Diurnal Fluctuations of Ozone Concentrations and its Precursors and Prediction of Ozone Using Multiple Linear Regressions

NOR AZAM RAMLI, NURUL ADYANI GHAZALI & AHMAD SHUKRI YAHAYA

ABSTRACT

The chemical reaction of pollutants emitted into the atmosphere leads to a variety of oxidized products, which are commonly referred to as secondary pollutants. Ground level ozone is a known secondary photochemical pollutant of major importance possessing detrimental effects on health, agriculture, natural/urban ecosystems and materials. Ozone ($O_3$) can irritate lung airways and cause inflammation much like sunburn. Hourly and monthly variations of $O_3$ and their precursors – nitrogen oxides ($NO_x$) and meteorological parameters (temperature and wind speed) were presented using time series plots. Possibility of employing multiple linear regression models as a tool for prediction of $O_3$ concentration was also tested. Measurement was performed continuously in 2005 at two sampling stations located in the metropolitan area of Malaysia. Results indicated that the formation of $O_3$ in the study area was influenced by $NO_x$ precursors and meteorological conditions. The hourly variation showed maximum $O_3$ concentrations were recorded between 1300 to 1400 hours, while $NO_x$ and nitrogen dioxide ($NO_2$) exhibited two maxima, at 0800-1000 and 2000-2100 hours. The daily cycle of highest $O_3$ concentrations were revealed a lower night level and inverse relations between $O_3$ and $NO_x$. This is clear evidence of photochemical formation of $O_3$. Temperature has the highest influence to the high $O_3$ concentrations.

Key words: Ground-level ozone, time series plot, multiple linear regression

ABSTRAK


Kata kunci: Ozon aras permukaan, plot siri masa, regresi linear berganda
INTRODUCTION

Ozone ($O_3$) is a gas composed of three atoms of oxygen and occurs both in the Earth’s upper atmosphere and at ground level. This secondary pollutant, is generated in the presence of solar ultraviolet radiation through a complex series of photochemical reactions involving VOC combining with a group of air pollutants known as nitrogen oxides ($NO_x$), originating from both anthropogenic (e.g., industry and vehicle emissions) and biogenic sources (e.g., forest, and soil) (EPA 2007). There are no significant primary emissions of $O_3$ into the atmosphere and all the $O_3$ found has been formed by chemical reactions that occur in the air (WHO 2009). Ozone is mainly formed indirectly by the action of sunlight on nitrogen dioxide. As a result of the various reactions that take place in the atmosphere, $O_3$ tends to build up downwind of urban centers where most of the $NO_x$ is emitted from vehicles (Andrew & Thomas 2003). Ozone, therefore, is a secondary photochemical pollutant that is not polluting in its own right. This is the main reason why the presence of $O_3$ is such a serious environmental problem that is difficult to control and predict (Abdul-Wahab et al. 2005).

Ozone is naturally present in relatively large concentrations in the stratosphere, an upper atmospheric layer. Stratospheric $O_3$ should not be considered a photochemical air pollutant; a phrase that’s used which should be restricted to $O_3$ issues associated with the lower atmosphere or troposphere (Freedman 1995). It is known that ground level ozone is responsible for various adverse effects on human health, climate and vegetation. In addition, high $O_3$ levels not only play a role in causing damage to plant species, various natural materials and manufactured goods but also lead to the damage of lung tissues in humans (Abdul-Wahab et al. 2005). Therefore, understanding on the formation of ground level ozone is very important. It can be used as a basis for issuing advance warning to the public prior to $O_3$ reaching the peak concentration of a day (Ghazali et al. 2009).

CHARACTERISTICS AND TRANSFORMATION OF OZONE

The formation of $O_3$ in the upper atmosphere was explained as a chemical process involving radiant energy ($h\nu$) from the sun. Certain wavelengths in the ultraviolet range are able to break oxygen ($O_2$) into two low-energy oxygen atoms, $O$ ($^3P$) (Ghazali et al. 2009).

Radiant energy was shown to break nitrogen dioxide ($NO_2$) into nitric oxide (NO) and $O$ ($^3P$). When $NO_2$ is exposed to UV light ($\lambda<424$ nm), an oxygen atom and a NO molecule are generated via the $NO_2$ photolysis reaction (Lawrence 2003). As a result, $O_3$ is formed as the product of the reaction between $O$ ($^3P$) and $O_2$. The $O_1$ that forms, however, can react quickly with NO to produce $NO_2$ and $O_2$.

The steady-state concentration of $O_3$ as is defined by Seinfeld and Pandis (2006) are:

$$[O_3] = \frac{J[NO_2]}{K^*[NO]} \quad (1)$$

The characteristics and transformation of ozone have been discussed in details in Ghazali et al. (2009).

STUDY AREA

Selangor is one of the most fascinating states in Malaysia, being the most developed in the country and among the best in infrastructure and telecommunication facilities in the region. The state covers about 125,000 km$^2$ and has an average temperature of 26$^\circ$ C. Selangor's geographical position in the centre of Peninsular Malaysia contributed to the state's rapid
development as Malaysia's transportation and industrial hub, which in turn attracts migrants from other country. The influx of illegal immigrants, particularly from Indonesia, has further contributed to Selangor's rapid population growth.

Shah Alam (Coordinates: 3°5′00″N 101°32′00″E) is a city in Malaysia, about 25 kilometers west of the country's capital, Kuala Lumpur. Shah Alam is the state capital of Selangor (Figure 1). The total area for Shah Alam covers about 290.3 km² with a total population of 584,340 (July 2006). The monitoring station for Shah Alam was located at Sekolah Menengah Kebangsaan TTDI Jaya, Shah Alam, with coordinate N 03° 05′ 19.819″ and E 101° 34′ 9.49″. The station is surrounded by residential areas. However, it is not far from industrial zone (Ghazali 2009). The station of Shah Alam is located at latitude 3°06′ 17″ North of the equator and longitude 101° 33′ 22″ East of the prime meridian.

Figure 1. The location of monitoring station for Shah Alam

Figure 2. The location of monitoring station for Nilai

Nilai is a town located in Negeri Sembilan (Figure 2). It is geographically located at latitude 2°49′ 00″ North of the equator and longitude 101°48′ 00″ East of the prime meridian. Nilai covers an area of 738 km² from 6645 km² of the state of Negeri Sembilan. It is located in
the south part of Malaysia and is a heavily industrialized area. Climatically, Negeri Sembilan represents a typical tropical monsoon characterized by uniformly high temperatures between 27°C – 30°C in daytime and 22°C – 24°C in nighttime. Due to its proximity, and connection through the Malayan Railway Administration or known as ‘Keretapi Tanah Melayu Berhad’ to Kuala Lumpur, Putrajaya, and Kuala Lumpur International Airport (KLIA), it is a rapidly growing town.

Development projects can be seen as one drives around Nilai. The growth in industrial sectors have increased air pollution problem in Nilai. Apart from that, the growth in industrial activities is normally accompanied by the rapid increase in population and traffic density. The population distributions in Nilai had also growth every year due to the increase in vacancies offered by the industries. The total number of population estimated in Nilai for 2006 is about 458,300 from 1,004,807 total population of the state of Negeri Sembilan.

In addition to that, numbers of mobile sources that include passenger cars, taxis, buses, motorcycles, vans and lorries have also increased every year. The total number of mobile sources in 2005 is 640,322. In Nilai, the air quality was moderate for 84 percent of the time, 11 percent good and 5 percent unhealthy (DoE 2007). The station of Nilai is located at latitude 2°49’ 18” North of the equator and longitude 101º 48’ 41” East of the prime meridian.

RESEARCH METHOD

Multiple regression analysis is one of the most widely used methodologies for expressing the dependence of a response variable on several predictor variables (Abdul-Wahab et al. 2004; Ghazali et al. 2008; Ghazali et al. 2009). The application of multiple linear regression techniques allows formulation of explicit equations that are simple and can be used to better understand the processes involved in O₃ formation (Barrero et al. 2005). Our model is aimed at predicting hourly O₃ concentration maxima several hours in advance.

Data were collected in an urban area by continuous air monitoring stations in Shah Alam, Selangor and Nilai, Negeri Sembilan from January 2005 to December 2005. These facilities provided hourly concentrations of pollutants such as O₃, NO₂, NO, NOₓ, non-methane hydrocarbons (NmHC) and total hydrocarbons (THC) as well as meteorological parameters such as temperature, wind speed, wind direction and humidity. Samples of O₃ concentrations were collected by using UV Absorption Ozone Analyzer Model 400A (EPA Approved EQOA 0992-087). The model 400A UV Absorption Ozone Analyzer is a microprocessor controlled analyzer that uses a system based on the Beer-Lambert law for measuring low ranges of ozone in ambient air. Samples of NOₓ concentrations were collected using chemiluminescent NO/NO₂/NOₓ Analyzer Model 200A (EPA Approval RFNA 1194-099). The model 200A uses the proven chemiluminescence’s detection principle, coupled with state of the art microprocessor technology to provide the sensitivity, stability and ease of use needed for ambient or dilution CEM monitoring requirement.

Time series plots of diurnal cycle for both O₃ and NO₂ concentrations were plotted using Microsoft Excel. The mathematical model was developed using multiple linear regression for variable introducing using SPSS Version 11.5. Coefficient of determination (R²) and root mean squared error (RMSE) were used as performance indicators. R² measured the variability of the observed data explained by the model while RMSE indicated the difference between observed and predicted concentrations (Ghazali et al. 2009).

RESULTS AND DISCUSSIONS

Descriptive Statistics

Table 1 gives summaries of the descriptive statistics for each pollutant and meteorological parameters for Shah Alam and Nilai.
Table 1. Summary of descriptive statistics for Shah Alam and Nilai

<table>
<thead>
<tr>
<th>Descriptive Statistics</th>
<th>O\textsubscript{3} (ppm)</th>
<th>NO\textsubscript{2} (ppm)</th>
<th>Temperature (°C)</th>
<th>Wind Speed (m/s)</th>
<th>Humidity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SHAH ALAM</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.022</td>
<td>0.017</td>
<td>27.8</td>
<td>4.51</td>
<td>78.1</td>
</tr>
<tr>
<td>Median</td>
<td>0.011</td>
<td>0.016</td>
<td>26.3</td>
<td>3.69</td>
<td>83.1</td>
</tr>
<tr>
<td>Std.Deviation</td>
<td>0.025</td>
<td>0.011</td>
<td>3.87</td>
<td>2.98</td>
<td>14.3</td>
</tr>
<tr>
<td><strong>Maximum</strong></td>
<td><strong>0.152</strong></td>
<td><strong>0.080</strong></td>
<td><strong>21.7</strong></td>
<td><strong>17.7</strong></td>
<td><strong>35.0</strong></td>
</tr>
<tr>
<td><strong>NILAI</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.019</td>
<td>0.012</td>
<td>27.1</td>
<td>6.98</td>
<td>75.9</td>
</tr>
<tr>
<td>Median</td>
<td>0.009</td>
<td>0.010</td>
<td>25.9</td>
<td>6.15</td>
<td>78.1</td>
</tr>
<tr>
<td>Std.Deviation</td>
<td>0.019</td>
<td>0.008</td>
<td>3.55</td>
<td>4.19</td>
<td>12.5</td>
</tr>
<tr>
<td><strong>Maximum</strong></td>
<td><strong>0.119</strong></td>
<td><strong>0.062</strong></td>
<td><strong>37.6</strong></td>
<td><strong>26.3</strong></td>
<td><strong>95.0</strong></td>
</tr>
</tbody>
</table>

The mean for both 2 pollutants are higher than the median, indicated that there were extreme event occurred. Mean O\textsubscript{3} concentration for Shah Alam and Nilai are 0.022 ppm and 0.019 ppm respectively and it is below the annual O\textsubscript{3} guideline value of 0.1 ppm (for 1 hour averaging time). However, the maximum hourly O\textsubscript{3} concentration recorded for the whole year is 0.152 ppm for Shah Alam and 0.119 ppm for Nilai. This value is higher than the Malaysia Ambient Air Quality Guidelines for 1-hour averaging time which is 0.1 ppm. This value also higher than other countries guideline such as USA with 0.12 ppm for 1-hour averaging time (US EPA 2009) and WHO guideline with 0.05 ppm for 8-hour averaging time (WHO 2009). The mean and maximum hourly NO\textsubscript{2} is also below the guideline which is 0.080 ppm (0.17 ppm for 1-hour averaging time).

**Monthly Average Mean Concentrations**

The monthly average mean concentrations for both O\textsubscript{3} and NO\textsubscript{2} from January to December 2005 are plotted using multiple line plots for Shah Alam and Nilai are shown in Figure 3.

![Figure 3. Plot of monthly average mean O\textsubscript{3} and NO\textsubscript{2} concentrations in Shah Alam and Nilai 2005](image-url)
From the plot we can clearly shows that the highest monthly O\textsubscript{3} concentration recorded in 2005 for Shah Alam were found in March (0.028 ppm) while the lowest O\textsubscript{3} concentration are found in June and July (0.02 ppm). On the other hand, the highest monthly NO\textsubscript{2} concentrations are recorded in March (0.021 ppm) and the lowest monthly NO\textsubscript{2} concentrations are found in December (0.015 ppm). However, the highest monthly O\textsubscript{3} concentration for Nilai was recorded in February (0.0268 ppm) while the lowest monthly O\textsubscript{3} concentration is found in November (0.0139 ppm). Besides, the highest monthly NO\textsubscript{2} concentrations are recorded in June (0.0143 ppm) and the lowest monthly NO\textsubscript{2} concentrations are found in January (0.0084 ppm).

**Roll-back Means**

There are occasions where the magnitude of emission reductions needs to be quickly estimated. The ‘roll-back’ equation can be used to estimate a factor R which is by how much emissions should be reduced to achieve a given reduction in the annual mean concentration (it assumes the background concentration is constant) (Seinfeld & Pandis 2006).

$$R = \frac{\text{mean} - \text{guidelines}}{\text{mean} - \text{background}}$$

where mean is the current annual mean of the pollutant concentration, guidelines is the annual mean corresponding to the air quality standard and background concentration assumed to be constant. This approach may find application in assessing the control needed for sources influencing an air quality management area. The equation means that a measure of the reduction can be obtain without dispersion modeling. Table 2 presents the roll-back means for Shah Alam and Nilai.

<table>
<thead>
<tr>
<th>Sites</th>
<th>Pollutants</th>
<th>Ozone</th>
<th>Nitrogen Dioxide</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHAH ALAM</td>
<td>Mean</td>
<td>0.0220</td>
<td>0.0170</td>
</tr>
<tr>
<td></td>
<td>Guidelines</td>
<td>0.1000</td>
<td>0.1700</td>
</tr>
<tr>
<td></td>
<td>Background</td>
<td>0.0311</td>
<td>0.0026</td>
</tr>
<tr>
<td>Roll-back means</td>
<td><strong>8.5714</strong></td>
<td>-10.6250</td>
<td></td>
</tr>
<tr>
<td>NILAI</td>
<td>Mean</td>
<td>0.0190</td>
<td>0.0115</td>
</tr>
<tr>
<td></td>
<td>Guidelines</td>
<td>0.1000</td>
<td>0.1700</td>
</tr>
<tr>
<td></td>
<td>Background</td>
<td>0.0311</td>
<td>0.0026</td>
</tr>
<tr>
<td>Roll-back means</td>
<td><strong>6.6942</strong></td>
<td>-17.8089</td>
<td></td>
</tr>
</tbody>
</table>

Mean concentration for ozone in Shah Alam and Nilai are almost same and it is indicate that the value is at alert level. From table, the roll-back means for Shah Alam are (O\textsubscript{3} = 8.5714, NO\textsubscript{2} = -10.6250) and for Nilai are (O\textsubscript{3} = 6.6942, NO\textsubscript{2} = -17.8089). The positive roll-back means of O\textsubscript{3} for both sites indicate that the background value is above the mean. Hence, care should be taken to ensure it does not go near to the guideline value more than the permissible occurrences in one year. The negative roll-back means of NO\textsubscript{2} for both sites indicate that the mean concentration is well below guideline value; hence pollution is not serious in this area.
Diurnal Fluctuations of Ozone and its Precursors Nitrogen Dioxide

The average diurnal profiles of \( \text{O}_3 \) and \( \text{NO}_2 \) concentrations for Shah Alam are shown in Figure 4. Diurnal profile for March 2005 and June 2005 represent the dry season while diurnal profile for September 2005 and December 2005 represent the wet season.

![Diurnal Fluctuations of Ozone and its Precursors Nitrogen Dioxide](https://via.placeholder.com/150)

Figure 4. Diurnal fluctuations of ozone concentrations and its precursors Nitrogen Dioxide concentrations for Shah Alam.
The diurnal variation of $O_3$ shows almost the same pattern at all five sites indicating a typical pattern for polluted urban areas characterized by high concentrations during daytime and low concentrations during late night and early morning. In contrast, the precursors, NO and NO$_2$ show an almost opposite diurnal variation pattern to $O_3$, which is characterized by high concentrations during night and early morning and low concentrations during daytime, especially noon and afternoon. An increase in the concentration of $O_3$ is closely corresponding to decrease in the concentrations of the precursors. The similar diurnal pattern in $O_3$ and precursors was observed in numerous urban areas around the world (Tu et al. 2007; Mazzeo et al. 2005; Dueñas et al. 2002; Zhang & Oanh 2002; Lal et al. 2000).

Ozone exhibits strong day-to-day variation and is sometimes virtually undetectable. Low wind speeds (<3 m/s) promote the build-up of high local ozone concentrations. The availability of NO and NmHC from local emissions such as vehicles and other anthropogenic activities must be considered to be among the contributing factors. Changes in $O_3$ amounts are closely linked to temperature, with colder temperatures resulting in more polar stratospheric clouds and lower $O_3$ levels. $O_3$ generates heat in the stratosphere, both by absorbing the sun’s ultraviolet radiation and by absorbing upwelling infrared radiation from the lower atmosphere. Besides the photochemical reactions effected by solar radiation and variations in anthropogenic emissions, boundary layer processes and meteorological parameters also play important roles in the process of diurnal variations in $O_3$ and its precursors.

The average diurnal profiles of $O_3$ for both dry and wet season shows almost the same pattern with suggest that $O_3$ concentrations have lower concentrations at night. From Figures 4 and Figure 5 it was observed that, at 0100 hour until 0800 hour, the concentrations of $O_3$ are low due to the absence of sunshine and low NO$_2$. During morning hours, lower boundary layer height reduces the mixing processes between the $O_3$ surface layer (the bottom 10% of the boundary layer) and the $O_3$ upper layer largely in comparison to the noontime and contribute to early morning low $O_3$ levels (Lal et al. 2000). $O_3$ will increase following the high NO$_2$ due to traffic emissions during morning peak hours around 0800 to 0900 hour, the accumulated NO$_2$ concentration reached the maximum at 1000 hour. This is then followed by marked decrease in NO$_2$ concentrations towards midday until 1600 hour and then increased again in the late evening before reached second peak around 2100-2300 hours.

Conversely, after sunrise, sunlight intensity starts to increase and $O_3$ concentrations reaches peak concentrations at 1300 to 1500 hour coincided with the high UVB intensity and temperature. This was expected for a photochemical produced species such as $O_3$. Hence, production of $O_3$ by photochemical reaction is evidence. Moreover, $O_3$ concentrations will decreasing rapidly after peak until evening and then keeps decreasing gradually, maintaining low values over night due to the lack of solar radiation.

The diurnal variations in NO, NO$_2$ and $O_3$ concentrations during the four season observed exhibited a typical photochemical pollution situation. The peak was first reached by NO followed by NO$_2$ and finally $O_3$ several hours later. Subsequently, the after office traffic injected an additional burden of NO into the atmosphere which scavenged the remaining traces of $O_3$ by early evening and followed by the NO and other primary pollutants reaccumulated for the remainder of the night. This phenomenon seemed to suggest that the photochemical smog formation was able to occur in the city. Nonetheless, while humidity may affect oxidant formation both directly and indirectly, the magnitude of these effects has not been conclusively established.

The average diurnal profiles of $O_3$ and NO$_2$ concentrations for Nilai are shown in Figure 5. Diurnal profile for March 2005 and June 2005 represent the dry season while diurnal profile for September 2005 and December 2005 represent the wet season.

The overall average diurnal profile of $O_3$ for Nilai shows the same pattern with Shah Alam. Ozone production is strongly influenced by available of UVB radiation and air temperature. Night time concentration of $O_3$ is relatively low because the production of $O_3$ depends on the intensity of sunlight. After sunrise, sunlight intensity starts to increase, and
reaches a maximum value between hour 1300 hours and 1400 hours. From 0800 -1000 hours and 2000 – 2100 hours, the number of cars driven on the road is at its highest for the day, so the emission of pollutants from automobiles is maximized during these two periods. The maximum concentration of $O_3$ occurs about 1400 hours and drop just steadily to a low in the evening. On the other hand, the $NO_2$ concentrations exhibited two maxima, at about 0800-1000 hours and 2000-2100 hours, in correspondence with peaks in vehicular traffic density.

Figure 5. Diurnal fluctuations of ozone concentrations and its precursors Nitrogen Dioxide concentrations for Nilai
Comparison of diurnal $O_3$ profiles with $NO_2$ shows that minimum $O_3$ concentrations were coincident with concentration maxima of the $NO_2$ concentrations. In a typical day, primary pollutant concentrations exhibited two maxima in correspondence with peak in vehicular traffic density and minimal concentrations are observed in the first hours of the afternoon, reflecting favourable dispersion conditions of the atmosphere coupled with reduced vehicular traffic input. $NO_2$ have been destroyed from further chemistry by the present of sunlight to form $O_3$. This shows that the formation of ozone is chemically linked to the emissions of nitrogen oxides through oxidation of gaseous precursors such as $NO_2$. The occurrences of factors conducive for $O_3$ production, such as high atmospheric stability, high insolation and temperature, coupled with the availability of $NO_x$ (> 90% constituted $NO_2$) and NmHC might have led to a higher rate of ozone formation.

### Multiple Regression Analysis

A multiple linear regression modeling of ozone concentrations was carried out in order to find predictive equations for the concentration with the pollutant and meteorological variables as the predictor variables. Based on the information extracted from the correlation plots, a number of linear regression models containing several different combinations of predictor variables were developed and evaluated. Coefficient of determination ($R^2$) is used to determine the best model. The model with a high value of $R^2$ that does not contain too many variables is considered to be the best model.

Using regression analysis, only 5 factors, namely nitrogen oxides concentration, hydrocarbons concentrations (THC and NmHC), temperature, wind speed and previous hour $O_3$ concentrations ($O_{3,t-1}$) are observed to be statistically significant and theoretically meaningful and are therefore selected. Table 3 presents, for Shah Alam and Nilai, the results of the regression analysis with $O_3$ concentrations as the dependent variable and other pollutants and meteorological parameters as independent variables.

**Table 3. Linear regression model for prediction of $O_3$ using the original independent variables for Shah Alam and Nilai**

<table>
<thead>
<tr>
<th>Predictors</th>
<th>Constant</th>
<th>$[O_{3,t-1}]$</th>
<th>$[NO_2]$</th>
<th>$[NmHC]$</th>
<th>Temp</th>
<th>WS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SHAH ALAM</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R square (0.897)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted R square (0.897)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standardized coefficient</td>
<td>0.661</td>
<td>-0.084</td>
<td>-0.010</td>
<td>0.451</td>
<td>-0.081</td>
<td></td>
</tr>
<tr>
<td>Estimated regression coefficient</td>
<td>-0.075</td>
<td>0.662</td>
<td>-0.172</td>
<td>-0.001</td>
<td>0.003</td>
<td>-0.001</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.004</td>
<td>0.007</td>
<td>0.011</td>
<td>0.001</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td><strong>NILAI</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R square (0.890)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted R square (0.890)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standardized coefficient</td>
<td>0.745</td>
<td>-0.141</td>
<td>-0.018</td>
<td>0.222</td>
<td>-0.081</td>
<td></td>
</tr>
<tr>
<td>Estimated regression coefficient</td>
<td>-0.022</td>
<td>0.737</td>
<td>-0.321</td>
<td>-0.001</td>
<td>0.003</td>
<td>-0.001</td>
</tr>
<tr>
<td>Standard Error</td>
<td>0.002</td>
<td>0.005</td>
<td>0.012</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

The $R^2$ gives the proportion of the variation in the ozone concentration explained by the independent variables in the model. The table shows that, when the six best variables are fitted to the ozone data, the values of $R^2$ and adjusted $R^2$ are about the same for both sites (Shah Alam = 0.897 and 0.897, Nilai = 0.890 and 0.890). This means that, close to 90% of the variation in the ozone concentrations are explained by the independent variables listed in the tables.

From Table 3, the highest value of standardized coefficient that shows the greater effect to $O_3$ for model Shah Alam is a previous hour $O_3$ concentration which is 0.661 followed by temperature (0.451), $NO_2$ (-0.084), $ws$ (-0.081) and NmHC (-0.010). While for Nilai, the
highest value of standardized coefficient is a previous hour \( O_3 \) concentration (0.745) followed by temperature (0.222), \( NO_2 \) (-0.141), \( ws \) (-0.081) and \( NmHC \) (-0.018). It is shows that a previous hour \( O_3 \) concentration is the strongest variables that effect \( O_3 \) concentrations for both models. The final revised model was selected and the prediction model for Shah Alam and Nilai are shown in equation (4) and (5) respectively:

\[
O_3 = -0.075 - 0.172NO_2 + 0.003Temp - 0.001ws - 0.001NmHC + 0.662O_{3,t-1} \quad (4)
\]

\[
O_3 = -0.056 - 0.315NO_2 + 0.001NmHC - 0.001THC + 0.002Temp + 0.699O_{3,t-1} \quad (5)
\]

The fact that nitrogen dioxide and previous hourly ozone concentration influenced the ozone prediction values was evident. The model also suggested the influence of weather parameter i.e. windspeed and temperature on ozone concentrations. High insolation will result in an increase in temperature which will eventually provide suitable conditions for conversions of \( NO_2 \) into \( O_3 \). Increase in windspeed will promote dispersions hence reducing the ability of \( O_3 \) formations in ambient urban air.

Predictions of \( O_3 \) concentration have been done by using the predicted models to determine the accuracy of the models. The predicted hourly \( O_3 \) concentrations against the observed hourly \( O_3 \) concentrations for the model derived for Shah Alam and Nilai were plotted in Figure 6. This plotted were used to determine a goodness-of-fit of the model. The regression lines (A and C) indicate 95% confidence interval of the model was also drawn.

From Figure 6, it was found that, 89.7% of the point falls within this range suggesting the accuracy of the model developed at \( R^2 = 0.897 \). While the accuracy of the predicted model for Nilai is 89.0%. The \( R^2 \) values are 0.890.

![Figure 6. Scatter plot of predicted \( O_3 \) against observed \( O_3 \) for Shah Alam and Nilai 2005](image)

CONCLUDING REMARKS

This study presents a multiple linear regression model for hourly \( O_3 \) concentrations. The model allows predicting \( O_3 \) concentrations maxima several hours in advance. Our model based on the
hour of the day and the day of the year, the relations between \( O_3 \) temperature, nitrogen oxides and hydrocarbons and previous hour \( O_3 \) levels. These variables take into account for the seasonal influences, photochemical and inertial components of \( O_3 \) levels, respectively. A traditional pattern of transformation of \( NO_2 \) into \( O_3 \) in atmosphere was justified through this study. It was observed that \( O_3 \) concentrations were negatively correlated to \( NO_2 \). This result was expected and confirmed the theory that \( NO_2 \) act as the precursors of \( O_3 \). This negative correlation is typical of on urban areas where the diurnal cycle of \( O_3 \) is impacted by traffic which releases \( NO_2 \), its precursor.

The result of fitting the best regression models on the \( O_3 \) data using six of the independent variables gave about the same values of adjusted coefficient of determination of around 0.89 for both sites. The prediction model shows the importance of previous hour \( O_3 \) concentrations, \( NmHC \) and temperature on determining the concentrations of \( O_3 \). These variables were used as a predictor variable in two separates multiple regression analyses to obtain the final models for the ozone concentrations.

Despite being linear model and therefore unable to account for non-linear behavior, it has an acceptable performance which is comparable to other, more complex, models found elsewhere in the literature. The model is easy to implement in association with automated pollution monitoring station systems, since the involved variables are commonly automatically measured therefore it will be a useful tool to public health protection because it can provide early warnings to the population. The availability of accurate and real time predictions of air pollutant concentrations level and limit value exceedences would give an early warning system for providing air quality information. It would support the action taken by government to achieve compliance with the air quality standards and to protect populations.

ACKNOWLEDGEMENT

This study was funded by Universiti Sains Malaysia under ST Grant 6035223 and ST Grant 6039013. Thanks to Universiti Sains Malaysia and Fellowship Scheme for providing financial support to carry out this study and also thanks to Department of Environment Malaysia for their support and cooperation.

REFERENCES


Clean Air Research Group
School of Civil Engineering
Engineering Campus, Universiti Sains Malaysia
14300 Nibong Tebal, Penang, MALAYSIA.
Tel: 04-5996227 Fax: 04-5941009
E-mail: ceazam@eng.usm.my