The Factors Affecting Maternal Mortality in Java: Spatial Modeling Application on Demographic Data

C H P Tatipatta¹, Y Widyantingsih² and S Nurrohmah³

¹,²,³Departement of Mathematics, Universitas Indonesia, KampusBaru UI, Depok, 16424, Jawa Barat, Indonesia.

Abstract. Maternal deaths are considered very serious issue worldwide, according to World Health Organization (WHO) everyday approximately 830 women die from preventable causes related to pregnancy and childbirth. According to Indonesia’s health profile in 2014, West Java, Central Java and East Java provinces contributed 26 percent of the total number of maternal deaths in Indonesia. The objective of this study is to model maternal deaths in West Java, Central Java, and East Java using Spatial Autoregressive Model (SAM) and Spatial Error Model (SEM), inspect the factors affecting maternal deaths, and compare the models. The results showed that number of health personnel, the number of residents received Raskin (BerasMiskin), and number of illiterate residents were significant factors that affect maternal deaths. Based on the result of the Moran’s I test, spatial autocorrelation in the residual of linear regression model, and the response variable (maternal deaths) are significant. The likelihood ratio test showed that the spatial regression models were better than multiple linear regression model in predicting the maternal deaths. The AIC value of the OLS model is 243.1, SAM is 240.3, and SEM is 234.4. Based on the AIC values, SEM is better than the SAM in predicting maternal deaths.

1. Introduction
Maternal health problems occurring in a country are very serious, as maternal health is an important indicator in the development of the health sector to describe the state of health in a region. One of most important maternal health indicators are maternal mortality ratio. One of the Millennium Development Goals (MDGs) of Indonesia is to reduce maternal mortality ratio from 446 deaths per one hundred thousand live births to 110 deaths per one hundred thousand live births during the period 1990 to 2015.

The maternal mortality ratio of Indonesia has not reached the target of the established MDGs, where the ratio is 133 deaths per one hundred thousand live births and the maternal mortality ratio is higher compared to the Philippines, 117 deaths per one hundred thousand live births, Vietnam is at 54 deaths per one hundred thousand births life, Malaysia is at 41 deaths per one hundred thousand live births, Brunei Darussalam is at 23 deaths per hundred thousand live births, Thailand is at 21 deaths per one hundred thousand live births, and Singapore is at the ratio of 10 maternal deaths per one hundred thousand live births. That puts Indonesia as the country with the fourth highest maternal mortality ratio in Southeast Asia.

Maternal mortality ratio depends on the number of maternal death. Maternal death is the death of a woman while pregnant or within 42 days of termination of pregnancy, irrespective of the duration and site of the pregnancy, from any cause related to or aggravated by the pregnancy or its management but not from accidental or incidental causes. To facilitate the identification of maternal deaths in circumstances in which cause of death attribution is inadequate, a new category has been introduced:
Pregnancy-related death is defined as the death of a woman while pregnant or within 42 days of termination of pregnancy, irrespective of the cause of death. According Indonesia’s health profile in 2014, West Java, Central Java, and East Java contribute 26 percent of maternal deaths in Indonesia.

Based on literature study, there are 5 variables suspected to affect maternal deaths. The first variable is the number of residents received Raskin (Beras Miskin) which is involved based on [5], that nutritional support provided by the government for the poor affects the number of maternal deaths occurring in an area. The second variable is the proportion of people that have access to clean water [4], that the availability of clean water affects the number of maternal deaths in an area. The third and fourth variables are number of health personnel and the number of births without health personnel [9], that the number of health personnel and the number of births without health personnel affect the number of maternal deaths in an area. The fifth variable is the number of illiterate residents [8], that the number of illiterate residents affects the number of maternal deaths in an area. Therefore, the objective of this study is to modelling maternal deaths using spatial autoregressive model and spatial error model, inspect the factors affecting maternal death, and determining the better model based on AIC values.

2. Study Area and Methods

2.1 Study Area
The study area is located in West Java, Central Java, and East Java, Indonesia with approximately 117,775km² (Figure 1). Total number of pregnant woman in the study area is approximately 2,225,662.

![Figure 1. Study Area West Java, Central Java, and East Java](image)

2.2 Methods
In this study, multiple linear regression is used to describe the relationship between the dependent and explanatory variable. There are four major assumption for regression analysis, the error term has zero mean, the error term has constant variance $\sigma^2$, the error are uncorrelated and the errors are normally distributed. When one or more of those major assumption are violated, we need to transform the dependent variables or explanatory variables. If the error is still correlated after the transformation, it is suspected that spatial relationship is exist. The spatial relationships between the dependent and the error terms were analyzed (modelled) and spdep library in R Software v.3.3.3 and Geoda v1.12.1.59 are used for computation. The output of the analysis is a single regression equation describing the relationship between the dependent and explanatory variables across the whole study area Montgomery [7]. The general model of multiple linear regression is:

$$ y = X\beta + \varepsilon $$

where $y$ is an $n \times 1$ vector of dependent (response) variable, $X$ is an $n \times (k + 1)$ matrix of explanatory variables, $\beta$ is an $(k + 1) \times 1$ vector of regression parameter, $\varepsilon$ is an $n \times 1$ vector of error
term, \( k \) is the number of explanatory variable, and \( n \) is the number of observations. The regression model in a matrix form is:

\[
\begin{bmatrix}
  y_1 \\
  y_2 \\
  \vdots \\
  y_n
\end{bmatrix} = 
\begin{bmatrix}
  1 & x_{11} & \ldots & x_{1k} \\
  1 & x_{21} & \ldots & x_{2k} \\
  \vdots & \vdots & \ddots & \vdots \\
  1 & x_{nk} & \ldots & x_{nk}
\end{bmatrix} 
\begin{bmatrix}
  \beta_0 \\
  \beta_1 \\
  \vdots \\
  \beta_k
\end{bmatrix} + 
\begin{bmatrix}
  \epsilon_1 \\
  \epsilon_2 \\
  \vdots \\
  \epsilon_n
\end{bmatrix}
\]

The parameter of \( \beta \) were estimated using Ordinary Least Square (OLS) method. OLS is a method to find the estimator of parameters of the linear regression by minimizing the sum of square error (SSE) of the linear regression model. The least square function is written as:

\[S(\beta) = \epsilon'\epsilon\]

by substituting \( \epsilon = (y - X\beta) \) to equation (1), obtained

\[S(\beta) = (y - X\beta)'(y - X\beta)\]

\[= y'y - 2\beta'X'y + \beta'XX\beta\]

and minimizing the least square function by using first partial derivative of \( \beta \),

\[
\frac{\partial S(\beta)}{\partial \beta} = 0
\]

\[-2X'y + 2XX\hat{\beta} = 0\]

\[XX\hat{\beta} = X'y\]

\[\hat{\beta} = (X'X)^{-1}X'y\]

so, the estimator of \( \beta \) is:

\[\hat{\beta} = (X'X)^{-1}X'y\]

The response variable have a tendency to spatially auto correlated where the value of a variable in a location correlated to the value of the variable in the location around it, and the problem that usually arise from this condition is the error term of the linear regression is auto correlated. Therefore, spatial regression models were performed to improve the prediction of maternal death. Starting with a standard linear regression model, two spatial models can be distinguished, which is Spatial Autoregressive Model (SAR) and Spatial Error Model (SEM).

The general model for Spatial Autoregressive is

\[y = \rho Wy + X\beta + \epsilon\]

\[\epsilon \sim \text{NIID}(0, \sigma^2 I_n)\]

and the general model for Spatial Error Model is

\[y = X\beta + u\]

\[u = \lambda Wu + \epsilon\]

\[\epsilon \sim \text{NIID}(0, \sigma^2 I_n)\]

Where \( y \) denotes an \( n \times 1 \) vector of dependent variable, \( X \) is an \( n \times p \) matrix of independent variable, \( \epsilon \) is a normally distributed error, \( \beta \) is a \( p \times 1 \) vector of regression parameter, \( \rho \) is the coefficient of spatial lag, \( \lambda \) is the coefficient of spatial autocorrelation in error term, \( Wy \) denotes the endogenous interaction effects among the response variable, and \( Wu \) denotes interaction effects among the error term.

3. Results and Discussion
The observed maternal death, proportion of people who have access to clean water, the number of health personnel, the number of resident who received Raskin, the number of illiterate residents and the number of births without health personnel in West Java, Central Java, and East Java are shown in Figure 2.

Figure 2. Map of Maternal Death and Explanatory Variables

The observed value of maternal death in West Java, Central Java, and East Java are vary from 1 to 73 with a mean of 21. Most of the area in the west and central area of Java have maternal death above the mean. The number of births without health personnel varies from 0 to 12900 with a mean of 982. Based on the map, the areas which have more number of births without health personnel have high maternal death. There are 14 Districts/City that have only 36% – 55% people have access to clean water, while only 25 Districts/City that have 55% – 66.7% people have access to clean water, 20 Districts/City that have 66.7% – 75.7% people have access to clean water, 20 Districts/City have 75.7% – 85% people have access to clean water, and 16 Districts/City have 75.7% – 85% people have access to clean water (Figure 2). The map of the other land use variable can be seen in Figure 2. To determine the explanatory variables that have significant effect to the maternal death ($Y$), OLS model were used. The explanatory the number of health personnel ($HP$), the number of residents received RASKIN ($R$), proportion of people have access to clean water ($CW$), the number of births without health personnel ($NHP$), and the number of illiterate residents ($IR$). The proposed model is:

$$Y = \beta_0 + \beta_1 HP + \beta_2 R + \beta_3 CW + \beta_4 NHP + \beta_5 IR + \epsilon,$$

Table 1. Estimated Values and the Result of Residual Analysis

| Variable | Estimation | SD     | t-value | p-value | VIF  |
|----------|------------|--------|---------|---------|------|
| (Intercept) | 1.628 | 7.977 | 0.204 | 0.83885 |      |
| $HP$ | $1.387 \times 10^{-2}$ | $6.282 \times 10^{-4}$ | 2.207 | 0.03038* | 1.277 |
| $R$ | $1.555 \times 10^{-4}$ | $2.119 \times 10^{-6}$ | 7.341 | $2.27 \times 10^{-10}$*** | 2.049 |
| $CW$ | $-1.041$ | 9.605 | $-0.108$ | 0.91402 | 1.206 |
The F-statistic on Table 1 show that the model is sufficient to predict maternal death (p-value = 6.477 × 10^{-14} < 0.05). The explanatory variable that significantly affect the maternal death at level 0.05 are the number of health personnel, the number of residents received Raskin, and the number of illiterate people. VIF test were performed to analyze the multicollinearity, and the result is all of the explanatory variables chosen have no multicollinearity problem (VIFs value < 10). Shapiro-Wilk, Breusch-Pagan, and Durbin-Watson test was performed to analyze the normality, homogeneity, and autocorrelation of the residual of the regression model. It is found that the residual have a heteroskedasticity problem (p-value = 0.001 < 0.05) and auto correlated (p-value = 0.007522 < 0.005).

Since the residual model is not normally distributed, one way to handle the misspecification of the model is to transform one or more explanatory variables or the dependent (response) variable using square root, log, arcsin, or inverse transformation. The best possible transformation is to transform the dependent variable (Y) using root transformation. The proposed model is:

\[ \sqrt{Y} = \beta_0 + \beta_1 HP + \beta_2 R + \beta_3 CW + \beta_4 NHP + \beta_5 IR + \epsilon. \]

| Variable          | Estimation | SD       | t-value | p-value | VIF |
|-------------------|------------|----------|---------|---------|-----|
| (Intercept)       | 2.815      | 8.028 × 10^{-1} | 3.507   | 0.000775 |     |
| HP                | 1.639 × 10^{-3} | 6.322 × 10^{-4} | 2.593   | 0.011474* | 1.26 |
| R                 | 1.427 × 10^{-5} | 2.133 × 10^{-6} | 6.691   | 3.68 × 10^{-9}*** | 1.96 |
| CW                | −6.666 × 10^{-1} | 9.667 × 10^{-1} | −0.690  | 0.492653 | 1.20 |
| NHP               | −8.494 × 10^{-6} | 7.664 × 10^{-5} | −0.111  | 0.912043 | 1.36 |
| IR                | −3.331 × 10^{-5} | 1.427 × 10^{-5} | −2.334  | 0.022307* | 1.26 |
| F-statistics      | 21.47      |          | 3.287 × 10^{-13} |     |     |
| AIC               | 243.135    |           |         |         |     |
| Adjusted R-Squared| 0.5644    |           |         |         |     |
| Log Likelihood    | −116.9259  |           |         |         |     |
| Breusch-Pagan Test| 8.4468    |           | 0.133   |         |     |
| Shapiro-Wilk      | 0.9831     |           | 0.3736  |         |     |
| Durbin-Watson     | 1.6037     |           | 0.02184 |         |     |

After transformed the maternal death (Y), the residual of the regression model have no heteroskedasticity problem (BP = 8.44, p-value = 0.133> 0.05). However, the Durbin-Watson test result showed that the residual auto correlated even after transformed using root transformation (p-value = 0.0021< 0.05). It might caused by the existence of spatial autocorrelation in the data, as we
can see in the Figure 2 that the area with a high value of maternal death have a tendency to adjacent to the area that have high value. The alternative solution is to modelling the data using spatial regression model.

The spatial auto correlated in the data was analyzed by using Moran’s I test. The spatial weight matrix that was used in this study is queen contiguity where two areas are considered to have spatial relation one another if they share a side or a vertex. The weight matrix has $80 \times 80$ elements.

The Moran’s I test was performed at the dependent variable, explanatory variables, and the residual of the regression model, and the result written in Table 3.

**Table 3. Moran’s I test result.**

| Moran’s I | $E[I]$ | Mean | Sd  | z-value | p-value |
|-----------|--------|------|-----|---------|---------|
| $\sqrt{Y}$ | 0.2643 | -0.0128 | 0.0797 | 3.4941  | 0.001   |
| Residual  | 0.2400 | -0.0128 | 0.0853 | 2.9794  | 0.003   |

Based on Moran’s I test result in Table 3, it is found that dependent variable as well as the residual of multiple linear regression model are spatially auto correlated ($p$-values < 0.05). We tried to model the data using spatial regression models that have been explained in Figure 2. The result is given in Table 4.

**Table 4. Spatial Autoregressive Model and Spatial Error Model Test Result**

| Variable | Spatial Autoregressive (SAM) | Spatial Error (SEM) |
|----------|-----------------------------|---------------------|
| (Intercept) | 1.59871* | 2.51665* |
| $HP$ | 0.0016* | 0.0018* |
| $R$ | $1.387 \times 10^{-5}$* | $1.37031 \times 10^{-5}$* |
| $CW$ | $-0.3475$ | $-0.353$ |
| $NHP$ | $-3.1726 \times 10^{-5}$ | $9.18513 \times 10^{-6}$ |
| $IR$ | $-2.6414 \times 10^{-5}$ | $-3.182 \times 10^{-5}$* |
| $\rho$ | 0.20458* | - |
| $\lambda$ | - | 0.4212* |
| LR-Test | 0.02943* | 0.0032* |
| Shapiro-Wilk Test | 0.7654 | 0.2251 |
| BP-Test | 0.44623 | 0.30661 |
| Moran’s I | 0.1163 | 0.4098 |
| AIC | 240.393 | 234.486 |

*: $p$-value < 0.05

Likelihood ratio test was conducted to see if the spatial regression models is better than OLS to model the maternal death in West Java, Central Java, and East Java. Based on AIP value, the SEM model is slightly better than SAM. And the significant factors affecting maternal death in West Java, Central Java, and East Java are the number of health personnel, the numbers of residents received Raskin, and the number of illiterate residents.

The Spatial Error Model for the maternal death at i-th location is as the following.

$$
\sqrt{Y} = 2.51665 + 0.0018*HP + 1.37031*10^{-5}*R - 0.353*CW + 9.18513*10^{-6}*NHP - 3.182*10^{-5}*IR + 0.4212 \sum_{j \neq i}^{80} w_{ij} \mu_j + \varepsilon_i
$$

4. **Conclusion**

Maternal death in West Java, Central Java, and East Java is spatially auto correlated, make it difficult to construct the model using the ordinary least square (OLS). Since, the residual of the OLS is auto correlated even after the data was transformed, the alternative is spatial regression model. Based on the result SEM is better than SAM to model maternal death in West Java, Central Java, and East Java.
Based on SEM the explanatory variables that significantly affect maternal death are the number of health personnel, the number of residents who received Raskin, and the number of illiterate residents.

5. References
[1] Anselin, L 1988 Spatial Econometrics: Methods and Models (Dordrecht: Kluwer Academics Publisher)
[2] Cressie N A C 2015 Statistics for Spatial Data (New Jersey: John Wiley & Sons Inc)
[3] Hogg R V and Craig A T 1995 Introduction to Mathematical Statistics: Fifth Edition (Upper Saddle River, New Jersey: Prentice-Hall, Inc.)
[4] Benova L, Cumming O and Campbell O M 2014 Systematic Review and Meta-analysis: Association Between Water and Sanitation Environment and Maternal Mortality Tropical Medicine & International Health 19 p 388
[5] Lanree-Abass B A 2008 Poverty and Maternal Mortality in Nigeria: Towards a More Viabel Ethics of Modern Medical Practice International Journal for Equity in Health 7 p 11
[6] Lee J and Wong D W S 2001 Statistical Analysis with Arcview GIS (New York: John Wiley & Sons, Inc)
[7] Montgomery D C, Peck E A and Vinning G G 2012 Introduction to Linear Regression Analysis (New York: John Wiley & Sons, Inc)
[8] Pillai V K, Maleku A and Fangshun W 2013 Maternal Mortality and Female Literacy Rates in Developing Countries during 1970-2000: A Latent Growth Curve Analysis International Journal of Population Research 163292
[9] Yego F, D’Este C, Byles J, Williams J S and Nyongesa P 2014 Risk Factors for Maternal Mortality in a Tertiary Hospital in Kenya: a Case Control Study BMC Pregnancy Childbirth 14 p 38
[10] Astuti W D, Khaqiqi Z and Lestari, D 2011 Perilaku Hidup Bersih dan Sehat (PHBS) dalam Rumah Tangga Ibu Hamil dan Ibu Pernah Hamil di Indonesia Bulletin Penelitian Sistem Kesehatan 14 p 4

Acknowledgements
This research was funded by Directorate of Research and Community Service of Universitas Indonesia (DRPM UI) as a grant of PITTA (Publikasi Internasional Terindeks untuk Tugas Akhir) 2018.