Modeling of Infant Mortality Rate in East Java Province Using Mixed Geographically Weighted Regression Approach for Improving Quality of Health Services

Ninda Ayu Puspitasari[1], Nadia Murbarani[1], Nopiyanti[1], Sartika Aprilia[1], Nur Chamidah[2]
[1]*Student of Program Study of Statistics, Department of Mathematics, Universitas Airlangga, Surabaya, Indonesia*
[2]*Department of Mathematics, Faculty of Sciences and Technology, Universitas Airlangga, Surabaya, Indonesia*

Keywords: East Java Province , Infant Mortality Rate, Mixed Geographically Weighted Regression

Abstract: Infant mortality is defined as the death of a baby aged less than one year old. Sustainable Development Goals (SDG) program is a sustainable development program in which there are 17 goals and 169 measurable targets with specified deadlines. One of the targets to be achieved by the year 2030 is to be able to reduce infant mortality at least up to 12 every 1,000 live births. Infant Mortality Rate in East Java Province still shows a high number of 5,196 babies die every year. It shows that as many as 13 babies die every day. The mixed geographically weighted regression (MGWR) model is a combination between global regression and local regression models where some predictor variables affect globally and the others affect locally to the response. This study aims to select the best model based on the smallest AICc (Akaike Information Criterion corrected) value using fixed Gaussian weighted. The variable that affect globally is percentage of pregnant mothers visit. While other predictor variables that affect locally are percentage of Integrated Health Services (*Posyandu*), percentage of number of households with decent sanitation, and percentage of low birth weight infants.

1 INTRODUCTION

Infant mortality is the death of a baby aged less than one year old. Infant mortality is measured as the infant mortality rate, which is the number of child deaths under one year per 1000 births in one year (Central Bureau of Statistics, 2012). SDGs is a sustainable development program in which there are 17 goals with 169 measurable targets with specified deadlines. SDGs is a world development agenda aimed at human and planetary welfare. One of the goals of SDGs is good health, which ensures a healthy life and promotes well-being for all people of all ages. The goal has 13 targets, one of which is by 2030 ending preventable infant and toddler deaths with all countries trying to reduce infant mortality by at least 12 per 1,000 live births (Annisa, 2013). The condition of Infant Mortality Rate (IMR) and Neonatal Mortality Rate in East Java Province is relatively small, however, the absolute mortality rate still shows a high number of 5,196 toddlers per year. It shows that as many as 13 babies died and 14 toddlers die every day. The Mixed Geographically Weighted Regression (MGWR) model is a combination model of global regression with Global Weighted Regression (GWR) considering the situation where some predictor variables affecting the response are global and other predictor variables are localized according to the location of the data observation (Asih et al, 2013).

2 LITERATURE REVIEW

2.1. Infant Mortality Rate

Infant Mortality Rate (IMR) describes the number of infant mortality of less than one year per 1000 live births in a given year. IMR is one indicator of health development successes that has been declared in National Health System and even used as a central indicator of success of health development in Indonesia (Azizah, 2013). Besides, IMR also reflects the level of maternal health, environmental health, and general level of socio-economic development of the community because the IMR is very sensitive to

Puspitasari, N., Murbani, N., Nopriyanti, ., Aprilia, S. and Chamidah, N.
Modeling of Infant Mortality Rate in East Java Province Using Mixed Geographically Weighted Regression Approach for Improving Quality of Health Services.
DOI: 10.5220/0007546105150519
In *Proceedings of the 2nd International Conference Postgraduate School (ICPS 2018)*, pages 515-519
ISBN: 978-989-758-348-3

changes in the level of health and welfare of the community.

2.2. Mixed Geographically Weighted Regression Model

Mixed Geographically Weighted Regression (MGWR) is a combination of global linear regression model with GWR model. Thus, the MGWR model will produce parameter estimators that are partly global and some others are localized according to the location of observation (Fotheringham et al 2002). The parameter estimation on MGWR model can be done by WLS (Weighted LeastSquare) method as in GWR model (May et al, 2004). The general MGWR model can be written in the form of:

$$y_i = \sum_{k=1}^{p} \beta_g x_{ig} + \sum_{i=1}^{p_l} \beta_i(u_i, v_i) x_{il} + \varepsilon_i; i = 1,2,\dots,n$$

……………………………………………………………(1)

Estimation of MGWR model parameters with Weighted Least Square (WLS) approach. Estimation of MGWR model parameters is as follows:

$$\boldsymbol{y} = \boldsymbol{X}_l \boldsymbol{\beta}_l(u_i, v_i) + \boldsymbol{X}_g \boldsymbol{\beta}_g + \boldsymbol{\varepsilon} \text{ …………………..(2)}$$

given that $\boldsymbol{X}_g$ is the global predictor variable matrix, $\boldsymbol{X}_l$ is the local predictor variable matrix, β_g is the global predictor variable parameter vector and $\boldsymbol{\beta}_l(u_i, v_i)$ is the local predictor variable parameter matrix.

2.3. Test The Suitable of Model MGWR

Test the suitability of the MGWR model is aimed to find out which model is suitable or not. The hypothesis is as follows:

$$H_0 : \beta_q(u_i, v_i) = \beta_k \text{……………………………(3)}$$

Sustainability test of the MGWR model is aimed to find out which model is suitable or not. The hypothesis is as follows:

$$\mathrm{F(1)} = \frac{\mathrm{y^T[(I-H)-(I-S)^T(I-S)]y/v_1}}{\mathrm{y^T(I-S)^T(I-S)y/u1}} \text{ …………………(4)}$$

2.4. Concurrent Test of Global Parameter

The simultaneous hypothesis test is intended to know the significance of predictor variables in the MGWR model simultaneously on the parameter of global predictor variable $x_k(q+1 \le k \le p)$. The collision hypothesis in this test is:

$$H_0 : \beta_{q+1} = \beta_{q+2} = \cdots = \beta_p = 0 \text{ ……………..(5)}$$

Statistical test used in the simultaneous test of global parameters is as follows:

$$F(2) = \frac{y^T[(I-S_l)^T(I-S_l)-(I-S)^T(I-S)]y/r1}{y^T(I-S)^T(I-S)y/u1} \text{ …….(6)}$$

2.5. Concurrent Test of Local Parameters $\mathbf{x_k}(1 \le \mathbf{k} \le \mathbf{q})$

In the simultaneous test, this local parameter uses the following hypothesis:

$$H_0 : \beta_1(u_i, v_i) = \beta_2(u_i, v_i) = \cdots = \beta_q(u_i, v_i) = 0$$

…………………………………………………..(7)

statistical test used in the simultaneous test of local parameters is as follows:

$$F(3) = \frac{y^T\left[(I-S_g)^T(I-S_g)-(I-S)^T(I-S)\right]y/t_1}{y^T(I-S)^T(I-S)y/u1} \text{ …………..(8)}$$

2.6. Partial Test of Global Parameters $\boldsymbol{x_k}(\boldsymbol{q+1 \le k \le p})$

In partial test, this local parameter uses hypothesis as follows:

$$H_0 : \beta_k = 0 \text{…………………………………….(9)}$$

statistical test used in the partial test of this global parameter is:

$$T_g = \frac{\hat{\beta}_k}{\hat{\sigma}\sqrt{g_{kk}}} \text{…………………………….............. (10)}$$

2.7. Partial Test of Local Parameters

The hypothesis in this local partial test is as follows:

$$H_0 : \beta_k(u_i, v_i) = 0 \text{ ………………………….…(11)}$$

statistical test and partial test of local parameters can be calculated using the following formula:

$$T_i = \frac{\hat{\beta}_k(u_i, v_i)}{\hat{\sigma}\sqrt{m_{kk}}} \text{ ………………………………………..(12)}$$

3 MATERIAL AND METHODS

3.1. Source of Data

The data used in this study is secondary data obtained from the publication Health Profile of East Java Province 2016. The observation unit used is 38 administrative areas in East Java Province consisting of 29 districts and 9 cities.

3.2. Research Variable

Variables used in this study include 8 variables consisting of one response variable and 7 predictor variables as follows:

Y : Infant Mortality Rate
X_1 : Percentage of Infants Gaining Exclusive Breast Milk
X_2 : Percentage of Integrated Health Services (*Posyandu*)
X_3 : Percentage of Pregnant Mothers Visit
X_4 : Percentage of Infants Gaining Health Services
X_5 : Percentage of Households Gaining Clean Drinking Water
X_6 : Percentage of Number of Households with Decent Sanitation
X_7 : Percentage of Low Birth Weight Infants

4 RESULT AND DISCUSSION

4.1 Modeling of Infant Mortality Rate in East Java with MGWR

4.1.1 Classic Assumption Test

Classic assumption test consists of 4 subjects namely normality testing, multicollinearity testing, heterogeneity testing, and spatial dependency testing.

Based on the results obtained, the data used meet the classical assumption

4.1.2 Geographically Weighted Regression Modeling (GWR) Modeling

The selection of the kernel function weights used was conducted by doing a comparison of the smallest AICc values based on the four kernel function weights (Fixed Gaussian, Fixed Bisquare, Adaptive Gaussian and Adaptive Bisquare) to obtain the best weights. The best weights were used to estimate the parameters in each regency / city in East Java Province.

Table 1 : Value of AICc and R2 GWR model

Kernel Weighted Funciton	R^2	*Minimum AICc*
Fixed Gaussian	0,47061	256,523715

Based on Table 4.1, the best kernel-weighted function is Fixed Gaussian which has the smallest AICc value that is 256,523715 so that Fixed Gaussian kernel weighting function was used to estimate the best GWR model in this research.

4.1.3 Mixed Geographically weighted Regression (MGWR) Modeling

The determination of the kernel function weights is do the smallest Minimum AICc value comparison based on the three kernel functional weights that are Fixed Gaussian, Fixed Bisquare, and Adaptive Bisquare to obtain the best functional weights.

Tabel 2 : Value of AICc dan R^2 MGWR model

Kernel Weighted Function	R^2	*Minimum AICc*
Fixed Gaussian	0,572701	254,705810

Based on Table 4.2 above, it can be seen that the best Kernel functional weights is Fixed Gaussian because it has the smallest minimum AICc value. It is equal to 254,705810, with R2 equals to 57,27%.

After obtaining the best Kernel weighted functionality on the GWR model and the MGWR model, the next step was to compare the best model of both models from the minimal AICc value.

Table 3 : Comparison of AICc value GWR and MGWR model

Model	*AICc* minimum
GWR	256,523715
MGWR	254,705810

Based on Table 4.5, the AICc value of the MGWR model is smaller than the AICc GWR model value of 254,705810. Thus, the MGWR model is the most appropriate for modeling infant mortality rate in East Java. Based on testing using GWR 4.0 software obtained the test results of conformity MGWR model.

Table 4: MGWR conformity model test results

Source	DF	SS	MS	F
Global Residual	30	24664,70		
GWR Improvement	2,290	8503,542	3713,337	
GWR Residual	27,710	16161,16	583,225	6,368

Table 4.4 shows that the F value of 6.3688 > F _ ((30,27,710; 0,05)) is 1.88. Then, the decision taken was Reject so that it can be concluded that the obtained MGWR spatial regression model is appropriate. Once the most appropriate model was obtained, the next step was to examine the global parameter significance using GWR 4.0. software.

Table 5 : Global parameter estimation results

Variable	Estimation	Standard Residual	*t*
X_3	-1,533768	0,69718	-2,19996

Based on Table 4.5, the global parameter X_3 has significant effect on infant mortality rate in East Java since the value of ❘t_countage = big ❘ -2,19996 ❘ > 2,04227. Local parameter testing of the spatial regression MGWR model is partially performed at each i-location. This test was conducted to determine the regression parameters that affect infant mortality rate in each district/city in East Java. For example, we will analyze the estimation model in areas with the highest infant mortality rate namely Probolinggo and the lowest infant mortality rate, Situbondo. The MGWR model for Probolinggo City is:

$$Y = 359{,}517 - 3{,}069231X_3 - 3.17522X_6$$

MGWR model for Situbondo Regency are:

$$Y = 359{,}401 - 1{,}70053X_2 - 1{,}533768X_3 - 3{,}17469\ X_6 + 1{,}83614\ X_7$$

5 CONCLUSION

Based on the analysis and discussion that has been performed, the conclusions as follows:

1. The best model obtained for Infant Mortality Rate in East Java 2016 is Mixed Geographically Weighted Regression Model (MGWR) with Gaussian Kernel weighted function. The MGWR model can explain the diversity of response variables of 57.27% with an AICc value of 254,705810.
2. Globally affecting factors of Infant Mortality Rate in East Java Province is the Percentage of Pregnant Mothers Visit. While the factors that influence locally is the Percentage of Integrated Health Services (*Posyandu*), Percentage of Number of Households with Decent Sanitation and Percentage of Low Birth Weight Infants.

REFERENCES

Azizah, L N. 2013. Geographically Weighted Regression (GWR) Significance Testing with Test F statistics and Test t. Thesis not published. Malang: UIN Malik Ibrahim Malang.

Febrina, Ika et al. 2011. Modeling of Mixed Geographically Weighted Regression (MGWR) In Underfives Mortality Rate in Bojonegoro Regency in 2011. Thesis. Surabaya: Ten November Institute of Technology.

Fortheringham, A.S., Brunsdon C, Chartlon M., 2002. Geographically Weighted Regression: the analysis of spatially varying relationship. Chichester: Willey.

Diah, Widyatun. 2012. Jurnal Bidan Diah. *http://jurnalbidandiah.blogspot.co.id/2012/06/*

pelayanan-kesehatan-pada-bayi-and.html. Retrieved May 26, 2018.

East Java Provincial Health Office. 2016. Health Profile of East Java Province 2016. Surabaya: East Java Provincial Health Office.

Sulistyawati, A. 2009. Midwifery Care During Pregnancy. Jakarta: Salemba Medika.

Yasin, H. 2011. Variable Selection in Geographically Weighted Regression Model. Journal Statistics. *http://eprints.undip.ac.id/33669/1/3/Hasbi.pdf*. Retrieved April 18, 2018.

Yasin, H. 2013. Hypothesis Test Model Mixed Geographically Weighted Regression, Journal Statistics *http://eprints.undip.ac.id/40352/1/D03_Hasbi_Yasin.pdf*. Accessed April 18, 2018.

Mahmuda. 2012. Partial Test Statistics on Mixed Geographically Weighted Regression Model. Malang: UIN Maulana Malik Ibrahim Malang.