote Sensing of Environment 200: 140-153.
[48] Propastin, P. (2012): Modifying geographically weighted regression for estimating aboveground biomass in tropical rainforests by multispectral remote sensing data. – International Journal of Applied Earth Observation and Geoinformation 18: 82-90.
[49] Qin, H., Huang, Q., Zhang, Z., Lu, Y., Li, M., Lang, X., Chen, Z. (2019): Carbon dioxide emission driving factors analysis and policy implications of Chinese cities: combining geographically weighted regression with two-step cluster. – Science of the Total Environment 684: 413-424.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1805_72157232

[50] Rodríguez-Veiga, P., Saatchi, S., Tansey, K., Balzter, H. (2016): Magnitude, spatial distribution and uncertainty of forest biomass stocks in Mexico. – Remote Sensing of Environment 183: 265-281.
[51] Sagar, R., Singh, J. S. (2006): Tree density, basal area and species diversity in a disturbed dry tropical forest of northern India: implications for conservation. – Environment Conservation 33: 256-262.
[52] Shi, H., Zhang, L., Liu, J. (2006): A new spatial-attribute weighting function for geographically weighted regression. – Canadian Journal of Forest Research 36: 996-1005.
[53] Song, P., Huang, J., Mansaray, L. R. (2019): An improved surface soil moisture downscaling approach over cloudy areas based on geographically weighted regression. – Agricultural and Forest Meteorology 275: 146-158.
[54] Tian, X. L., Xia, Q., Xia, H. B., Ni, J. (2011): Forest biomass and its spatial pattern in Guizhou province. – Chinese Journal of Applied Ecology 22: 287-294 (in Chinese).
[55] Wang, S., Liu, Y., Zhao, C., Pu, H. (2019): Residential energy consumption and its linkages with life expectancy in mainland China: a geographically weighted regression approach and energy-ladder-based perspective, – Energy 177: 347e357.
[56] Wang, D., Li, X., Zou, D., Wu, T., Xu, H., Hu, G., Li, R., Ding, Y., Lin, Z., Li, W., Wu, X. (2020): Modeling soil organic carbon spatial distribution for a complex terrain based on geographically weighted regression in the eastern Qinghai-Tibetan Plateau. – Catena 187: 104399.
[57] Zhang L., Shi, H. (2004): Local modeling of tree growth by geographically weighted regression. – Forest Science 50: 225-244.
[58] Zhang, L., Bi, H., Cheng, P., Davis, C. J. (2004): Modeling spatial variation in tree diameter–height relationships. – Forest Ecology and Management 189: 317-329.
[59] Zhang, L., Ma, Z., Guo, L. (2009): An evaluation of spatial autocorrelation and heterogeneity in the residuals of six regression models. – Forest Science 55: 533-548.
[60] Zeng, W. S., Tang, S. Z., Xia, Z. S., Zhu, S., Luo, H. Z. (2011): Using linear mixed model and dummy variable model approaches to construct generalized single-tree biomass equations in Guizhou. – Forest Research: 24: 285-291 (in Chinese).
[61] Zhen, Z., Li, F. R., Liu, Z. G., Liu, C., Zhao, Y. L., Ma, Z. H., Zhang, L. (2013): Geographically local modeling of occurrence, count, and volume of downwood in Northeast China. – Applied Geography 37: 114-126.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1805_72157232