Spatial Analysis of Air Pollution in

Bangkok Metropolitan Region

Sunee Sammatat and Krisada Lekdee

Department of Mathematics and Statistics, Faculty of Science and Technology
Rajamangala University of Technology Phra Nakhon, Bangkok, Thailand

This article is distributed under the Creative Commons by-nc-nd Attribution License.
Copyright © 2019 Hikari Ltd.

Abstract

The objectives of this research are to propose a linear mixed model including spatial effects, to present the distribution of air pollutants including Sulphur dioxide (SO2), Nitrogen dioxide (NO2), Carbon monoxide (CO) and Ozone (O3) and Particulate matter (PM10) and to investigate the factors related to those air pollutants in Bangkok Metropolitan Region. The spatial effects were assumed to be a conditional autoregressive (CAR) model. The dependent variables were the levels of SO2, NO2, CO, O3 and PM10 in each area and the independents considered were industrial area, business area and suburban area. The results indicated that the factors influencing the levels of air pollution ranking from the highest to the lowest level were industrial area, suburban area and business area, respectively.

Keywords: Linear mixed model (LMM), Air pollution, Conditional autoregressive model (CAR) model, Spatial effects

1 Introduction

Air pollution refers to atmospheric conditions which substances or contaminants are contaminated in the form of gas, liquid or solid particles where either naturally occurring or caused by human actions. The amount of concentration higher than normal level affects health of people, animals, plants or other materials. Some important pollutants are Ozone (O3), Nitrogen dioxide (NO2), Carbon monoxide (CO), Sulfur dioxide (SO2) and Particulate matter smaller than 10 microns (PM10).
Ozone ($O_3$) is a toxic gas. If some areas are full of abnormally high ozone, it can affect human health. The ozone is dangerous for lungs, genetic systems, fetuses, decreased respiratory immunity, asthma and heart disease recurrence, reduced the amount of breath. Moreover, it increases amount of fluid in the lungs, causing to be irritation in the respiratory system. Nitrogen dioxide ($NO_2$) mostly comes from cars. It is dangerous for lungs and trachea. Carbon monoxide (CO) is the gas produced by incomplete combustion of carbon compounds. Moreover, it is the gas without color, taste and smell and lighter than general air. Red blood cells cannot be able to receive $O_2$ normally. Then the body gets less $O_2$. The heart must pump more blood to increase through the lungs. In general symptoms when receiving Carbon monoxide regularly are such as dizziness, uncomfortable breathing Nausea, vomiting headache and numbness. If someone receives a lot of Carbon dioxide, they will be shocked, unconscious or dead. Sulfur dioxide (SO2) is the toxic gas found in the atmosphere. It comes from burning of fuel in the transportation sector, industry, household and from nature. PM10 refers to atmospheric particulate matter (PM) that have a diameter of less than 10 micrometers. It is a small solid particle floating in the air. It comes from crushed objects. It is raised by the wind and falls to the ground. The time to fall down depends on the size and the weight of the dust particles. The source of dust shows the toxicity properties of dust, such as asbestos, lead, hydrocarbon, and radioactive substance. It effects health both directly and indirectly.

[1] stated that air pollution was a complex and difficult problem. The best way to solve the problems was to control the source of pollutants. If there was an appropriate level of pollution to the atmosphere, the air pollution was naturally diluted from the atmosphere. [2] found that if air pollution was released into the atmosphere in large quantities beyond the ability that the atmosphere could accommodate, the accumulation of pollution in the atmosphere caused harm to human and animal health. [3] found that air pollution affected the climate and harmed to the environment and ecology. For example, Sulfur dioxide (SO2) caused acid rain, Ozone (O3) caused by the reaction between hydrocarbon gases (HC) and nitrogen oxides (NO) from the exhaust of cars with sunlight as catalysts had a direct impact on the respiratory system of people and animals and plant growth. [4] stated that pollution was caused by nature or from humans. The pollution existed throughout the atmosphere in which the one near the earth surface would be relatively high affecting our lives.

Bangkok metropolitan is a large city with heavy traffic. With the construction of high-rise buildings, residential, public transportation systems, the growth of industrial plants in the metropolitan area, air pollution has increased. Therefore, the researcher is interested in studying the distribution of toxic gases and particulate matter to see that what areas in Bangkok metropolitan have high air pollution. A linear mixed model (LMM) was used and the factors to be considered were industrial area, business area and residential area. The benefits
Spatial analysis of air pollution in Bangkok metropolitan region

from this study is for planning, managing, campaigning and preventing air pollution in Bangkok metropolitan.

2 Methodology and application

2.1 Linear Mixed Model (LMM)

[5-6] described a linear mixed model as follows. Under a condition of knowing a random variable \( b_i \), assume a variable \( y_{it} \), \( i = 1, \ldots, m \) and \( t = 1, \ldots, n_i \) is independent and has a normal distribution with mean \( \mu_{it} \) and variance \( \sigma_{it}^2 \). That is \( y_{it} | b_i \sim N(\mu_{it}, \sigma_{it}^2) \).

A linear mixed model in which dependent variables are assumed to be a normal distribution is defined as

\[
\mu_{it} = x_i^T \beta + z_i^T b_i
\]

(1)

Where \( E(y_{it} | b_i) = \mu_{it} \), \( \beta \) is a fixed vector of size \( p \times 1 \) corresponding to the fixed effect (covariate) \( x_i \), \( b_i \) is the random effects of size \( q \times 1 \) which is related to the random variable \( z_i \) and \( y_{it} \) is the observed value.

2.2 Conditional autoregressive model (CAR)

[7] described the pattern of the CAR as follows. Let \( v = (v_1, \ldots, v_m)^T \) be the vector of random influences that vary by area \( i, \ i = 1, \ldots, m \). The conditional probability distribution of \( v_i \) is defined as

\[
v_i | v_{(-i)} \sim N\left( \frac{\sum_{j=1}^{m} w_{ij} v_j}{w_{ii} + \tau^2}, \frac{\tau^2}{w_{ii}} \right)
\]

(2)

where \( v_{(-i)} = \{v_j : j \neq i\} \), \( \tau_i^2 \) is the conditional variance. \( w_{ij} = 1 \) if \( i \) and \( j \) are adjacent where \( i \neq j \), \( w_{ij} = 0 \) if \( i \) and \( j \) are not adjacent and \( w_{ii} = \sum_j w_{ij} \).


2.3 Bayesian Model

[11] described the Bayesian method as follows. Let \( y = (y_1, \ldots, y_m)^T \) be the vector of observed values. \( \theta = (\theta_1, \ldots, \theta_i)^T \) is the vector of parameters. \( f(y | \theta) \) is the conditional probability density function of \( y \). When the conditions \( y \) are known, determine \( \pi(\theta) \) the priority distribution of \( \theta \). According to Bayes' rule, the posterior distribution function is
\[ p(\theta | y) = \frac{f(y | \theta) \pi(\theta)}{m(y)}. \]  

(3)

where \( m(y) = \int f(y | \theta) \pi(\theta) d\theta \) is a constant. The value of \( m(y) \) does not depend on parameters; therefore, Equation (3) can be written as \( p(\theta | y) \propto f(y | \theta) \pi(\theta) \). For the parameter estimation in (3), the numerical method is required because of the multidimensional dimension problem. The well-known numerical method is a Markov Chain Monte Carlo method (MCMC) using Gibbs sampling techniques [12].

2.4 Application
The data were extracted from the Pollution Control Department, Ministry of Natural Resources and Environment [13]. The monthly air pollutant data were collected from the 28 areas having monitoring stations in 2015. The air pollutants consisted of Ozone (O3), Nitrogen dioxide (NO2), Carbon monoxide (CO), Sulfur dioxide (SO2) and Particulate matter smaller than 10 microns (PM10). The dependent variables were the levels of each pollutant, and the independent variables were industry, business, suburban areas. The LMM with CAR spatial effects was used. The parameters were estimated using a Bayesian method via programming in OpenBUGS.

The LMM including spatial relationships was applied. Assume that \( y_{it} | b_i, v_i \sim N(\mu_i, \sigma_i^2) \) is the level of pollutant in the station \( i, i = 1, \ldots, 28 \) at month \( t = 1, \ldots, 60 \). The LMM model in (1) can be extended to include the CAR spatial relationship in (2) and a linear trend as follows.

\[ \mu_i = \beta_0 + \beta_1 \ast \text{Industry} + \beta_2 \ast \text{Business} + \beta_3 \ast \text{Suburb} + b_{it} + b_{2it} + v_i \quad (4) \]

where \( \beta_0, \beta_1, \beta_2, \beta_3 \) are coefficients of regression related to the factors, industry, business and suburban area, respectively. \( v_i \) is a spatial effect having CAR model in (2). \( b_{it} \) is a random effect between provinces and \( b_{2it} \) is a random effect within province. Under the Bayesian method, prior distributions are assumed to be non-informative which does not affect the posterior. \( b_{it} \sim N(0, \tau_{bi}^{-2}) \), \( b_{2it} \sim N(0, \tau_{b2i}^{-2}) \), \( \tau_y^2, \tau_{b1}^2, \tau_{b2}^2, \tau_v^2 \sim \text{InvGamma}(0.01,0.01) \). For parameter estimation, the Bayesian method was used via programming in OpenBUGS.

3. Results
The general characteristic of the data shows that the average monthly level of SO2, NO2, CO, O3 and PM10 are 2.88, 19.18, 0.71, 18.53 and 42.06, respectively, as shown in Table 1.
Spatial analysis of air pollution in Bangkok metropolitan region

Table 1 Mean of air pollution per month

<table>
<thead>
<tr>
<th>Types of pollution</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>SO2</td>
<td>2.88</td>
<td>2.62</td>
</tr>
<tr>
<td>NO2</td>
<td>19.18</td>
<td>9.97</td>
</tr>
<tr>
<td>CO</td>
<td>0.71</td>
<td>0.30</td>
</tr>
<tr>
<td>O3</td>
<td>18.53</td>
<td>8.48</td>
</tr>
<tr>
<td>PM10</td>
<td>42.06</td>
<td>15.93</td>
</tr>
</tbody>
</table>

3.1 Factors affecting the level of SO2

The factors affecting the level of SO2 are shown in Table 2.

Table 2 Factors affecting the level of SO2

<table>
<thead>
<tr>
<th>Factor</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>95% Credible Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$ (Intercept)</td>
<td>-0.7141</td>
<td>3.884</td>
<td>-10.07 7.913</td>
</tr>
<tr>
<td>$\beta_1$ (Industry)</td>
<td>23.07</td>
<td>7.838</td>
<td>6.489 37.02</td>
</tr>
<tr>
<td>$\beta_2$ (Business)</td>
<td>-4.09</td>
<td>8.46</td>
<td>-19.91 7.721</td>
</tr>
<tr>
<td>$\beta_3$ Suburb (Reference)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

From Table 3, factors that influence SO2 level are industrial area and business area where the suburban area is a reference area. In the industrial area the level of SO2 is 23.07 higher than the one in the suburban area. In the business area the level of SO2 is 4.09 less than the one in the suburban area. The estimation of the influence of the region on SO2 ranking from the highest level to the lowest level are industrial area, suburban area and business area, respectively.

3.2 Factors affecting the level of NO2

The factors affecting the level of NO2 are shown in Table 3.

Table 3 Factors affecting the level of NO2

<table>
<thead>
<tr>
<th>Factor</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>95% Credible Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_0$ (Intercept)</td>
<td>15.28</td>
<td>5.138</td>
<td>7.553 24.51</td>
</tr>
<tr>
<td>$\beta_1$ (Industry)</td>
<td>16.69</td>
<td>7.602</td>
<td>0.7097 28.31</td>
</tr>
<tr>
<td>$\beta_2$ (Business)</td>
<td>-2.381</td>
<td>7.034</td>
<td>-14.34 6.874</td>
</tr>
<tr>
<td>$\beta_3$ Suburb (Reference)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

From Table 3, factors that influence the NO2 level are industrial area and business area where the suburban area is a reference area. In the industrial area the
level of NO₂ is 16.69 higher than the one in the suburban area. In the business area the level of NO₂ is 2.381 less than the one in the suburban area. The estimation of the influence of the area on NO₂ ranking from the highest level to the lowest level are industrial area, suburban area and business area, respectively.

3.3 Factors affecting the level of CO

The factors affecting the level of CO are shown in Table 4.

**Table 4 Factors affecting the level of CO**

<table>
<thead>
<tr>
<th>Factor</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>95% Credible Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_0 ) (Intercept)</td>
<td>-0.1581</td>
<td>0.4915</td>
<td>-1.133</td>
</tr>
<tr>
<td>( \beta_1 ) (Industry)</td>
<td>4.694</td>
<td>1.692</td>
<td>2.161</td>
</tr>
<tr>
<td>( \beta_2 ) (Business)</td>
<td>-3.847</td>
<td>1.013</td>
<td>-5.926</td>
</tr>
<tr>
<td>( \beta_3 ) Suburb (Reference)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

From Table 4, factors that influence CO level are industrial area and business area where the suburban area is a reference region. In the industrial area the level of CO is 4.694 higher than the one in the suburban area. In the business area the level of CO is 3.847 less than the one in the suburban area. The estimation of the influence of the area on CO ranking from the highest level to the lowest level are industrial area, suburban area and business area, respectively.

3.4 Factors affecting the level of O₃

The factors affecting the level of O₃ are shown in Table 5.

**Table 5 Factors affecting the level of O₃**

<table>
<thead>
<tr>
<th>Factor</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>95% Credible Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_0 ) (Intercept)</td>
<td>16.22</td>
<td>5.399</td>
<td>7.764</td>
</tr>
<tr>
<td>( \beta_1 ) (Industry)</td>
<td>23.43</td>
<td>8.568</td>
<td>8.31</td>
</tr>
<tr>
<td>( \beta_2 ) (Business)</td>
<td>-9.866</td>
<td>6.887</td>
<td>-22.77</td>
</tr>
<tr>
<td>( \beta_3 ) Suburb (Reference)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

From Table 5, factors that influence O₃ level are industrial area and business area where the suburban area is a reference region. In the industrial area the level of O₃ is 23.43 higher than the one in the suburban area. In the business area the level of O₃ is 9.866 less than the one in the suburban area. The estimation of the influence of the region on O₃ ranking from the highest level to the lowest level are industrial area, suburban area and business area, respectively.
3.5 Factors affecting the level of PM10

The factors affecting the level of PM10 are shown in Table 6. From Table 6, factors that influence PM10 level are industrial area and business area where the suburban area is a reference region. In the industrial area the level of PM10 is 25.52 higher than the one in the suburban area. In the business area the level of PM10 is 12.32 less than the one in the suburban area. The estimation of the influence of the region on O3 ranking from the highest level to the lowest level are industrial area, suburban area and business area, respectively.

Table 6 Factors affecting the level of PM10

<table>
<thead>
<tr>
<th>Factor</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>95% Credible Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_0 ) (Intercept)</td>
<td>39.42</td>
<td>8.806</td>
<td>26.32</td>
</tr>
<tr>
<td>( \beta_1 ) (Industry)</td>
<td>25.52</td>
<td>12.21</td>
<td>3.352</td>
</tr>
<tr>
<td>( \beta_2 ) (Business)</td>
<td>-12.32</td>
<td>13.7</td>
<td>-32.77</td>
</tr>
<tr>
<td>( \beta_3 ) Suburb (Reference)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

4. Discussion

For this paper, the LMM model with spatial effects was used to analyze air pollution data. The model was suitable because it considered the spatial relationship of the data. The Spatial relationship arises from the fact stating that anything that is close together will have more relevance than what is far away. The level of the air pollution in the areas which are near each other inevitably results from the spatial relationship.

When the spatial influence is added into the model, it makes the model more complicated. Estimating parameters commonly used such as maximum likelihood (ML) cannot be used. Therefore, the Bayesian method and Markov Chain Monte Carlo process (MCMC) were used in this study. The Bayesian method has several advantages. The first one the parameter can be estimated even though the sample is small. The ML method is used with large samples only. The MCMC process is a sampling by using a computer. Although the form of probability distribution of functions was not known, the parameters such as mean and standard deviation can be estimated by using a random number from various functions. One of the well known method of getting the random numbers is Gibb sampling technique.

The factors influencing on the level of SO2, NO2, CO, O3 and PM10 are industrial area, business area suburban area. The estimation of the influence of the area on the level of all pollutants ranking from the highest level to the lowest level are industrial area, suburban area and business area, respectively. Air pollution from industrial sources is caused by burning fuel and production process affecting
the air quality in the atmosphere. It may result in affecting the health of people in the community in general or causing annoyance. There are 3 types of fuel used in the industry which are solid fuels, liquid fuels such as fuel oil and diesel and gaseous fuels including natural gas and liquefied petroleum gas (LPG).

Air pollutants arise from the combustion of various fuels, including Carbon monoxide (CO), Sulphur dioxide (SO2) and Nitrogen dioxide (NO2). It is found that the amount of air pollutants released into the atmosphere increases every year in accordance with the increased fuel consumption in the industrial production processes. To solve the problem from the pollution sources, accelerating the awareness of industrial plant operators and managing the factory to be environmentally friendly are needed. To be responsible to the environment, there is a need to integrate both directly and indirectly in a systematic way.

5 Conclusion

A linear mixed model including spatial effects was used to present the distribution of air pollutants including SO2, NO2, CO, O3 and PM10. The factors related to those air pollutants in Bangkok Metropolitan Region were investigated. The spatial effects were assumed to be the CAR model. The results indicated that the factors influencing the level of air pollution ranking from the highest to the lowest level were industrial area, suburban area and business area, respectively.

Acknowledgements. We gratefully thank the Faculty of Science and Technology, the Institute of Research and Development, and Rajamangala University of Technology Phra Nakhon (rmutp.ac.th) for their technical and financial supports.

References


**Received: November 27, 2019; Published: December 20, 2019**