Development of Air Polluter Model for the Carbon Monoxide (CO) Element Based on Mixed Geographically Temporal Weighted Regression (MGTWR) Kriging

by Fatmawati Fatmawati

Applied Mathematical Sciences, Vol. 8, 2014, no. 118, 5863 – 5873
HIKARI Ltd, www.m-hikari.com
http://dx.doi.org/10.12988/ams.2014.45334

Development of Air Polluter Model for the Carbon Monoxide (CO) Element Based on Mixed Geographically Temporal Weighted Regression (MGTWR) Kriging

Kukuh Winarso

Post Graduate Student of MIPA, Universitas Airlangga and Departement of Industrial Engineering, Faculty of Engineering, Universitas Trunojoyo Madura, JL. Raya Telang P.O. Box 2 Bangkalan, Madura, Indonesia

Hari Basuki Notobroto

Department of Biostatistics and Demography, Faculty of Public Health, Airlangga University, JL. Mulyorejo, Kampus C, Unair Surabaya, Indonesia

Fatmawati

Department of Mathematics, Faculty of Sciences and Technology, Airlangga University, JL. Mulyorejo, Kampus C, Unair Surabaya

Abstract

Motor vehicle exhaust is the main source of Carbon Monoxide (CO). In urban areas, where motor vehicles are easily found, CO is commonly identified as one of the air polluter sources. CO is hazardous, and human being will suffer from serious health problems if they are exposed to this polluting agent in a prolonged period. This study aims to develop a model that will aid the effort to minimize the negative effects of CO. The model is built in which it will be able to show the cause-factors and the preventing-factors of CO generation. The Mixed Geographically Temporal Weighted Regression (MGTWR) approach is used as the basis of the model. MTGWR is a spatial-temporal regression model, which takes into account the geographical and temporal aspects of the pollution. However, the MGTWR model cannot be used to predict the effect if it is used outside the sample location of the research, unless we predict the associated regression coefficients in the respected area beforehand. In this case, we use the estimated predictor parameter based on the Kriging method to predict the regression parameters outside the research location. As the result, the Kriging Predictor-based MTGWR model can be used to estimate the pollution level caused by CO outside the sample location of the research.

Keywords: Air Polluter, Carbon Monoxide, Spatio-Temporal, MGTWR, Kriging-predictor

1. INTRODUCTION

Previous study shows that there is a parallel relationship between the number of polluted air-generated disease sufferers and the intensity of industrialization and urbanization in countries whose level of air pollution are severely high [1]. Carbon Monoxide (CO) is one of the air polluters, which is generated from the imperfect fuel combustion of motor vehicles and industrial machines. Motor vehicle exhaust is the main generator of Carbon Monoxide and it is common to be found in big cities. Based on data, about 60% of air pollution in major urban areas is related with public transportation [2]. Within the context of human health, Carbon Monoxide is toxic. It causes lower fetal weight, higher rate of infant death, and leads to human brain damage. The air quality standard that is related with Carbon Monoxide is 10.000 ug/Nm3 [3]

Demographically, the potential of the effect and causes of air pollution will be different one to another in different regions. Also, it is inappropriate to analyze the effect and causes of air pollution based on a global approach, as we will be unable to examine the influence of local variations if we are to use global approach [4-5]. The common spatial regression method known is the Geographically Weighted Regression (GWR). GWR is a weighted-regression technique that is based on the simple regression approach. It is a common statistical technique used for spatial heterogeneity. The term heterogeneity in this case is the measurement relationship of different variables among various locations [6]. Spatial heterogeneity occurs when an independent variable gives different responses to different locations within the same research location [7]. The crux of GWR approach is determining the regression model for each of the location nodes so that the regression models obtained are unique, in which a model for a node will be different to the others [8]. There is another model aside from GWR, which is Mixed Geographically Temporal Weighted regression (MGTWR). MTGWR is a combined modeling approach between global regression and local regression (GWR) [9]. This paper takes into consideration the element of location, as well as the time into the local variation based on MGTWR approach. The usefulness of global regression is that it can be used to predict in each location. However, in the spatial regression models such as GWR, MGWR (Mixed Geographically Weighted Regression), and MGTWR, the global regression cannot be used for predicting outside the research sample location, unless we predict the respected location's regression coefficients beforehand [10]. In order to eliminate this obstacle, we estimate the predictor parameter using the Kriging method to predict the regression parameter outside the research location.

2. RESEARCH OBJECTIVE

This research aims to develop the air polluter model for the Carbon Monoxide (CO) element based on the Mixed Geographically Temporal Weighted Regression (MGTWR) approach, in which the developed model can be used to predict outside the research location using the Kriging predictor estimation. This way, the developed model can be used to predict the regression parameter outside the research location.

3. METHODOLOGY

3.1 Mixed Geographically Temporal Weighted Regression (MGTWR)

MGTWR is the advancement of the MGWR model [11] with the incorporation of the temporal element into the model. The temporal factor is aimed to predict the observation time, and to complement the location (location coordinate) factor. Mathematically, the temporal model can be stated as the following:

$$y_i = \sum_{j=1}^{q} \beta_i x_{ij} + \beta_j \left(u_i, v_i, t_i\right) x_{ij} + \varepsilon_i \tag{1}$$

Where: $i = 1,2,3\ldots.n$

β_i is global variable x_i

β_j is local variable x_j

y_i is Response of variables

x_i is predictors of variables

u_i is longitude at the time-i

v_i is latitude at the time-i

t_i is the time length at the time-i

3.2. Kriging Spatial Predictor Model

Kriging method is used to predict the regression parameter outside the research location [12]. The assumption that is associated with the predictor is:

$$p(b;s_0) = \sum_{i=1}^{n} \lambda_i \mathrm{b}(\mathrm{s}_\mathrm{i}), \quad \sum_{i=1}^{n} \lambda_i = 1 \tag{2}$$

Where:

λ_i = the weight $\mathrm{b}(\mathrm{s_i})$ for the estimation of location s, with the value of $\mathrm{b}(\mathrm{s_i})$ has different weighted-coefficients for the estimation in different locations

s_i = location vector

s_0 = Location of the data to be estimated

n = number of the sample data

The Kriging spatial predictor matrix notation can be stated as the following [12]:

$$\begin{bmatrix} \gamma(s_i - s_j) & f_t(s_i) \\ f_t(s_j) & 0 \end{bmatrix} \begin{bmatrix} \lambda_j \\ m_t \end{bmatrix} = \begin{bmatrix} \gamma(s_i - s_j) \\ f_t(s_0) \end{bmatrix}, \text{ for } i = 1, 2, \ldots, n \text{ and } t = 1, 2, \ldots, p$$

Based on the previous matrix, we can compute the weight matrix for the Universal Kriging, which is:

$$\begin{bmatrix} \lambda_j \\ m_t \end{bmatrix} = \begin{bmatrix} \gamma(s_i - s_j) & f_t(s_i) \\ f_t(s_j) & 0 \end{bmatrix}^{-1} \begin{bmatrix} \gamma(s_i - s_j) \\ f_t(s_0) \end{bmatrix} \tag{3}$$

Notations :

$\gamma(s_i - s_j)$ = *Semivariogram* between sampled nodes
$\gamma(s_0 - s_i)$ = *Semivariogram* between sampled node and estimated node
$f_t(s_i), f_t(s_j)$ = Location coordinate of the sampled data
λ_j = Weighting value to be computed
m_t = Value of *lagrange* parameter
(s_i, s_j) = Location of the sampled data
s_0 = Location of the data to be estimated
p = Number of order in the trend equation

Semivariogram is used to determine the distance where the observation data values are not dependent one to another or have no correlation among them [12]. The equation in the Kriging model is used to model the trend of the low order polynomial, which is the first or the second order [13]. If the obtained trend has the first order at R^3, then the equation is:

$m(s) = m(x,y,z) = a_0 + a_1x + a_2y + a_3z$

where x,y,z are the sampled location coordinates. Thus, the equation (3) becomes as the following:

$$\sum_{j=1}^{n} \lambda_j \gamma(s_i - s_j) + m_0 + m_1 x_i + m_2 y_i + m_3 z_i = \gamma(s_0 - s_i), \text{ for } i = 1,2,\ldots,n$$

With assumption:

$$\sum_{j=1}^{n} \lambda_j x_j = x, \quad \sum_{j=1}^{n} \lambda_j y_j = y, \quad \sum_{j=1}^{n} \lambda_j z_j = z, \quad \sum_{j=1}^{n} \lambda_j = 1$$

If we want to transform the equation above into the Kriging matrix in the following first order Kriging trend (KT):

$$K_{KT} = \begin{bmatrix} \gamma(s_1 - s_1) & \cdots & \gamma(s_1 - s_n) & 1 & \mathrm{x}_1 & \mathrm{y}_1 & \mathrm{z}_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \gamma(s_n - s_1) & \cdots & \gamma(s_1 - s_n) & 1 & \mathrm{x}_n & \mathrm{y}_n & \mathrm{z}_n \\ 1 & \cdots & 1 & 0 & 0 & 0 & 0 \\ x_1 & \cdots & x_n & 0 & 0 & 0 & 0 \\ y_1 & \cdots & y_n & 0 & 0 & 0 & 0 \\ z_1 & \cdots & z_n & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \lambda_{KT} = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \\ m_0 \\ m_1 \\ m_2 \\ m_3 \end{bmatrix}$$

$$k_{KT} = \begin{bmatrix} \gamma(s_0 - s_1) \\ \vdots \\ \gamma(s_0 - s_n) \\ 1 \\ x \\ y \\ z \end{bmatrix}$$

where : $K_{KT}\,\lambda_{KT} = k_{KT}$

then the weighting matrices can be stated as the followings [13]:

$$\lambda_{KT} = K_{KT}^{-1} k_{KT}$$

The case study in this research uses the Carbon Monoxide (CO) element as the response variable Y, while the predictor variables are the air temperature (X_1), the wind velocity (X_2), the air humidity (X_3), the traffic velocity (X_4), the area size of the urban forest (X_5), the population density (X_6), and the business center aspect (X_7) [14]

4. RESULTS AND DISCUSSION

The first MGTWR model test is the model fitness, in which the hypothesis can be stated as the followings:

H$_0$: $\beta_k\left(u_i, v_i, t_i\right) = \beta_k\left(u_i, v_i\right) \quad k = 0, 1, \cdot 2 \cdot, \quad, 4\ i = d\ a\ n \cdots$

(MGTWR model is not different from the MGWR model)

H$_1$: At least one $\beta_k\left(u_i, v_i, t_i\right) \neq \beta_k\left(u_i, v_i\right)$

(MGTWR model is significantly different from the MGWR model)

The test results are summarized in Table 1:

Tabel 1 Fitness Test of MGTWR Model for the CO Response

Source of Error	Sum Square	Degree of Fredom	Mean Square	F	p-value
Improvement	7,7679	15,9260	0,4877	4,2433	0,0000
MGTWR	46,1362	401,3706	0,1149		
MGWR	53,9041	417,2966			

Table 1 shows that the F test statistical value is 4,2433, and the the p-value of 0,0000. Using the significance value (α) of 5%, we must reject H$_0$, and conclude that the MGTWR model is significantly different from the MGWR model. Therefore, we can further conclude that the MGTWR model is more proper to model the Air Polluter Standard Index (APSI) for Carbon Monoxide (CO). This means that the time element is influential in the APSI modeling for CO so that not only the location factor is considered, but the observation time is taken as the influential factor to the APSI model for CO as well.

Next step is the global parameter test that is conducted to identify which global predictor variables significantly influence the response. In this test, we perform the partial global parameter test using the hypothesis below:

H$_0$: $\beta_k = 0$ (global variable x_k is not significant)

H$_1$: $\beta_k \neq 0$ (global variable x_k is significant)

Table 2 shows that by using the level of significance (α) 5%, we can conclude that the global predictor variable that significantly influence the APSI for CO are the area size of the urban forest (X$_5$) and the business center aspect (X$_7$), because these variables have the p-value that are less than 0.05.

Table 2 Parsial Test of Global Variable of MGTWR Model

Global Parameter			
Variable	Beta	T	p-value
X_5	-0,0445	-0,4171	0,0080*
X_6	0,0088	0,3556	0,3613
X_7	-0,0456	2,0045	0,0228*

Note: *) significant at $\alpha = 5\%$

By using α = 5%, the location and the observation time of APSI for CO in Surabaya city can be classified based on the significant variables for APSI values for CO. The detailed significant variable list in each location and in each observation time is presented in Table 3:

Table 3 MGTWR Model for CO at Five Observation Location Nodes

No	Observation Location	Observation Time	Significant Variables
1	SUF 1 (Location 1)	Morning	X_3, X_4
		Noon	X_3, X_4
		Evening	X_1, X_2, and X_4
2	SUF 3 (Location 2)	Morning	X_3, X_4
		Noon	X_3, X_4
		Evening	X_1, X_2, X_3, X_4
3	SUF 4 (Location 3)	Morning	X_3, X_4
		Noon	X_4
		Evening	X_1, X_2, X_3, X_4
4	SUF 5 (Location 4)	Morning	X_3, X_4
		Noon	X_3, X_4
		Evening	X_1, X_2, X_3, X_4
5	SUF 6 (Location 5)	Morning	X_3, X_4
		Noon	X_2
		Evening	X_2, X_3, and X_4

The MGTWR model obtained in each of the location and observation time will be different one to another, depending on the parameter value of MGTWR and on which predictor variables that significantly influence the response variable of APSI (CO) with the R^2 value of 36.60%.

To predict the MGTWR model at a certain location node, we use the Kriging approach that is presented in the following Table 4:

Table 4 Kriging-based MGTWR Model with Local Parameter Lokal for CO Element

No	Sub-district	Time	Constant	X_1	X_2	X_3	X_4
1	Location z	Morning	4.5498	-0.0541	0.0690	-0.1437	0.0614
		Noon	4.4902	0.0164	0.0180	0.0932	0.2797
		Evening	4.5415	0.1745	0.0870	0.1284	0.2571
2	Location y	Morning	4.5523	-0.0425	0.0667	-0.1441	0.0636
		Noon	4.4903	0.0145	0.0192	0.0833	0.2773
		Evening	4.5366	0.1706	0.0840	0.1233	0.2567
3	Location x	Morning	4.5457	-0.0618	0.0665	-0.1335	0.0693
		Noon	4.4897	0.0202	0.0206	0.0923	0.2775
		Evening	4.5441	0.1686	0.0874	0.1262	0.2612

In Table 4, if we are to estimate for location x, the Kriging-MGTWR with the CO response for evening is:

$$\text{Ln (CO)} = 4.544 + 0.0874X_2 + 0.2612X_4 + 0.0445X_5 - 0.0456X_7$$

The classification map in Surabaya based on CO polluter can be seen in the following:

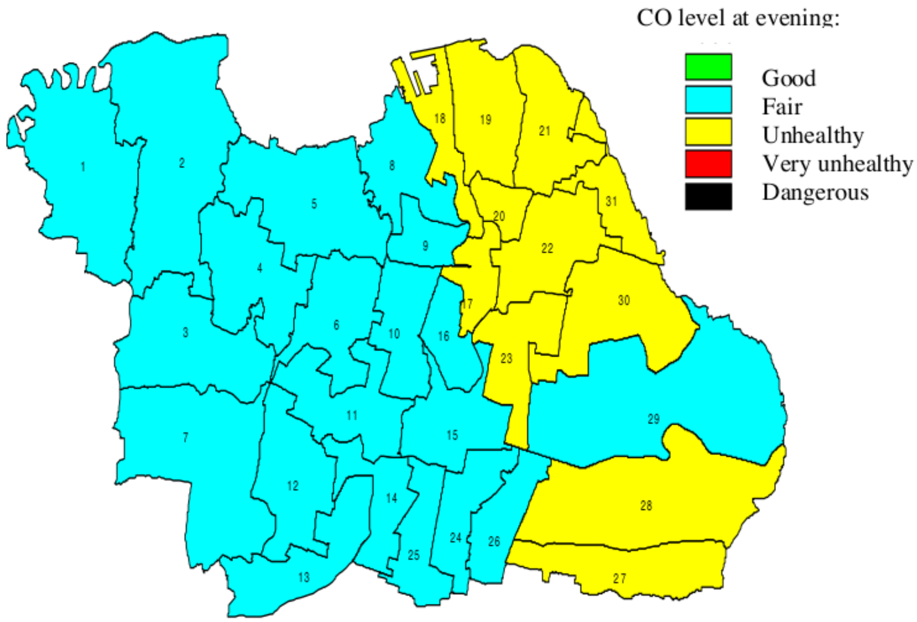

Figure 1. *Classification of Sub-districts in Surabaya Based on the Air Polluter CO*

There are eleven sub-districts that are classified as unhealthy (yellow-colored) and twenty sub-districts that are classified as fair (blue-colored).

5. CONCLUSION

Aside from the global influence, the MGTWR model of the air pollution is also influenced significantly by the location factor (geographical factor). Therefore, in this study, the mixed model is suitable to model the air polluter. The local influence of MTGWR for CO element for observation time morning and noon show that the two significantly influencing predictor variables are the air humidity (x_3) and the traffic velocity (x_4). This indicates that in the morning and noon, the influence of motor-vehicle exhaust and air humidity are very dominant in polluting the air with CO. For the night observation, almost all of the predictor variables influence the air pollution level, with the global influence factors of it are the area size of the urban forest (x_5) and the business center (x_7). To predict an

MGTWR model outside the observation nodes, we can use the Kriging estimation approach.

REFERENCES

[1] Ebtekar, 2006. Air Pollution induced asthma and aleterations in cytokine pattern. Allergy Asthma Immunol, 5(2):47-56. ISSN 1735-1502

[2] Environment Minister, (1997). Air Polluter Standard Index - APSI. Jakarta, Indonesia.

[3] Fahimi, M., Dharma, B., Fetararayani, D., Baskoro, (2012). Asosiasi antara polusi udara dengan IgE total serum dan tes faal paru pada polisi lalul intas. *Jurnal Penyakit Dalam* , 1-9.

[4] Robinson, D. and Lloyd, J. M. 2011. Increasing the accuracy of nitrogen dioxide (NO2) pollution mapping using geographically weighted regression & geostatistic . International Journal of Applied Earth Observation and Geoinformation, In Press 1-10. ISSNs: 0303-2434

[5] Gilbert, A. and Chakraborty, J. 2011. Using geographically weighted regression for environmental justice analysis: Cumulative cancer risks from air toxics in Florida. Social Science Research, 40 (2011): 273-286. DOI:10.1016/j.ssresearch.2010.08.006

[6] Mennis, J. 2006. Mapping the results of geographically weighted. The Cartographic, 43(2): 171-179. 171-179. DOI: 10.1179/000870406X114658

[7] Fotheringham A.S., Brunsdon C., Chartlon (2002). *Geographically Weighted Regression. the analysis of spatially varying relationship*. John Wiley and Sons, LTD. ISBN 0-471-49616-2

[8] Brunsdon, C., Fotheringham, A. and Charlton. 1999. Some notes on parametric significance tests for Geographically Weighted Regression. Journal of Regional Science, 39 (3): 497-524. DOI: 10.1111/0022-4146.00146

[9] Purhadi and Hasbi, Y. 2012. Mixed geographically weighted regression model (Case study: The percentage of poor households in Mojokerto 2008). European Journal of Scientific Research, 69 (2): 188-196. ISSN: 1450202X

[10] Walter J., Carsten R. and Jeremy W. Lichstein (2005) *Local and global approaches to spatiadata analysis in ecology*, Global Ecology and Biogeography 14 (1), 97-98. DOI: 10.1111/j.1466-822X.2004.00129.x

[11] Leung Y., Mei Chang L., and ZhangWen X. (2000a), Statistical test for spatial nonstationarity based on the geographically weighted regression model,Environment and Planning, 32 (5): 871–890. DOI:10.1068/a32117

[12] Cressie, N, A, C. (1992). Statistics For Spatial Data. Terra nova. 4 (5): 613-617. DOI: 10.1111/j.1365-3121.1992.tb00605.x

[13] Goovaerts, P. (1997). Geostatistics for Natural resource Evaluation. Oxford university Press, New York. ISBN-13 978-0-19-511538-3. ISBN 0-19-511538-4

[14] K Winarso, H Basuki, Fatmawati, 2014, Development Of Air Polluter Model For The Carbon Monoxide (CO) Element Based On Mixed Geographically Temporal Weighted Regression (MGTWR) Kriging, international Of Journal Academics Research, 6 (2): 193-198, DOI: 10.7813/2075-4124.2014/6-2,

Received: May 1, 2014

Development of Air Polluter Model for the Carbon Monoxide (CO) Element Based on Mixed Geographically Temporal Weighted Regression (MGTWR) Kriging

ORIGINALITY REPORT

7%	7%	4%	1%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	ijias.issr-journals.org Internet Source	1%
2	es.scribd.com Internet Source	1%
3	pdm-mipa.ugm.ac.id Internet Source	1%
4	docplayer.net Internet Source	1%
5	iopscience.iop.org Internet Source	1%
6	heanoti.com Internet Source	<1%
7	onlinelibrary.wiley.com Internet Source	<1%
8	fr.scribd.com Internet Source	<1%

9	Ghislain Geniaux. "Semi-Parametric Tools for Spatial Hedonic Models: An Introduction to Mixed Geographically Weighted Regression and Geoadditive Models", Hedonic Methods in Housing Markets, 2008 Publication	<1%
10	Xu, Yanqing, and Lei Wang. "GIS-based analysis of obesity and the built environment in the US", Cartography and Geographic Information Science, 2014. Publication	<1%
11	www.tandfonline.com Internet Source	<1%

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	On		

Development of Air Polluter Model for the Carbon Monoxide (CO) Element Based on Mixed Geographically Temporal Weighted Regression (MGTWR) Kriging

GRADEMARK REPORT

FINAL GRADE

/0

GENERAL COMMENTS

Instructor

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9

PAGE 10

PAGE 11