Media Statistika

p-ISSN 1979 – 3693 e-ISSN 2477 – 0647
MEDIA STATISTIKA 13(2) 2020: 125-135
http://ejournal.undip.ac.id/index.php/media_statistika

IMPLEMENTATION OF LOCALLY COMPENSATED RIDGE-GEOGRAPHICALLY WEIGHTED REGRESSION MODEL IN SPATIAL DATA WITH MULTICOLLINEARITY PROBLEMS
(Case Study: Stunting among Children Aged under Five Years in East Nusa Tenggara Province)

Alfi Fadliana[1], Henny Pramoedyo[2], Rahma Fitriani[2]
[1] Department of Informatics Engineering, Universitas Islam Raden Rahmat
[2] Department of Statistics, Universitas Brawijaya

e-mail: alfifadliana@gmail.com

DOI: 10.14710/medstat.13.2.125-135

Article Info:
Received: 12 June 2019
Accepted: 19 October 2020
Available Online: 28 December 2020

Keywords:
Stunting, Multicollinearity, Geographically Weighted Regression, Locally Compensated Ridge

Abstract: East Nusa Tenggara Province, according to the findings of 2013 Baseline Health Research and 2016 and 2017 Nutritional Status Surveys, was recorded as the province with the highest prevalence of stunting in Indonesia. Efforts should be made to formulate policies that are integrated with spatial aspects in order to reduce the prevalence of stunting. The LCR-GWR model approach is used by using locally compensated ridge, which were meant to adjusts to the effect of collinearity between predictor variables (i.e., the factors affecting the prevalence of stunting) in each area. Results of the analysis showed that factors affecting the prevalence of stunting in all districts/cities in East Nusa Tenggara Province are the percentage of children aged under five who were weighed ≥ 4 times, the percentage of children aged under five who receive complete basic immunization, the percentage of households consuming iodized salt, the percentage of households with decent source of drinking water and the real per capita expenditure. The analysis showed that LCR-GWR is able to produce a better model than the GWR model in overcoming local multicollinearity problems in stunting in East Nusa Tenggara Province, with lower RMSE value (0.0344) than the GWR RMSE model (3.8899).

1. INTRODUCTION

Stunting is a condition in which children aged under five years fail to reach full potential for growth as a result of chronic malnutrition; their body height is below the standard height of their age. The results of the 2013 Baseline Health Research showed that East Nusa Tenggara was recorded as the province with the highest national prevalence of stunting among children aged under five years with a percentage of 51.7%. Results of the 2016 and 2017 Nutritional Status Survey by the Indonesian Ministry of Health showed that

East Nusa Tenggara Province was the province with the highest prevalence of stunting among children aged under five, with percentage of 41.2% and 40.3% respectively (Indonesian Ministry of Health, 2013, 2016, 2017).

The Geographically Weighted Regression (GWR) model may be the right option to study the factors affecting the prevalence of stunting because this approach has the ability to overcome spatial diversity/heterogeneity. However, due to several factors that are thought to have an effect on the prevalence of stunting, such as location of residence (geographic); maternal conditions; conditions of infants/children aged under five; household environmental conditions; clean living habits; quality of human resources; and economic level, which is very likely to be correlated or linearly related in each region, the use of the GWR model would be less effective. This is because GWR ignores any dependencies that could possibly occur on local regression coefficients between different predictor variables, which technically known as local multicollinearity (Páez et al., 2011; Wheeler, 2007, 2009; Wheeler & Calder, 2007; Wheeler & Tiefelsdorf, 2005).

In spatial regression, local multicollinearity can be overcome using the concept of the ridge regression method into GWR which is known as Geographically Weighted Ridge Regression (GWRR). The parameter estimation solution for the GWRR model is obtained using the Weighted Least Square (WLS) method, namely by giving different weights to each location and adding the coefficient $\lambda\boldsymbol{I}$ to the matrix $\boldsymbol{X}^{*T}\boldsymbol{W}(u_i, v_i)\boldsymbol{X}^*$, where λ is the magnitude of the bias coefficient of the parameter estimator located at the interval $0 < \lambda < 1$, $\boldsymbol{I}$ is the $k \times k$ identity matrix, and $\boldsymbol{X}^*$is the $\boldsymbol{X}$ matrix that has been centered-scaling.

The GWRR has its shortcoming: this model uses a bias coefficient, λ, for the entire observation area. In fact, not all observation areas may experience local multicollinearity problems. Adding a ridge parameter to the matrix $\boldsymbol{X}^{*T}\boldsymbol{W}(u_i, v_i)\boldsymbol{X}^*$ which in fact does not have a local multicollinearity problem between X variables can actually reduce the effectiveness of this model. Gollini et al. (2015) introduced a Locally Compensated Ridge-Geographically Weighted Regression (LCR-GWR) model that uses one bias coefficient for a certain area, which means that if there are N areas of observation, there are n different ridge bias coefficients. This method produces a ridge bias coefficient locally. The ridge parameter is allowed to vary across areas to adjust to the influence of collinearity between the predictor variables in each area so that it is expected that more accurate estimation solutions of the parameter coefficients in the model could be obtained (Fadliana et al., 2019).

This study specifically discusses the implementation of the LCR-GWR model to analyze the factors that affect the prevalence distribution of stunting among children aged under five in East Nusa Tenggara Province, which indicates local multicollinearity problems.

2. THEORETICAL REVIEW

2.1. *Geographically Weighted Regression* (GWR)

The Geographically Weighted Regression (GWR) model is a development of a global linear regression model with regard to regional or spatial aspects. In matrix notation, the GWR model can be written as

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta}(u_i, v_i) + \boldsymbol{\varepsilon} \tag{1}$$

The local parameter $\widehat{\boldsymbol{\beta}}(u_i, v_i)$ are estimated by Weighted Least Square (WLS), namely by giving different weights for each observation area

$$\widehat{\boldsymbol{\beta}}(u_i,v_i) = [\boldsymbol{X}^T\boldsymbol{W}(u_i,v_i)\boldsymbol{X}]^{-1}\boldsymbol{X}^T\boldsymbol{W}(u_i,v_i)\boldsymbol{y} \tag{2}$$

where $\boldsymbol{X} = [\boldsymbol{X}^T(1); \boldsymbol{X}^T(2); \dots \boldsymbol{X}^T(n)]^T$ is the design matrix of predictor variables, which typically includes a column of 1s for the intercept, $\boldsymbol{W}(u_i,v_i) = diag\big(w_1(u_i,v_i), w_2(u_i,v_i), \dots, w_n(u_i,v_i)\big)$ is the diagonal weights matrix that varies by calibration location i, $\boldsymbol{y}$ is the $n \times 1$ vector of response variables, and $\widehat{\boldsymbol{\beta}}(u_i,v_i) = \left(\hat{\beta}_0(u_i,v_i), \hat{\beta}_1(u_i,v_i), \dots, \hat{\beta}_p(u_i,v_i)\right)^T$ is the vector of $(p+1)$ local regression coefficient at location i for p predictor variables and an intercept term (Wheeler, 2009).

The weights matrix $\boldsymbol{W}(\boldsymbol{u_i},\boldsymbol{v_i})$, is calculated from Adaptive Gaussian Kernel function, given by

$$w_{ij} = exp\left[-\frac{1}{2}\left(\frac{d_{ij}}{h_i}\right)^2\right] \tag{3}$$

where d_{ij} is the euclidean distance between the calibration location i and location j, and h_i is referred to as the bandwidth. The search for the optimum bandwidth value is obtained through an iteration process by changing the h value until the minimum Cross Validation (CV) is obtained

$$CV = \sum_{i=1}^{n}\left(y_i - \hat{y}_{\neq i}(h)\right)^2 \tag{4}$$

where $\hat{y}_{\neq i}(h)$ is the fitted value of y_i with the observations for point i omitted from the calibration process (Fotheringham et al., 2002).

2.2. Local Multicollinearity

Local multicollinearity is defined as a condition where there is a perfect or nearly perfect linear relationship between predictor variables at each observation location. One measure that can be used to test for multicollinearity is Variance Inflation Factors (VIF). In GWR modeling, the VIF value is calculated using the following formula

$$VIF(u_i,v_i) = \frac{1}{1 - R_k^2(u_i,v_i)} \tag{5}$$

where $R_k^2(u_i,v_i)$ is local R^2 or determination coefficient between x_k other predictor variables for each location (u_i,v_i) (Wheeler, 2007).

According to Fotheringham et al. (2002), local R^2 is calculated using the following formula

$$R_k^2(u_i,v_i) = 1 - \frac{RSS^w}{TSS^w} \tag{6}$$

where TSS^w is the geographically weighted total sum of squares, defined as

$$TSS^w = \sum_j w_j(u_i,v_i)\left[y_j - \bar{y}\right]^2 \tag{7}$$

and RSS^w is the geographically weighted residual sum of squares, defined as

$$RSS^w = \sum_j w_j(u_i, v_i)\left[y_j - \hat{y}_j\right]^2 \tag{8}$$

2.3. Locally Compensated Ridge-Geographically Weighted Regression (LCR-GWR)

Locally Compensated Ridge-Geographically Weighted Regression (LCR-GWR) model is a development of the GWRR model using one bias coefficient for a particular region. That is, if there are N observation regions, then there are n different ridge bias coefficients. This method produces a ridge bias coefficient locally. The parameters of the ridge are left to vary in each region adjusting to the effect of collinearity between predictor variables in each region so that the expected parameter coefficients on the model could be more accurate (Fadliana et al., 2019).

The solution of parameter estimation for the LCR-GWR model is done using the WLS method on the GWR model by first centering on the y variable and centering-scaling on X variables, and then adding the coefficient $\lambda\boldsymbol{I}(u_i, v_i)$ which is the Locally Compensated (LC) value of λ in the (u_i, v_i) region. So that the estimator of the LCR-GWR model, $\widehat{\boldsymbol{\beta}}(u_i, v_i)$, is obtained at the specified value λ for each location as follows

$$\widehat{\boldsymbol{\beta}}(u_i, v_i, \lambda_i) = \left[\boldsymbol{X}^{*T}\boldsymbol{W}(u_i, v_i)\boldsymbol{X}^* + \lambda_i\boldsymbol{I}(u_i, v_i)\right]^{-1}\boldsymbol{X}^{*T}\boldsymbol{W}(u_i, v_i)\boldsymbol{y}^* \tag{9}$$

where, $\widehat{\boldsymbol{\beta}}(u_i, v_i, \lambda_i) = \begin{bmatrix} \hat{\beta}_0(u_0, v_0, \lambda_0) \\ \hat{\beta}_1(u_1, v_1, \lambda_1) \\ \hat{\beta}_2(u_2, v_2, \lambda_2) \\ \vdots \\ \hat{\beta}_p(u_p, v_p, \lambda_p) \end{bmatrix}$

The ridge bias coefficient in the LCR-GWR model is determined by the equation $= \left((\epsilon_1 - \epsilon_p)/(c-1)\right) - \epsilon_p$, which is obtained by connecting the eigenvalue and conditional number (c) of matrix multiplication $\boldsymbol{X}^T\boldsymbol{W}(u_i, v_i)\boldsymbol{X}$. Conditional number (c) is defined as ϵ_1/ϵ_p, where ϵ_1 is the largest eigenvalue and ϵ_p is the smallest eigenvalue (Gollini et al., 2015).

3. RESEARCH METHOD

3.1. Data

This study used secondary data sourced from the published results of Nutritional Status Survey by the Indonesian Ministry of Health in 2017 (Indonesian Ministry of Health, 2018) and the publication of Indonesian Central Bureau of Statistics in the form of People's Welfare Statistics which is based on the National Socio-Economic Survey in 2017 (East Nusa Tenggara Province Central Bureau of Statistics, 2017), poverty data for 2017 (Central Bureau of Statistics, 2018a), and publication of the Human Development Index (HDI) in 2017 (Central Bureau of Statistics, 2018b). The research area covered 21 districts and 1 city in East Nusa Tenggara Province.

3.2. Variables

The research variables used in this study are shown in Table 1.

Table 1. Research Variables

Variables	Units
Stunting among children under five (Y)	Percentage (%)
Pregnant women at chronic energy deficiency risk (X_1)	Percentage (%)
Pregnant women who received <90 blood booster tablets (X_2)	Percentage (%)
Low-birthweight babies (X_3)	Percentage (%)
Babies given exclusive breast milk (X_4)	Percentage (%)
Children under five who were weighed ≥4 times (X_5)	Percentage (%)
Children under five who got complete basic immunization (X_6)	Percentage (%)
Households consuming iodized salt (X_7)	Percentage (%)
Households with no defecation facilities (X_8)	Percentage (%)
Households with decent source of drinking water (X_9)	Percentage (%)
Human Development Index (HDI) (X_{10})	-
Real per capita expenditure (X_{11})	Thousands rupiah/year
Poor population (X_{12})	Percentage (%)
Easting coordinate of i district/city (u_i)	Meters (m)
Northing coordinate of i district/city (v_i)	Meters (m)

3.3. Analysis Steps

The data analysis techniques used in this study are as follows:

1. Mapping the characteristics of the districts/cities based on the prevalence of stunting.
2. Conducting multicollinearity testing.
3. Performing LCR-GWR modeling with the following steps:
   a. Calculating euclidean distances.
   b. Determining the optimum bandwidth with the Cross Validation (CV) approach (Equation 4).
   c. Determining the weighting matrix with Adaptive Gaussian Kernel function (Equation 3).
   d. Determining the ridge bias coefficient, λ, for each location.
   e. Estimating the parameters of LCR-GWR model (Equation 9).
   f. Testing the significant of the LCR-GWR model parameters.
   g. Interpreting dan making conclusion.

4. RESULTS AND DISCUSSION

4.1 General Characteristics of Districts/Cities in East Nusa Tenggara Province based on the Prevalence of Stunting among Children Aged Under Five

The prevalence of stunting among children under five can be used as a parameter of the nutritional status of children under five based on indicators of height for age (height/age) of a country or region. Figure 1 shows the prevalence distribution of stunting among children under five in districts/cities in East Nusa Tenggara in 2017.

4.2 Multicollinearity Testing

Multicollinearity examination in this study was carried out using the local Variance Inflation Factor (VIF) criteria (Equation 5). Predictor variables that have high local VIF values include HDI (X_{10}) and real per capita expenditure (X_{11}). The local VIF value of the HDI variable and real per capita expenditure for 22 districts/cities are greater than 10. This indicates that several predictor variables in this study present local multicollinearity problems. The presence of multicollinearity can make it possible to estimate the parameters of the GWR model but the resulting standard error tends to be large. Consequently, the population value of the coefficient cannot be estimated at the high level of precision or accuracy. Local multicollinearity in the weighted predictor variables can lead to estimates of GWR coefficients that are correlated locally and across space, have increased variability, and are sometimes counterintuitive and contradictory to global regression estimates (Czarnota et al., 2015; Tu et al., 2008; Wheeler, 2007).

Mapping of stunting prevalence data (Figure 1) shows there are 11 districts/cities that fall into the very high prevalence category, 8 districts/cities in the high prevalence category, and 3 districts/cities classified as medium prevalence. This mapping category refers to the prevalence cut-off values public health significance set by World Health Organization (WHO) (2010), namely: low prevalence (<20%), medium prevalence (20-29%), high prevalence (30-39), dan very high prevalence (≥40%).

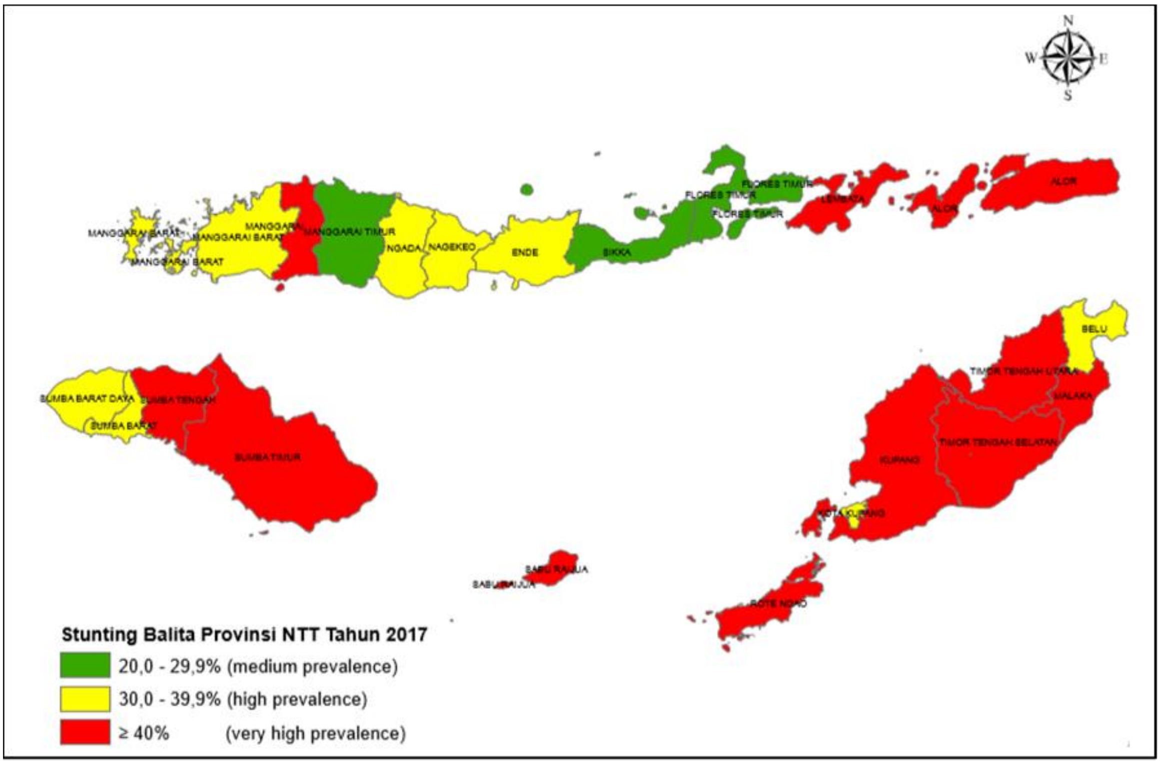

Figure 1. Prevalence Map of Stunting among Children Under-Five in East Nusa Tenggara in 2017

4.3 Locally Compensated Ridge-Geographically Weighted Regression (LCR-GWR) Modelling

The first step that must be taken in forming the LCR-GWR model is to construct a weighting matrix through a function that involves the euclidean distance component between districts/cities. The matrix weighting function used in this study is the Adaptive Gaussian Kernel weighting function (Equation 3). The Adaptive Gaussian Kernel weighting function requires a certain bandwidth value as the basis for determining the weight at each observation area. The determination of the optimal bandwidth is performed by an iteration process so that the minimum value of the Cross Validation (CV) is obtained (Equation 4).

By substituting the optimum bandwidth value and euclidean distance into the Adaptive Gaussian Kernel weighting function as in Equation 3, a weighting matrix $\boldsymbol{W}_i$ will be obtained. For example, Table 2 presents the weighted values for each district/city against Kupang City.

Table 2. Adaptive Gaussian Kernel Weighting against to Kupang City

District/City	Weighting Value ($W_{Kota\ Kupang}$)	District/City	Weighting Value ($W_{Kota\ Kupang}$)
West Sumba	0.49729	Ngada	0.70224
East Sumba	0.64712	Manggarai	0.60656
Kupang	0.99437	Rote Ndao	0.98066
South Central Timor	0.97099	West Manggarai	0.54649
North Central Timor	0.93962	Central Sumba	0.54147
Belu	0.89199	Southwest Sumba	0.46279
Alor	0.83811	Nagekeo	0.74138
Lembata	0.87811	East Manggarai	0.64955
East Flores	0.86166	Sabu Raijua	0.88412
Sikka	0.84815	Malaka	0.92192
Ende	0.79776	Kupang City	1.00000

The weighting matrix used in the LCR-GWR modelling is formed as a diagonal matrix as follows

$$\begin{aligned}\boldsymbol{W}_{Kupang\ City} &= diag\left[\boldsymbol{W}_{Kupang\ City(1)}, \boldsymbol{W}_{Kupang\ City(2)}, \dots, \boldsymbol{W}_{Kupang\ City(22)}\right] \\ &= diag[0.49729, 0.64712, \dots, 1]\end{aligned}$$

Next, the analysis step proceeds by determining the local ridge bias coefficient that varies across the observation locations, $\lambda_i(u_i, v_i)$.

This ridge bias coefficient is obtained by connecting the eigenvalue and conditional number (c) of matrix multiplication $\boldsymbol{X}^T\boldsymbol{W}(u_i, v_i)\boldsymbol{X}$. Conditional number (c) of matrix multiplication $\boldsymbol{X}^T\boldsymbol{W}(u_i, v_i)\boldsymbol{X}$ is defined as ratio of the largest eigenvalues to the smallest eigenvalues of the matrix $\boldsymbol{X}^T\boldsymbol{W}(u_i, v_i)\boldsymbol{X}$, defined as $\epsilon_1 + \lambda/\epsilon_p + \lambda$ where the eigenvalues of the matrix $(\boldsymbol{X}^T\boldsymbol{W}(u_i, v_i)\boldsymbol{X} + \lambda\boldsymbol{I})$ are $\epsilon_1 + \lambda, \epsilon_2 + \lambda, \dots, \epsilon_p + \lambda$. The ridge bias coefficient can be obtained based on the equation $\lambda = \left((\epsilon_1 - \epsilon_p)/(c-1)\right) - \epsilon_p$.

The LCR-GWR model is compatible with local ridge regression with their own ridge parameters (i.e., ridge parameters vary across observation areas), and only matches those ridge regressions in areas where the local conditional number is above the specified threshold defined by users. Thus, the addition of ridge bias coefficient is not used in all observation areas, but only in areas where multicollinearity tends to be a problem, so as to produce a more accurate model with the problem of spatial heterogeneity and local multicollinearity.

Furthermore, by using the Adaptive Gaussian Kernel weighted diagonal matrix formed by the optimum bandwidth which minimizes CV and by adding the coefficient $\lambda\boldsymbol{I}(u_i, v_i)$ which is the Locally Compensated (LC) value of λ in the region (u_i, v_i), the parameter estimator (coefficient) of the LCR-GWR model will be obtained for each district/city. The value of the bias ridge coefficient and the parameter coefficient of the LCR-GWR model for each district/city in the East Nusa Tenggara Province can be seen in the attachment.

The parameter estimation solution for the LCR-GWR model is then partially tested to show that the parameters have a significant or insignificant effect. Partial testing is done using the t-test statistic. If the statistical value of the test $|t| > t_{(0.0025)(22-12-1)} = 1.83311$, then it is decided that H_0 is rejected or the parameter has a significant effect. For example,

Table 3 shows the results of a partial test of the LCR-GWR model parameters for the Kupang City.

Table 3 shows that the predictor variables: X_5 (children under five were weighed ≥4 times), X_6 (children under five get complete basic immunization), X_7 (households consume iodized salt), X_9 (households with decent source of drinking water), and X_{11} (real per capita expenditure) in the LCR-GWR model with the Adaptive Gaussian Kernel weighting function have a significant effect on the response variable, Y, (stunting among children under-five) in Kupang City. Table 4 presents the predictor variables that have a significant effect on the prevalence of stunting among children under-five for each district/city of East Nusa Tenggara Province.

Table 3. The Results of a Partial Test of the LCR-GWR Model Parameters for the Kupang City

Parameters	**Estimation Value**	***Standard Error***	***t-Test Statistics***	**Decision**
β_1	0.25297	0.276396	0.915246	H_0 accepted
β_2	-0.11761	0.145111	-0.810448	H_0 accepted
β_3	-0.02400	0.327273	-0.073324	H_0 accepted
β_4	-0.07561	0.138325	-0.546607	H_0 accepted
β_5	-0.49167	0.250346	-1.963962	H_0 rejected
β_6	0.61039	0.257525	2.370219	H_0 rejected
β_7	-0.51755	0.106227	-4.872138	H_0 rejected
β_8	-0.04578	0.192207	-0.238163	H_0 accepted
β_9	0.40122	0.158493	2.531486	H_0 rejected
β_{10}	-0.85077	0.800955	-1.062191	H_0 accepted
β_{11}	0.51518	0.002794	184.412906	H_0 rejected
β_{12}	0.50102	0.416946	1.201636	H_0 accepted

Table 4. District/City Grouping based on Predictor Variables with Statistically Significant Effect

Variables	**Districts/Cities**
X_1	-
X_2	-
X_3	-
X_4	-
X_5	All districts/cities in East Nusa Tenggara Province
X_6	All districts/cities in East Nusa Tenggara Province
X_7	All districts/cities in East Nusa Tenggara Province
X_8	-
X_9	All districts/cities in East Nusa Tenggara Province
X_{10}	-
X_{11}	All districts/cities in East Nusa Tenggara Province
X_{12}	-

The estimation results of the LCR-GWR model parameters show that only the GWR coefficient β_8 has a variable estimation value which is in positive and negative numbers with the average estimated coefficient of GWR located in negative numbers. The LCR-GWR coefficients β_1, β_6, β_9, β_{11}, and β_{12} have the same variable estimation values for positive numbers. Whereas the other LCR-GWR coefficients, $\beta_2, \beta_3, \beta_4, \beta_5, \beta_7, \beta_7$, and β_{10} variations in the estimation values lie in negative numbers.

Based on the results of the LCR-GWR model, it can be concluded that the factors causing the high prevalence of stunting among children aged under five for all districts/cities in East Nusa Tenggara Province include: (1) the percentage of children aged under five who were weighed ≥4 times, (2) the percentage of children aged under five who get complete basic immunization, (3) the percentage of households consuming iodized salt, (4) the percentage of households with decent source of drinking water and (5) the real per capita expenditure. The higher the percentage of children aged under five who were weighed ≥4 times and the percentage of households consuming iodized salt, the lower the prevalence of stunting among children aged under five. In addition, the decreasing percentage of children aged under five who receive a complete basic immunization, the percentage of households with decent source of drinking water and real per capita expenditure will increase the prevalence of stunting among toddlers. Meanwhile, the remaining variables did not have a statistically significant effect on the prevalence of stunting among children aged under five in any district/city of East Nusa Tenggara Province.

Table 5. Minimum, Maximum, and Average Value of $\hat{\beta}(u_i, v_i)$ LCR-GWR Model

Parameters	Min	Max	Average
β_1	0.148212	0.257244	0.202414
β_2	-0.223639	-0.110505	-0.169719
β_3	-0.086204	-0.017222	-0.052270
β_4	-0.157071	-0.073620	-0.117766
β_5	-0.545428	-0.465231	-0.501763
β_6	0.517246	0.630520	0.578419
β_7	-0.528042	-0.501588	-0.520363
β_8	-0.045777	0.046066	-0.001780
β_9	0.394338	0.472187	0.438286
β_{10}	-0.881632	-0.835494	-0.863963
β_{11}	0.464781	0.541917	0.518352
β_{12}	0.492739	0.539530	0.510100

Root Mean Square Error (RMSE) can be used to demonstrate that the LCR-GWR model is able to solve the local multicollinearity problem in the GWR model. Outcome of the calculation results show that the RMSE value of the LCR-GWR model is 0.0344, which is lower than that of the RMSE model GWR (3.8899), so it can be said that the LCR-GWR model is better at overcoming local multicollinearity problems in case data of stunting in East Nusa Tenggara Province compared to the GWR model.

5. CONCLUSION

Based on results and discussion, it can be concluded that the LCR-GWR model with the Adaptive Gaussian Kernel weighting function that is formed shows that the percentage of children aged under five who were weighed ≥4 times, the percentage of children aged under five who get complete basic immunization, the percentage of households consuming iodized salt, the percentage of households with decent source of drinking water, and the real per capita expenditure have a statistically significant effect on the prevalence of stunting among children aged under five in all districts/cities in East Nusa Tenggara Province. The results of the LCR-GWR analysis with the Adaptive Gaussian Kernel weighting function are more effective or able to produce a better model than the GWR model in overcoming local multicollinearity problems in case data of stunting in East Nusa Tenggara Province, with a lower RMSE value (0.0344) compared to RMSE GWR model (3.8899).

REFERENCES

Central Bureau of Statistics. (2018a). *District/City Poverty Data and Information in 2017*. Jakarta: Central Bureau of Statistics.

Central Bureau of Statistics. (2018b). *Human Development Index in 2017*. Jakarta: Central Bureau of Statistics.

Czarnota, J., Wheeler, D. C., & Gennings, C. (2015). Evaluating Geographically Weighted Regression Models for Environmental Chemical Risk Analysis. *Cancer Informatics*, *14*, 117–127.

East Nusa Tenggara Province Central Bureau of Statistics. (2017). *East Nusa Tenggara Province People's Welfare Statistics in 2017*. Kupang: East Nusa Tenggara Province Central Bureau of Statistics.

Fadliana, A., Pramoedyo, H., & Fitriani, R. (2019). Parameter Estimation of Locally Compensated Ridge-Geographically Weighted Regression Model. *IOP Conference Series: Materials Science and Engineering*, *546*(5).

Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2002). *Geographically Weighted Regression: the Analysis of Spatially Varying Relationships*. Chichester, West Sussex: John Wiley & Sons Ltd.

Gollini, I., Lu, B., Charlton, M., Brunsdon, C., & Harris, P. (2015). GWmodel : an R Package for Exploring Spatial Heterogeneity. *Journal of Statistical Software*, *63*(17), 1–50.

Indonesian Ministry of Health. (2013). *Baseline Health Research in 2013*. Jakarta: Indonesian Ministry of Health.

Indonesian Ministry of Health. (2016). *Situtation of Stunted Under-Fives*. Jakarta: Indonesian Ministry of Health.

Indonesian Ministry of Health. (2017). *Results of Nutritional Status Survey and Its Explanation in 2016*. Jakarta: Indonesian Ministry of Health.

Indonesian Ministry of Health. (2018). *Pocket Book of Nutritional Status Survey in 2017*. Jakarta: Indonesian Ministry of Health.

Páez, A., Farber, S., & Wheeler, D. (2011). A Simulation-Based Study of Geographically Weighted Regression as A Method for Investigating Spatially Varying Relationships. *Environment and Planning A*, *43*(12), 2992–3010.

Tu, Y. K., Gunnell, D., & Gilthorpe, M. S. (2008). Simpson's Paradox, Lord's Paradox, and Suppression Effects are the Same Phenomenon - The Reversal Paradox. *Emerging Themes in Epidemiology*, *5*, 1–9.

Wheeler, D. C. (2007). Diagnostic Tools and A Remedial Method for Collinearity in Geographically Weighted Regression. *Environment and Planning A*, *39*(10), 2464–2481.

Wheeler, D. C. (2009). Simultaneous Coefficient Penalization and Model Selection in Geographically Weighted Regression: The Geographically Weighted Lasso. *Environment and Planning A*, *41*(3), 722–742.

Wheeler, D. C., & Calder, C. A. (2007). An Assessment of Coefficient Accuracy in Linear Regression Models with Spatially Varying Coefficients. *Journal of Geographical Systems*, *9*(2), 145–166.

Wheeler, D. C., & Tiefelsdorf, M. (2005). Multicollinearity and Correlation among Local Regression Coefficients in Geographically Weighted Regression. *Journal of Geographical Systems*, *7*(2), 161–187.

World Health Organization (WHO). (2010). *Nutrition Landacape Information System (NLIS) Country Profile Indicators: Interpretation Guide*. Geneva: WHO Press.

ATTACHMENT

Ridge Bias Coefficient Value and LCR-GWR Model Parameter Coefficient

No	District/City	λ	β_0	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}	β_{11}	β_{12}
1	West Sumba	0.0338	-0.0275	0.1544	-0.2143	-0.0833	-0.1469	-0.4780	0.5172	-0.5188	0.0422	0.4619	-0.8807	0.5411	0.4963
2	East Sumba	0.0364	-0.0271	0.1632	-0.2048	-0.0733	-0.1458	-0.4778	0.5192	-0.5152	0.0351	0.4573	-0.8752	0.5349	0.4973
3	Kupang	0.0520	0.0143	0.2502	-0.1215	-0.0275	-0.0757	-0.4975	0.6161	-0.5204	-0.0450	0.4044	-0.8557	0.5182	0.5043
4	South Central Timor	0.0500	0.0157	0.2482	-0.1241	-0.0315	-0.0736	-0.4973	0.6181	-0.5226	-0.0440	0.4048	-0.8614	0.5254	0.5039
5	North Central Timor	0.0479	0.0146	0.2445	-0.1289	-0.0347	-0.0760	-0.5028	0.6221	-0.5247	-0.0418	0.4088	-0.8644	0.5263	0.5074
6	Belu	0.0464	0.0146	0.2421	-0.1317	-0.0375	-0.0764	-0.5030	0.6226	-0.5259	-0.0400	0.4100	-0.8673	0.5298	0.5076
7	Alor	0.0436	0.0104	0.2348	-0.1407	-0.0414	-0.0845	-0.5142	0.6281	-0.5280	-0.0347	0.4188	-0.8681	0.5249	0.5153
8	Lembata	0.0439	0.0043	0.2314	-0.1449	-0.0387	-0.0948	-0.5254	0.6305	-0.5265	-0.0315	0.4264	-0.8603	0.5090	0.5231
9	East Flores	0.0433	-0.0036	0.2241	-0.1529	-0.0382	-0.1089	-0.5352	0.6289	-0.5246	-0.0240	0.4360	-0.8524	0.4933	0.5305
10	Sikka	0.0441	-0.0186	0.2130	-0.1636	-0.0338	-0.1339	-0.5454	0.6160	-0.5174	-0.0109	0.4500	-0.8355	0.4648	0.5395
11	Ende	0.0384	-0.0308	0.1830	-0.1920	-0.0517	-0.1533	-0.5323	0.5802	-0.5183	0.0164	0.4640	-0.8525	0.4862	0.5327
12	Ngada	0.0336	-0.0336	0.1593	-0.2139	-0.0732	-0.1568	-0.5121	0.5487	-0.5219	0.0364	0.4703	-0.8726	0.5166	0.5187
13	Manggarai	0.0320	-0.0333	0.1505	-0.2220	-0.0829	-0.1564	-0.5033	0.5369	-0.5235	0.0439	0.4722	-0.8793	0.5278	0.5123
14	Rote Ndao	0.0587	0.0133	0.2572	-0.1105	-0.0172	-0.0775	-0.4757	0.5941	-0.5104	-0.0453	0.3943	-0.8388	0.5087	0.4927
15	West Manggarai	0.0317	-0.0324	0.1482	-0.2236	-0.0862	-0.1548	-0.4974	0.5311	-0.5235	0.0461	0.4715	-0.8816	0.5330	0.5082
16	Central Sumba	0.0340	-0.0284	0.1546	-0.2143	-0.0820	-0.1482	-0.4807	0.5188	-0.5187	0.0418	0.4628	-0.8800	0.5388	0.4981
17	Southwest Sumba	0.0333	-0.0276	0.1530	-0.2160	-0.0850	-0.1472	-0.4786	0.5177	-0.5196	0.0433	0.4628	-0.8816	0.5419	0.4965
18	Nagekeo	0.0349	-0.0331	0.1659	-0.2078	-0.0668	-0.1563	-0.5179	0.5575	-0.5209	0.0309	0.4686	-0.8674	0.5085	0.5229
19	East Manggarai	0.0325	-0.0338	0.1534	-0.2195	-0.0794	-0.1571	-0.5080	0.5423	-0.5232	0.0413	0.4721	-0.8769	0.5230	0.5155
20	Sabu Raijua	0.0526	-0.0090	0.2244	-0.1406	-0.0262	-0.1169	-0.4652	0.5484	-0.5016	-0.0120	0.4169	-0.8390	0.5067	0.4931
21	Malaka	0.0478	0.0156	0.2448	-0.1283	-0.0355	-0.0743	-0.4993	0.6202	-0.5247	-0.0418	0.4071	-0.8657	0.5298	0.5052
22	Kupang City	0.0542	0.0143	0.2530	-0.1176	-0.0240	-0.0756	-0.4917	0.6104	-0.5176	-0.0458	0.4012	-0.8508	0.5152	0.5010