# Evaluating the Influence of Meteorological Parameters on Ozone Concentration Levels

**Article** *in* International Journal of Engineering & Technology · July 2018

DOI: 10.14419/ijet.v7i3.7.16211



**4 authors**, including:

Amina Nazif
Universiti Teknologi Petronas


Izma Mohammed
PETRONAS


Motasem S. Abualqumboz
Utah State University

*International Journal of Engineering & Technology*, 7 (3.7) (2018) 65-67

International Journal of Engineering & Technology

*Website: www.sciencepubco.com/index.php/IJET*

***Research paper***

# Evaluating the Influence of Meteorological Parameters on Ozone Concentration Levels

**Amina Nazif*, Nurul Izma Mohammed, Amirhossein Malakahmad, and Motasem S. Abualqumboz**

*Civil and Environmental Engineering Department, Universiti Teknologi PETRONAS, 32610 Bandar Seri Iskandar, Perak Darul Ridzuan, Malaysia.*
**Corresponding author E-mail: aminanazif@yahoo.co.uk*

**Abstract**

Over the years, anthropogenic activities have led to the increase in air pollution concentration levels in the atmosphere, this persistent increase in pollution levels can be influenced by meteorological parameters. These parameters assist in the formation and transportation of air pollutants in the atmosphere. Hence, this study aims at evaluating the association between meteorological parameters and air pollutants. The analysis was carried out using Ozone ($O_3$), Particulate matter ($PM_{10}$), Nitrogen dioxide ($NO_2$), temperature, humidity, wind speed, and wind direction data from 2006 to 2010, from two industrial air quality monitoring stations. Stepwise regression (SR) analysis was used to assess the influence of meteorological parameters in accounting for the variability of $O_3$ concentration levels. The SR analysis showed that meteorological parameters accounted for more than 50 % of $O_3$ variability. It can be concluded that different relationship between meteorological parameters and $O_3$ can exist in different locations in the same region.

***Keywords***: *Air pollution; Ozone; Particulate matter; Multiple linear regression; Stepwise regression.*

## 1. Introduction

The presence of air pollutants particularly Ozone ($O_3$) due to anthropogenic activities tend to increase the pollution in the atmosphere [1]. Industrial activities and mobile emission are the main sources of air pollutants [1]. These activities assist in the release of $O_3$ precursor pollutants. The formation of $O_3$ is achieved through the chemical reaction of precursor pollutants particularly, Oxides of Nitrogen ($NO_x$), Carbon Monoxide (CO) and Volatile Organic Compounds (VOC's) with the aid of suitable atmospheric conditions. This in turn increases $O_3$ concentration in the atmosphere [2]. Also, the formation and transportation of air pollutants can be influenced by meteorological parameters [3].

$O_3$ has resulted in tremendous health effects to the exposed citizens, especially the young and the elderly [4, 5]. Furthermore, the severity of these effects can be peculiar in different regions, these can be attributed to seasonal variation, geographical location and topographic conditions [6].

Various statistical methods have been used in evaluating $O_3$ concentration levels in different parts of the world [7, 8]. The distinction in seasonal variations of regions makes the influence of meteorological parameters of each region distinct. As such the absolute understanding of the peculiar pollution pattern should be assessed. Hence, the aim of this paper is to assess the relationship and evaluate the influence of four meteorological parameters namely; temperature, humidity, wind speed and wind direction with $O_3$ using stepwise regression (SR). In addition, Nitrogen dioxide ($NO_2$) was also included in the analysis.

## 2. Methodology

### 2.1 Study Area

Selangor: State of Selangor in Malaysia is growing rapidly due to urbanization and industrialization. Having a population of 5.4 million people, covering an area of 8,104 sq km. This area is constantly exposed to air pollution problems [1, 9]. The industrial air quality monitoring station in the Klang valley area is located in Petaling Jaya. This sampling station is the nearest station to the Kuala Lumpur city center and it is surrounded by industries, residential areas, commercial areas and congested roads.
Terengganu: The state of Terengganu is located in the North-eastern Peninsular, adjoined in the east by the southern China Sea. Covering an area of 13,035 sq.km, with a total populace of about a million people [10]. The industrial air quality monitoring station is in Kemaman. The area has ample industries, substantial traffic activities and residential areas [11].

### 2.2 Monitoring Records

For this study, hourly average $O_3$, $PM_{10}$, $NO_2$, temperature, humidity, wind speed and wind direction data from 2006 to 2010 was used. The data was acquired from the Department of Environment (DoE) Malaysia.

### 2.3 Methods

#### 2.3.1 Stepwise Regression

It is a step by step approach where insignificant variables are removed from the regression analysis, allowing only important variables to be present in the models [12]. SR analysis is used to evaluate the order of importance of variables and subsequently select a useful subset [13, 14]. SR equation is as shown in equation 1.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_3 X_3 + \varepsilon i \quad (1)$$

Where; Y is the dependent variable, $\beta_0$ is the constant, $\beta_1$, $\beta_2$,….$\beta_p$ are the regression coefficients of the independent variables $X_1$, $X_2$,…$X_p$ (predictors) and ε is the residual error (the difference between observations and predicted values).

### 2.3.2 Performance Indicators

The performance of the model was assessed using the coefficient of determination ($R^2$).

**Table** 1: Performance Indicators

| Performance Measure | Equation |
|---|---|
| Coefficient of Determination ($R^2$) | $R^2=\left[\frac{\sum_{i=1}^{n}(O_i-\bar{O})^2.(P_i-\bar{P})}{n.\sigma_O.\sigma_P}\right]^2$ |
| Root Mean Square Error (RMSE) | $RMSE=\sqrt{\frac{\sum_{i=1}^{n}(P_i-O_i)^2}{n}}$ |
| Mean Absolute Error (MAE) | $MAE = \frac{\sum_{i=1}^{n}\lvert P_i-O_i\rvert}{n}$ |
| Normalized Absolute Error (NAE) | $NAE = \frac{\sum_{i=1}^{n}\lvert P_i-O_i\rvert}{\sum_{i=1}^{n}O_i}$ |

Additionally, the performance indicators are Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Normalized Absolute Error (NAE) were used to evaluate the error levels of the SR models as shown in Table 1.

## 3. Result and discussion

### 3.1 Descriptive statistics

Descriptive statistics of the data used for this study is shown in Table 2, the analysis included mean, median, maximum, standard deviation, skewness and kurtosis of all data used. The descriptive statistics showed that $O_3$ was below the Malaysian ambient air quality guideline (MAAQG) levels of 0.20 ppm hourly average in Petaling Jaya but slightly higher in Kemaman. While, $NO_2$ was slightly above the hourly average level of 0.32 ppm in Petaling Jaya but lower in Kemaman. Additionally, high hourly $PM_{10}$ levels were recorded in both areas having maximum levels of 402 and 264 µg/m$^3$.

### Stepwise Regression of $O_3$

The SR analysis involves hourly average $O_3$, $NO_2$, temperature (T), humidity (H), wind speed (WS), and wind direction (WD). Based on Petaling Jaya analysis, the SR for $O_3$ shows that T accounts for 53 % of $O_3$ variability as shown in Table 3. The addition of WS, H, and WD accounted for 54 % of $O_3$ variability, establishing that these meteorological factors can account for 54 % variability of $O_3$. Temperature assist in the chemical formation of ozone establishing a positive dependence [1]. Meanwhile, $NO_2$ was insignificant in the analysis therefore it was not involved in the $O_3$ model. This stipulates that $NO_2$ has less significance in the formation of $O_3$ and there is a possibility of other precursor pollutants aiding in the formation of $O_3$ in the area, while meteorologial parameters were more significant [15]. Subsequently, in the SR analysis of Kemaman, the model showed that WS accounts for 52 % of $O_3$ variability. The addition of H, WS, and WD accounted for 60 % $O_3$ variability. The addition of $PM_{10}$ and $NO_2$ to the meteorological parameters accounted for 63 % $O_3$ variability.

**Table** 2: Descriptive statistics of hourly average levels for Petaling Jaya

| Petaling Jaya | | | |
|---|---|---|---|
| Variables | Mean | St Dev | Median |
| $O_3$ | 0.014 | 0.02 | 0.01 |
| PM | 46.14 | 23.03 | 43.00 |
| $NO_2$ | 0.52 | 3.66 | 0.029 |
| Temp | 27.92 | 4.11 | 27.00 |
| Humidity | 73.20 | 15.64 | 77.000 |
| Wind speed | 3.77 | 2.22 | 3.4 |
| Wind direction | 146.56 | 101.11 | 128.00 |
| Petaling Jaya | | | |
| Variables | Max | Skewness | Kurtosis |
| $O_3$ | 0.13 | 1.76 | 3.33 |
| PM | 402.00 | 1.99 | 10.86 |
| $NO_2$ | 0.12 | 7.57 | 56.61 |
| Temp | 39.40 | 0.51 | -0.71 |
| Humidity | 94.00 | -1.86 | 5.90 |
| Wind speed | 16.90 | 0.90 | 0.64 |
| Wind direction | 360.00 | 0.58 | -0.79 |

**Table** 3: Descriptive statistics of hourly average levels for Kemaman

| Kemaman | | | |
|---|---|---|---|
| Variables | Mean | St Dev | Median |
| $O_3$ | 0.02 | 0.015 | 0.02 |
| PM | 35.59 | 18.855 | 32.00 |
| $NO_2$ | 0.003 | 0.0026 | 0.002 |
| Temp | 26.97 | 4.90 | 25.20 |
| Humidity | 78.50 | 11.77 | 79.00 |
| Wind speed | 4.60 | 2.84 | 3.90 |
| Wind direction | 198.67 | 120.04 | 211.00 |
| Kemaman | | | |
| Variables | Max | Skewness | Kurtosis |
| $O_3$ | 0.09 | 0.77 | 0.14 |
| PM | 264 | 2.48 | 13.33 |
| $NO_2$ | 0.03 | 2.16 | 7.69 |
| Temp | 39.40 | 0.74 | -0.72 |
| Humidity | 100.00 | -0.29 | 0.09 |
| Wind speed | 18.60 | 1.12 | 0.95 |
| Wind direction | 360.00 | -0.23 | -1.45 |

**Table** 4: Stepwise Regression of hourly $O_3$

| No. | Models | $R^2$(%) |
|---|---|---|
| | Petaling Jaya | |
| 1 | $O_3$= -0.0719+0.0031T | 52.52 |
| 2 | $O_3$= -0.0671+0.0028T+0.0011WS | 53.76 |
| 3 | $O_3$= -0.0504+0.0025T+0.0010WS-0.0001H | 54.33 |
| 4 | $O_3$= -0.0505+0.0025T+0.0010WS-0.0001H-0.000005WD | 54.39 |
| 5 | $O_3$= -0.0504+0.0025T+0.0010WS-0.0001H-0.000001WD-0.00001PM | 54.39 |
| 6 | $O_3$= -0.0501+0.0025T-0.0001H+0.0010WS-0.000001WD | 53.89 |
| | Kemaman | |
| 1 | $O_3$= 0.0033+0.0361WS | 53.60 |
| 2 | $O_3$= -0.0153+0.0026WS+0.0009T | 58.77 |
| 3 | $O_3$= -0.0191+0.0026WS+0.00009PM | 60.74 |
| 4 | $O_3$= -0.0195+0.0026WS+0.0008T +0.00007PM+0.6160$NO_2$ | 61.84 |
| 5 | $O_3$= -0.0145+0.0026WS+0.0008T + 0.00007PM+0.5470$NO_2$-0.00001WD | 62.57 |
| 6 | $O_3$= -0.0044+0.00251WS+0.0007T +0.00008PM+0.558$NO_2$-0.00001WD - 0.00011H | 63.05 |
| 7 | $O_3$= -0.0032+0.0007T - | 60.20 |

| | | |
|---|---|---|
| | 0.00007H+0.0025WS-0.000012WD | |

**Table** 5: Performance Indicator result

| Study Area | Pollutant | RMSE | MAE | NAE |
|---|---|---|---|---|
| Petaling Jaya | $O_3$ | 0.02 | 0.03 | 0.54 |
| Kemaman | | 0.01 | 0.01 | 0.20 |

The performance indicators result for $O_3$ as shown in Table 4 established that the RMSE, MAE, and NAE results of Kemaman was slightly better, having low error levels than the Petaling Jaya SR model.

## 4. Conclusion

Overall, $O_3$ had a substantial relationship with humidity, temperature, wind speed and direction. In Petaling Jaya and Kemaman, temperature and wind speed accounted for >53% $O_3$ variability. This analysis would help atmospheric analyst and statisticians in assessing the peculiarity in air pollution assessment.

## Acknowledgement

Appreciation goes to Universiti Teknologi PETRONAS for making this study possible and the Department of Environment (DOE) Malaysia for providing the data used for this study.

## References

[1] Latif, M.T., L.S. Huey, and L. Juneng, *Variations of surface ozone concentration across the Klang Valley, Malaysia.* Atmospheric Environment, 2012. **61**: p. 434-445.
[2] Dominick, D., et al., *Spatial assessment of air quality patterns in Malaysia using multivariate analysis.* Atmospheric Environment, 2012. **60**: p. 172-181.
[3] Tarasova, O. and A.Y. Karpetchko, *Accounting for local meteorological effects in the ozone time-series of Lovozero (Kola Peninsula).* Atmospheric Chemistry and Physics, 2003. **3**(4): p. 941-949.
[4] Ebi, K.L. and G. McGregor, *Climate change, tropospheric ozone and particulate matter, and health impacts.* Environ Health Perspect, 2008. **116**(11): p. 1449-1455.
[5] Conti, S., et al., *Cardiorespiratory treatments as modifiers of the relationship between particulate matter and health: A case-only analysis on hospitalized patients in Italy.* Environmental research, 2015. **136**: p. 491-499.
[6] Peng, R.D., et al., *Seasonal analyses of air pollution and mortality in 100 US cities.* American journal of epidemiology, 2005. **161**(6): p. 585-594.
[7] Slini, T., K. Karatzas, and A. Papadopoulos, *Regression analysis and urban air quality forecasting: An application for the city of Athens.* Global Nest, 2002. **4**(2-3): p. 153-162.
[8] Pires, J.C., et al., *Comparison of several linear statistical models to predict tropospheric ozone concentrations.* Journal of Statistical Computation and Simulation, 2012. **82**(2): p. 183-192.
[9] Afroz, R., et al., *Benefits of air quality improvement in Klang Valley Malaysia.* International journal of environment and pollution, 2007. **30**(1): p. 119-136.
[10] D.o.S, *Population and housing census of Malaysia 2010*. 2010, Department of Statistics, MAlaysia: Malaysia.
[11] Azid, A., et al., *Source Apportionment of Air Pollution: A Case Study In Malaysia.* Jurnal Teknologi, 2014. **72**(1).
[12] Thomas, S. and R.B. Jacko, *Model for forecasting expressway fine particulate matter and carbon monoxide concentration: application of regression and neural network models.* Journal of the Air & Waste Management Association, 2007. **57**(4): p. 480-488.
[13] Huberty, C.J., *Problems with stepwise methods—better alternatives.* Advances in social science methodology, 1989. **1**: p. 43-70.
[14] Thompson, B., *Stepwise Regression and Stepwise Discriminant Analysis Need Not Apply.* 1995.
[15] Kovač-Andrić, E., J. Brana, and V. Gvozdić, *Impact of meteorological factors on ozone concentrations modelled by time series analysis and multivariate statistical methods.* Ecological Informatics, 2009. **4**(2): p. 117-122.