American Journal of Environmental Science, 2012, 8 (6), 661-667
ISSN: 1553-345X

doi:10.3844/ajessp.2012.661.667 Published Online 8 (6) 2012 (http://www.thescipub.com/ajes.toc)

On the Influence of Climate and Socio-Economic Condition to the Dengue Incidences: A Semiparametric Panel Regression Approach

[1]Mutiah Salamah, [2]Heri Kuswanto and [1]Nurma Yussanti

[1]Department of Statistics, Laboratory of Environmental and Health Statistics,
[2]Department of Statistics, Laboratory of Computational Statistics,
Institut Tekonologi Sepuluh Nopember (ITS), Surabaya, Indonesia

Received 2012-11-14, Revised 2013-01-03; Accepted 2013-02-14

ABSTRACT

Dengue is one of the most dangerous diseases in the worlds. In particularly in East Java province Indonesia, dengue has been identified as one of the major causes of death. Hence, it is important to investigate the factors that induce the number of dengue incidences in this region. This study examines climate and socio-economic conditions, which are assumed to influence the number of dengue in the examined region. The semiparametric panel regression approach has been applied and the results are compared with the standard panel regression. In this case, the socio-economic condition is treated parametrically while climate effect is modeled nonparametrically. The analysis showed that the number of dengue incidences is significantly influenced by the income per-capita and the number of inhabitant below 15 years. Furthermore, the dengue incidence is optimum under rainfall of 1500 to 3670 mm, temperature of 22 to 27 degree and humidity of 82 to 87%. The elasticity allows us to identify the most responsive and most irresponsive district towards the changes of climate variable. The study shows that Surabaya is the most responsive district with respect to the change of climate variables.

Keywords: Elasticity, Panel, Semiparametric, Dengue

1. INTRODUCTION

Dengue is one of the most dangerous diseases in the world. The spread of the dengue is infected by a vector called as Aides Aegepty. Dengue frequently arises as an extraordinary case that causes relatively high degree of mortality rate leading to economic loses. Indonesia becomes one of the countries with the highest rate of dengue incidence among ASEAN countries. Statistic data shows that there are about 1317 dengue cases in 2010. One of the provinces in Indonesia with high rate of dengue case is East Java. The number of dengue in East Java gradually increases from 8287 cases in 2004 and reaches the peak in 2007 with 25950 cases (DH, 2009). Therefore, study about the dengue incidence in East Java is an important issue.

Researches on identifying the causes of dengue in a region have been intensively studied. Khormi and Kumar (2011) investigate the influence of socio-economic condition to the number of dengue in Saudi Arabia. Ma *et al*. (2008) studied the similar case in Singapura. The study argues that the number of dengue incidence does not depend only on the socio-economic condition, however it is affected also by the climate condition. Some researches that study the impact of climate to dengue are Zhou *et al*. (2004); Hales *et al*. (2002) and Maslukha (2010). These researches assume that either the socio-economic or climate condition affects the dengue incidence separately. In fact, climate and socio-economic condition may affect the dengue incidence simultaneously. Mondzozo *et al*. (2011) is the

Corresponding Author: Mutiah Salamah, Department of Statistics, Laboratory of Environmental and Health Statistics, Institut Tekonologi Sepuluh Nopember (ITS), Surabaya, Indonesia

only study that models the dengue incidence by taking into account both climate and socio-economic conditions at the same time, using semiparametric panel regression approach. The study was conducted in Africa and it shows that the model can provide more information about the dengue incidence. In term of the goodness of the model, semiparametric panel regression approach outperforms the standard panel regression.

This study applies the similar approach as in Mondzozo *et al*. (2011) to model the dengue incidence in East Java Indonesia. The responsiveness degree of the examined districts towards dengue incidence will be performed. Moreover, elasticity representing the degree of vulnerability of each district towards dengue with refers to the climate condition will also be discussed.

2. MATERIALS AND METHODS

2.1. Panel Regression

Panel regression is a regression composed from cross-section and time series data. The panel regression is given by the following model:

$$y_{it} = a + \beta X_{it} + u_{it}; t = 1,\dots,T; i = 1,\dots,N$$

where, y_{it} is a response variable from region i at time t, X_{it} is the predictor, a is unobserved effect on every region, while u_{it} is error term. Prior to the modeling process, it is necessary to test the panel effect as well as to test the type of the effect whether fixed or random effect.

Testing panel effect is done by Brausch-Pagan Lagrange Multiplier test, with the following hypothesis:

$$H_0 : \sigma_\mu^2 = \sigma_\lambda^2 = 0 \text{(no panel effect)}$$
$$H_1 : \sigma_\mu^2 \neq \sigma_\lambda^2 \neq 0 \text{(panel effect)}$$

Statistical test:

$$LM = LM_1 + LM_2$$

Where:

$$LM_1 = \frac{NT}{2(T-1)}\left[1 - \frac{\tilde{u}(1_N \otimes J_T)\tilde{u}}{\tilde{u}'\tilde{u}}\right]^2$$

And:

$$LM_2 = \frac{NT}{2(N-1)}\left[1 - \frac{\tilde{u}(1_N \otimes J_T)\tilde{u}}{\tilde{u}'\tilde{u}}\right]^2$$

Where:

N = The number of region
T = The length of period
I_N = Identity matrix from N
J_T = Matrix of T with dimension of one
$\tilde{u}$ = Error of the OLS model. Reject the null if $LM \geq X_2^2$

Another way to test the panel effect is by F test as follow:

$$H_0 : \mu_i = 0 \text{ (no individual effect)}$$
$$H_1 : \mu_i \neq 0 \text{ (there is individual effect)}$$

Statistical test:

$$F = \frac{(RRSS - URSS/(N-1))}{URSS(NT-N-K)}$$

where, RRSS is sum square residual of OLS model and URSS is sum square residual of fixed model. Reject the H_0 if $F > F_{N-1,NT-N-k}$. The test to determine fixed or random effect is done by Hausman-test, with the following hypothesis:

$$H_0 : E(a_i \mid X_{it}) = 0 \text{ (random effect)}$$
$$H_1 : E(a_i \mid X_{it}) \neq 0 \text{ (fixed effect)}$$

Statistical test:

$$X_{hit}^2 = (b-\beta)'Var(b-\beta)^{-1}(b-\beta)$$

where, b is coefficient of random effect while β is coefficient of fixed effect. Reject the H_0 if $X_{hit}^2 > x_{(k,a)}^2$, where k is number of coefficient β (Baltagi, 1995).

2.2. Semiparametric Panel Regression

Semiparametric model combines parametric and nonparametric model. Semiparametric regression model can be written as follow Tseng *et al*. (2009); Li and Racine (2007) and Racine (2008):

$$y_{it} = f(X_{it}) + \beta Z_{it} + a_i + u_{it}; t = 1,\dots T; i = 1,\dots,N$$

Where:

y_{it} = The response variable in the i-th region on period t
X_{it} = A vector of nonparametric variable
Z_{it} = A vector of parametric variable
a_i = Unobserved effect for each region

u_{it} = The error term

Parameters estimated in this regression are f, β, a_i, where f is an unknown function while parametric variables are assumed to influence the spread of the dengue linearly Estimation method is done by two steps procedure following Robinson (1988).

First step: Form a conditional expection of response variable (y) on climate variables (X_{it}) such that:

$$E(y_{it} \mid X_{it}) = f(X_{it}) + \beta E(Z_{it} \mid X_{it})$$

where, the nonlinear variables X_{it} is assumed to be uncorrelated with regional individual effect and error. By substracting the first from second equation above we obtain:

$$y_{it} - E(y_{it} \mid X_{it}) = \beta[Z_{it} \mid X_{it}] + a_i + u_{it}$$

Conditional expectation is obtained by nonparametric kernel method such that $\tilde{y}_{it} = y_{it} - \hat{E}(y_{it} \mid X_{it})$, $\tilde{Z}_{it} = [Z_{it} - \hat{E}(Z_{it} \mid X_{it})]$ and $\hat{E}(.|.)$ is the kernel estimator. The semirapaetric model is then transformed into linear equation:

$$\tilde{y}_{it} = \beta\tilde{Z}_{it} + a_i + u_{it}$$

where, the parameter $\hat{\beta}$ is estimated by similar procedure as standard panel regression.

Second step: The second step will estimate the function f. After the value of $\hat{\beta}$ has been estimated, the function f is estimated by substituting $\hat{\beta}$ through:

$$y_{it} - \hat{\beta}Z_{it} = f(X_{it}) + a_i + u_{it}$$

By defining $y_{it} - \tilde{\beta}Z_{it} = \bar{\bar{y}}_{it}$, we obtain nonparametric form:

$$\bar{\bar{y}}_{it} = f(X_{it}) + a_i + u_{it}$$

And f is estimated locally by minimizing $\sum_i \sum_t [\bar{\bar{y}}_{it} = f(X_{it}) + a_i + u_{it}]^2 k\left(\frac{x_{it} - x}{h}\right)$ with k (.) is kernel density function. Ullah and Mudra (2002) showed that f can be expressed as $f(X_{it}) = V_{it}^{'}\theta(X_{it})$ where $V_{it} = (1, X_{it})$ and $\hat{\theta}(X_{it})$ can be estimated by:

$$\hat{\theta}(X_{it}) = \left[\sum_i \sum_t V_{it}^{*} V_{it}^{*}{}'K]^{-1}[\sum_i \sum_t V_{it}^{*} \bar{\bar{y}}_{it}^{*} k\left(\frac{X_{it} - X}{H}\right)\right]$$

This specification considers a fact that climate change causes a significant jump of dengue in a certain region. In general, elasticity is defined as change (in percentage) of response variable caused by one percent changing on the predictor variable. Elasticity is calculated by the following Equation 1:

$$e_{xy} = \left|\frac{\partial y}{\partial x} \cdot \frac{x}{y}\right| = \left|\hat{f}(x) \cdot \frac{x}{y}\right| \quad (1)$$

where, e_{xy} is elasticity of y variable on x. This equation is used to calculate the dengue projection based on the elasticity definition.

2.3. Methodology

2.3.1. Data

Data used in this research is secondary panel data spanning from 2003 to 2010. Data of climate variables are obtained from the climatology station Karang Ploso Malang, while socio-economic data is collected from Center of Statistic-East Java Province. Data about the number of dengue incidences are obtained from the Ministry of Public Health-East Java.

2.4. Variables

There are three kinds of data used in this research i.e., climate variable, socio-economic as the predictors and the number of dengue/100000 population as the response variable. The **Table 1** shows the variables used in the study.

All climate variables should be are measured as the average of annually data as the socio-economic data are available annually. The idea of rescaling the number of inhabitants below 15 years old and poor inhabitants is to reduce the lag of the data due to the scale of data measurement.

2.5. Analytical Method

This research uses dataset from 13 districts in East-Java region spanning from 2003-2010. The reason of using 13 districts data considering availability of climate data, in which complete temperature and humidity data are completely available only in 13 districts i.e., Blitar, Kediri, Malang, Jember, Banyuwangi, Pasuruan, Sidoarjo, Nganjuk, Magetan, Gresik, Sumenep, Surabaya and Batu. Nevertheless, these districts are representative enough for East Java province.

In summary, the steps of analysis can be described as follows:

- Perform panel regression between number of dengue incidents with socio economic and climate variables
- Parameters estimation on the semiparameteric panel regression
- Nonparametric estimation of function f
- Calculate the elasticity from smooth coefficient of the nonparametric kernel regression.

3. RESULTS

3.1. Descriptive Statistic of Variables

This section briefly describes the statistic of each variable observed over all districts during the observed periods (2003-2010). The mean of dengue incidence per 100000 inhabitants is 43.85 with the minimum of 3 incidents happened in Sidoarjo, while the maximum incidents happened in Batu 2010 with 167 events. The average of income per-capita is IDR 7748000 with standard deviation of 6411000. The mean of the number of inhabitants below 15 years old per 100000 inhabitants is 24257. It means that if there exists 100000 inhabitants in any district, thus 24257 of them are supposed to be below 15 years old. The average of annual rainfall is 1686.3 mm with the minimum is observed in Sumenep (2009) corresponding to 61 dengue incidents per 100000 inhabitants, while the maximum is 3670 mm in Blitar (2010) with dengue of 80 incidents per 100000.

3.2. Panel Regression

Panel regression analysis is done by regressing in panel the number of dengue per 100000 inhabitants with socio-economic variables (Z1 to Z4) and climate variables (X1 to X4). This model assumes that both groups of variable linearly affect the dengue incidences. **Table 2** presents the parameters of the regression.

The table shows the coefficients (estimate), standard error and statistical test of the panel regression. The value of the coefficient of determination is 42.698% which means that the 42.69% variability of dengue incidents can be explained by socio-economic and climate variables. Based on the value of t-statistic, it is known that there are two socio-economic variables that are significantly influence the number of dengue i.e., density and income per-capita. Density influences dengue with negative coefficient, meaning that the number of dengue incidence decreases with increasing the density, assuming that other variables are constant. Another fact is observed for income per-capita, where the number of dengue incidence increases with increasing income per capita in any districts.

For the climate variables, none of them influences the dengue incidence. This insignificant influence may happen due to incapability the model to overcome nonlinearity on the relationship between both variables. Therefore, standard panel regression may fail to generate a good result.

3.3. Semiparametric Panel Regression

The semiparametric panel regression that will be examined is:

$$y_{it} = f(X_{it}) + \beta Z_{it} + a_i + u_{it}$$
$$t = 1,...,T; i = 1,...,N$$

Where:

y_{it} = The dengue incidence per 100000 inhabitants in i-th distric on the period t
X_{it} = A vector of climate variables
Z_{it} = A vecor of socio-economic variables

Table 1. Predictor variables

Socio-economic variable	
Z_1	density (km^2)
Z_2	Income per Capita (IDR)
Z_3	Number of inhabitant <15 years old per 100000 inhabitants
Z_4	Number of poor inhabitants per 100000 inhabitants
Climate variables	
X_1	Temperature (°C)
X_2	Humidity (%)
X_3	Rainfall (mm)

Table 2. Parameter estimate of standard panel regression

Variabel	Estimate	Std. Error	Stat. test-t
Z_1	-1.54×10^{-2}	6.08×10^{-2}	-2.5337*
Z_2	1.40×10^{-5}	2.55×10^{-6}	5.4892**
Z_3	-1.3×10^{-3}	1.086×10^{-3}	-1.1972
Z_4	3.48×10^{-4}	7.487×10^{-4}	0.4646
X_1	9.74×10^{-3}	4.968×10^{-3}	1.9605.0
X_2	-4.6303	4.5836	-1.0102
X_3	6.41×10^{-1}	7.04×10^{-1}	0.9107
R^2 = 42,698%			

*; Significant at the level of α = 0.05, ** Significant at α = 0.01, ***; Significant at α = 0.001, Significant at α = 0.1

Table 3. Parameter estimate of parametric component of the semiparametric panel regression

Variable	Coefficient	Std. Error	t-Stat	Others
Z_1	-0.0011747	0.0065805	-0.1785	
Z_2	0.0000078	0.0000017	4.637***	
Z_3	-0.0023216	0.0008494	-2.7333**	
Z_4	-0.0004667	0.0007123	-0.6551	
LM test	60,5553*			
F test				2.8494***
Hausman test				13.2434*
R^2				34,47%

*; Significant at the level of α = 0.05, **; Significant at α = 0.01, ***; Significant at α = 0.001, Significant at α = 0.1

Table 4. Value of for each district

District	α
Blitar	14.180300
Kediri	1.143500
Malang	-13.969700
Jember	6.852300
Banyuwangi	0.266100
Pasuruan	-12.366100
Sidoarjo	-32.354800
Nganjuk	8.672800
Magetan	11.036900
Gresik	-7.010000
Sumenep	21.868000
Surabaya	8.931300
Batu	3.117400

Table 5. Elasticity of each district to the climate variables

District	Rainfall	Temperature	Humidity
Blitar	0.090	30.087	12.084
Kediri	-0.027	-98.953	8.023
Malang	0.035	3.911	-11.008
Jember	-0.065	2.558	-1.727
Banyuwangi	-0.113	-34.117	1.146
Pasuruan	0.003	0.122	0.540
Sidoarjo	-0.058	5.071	3.698
Nganjuk	0.157	-20.650	-1.032
Magetan	0.002	-0.869	-3.649
Gresik	0.002	-0.869	-3.649
Sumenep	-0.017	0.058	0.174
Surabaya	-0.319	136.825	50.185
Batu	0.035	18.418	1.386

The semiparametric panel regression model yields on two models i.e., nonparametric model with unknown function and parametric model with coefficients of β. **Table 3** summarizes the coefficients of parametric component. From the table, the LM test is significant on 0.01 level and hence we conclude that there is a panel effect in the examined dataset. Having tested the panel effect, we test the region effect by F test and it shows that the region effect is significant under 0.001 level. Hausman test is used to identify whether the effect will be fixed or random. As it is significant, thus we conclude that panel regression model with fixed effect is more appropriate for the case.

Based on the t statistic, it suggests that there are two variables that significantly influence the dengue i.e., income per-capita and number of inhabitants below 15 years old. For the income per capita, it shows a positive coefficient which means that region or district with higher income per-capita tends to have larger number of dengue incidents. In contrast to this, negative coefficient is obtained for number of inhabitants below 15 years old. In other words, we may say that dengue attacks adults more than children.

The fixed effect model leads to different values of a_i for every district as shown in the **Table 4**.

This a coefficient shows intercept of the panel model with fixed effect and hence this value can be used to show the vulnerability of the district if all districts have same socio-economic condition. The more negative the value, the more vulnerable the district towards dengue incidence. Regions that are least vulnerable are adalah Malang, Pasuruan dan Sidoarjo. While districts that are most vulnerable are Blitar, Nganjuk, Magetan, Sumenep dan Kota Surabaya. The remaining districs are moderate. The socio economic variables solely can explain the variability of dengue with only 34.47%. Although it is lower than parametric panel regression, the goodness of the semiparametric panel regression model should be combination of R^2 from both parametric and nonparametric component.

Having known the coefficients of the socio-economic variables, we proceed to the nonparametric model for the climate variables. The nonparametric component consists of three variables i.e., rainfall, temperature and humidity, by which the kernel function will be estimated and result on the plots shown in **Fig. 1**. The nonparametric model can explain the variability of the dengue incidence with value of 80.54%, shown by the R^2 value.

Figure 1 show plots resulted from nonparametric kernel regression using 95% confidence level. The upper panel of **Fig. 1** shows that dengue tends to be stable at the rainfall 500 to 1500 mm in a year, while if the rainfall reaches 1500 mm to 3500, it will increase the number of dengue. The middle panel of the figure shows that the relationship between dengue and temperature is nonlinear. It is shown by the unsystematic pattern of the plot. In general, dengue incidences will increase on the temperature of 22-27°C, while for temperature exceeds 27°C, the dengue incidence will decrease.

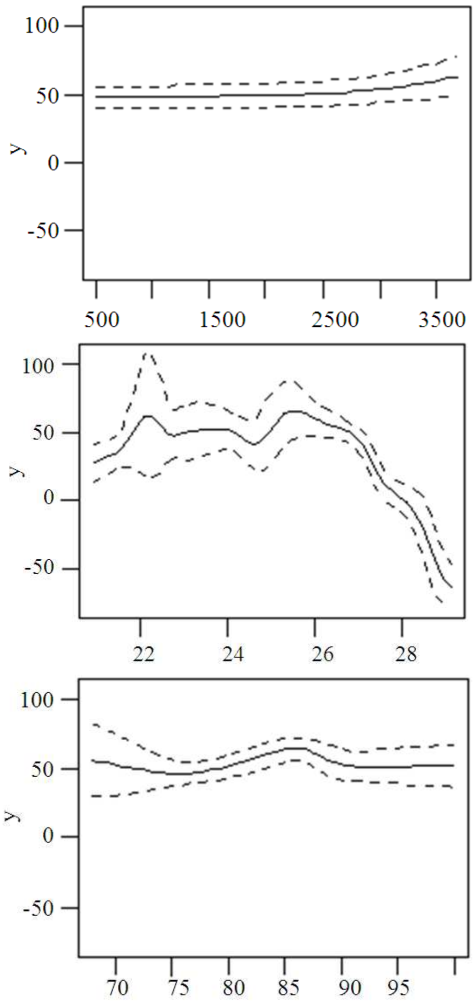

Fig. 1. Plots of nonparametric components

The last plot shows that humidity influences dengue with nonlinear pattern. With refers to the line plot, the dengue incidents will decrease with humidity of 70 to 75% and it will increase with humidity of 82 to 87%. The dengue will also increase under humidity of below 70%. Semiparametric panel regression yields on more flexible model as the effect of each region can be explained individually.

3.4. Elasticity of Dengue Incidence to the Climate Variables

The elasticity is obtained from the smooth coefficient of the nonparametric regression and the value represents the vulnerability of each district towards the climate condition. The **Table 5** below listed the elasticity.

The elasticity on the table shows the fluctuation of the dengue incidences due to the change of rainfall, temperature and humidity. The dengue in each district is responsive to the change of the temperature and humidity. Meanwhile, it is less responsive to the change of the rainfall intensity. Surabaya is the most responsive district towards the change of temperature and humidity, while Sumenep is the least responsive towards the change of both variables. Responsive in this case is defined as the vulnerability of the district to the dengue incidents if there is a change on the climate variables.

4. CONCLUSION

This study investigates the influence of socio-economic and climate variables to the dengue incidence using semiparametric panel regression approach. The analysis shows that the proposed model outperforms the standard panel regression. With this regression, the vulnerability and elasticity of each region in East Java province can be explained.

5. ACKNOWLEDGEMENT

The reachers gratefully acknowledge the financial support from the Institut Teknologi Sepuluh Nopember through the research grant Unggulan-decentralisation program.

6. REFERENCES

Baltagi, B.H., 1995. Econometric Analysis of Panel Data. 1th Edn., Wiley and Sons Ltd, Chichester, ISBN-10: 0471952990, pp: 257.

DH, 2009. Database Kesehatan per Propinsi. Bankdata Depkes RI, Diunduh.

Hales, S., N.D. Wet, J. Maindonald and A. Woodward, 2002. Potential effect of population and climate changes on global distribution of dengue fever: an empirical model. Lancet, 360: 830-834. DOI: 10.1016/S0140-6736(02)09964-6

Khormi, M.H. and L. Kumar, 2011. Modeling Dengue Fever risk Based on SocioEconomic Parameters, Nationality and Age Groups: GIS and remote sensing based case study. Sci. Total Environ., 409: 4713-4719. PMID: 21906782

Li, Q. and J.S. Racine, 2007. Nonparametric Econometrics: Theory and Practice. 1st Edn., Princeton University Press, ISBN-10: 0691121613, pp: 768.

Ma, S., E.E. Ooi and K.T. Goh, 2008. Socioeconomic determinants of dengue incidence in Singapore. Dengue Bull., 32: 17-28.

Maslukha, S., 2010. Hubungan perubahan parameter iklim akibat global warming dengan kasus demam berdarah dengue di kabupaten sidoarjo. Unpublished Thesis, Airlangga University.

Mondzozo, A.E., M. Musumba, B.A. McCarl and X. Wu, 2011. Climate change and vector-borne diseases: An economic impact analysis of malaria in Africa. Int. J. Environ. Res. Public Health, 8: 913-930. PMID: 21556186

Racine, J.S., 2008. Nonparametric econometrics: A primer. Foundat. Trends. Econom., 3: 1-88. DOI: 10.1561/0800000009

Robinson, P.M., 1988. Root-N-consistent semiparametric regression. Econometrica, 56: 931-954.

Tseng, W.C., C.C. Chen, C.C. Chang and Y.H. Chu, 2009. Estimating the economic impacts of climate change on infectious diseases: A case study on dengue fever in Taiwan. Climatic Change, 92: 123-140. DOI: 10.1007/s10584-008-9437-6

Ullah, A. and K. Mudra, 2002. Semiparametrik Panel Data Estimation: An Aplication to Immigrants Homelink Effect on US Producer Trade Flows. In: Handbook of Applied Econometrics and Statistika Inference, Ullah, A., A. Wan and A. Chaturvedi (Eds.), CRC Press, New York, ISBN-10: 082474411X, pp: 592-608.

Zhou, G., N. Minakawa, A.K. Githeko and G. Yan, 2004. Association between climate variability and malaria epidemics in the East African highlands. Proc. Natl. Acad. Sci. USA., 101: 2375-2380. PMID: 14983017