Journal of Theoretical and Applied Information Technology
31st May 2023. Vol.101. No 10

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

SINGLE BOOTSTRAP APPROACH WITH GEOGRAPHICALLY WEIGHT REGRESSION MODELING USING PARTICLE-SIZE FRACTION

HENNY PRAMOEDYO[1*], WIGBERTUS NGABU[2] , SATIVANDI RIZA[3] , LUTHFATUL AMALIANA[4]

[1,2,4]Statistic Department, Faculty of Mathematics and Science, University of Brawijaya, Malang, Indonesia
[3]Soil Science Department, Faculty of Agriculture, University of Brawijaya, Malang, Indonesia
E-mail: [1*]hennyp@ub.ac.id , [2]bertongabu@gmail.com

ABSTRACT

The challenge of the high need for soil spatial data information has led to the rapid development of spatial modeling for soil attributes in the last few decades. Soil texture is an essential attribute that determines the direction of soil management and must be modeled accurately. However, on the other hand, soil texture is a soil attribute that is relatively difficult to model because it is a compositional data set. The difficulty that arises from this compositional data set is the limitation of constant quantities; namely, the sum of the fractions of sand, silt, and clay must be 100%. Through DEM data, topographical variability can be obtained so that it will be a predictor or independent variable in predicting soil texture. In addition, Geographically Weighted Regression (GWR) was also used in this study to pay attention to the effect of spatial heterogeneity. It uses the bootstrap method with the GWR model to overcome bias in the model parameters. Residual bootstrap is a bootstrap method that is applied to the residual resampling process. The aims of this study: (1) To establish a soil texture prediction model using GWR with a single bootstrap approach, (2) To test the model's reliability in predicting surface soil texture. The results of this study are in the form of a prediction model and a map of the spatial distribution of PSF on surface soil which can later be used as a basis for determining sustainable soil management and supporting precision agriculture.
Keywords: *GWR, Single Bootsrap, Soil particle-size fractions*

1. INTRODUCTION

The development of spatial modeling for soil attributes has accelerated in the last few decades. It is a challenge to the high need for soil spatial data information to realize the acceleration of precision agricultural activities. Precision farming uses new technologies to increase yields and profitability while lowering standard inputs needed to grow crops. In other words, farmers who take advantage of precision farming use less to produce more. In this case, the role of modeling is as a provider of primary quantitative data to determine the direction of land management. Soil texture is an essential attribute that determines the direction of soil management and must be modeled accurately. The importance of soil texture cannot be overlooked. Soil texture, and of course, the particle size distribution, partly determines the flow of water, heat, and nutrients, the holding capacity of water and nutrients, and the shape and stability of the soil structure. However, on the other hand, soil texture is a soil attribute that is relatively difficult to model because it is a compositional set of data that describes the particle size of the soil mineral fraction with sand, silt, and clay variables [1].

Through DEM data, topographic variability can be obtained to be a predictor or independent variable in predicting soil texture [2]. In addition, Geographically Weighted Regression (GWR) was also used in this study to pay attention to the effect of spatial heterogeneity. Spatial heterogeneity is a condition in an area with different conditions from one location to another [3]. In Regression analysis, several assumptions must be met. However, suppose this regression is applied to model data that is influenced by spatial aspects or geographical conditions. In that case, several assumptions will be challenging to fulfill and cause spatial heterogeneity [4].

Outliers are an extreme data [5]. Outliers cannot be discarded, because they may provide important information that other data cannot provide. The result of the presence of outliers is that the parameter estimator produced by the MKT will be biased. Bias in the parameter estimator will result in the estimator losing the Best Linear Unbiased Estimator (BLUE) properties.

Bootstrap is a data-based simulation method used to estimate parameters and construct confidence intervals without knowing the sample population distribution. Computation-based bootstrap is an alternative to solving problems empirically. This method proved more accurate than the asymptotic method in small samples and unknown parameter distributions. The basic principle of the bootstrap method is to generate a new data set from the original data as many as B replications [6]. Based on the characteristics of the bootstrap method, which is a data-based simulation method, and without knowing the population distribution of the sample owned, the researcher wants to know whether the residual bootstrap method is capable of overcoming bias in parameter estimators due to outliers at the bottom of the data set in Geographically Weighted Regression (GWR) analysis. In predicting the soil texture of the Upper Kalikonto watershed. This study aims to form a soil texture prediction model using the GWR model with a single bootstrap approach. Both GWR methods employ a single bootstrap approach to overcome bias in parameter estimators predicting surface soil texture.

The Bootstrap Geographically Weighted Regression (GWR) method aims to improve the GWR model's predictive performance and parameter estimation accuracy.

In order to calculate inferential statistics (like confidence intervals and p-values) for each regression coefficient at each location in the dataset, the GWR bootstrap applies the bootstrap technique to the GWR spatial regression analysis process [7]. This involves resampling the data at random with replacement.

The GWR Bootstrap technique can assist in determining the local significance of each independent variable at each location as well as provide a more precise estimate of the range of GWR parameter values because the regression coefficient in the GWR spatial regression analysis can vary at each location and is not homogeneous across datasets.

We can improve interpretation and decision-making based on the outcomes of the GWR spatial regression analysis by employing the GWR Bootstrap approach, which yields more dependable, resilient, and accurate analysis results.

2. MATERIAL AND METHOD

Study Area and Datasets

The study area is located in Kalikonto watershed, Indonesia. The Anjasmoro Mountain in the north and the Kawi Mountain in the south are separated by these watersheds. The majority of the study area's land is used for agricultural purposes. There are 235.7 km2 of rolling hills and plains that compose the physiography. The Particle Size Fraction (Sand, Silt, and Clay percentages) at 50 sampling places for soil properties are found in the top 10 cm. The topsoil layer's sand, silt, and clay contents were the main focus of this investigation. To utilize as a predictor variable,

Elevation based on DEM data, slope, and local morphometric variables (LMV) were computed. These methods are predicated on the general idea that statistical connections between soil properties and morphometric variables are reasonably reliable [8].

Journal of Theoretical and Applied Information Technology
31st May 2023. Vol.101. No 10

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

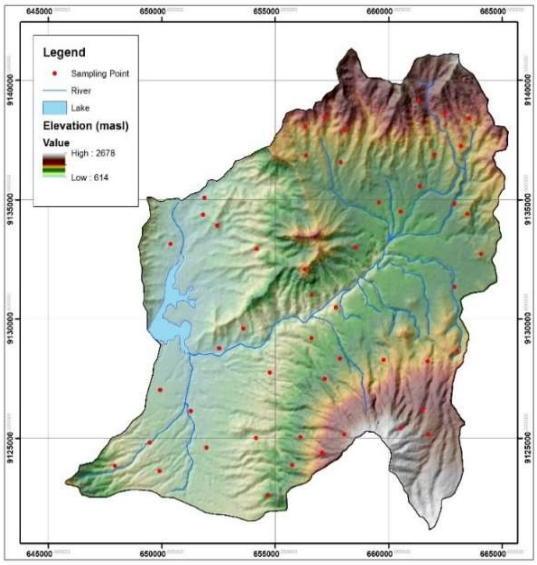


Figure 1. Sampling Point Location.

To calculate the parameter, we use 30 m SRTM DEM data from the USGS data site for the whole watershed. Prior to the computation of the topographic features, certain preprocessing, such as georeferencing, clipping/subsetting, and fill sink, was carried out as normal practice for DEM. Based on the Evans-Young method second-order polynomial, 16 Local Morphometric Variables (LMV) that are functions of partial derivatives of elevation were calculated [9]. The 16 LMV parameters will be paired with each PSF primary data to obtain a PSF estimation model, so that the final result will be three estimation models for each PSF (Sand, Silt, Clay).

The variables in this study consist of 8 Local Morphologic Variables (LMV) which show the curvature of a topography [10]. The LMV consists of:

1. Vertical Curvature (K_v)

$$K_v = \frac{p^2r + 2pqs + q^2t}{(p^2+q^2)\sqrt{(1+p^2+q^2)^3}}$$

2. Horizontal Curvature (K_h)

$$K_h = \frac{q^2r - 2pqs + p^2t}{(p^2+q^2)\sqrt{1+p^2+q^2}}$$

3. Accumulation Curvature (K_a)

$$K_a = \frac{(q^2r - 2pqs + p^2t)(p^2r + 2pqs + q^2t)}{[(p^2+q^2)(1+p^2+q^2)]^2}$$

4. Ring Curvature (K_r)

$$K_r = \left[\frac{(p^2-q^2)s - pq(r-t)}{(p^2+q^2)(1+p^2+q^2)}\right]^2$$

5. Slope (S)

$$S = arctan\sqrt{p^2+q^2}$$

6. Elevation (E_{lev})

However, to obtain these variables, an analysis of the DEM data is first carried out to obtain the derived value of the elevation which is the digital number value of the DEM data. To get the derived elevation value, the following formula is used:

$$p = \frac{z_3 + z_6 + z_9 - z_1 - z_4 + z_7)}{6w^2}$$

$$q = \frac{z_1 + z_2 + z_3 - z_7 - z_8 - z_9)}{6w^2}$$

$$r = \frac{z_1 + z_3 + z_4 + z_6 + z_7 + z_9 - 2(z_2 + z_5 + z_8)}{3w^2}$$

$$s = \frac{z_3 + z_7 - z_1 - z_9}{4w^2}$$

$$t = \frac{z_1 + z_2 + z_3 + z_7 + z_8 + z_8 - 2(z_4 + z_5 + z_6)}{3w^2}$$

Where z is the elevation and w is the cell size in pixels [11]. Technically, to get the z value, it is necessary to use the measurement window as shown below:

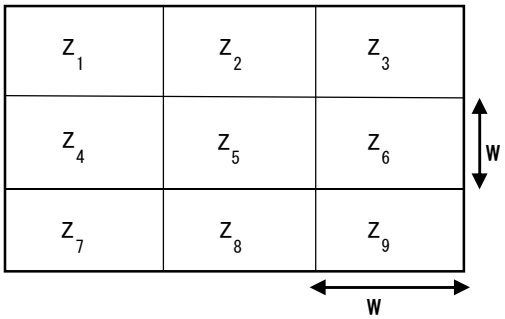


Figure 2. Illustration of the measurement window for obtaining elevation derived values (p,q,r,s and t).

PSF Modelling Using Geographically Weighted Regression

Estimation of each PSF uses Geographically Weighted Regression (GWR)

using the Weighted Least Square (WLS) method, which means giving different weights for each location. The GWR model is a development of the global regression model which was developed from the global regression model involving geographic factors [12]. Geographically Weighted Regression (GWR) is a spatial analysis using points which is also the development of linear regression analysis taking into account location (spatial) [13].

Parameter estimation of the Geographically Weighted Regression (GWR) model uses the Weighted Least Square (WLS) method which means it provides different weightings for each location [14]. The weighting for each location (u_i, v_i) is written as $w_j(u_i, v_i)$ where j = 1.2,....,n. Weighting with different values indicates that the nature of the location on the GWR model is different. The following is an estimation of the parameters for the GWR model for observation locations (u_i, v_i) based on the addition of weighting $w_j(u_i, v_i)$ which is written as follows.

$y_i w_j(u_i, v_i) = w_j(u_i, v_i)(\beta_0(u_i, v_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i) x_{ik} + e_i)$ (1)

Spatial Weight

A spatial weighting matrix is a matrix that can describe the location of an observation that is adjacent to other observations that can allow for relationships such as intersections and closeness of distances between observations. The diagonal for the spatial weighting matrix contains zero values, because the weighting matrix shows the relationship between the entire location [15]. The form of the spatial weighting matrix and its constituent elements are written as follows:

$$\boldsymbol{W} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{bmatrix}$$

Information:

$\boldsymbol{W}$: spatial weighting matrix of $n \times n$ size

w_{ij} : elements of the spatial weighting matrix between the I-th observation location and the J-th observation location, with i =1,2,..,n and j =1,2,…,n.

The spatial weighting matrix often used in GWR modeling is a kernel function. The spatial weighting matrix is useful as the weight of estimating different parameters between observation sites in GWR modeling [16]. This indicates the proximity between the centers of the observation points which has a great influence on the estimation of parameters for the i-th observation location. In this study, the weighting used was the Fixed Bisquare Kernel. A fixed kernel function is a function that has the same bandwidth for all points of location. Fixed Kernel can be used when the observation data points are located in order, with the following weighting formula:

$$w_{ij} = \begin{cases} \left[1 - \left(\frac{d_{ij}}{b}\right)^2\right]^2, & \text{if } d_{ij} < b \\ 0, & \text{if } d_{ij} \geq b \end{cases} \quad (2)$$

Testing the significance of the GWR model and selecting the best model were also carried out at this stage. Parameter coefficient testing is carried out using the t-test statistic. while the selection of the best model can use the AIC (Akaike's Information Criterion) score criteria [17]. The best

$AIC = 2n\,log(\hat{\sigma}) + nlog(2\pi) + n\left(\frac{n+tr(\boldsymbol{L})}{n-2-tr(\boldsymbol{L})}\right)$ (3)

$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{(n-[2tr(\boldsymbol{L}) - tr(\boldsymbol{L}^T\boldsymbol{L})])}$ (4)

model is the model with the smallest AIC value because the parameter estimators are close to the actual parameter values [18]. The AIC were calculated using formula below:

Were:

$\hat{\sigma}$: estimated *standart deviation*

n : amount of observation

Single Bootstrap

Bootstrap is a computation-based method which is a nonparametric and resampling technique for estimating the standard error $\hat{\tau}$.

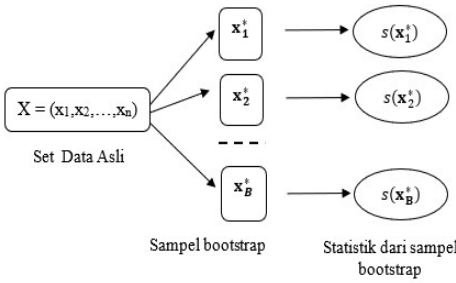


Figure 3. Bootstrap processing illustration.

The bootstrap algorithm begins by

generating B independent samples of each n, that is $\mathbf{x_1^*}$, $\mathbf{x_2^*}$, …, $\mathbf{x_B^*}$, then we got the statistic from the replication of the B $s(\mathbf{x}_b^*)$ with b = 1,2, . . , B. If $s(\mathbf{x})$ is the average of the observation sample then $s(\mathbf{x}_b^*)$ is the average of the bootstrap data sample [19]. We also need to calculate the residual bootstrap. The residual bootstrap is a data-based simulation method in which the resampling process is applied to the residuals generated by the regression analysis model [20]. The residual bootstrap sample is defined as a random sample of size n, with application to the GWR model using the residual model.

The GWR model that has been carried out will produce a residual data set. The residual data set is done by obtaining Single Bootstrap replication. The residual bootstrap data set is used to estimate the y_b^* (Bootsraped PSF value) with the following equation.

$$y_b^* = \beta_0(u_i, v_i) + \sum_{k=1}^{p} \beta_k(u_i, v_i) x_{ik} + e_b^* \quad (5)$$

The y_b^* that has been obtained in each replication data set is modeled with the GWR model to obtain parameter estimates for each replication data set. GWR Bootstrap model parameter estimation obtained by using the equation

$$\widehat{\boldsymbol{\beta}}(u_i, v_i) = [\boldsymbol{X'W}(u_i, v_i)\boldsymbol{X}]^{-1}\boldsymbol{X'W}(u_i, v_i)\boldsymbol{Y} \quad (6)$$

for each replication data set is $\hat{\beta}_b^*$, then the estimated value of the Bootstrap GWR parameter is

$$\hat{\beta}^* = \frac{1}{B}\sum_{b=1}^{B} \hat{\beta}_b^* \quad (7)$$

Then for each location is

$$\beta_{ij}^* = \frac{1}{B}\sum_{b=1}^{B} \hat{\beta}_{bij}^* \quad (8)$$

Step by step bootstrap Residual and Bootstrap standard error as follows :

The steps in the residual bootstrap are as follows:

1. Specifies $\hat{y}$ value of the parameter estimator generated by MKT.
2. Obtaining a Residual value. The residual value is obtained by calculating the difference $e_i = y_i - \hat{y}_i$
3. Takes n random samples with returns from $e_1, e_2, \dots, e_1$,, thus generating $e^* = (e_1^*, e_2^*, e_3^*, \dots, e_n^*)$,
4. Calculates bootstrap value for Y^* by adding e^*
5. Calculates the regression coefficient for the bootstrap sample Y^* with X so that it obtains β^*
6. Repeat steps 2, 3, and 4 according to the desired replication count.

Steps to estimate the default bootstrap error include:

1. Specifies bootstrap-free samples $\mathbf{x}_1^*$, $\mathbf{x}_2^*$,…, $\mathbf{x}_B^*$, where each sample consists of n data taken from its original dataset with returns.
2. Evaluate replication on each bootstrap sample formed. $\hat{\tau}_b^* = s(\mathbf{x}_b^*)$, b= 1,2,…, B where $s(\mathbf{x}_b^*)$ is the average of bootstrap result datasets with

$$s(\mathbf{x}_b^*) = \bar{\mathrm{x}}^* = \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_i^*$$

3. Estimates the default error $\hat{\tau}$ as much as B replication

$$se(\hat{\tau}_b^*) = \left\{ \frac{\sum_{b-1}^{B} (\hat{\tau}_b^* - \hat{\tau}^*)^2}{(B-1)} \right\}^{\frac{1}{2}},$$

b=1,2,…,B

$$\hat{\tau}^* = \frac{1}{B}\sum_{b=1}^{B} \hat{\tau}_b^*$$

3. RESULT AND DISCUSSION

Soil texture in this study was calculated quantitatively with laboratory analysis to obtain actual data at each observation point. The results of laboratory analysis showed that the study location was dominated by loam and sandy loam textures. Quantitative soil survey results, together with modeling, can provide a suitable paradigm for the spatial prediction of single soil characteristics such as soil texture, and can provide information about the terrain attributes

that affect the movement of soil particles.

Based on the results of Moran's I test using that can be seen on Table 1, the autocorrelation test results were obtained for the Sand variable where five variables were eliminated, namely X3, X4, X5, X8, X10, these five variables did not have spatial autocorrelation with respect to the Sand variable. Then the Spatial Autocorrelation test on the Silt variable obtained 4 variables which were eliminated, namely X3, X5, X8, X10, the four variables above did not have spatial autocorrelation on the Silt variable, then they were eliminated. And the Spatial autocorrelation test on the Clay variable shows that all variables can be used and nothing is eliminated. Spatial autocorrelation test with a confidence interval of 95%. Because the variables in the study contain spatial autocorrelation, the model will be better if you use a geographically weighted regression model instead of using a global regression model.

Table 1. Spatial Autocorrelation Testing Result

Parameters	Sand	Silt	Clay
	P-value	P-value	P-value
X1	0.047**	0.0391**	6.504e-05**
X2	0.0046**	0.0023**	4.788e-06**
X3	0.2584	0.1911	8.02e-05**
X4	0.6175	1.043e-13**	2.2e-16**
X5	0.1265	0.0921**	0.00011**
X6	0.0776**	0.04303**	5.823e-06**
X7	0.00117**	0.00019**	1.029e-08**
X8	0.3535	0.2871	0.000278**
X9	0.0253**	0.0134**	5.631e-06**
X10	0.2728	0.1945	4.26e-05**
X11	0.00016**	4.392e-09**	1.002e-07**
X12	9.22e-06**	4.82e-07**	3.021e-10**
X13	0.02646**	0.01585**	2.056e-05**
X14	8.48e-08**	3.056e-09**	1.624e-13**
X15	2.38e-10**	3.628e-12**	4.499e-15**
X16	9.41e-13**	2.2e-16**	2.2e-16**

Based on the prediction results using the GWR approach, the R-square values of the three PSFs are obtained as shown in Table 2. It can be seen in the table that the Sand variable can be explained by 10 independent variables of 55.02%, the remainder is explained by other variables. Then the Silt variable can be explained by the 12 independent variables of 65.61%, the rest is explained by other variables outside the research variables. And the Clay variable can be explained by the 16 independent variables at 59.84%, the rest is explained by other variables outside the predictors used.

Table 2. Gwr Modeling Result.

	Sand	Silt	Clay
R-square	0.5502	0.6561	0.5984
AIC	1138.719	960.3721	955.9912

Modeling using the GWR with the bootstrapping approach was carried out with 1000 repetitions of resampling. The results of the AIC value of the boostrap GWR modeling are shown in Table 3.

Table 3. AIC Value Of Bootstrapping's GWR

Parameter	Sand	Silt	Clay
AIC	1103.594	872.5418	901.704

As a comparison, modeling for PSF estimation has also been carried out. This is necessary to see which model has the best performance. The results show that the three predicted PSF values have the best performance with the GWR bootstrap approach as evidenced by the lowest AIC value among the other approaches.

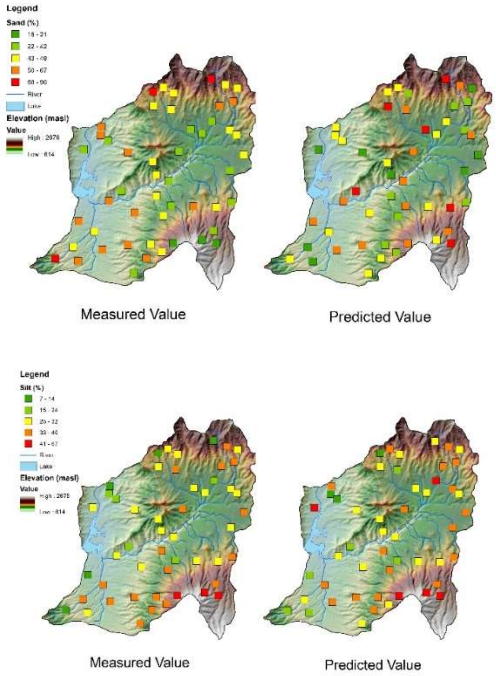
Journal of Theoretical and Applied Information Technology
31st May 2023. Vol.101. No 10

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 4. Comparison Between Modeling Approach

Modeling approach	Sand	Silt	Clay
Regresi	1175.038	1079.166	985.9001
GWR	1138.719	972.8993	955.9912
Bootstrap GWR	1103.594	872.8993	901.7104

Figure 4 shows the comparison map between measured and predicted value of PSF. We found that there were no significant differences between the measured and predicted. It can be conclude that the PSF's modeling using Bootsrap GWR have a good performance and can be use to predict the PSF in the study area. Overall, the proposed approach can produce results that are not significantly different from those obtained by conventional techniques, i.e., measured data.



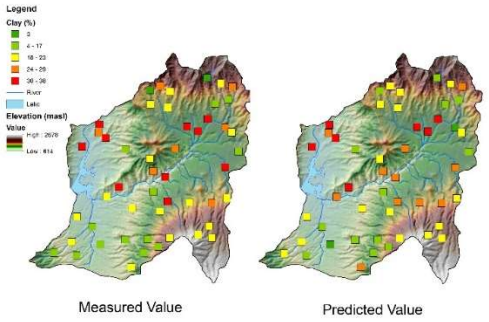


Figure 4. Comparison Between Measured And Predicted Value

4. CONCLUSION

Based on the research that has been done, it can be concluded that the Single Bootstrap GWR method has better performance than the general GWR. This is proven by the smallest AIC values for the three variables in the Prediction of Particle-Size Fraction on Top Soil. Although the predictive model completely explained the relationships between LMVs and soil PSF, it should be transformed to soil texture class in the future. In addition, modeling research and workflows can also be easily updated, with a view to improving them. We also think that being able to analyze data collected such as from soil surveys and combining collaboration between expert soil surveying expertise and statistical expertise, can make it easier for decision makers to plan precision agriculture in the future.

For further research it is necessary to proceed with the Fast Double Bootstrap (FDB) approach and the Bayesian approach, with the Fast Double Bootstrap and Bayesian approach will accommodate unfulfilled problems with normal distribution and small sample size as well as outlier data problems

5. ACKNOWLEDGEMENT

The authors wish to acknowledge the Department Statistics, Faculty of Mathematics and Sciences, University of Brawijaya for the funding support that made this research possible. We also want to acknowledge all the authors and colleagues for their thoughtful reviews during the drafting of this document.

REFERENCES:

[1] A. B. Mcbratney, *Pedometrics*. Cham: Springer International Publishing, 2018.

[2] V. Y.-J. Chen, T.-C. Yang, and H.-L. Jian, "Geographically weighted regression modeling for multiple outcomes," *Ann. Am. Assoc. Geogr.*, vol. 112, no. 5, pp. 1278–1295, 2022.

[3] G. Sun, H. Liu, D. Cui, and C. Chai, "Spatial heterogeneity of soil nutrients in Yili River Valley," *PeerJ*, vol. 10, p. e13311, 2022.

[4] E. Soroori, A. Mohammadzadeh Moghaddam, and M. Salehi, "Modeling spatial nonstationary and overdispersed crash data: Development and comparative analysis of global and geographically weighted regression models applied to macrolevel injury crash data," *J. Transp. Saf. Secur.*, vol. 13, no. 9, pp. 1000–1024, 2021.

[5] H. Wang, M. J. Bah, and M. Hammad, "Progress in outlier detection techniques: A survey," *Ieee Access*, vol. 7, pp. 107964–108000, 2019.

[6] J. Hahn and Z. Liao, "Bootstrap standard error estimates and inference," *Econometrica*, vol. 89, no. 4, pp. 1963–1977, 2021.

[7] P. Harris, C. Brunsdon, B. Lu, T. Nakaya, and M. Charlton, "Introducing bootstrap methods to investigate coefficient non-stationarity in spatial regression models," *Spat. Stat.*, vol. 21, pp. 241–261, 2017.

[8] A. . McBratney, M. . Mendonça Santos, and B. Minasny, "On digital soil mapping," *Geoderma*, vol. 117, no. 1–2, pp. 3–52, Nov. 2003, doi: 10.1016/S0016-7061(03)00223-4.

[9] P. L. Guth *et al.*, "Digital elevation models: terminology and definitions," *Remote Sens.*, vol. 13, no. 18, p. 3581, 2021.

[10] S. Riza, M. Sekine, A. Kanno, K. Yamamoto, T. Imai, and T. Higuchi, "Modeling soil landscapes and soil textures using hyperscale terrain attributes," *Geoderma*, vol. 402, p. 115177, 2021.

[11] A. Benito-Calvo *et al.*, "3D 360 surface morphometric analysis of pounding stone tools used by Hadza foragers of Tanzania: a new methodological approach for studying percussive stone artefacts," *J. Archaeol. Sci. Reports*, vol. 20, pp. 611–621, 2018.

[12] K. Cupido, A. S. Fotheringham, and P. Jevtic, "Local modelling of US mortality rates: A multiscale geographically weighted regression approach," *Popul. Space Place*, vol. 27, no. 1, p. e2379, 2021.

[13] D. C. Wheeler, "Geographically weighted regression," in *Handbook of Regional Science*, 2014.

[14] H. Yu, A. S. Fotheringham, Z. Li, T. Oshan, W. Kang, and L. J. Wolf, "Inference in multiscale geographically weighted regression," *Geogr. Anal.*, vol. 52, no. 1, pp. 87–106, 2020.

[15] E. Nuhn and S. Timpf, "Landmark weights-an alternative to spatial distances in shortest route algorithms," *Spat. Cogn. Comput.*, pp. 1–27, 2022.

[16] Z. Li and A. S. Fotheringham, "Computational improvements to multi-scale geographically weighted regression," *Int. J. Geogr. Inf. Sci.*, vol. 34, no. 7, pp. 1378–1397, 2020.

[17] G. Punzo, R. Castellano, and E. Bruno, "Using geographically weighted regressions to explore spatial heterogeneity of land use influencing factors in Campania (Southern Italy)," *Land use policy*, vol. 112, p. 105853, 2022.

[18] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature," *Geosci. Model Dev.*, 2014, doi: 10.5194/gmd-7-1247-2014.

[19] C. Jentsch and K. G. Lunsford, "Asymptotically valid bootstrap inference for proxy SVARs," *J. Bus. Econ. Stat.*, vol. 40, no. 4, pp. 1876–1891, 2022.

[20] J. Egbert and L. Plonsky, "Bootstrapping techniques," in *A practical handbook of corpus linguistics*, Springer, 2021, pp. 593–610.