



# A Review on Short-Term Prediction of Air Pollutant Concentrations

Ahmad Fauzi Raffee<sup>1</sup>, Siti Nazahiyah Rahmat<sup>1</sup>, Hazrul Abdul Hamid<sup>2\*</sup>, Muhammad Ismail Jaffar<sup>1</sup>

<sup>1</sup>Faculty of Civil and Environmental Engineering, Universiti Tun Hussein Onn Malaysia

<sup>2</sup>School of Distance Education, Universiti Sains Malaysia

\*Corresponding author E-mail: [hazrul@usm.my](mailto:hazrul@usm.my)

## Abstract

In the attempt to increase the production of the industrial sector to accommodate human needs; motor vehicles and power plants have led to the decline of air quality. The tremendous decline of air pollution levels can adversely affect human health, especially children, those elderly, as well as patients suffering from asthma and respiratory problems. As such, the air pollution modelling appears to be an important tool to help the local authorities in giving early warning, apart from functioning as a guide to develop policies in near future. Hence, in order to predict the concentration of air pollutants that involves multiple parameters, both artificial neural network (ANN) and principal component regression (PCR) have been widely used, in comparison to classical multivariate time series. Besides, this paper also presents comprehensive literature on univariate time series modelling. Overall, the classical multivariate time series modelling has to be further investigated so as to overcome the limitations of ANN and PCR, including univariate time series methods in short-term prediction of air pollutant concentrations.

**Keywords:** Multivariate time series, VAR, air pollution, forecasting

## 1. Introduction

The declining air quality has emerged as a major problem in environmental pollution. In fact, there are three types of air pollutants, which are natural, primary, and secondary pollutants. Natural pollutants find their way into the atmosphere as a result of a natural phenomenon [1]. Meanwhile, primary pollutant is defined as pollutant either in particulate or chemical composition directly emitted from the sources into the ambient atmosphere [2]. Other than that, pollutants that have changed form after being emitted into the atmosphere from sources due to oxidation or decay or reaction with other primary pollutants are known as secondary pollutants [3].

Moreover, the deteriorating air quality can be classified into short- and long-term effects to human health. The short-term effect includes acute effects, such as bronchoconstriction and cardiovascular changes, as well as asthma symptoms to chronic effects like respiratory issues. Being exposed to high level of pollutants within a short period of time can also lead to premature mortality due to respiratory and cardiovascular diseases [4],[5]. On the other hand, long-term human exposure can cause hypertensive disorder, numerous types of cancers, and increased risk of death among tuberculosis patients [6],[8]. For instance, about 300, 000 cases of premature deaths occurred in China in year 2010 due to air pollution, apart from the 0.8 million premature cases worldwide annually [9]. The main sources of air pollution come from industrial and transportation activities, especially in the urban area [10],[11]. For example, a study carried out by Kim *et al.*, [12] examined particulate matter (PM) and surface layer ozone (O<sub>3</sub>) as the two main pollutants that mostly influenced the reading of pollutant

concentrations and caused deaths around 2.1 and 0.47 million worldwide, respectively.

On top of that, statistical modelling is an important tool to monitor and to assess air quality. Furthermore, it is also useful for future prediction [13],[14]. In fact, statistical modelling in air quality can provide a big significance to the related agencies that may help to give early information to local community, especially to those vulnerable to low air quality level. Apart from that, such information is beneficial for the government in developing viable policies [15].

## 2. Air Pollution Modelling

This section focuses on prior studies related to air pollution statistical modelling for short-term prediction. The statistical modelling has been widely used in air pollution studies to predict air pollutant concentrations [16]. Table 1 shows some of the methods that are commonly used in this area.

### 2.1. Univariate Time Series Modelling

A set of observations measured sequentially through time is defined as time series. The univariate time series models that have been widely used in air pollutant modelling are autoregressive (AR), moving average (MA), autoregressive moving average (ARMA), and autoregressive integrated moving average (ARIMA) [21]. For instance, Lim *et al.*, [22] discovered that ARIMA and AFRIMA models can be applied for prediction of air quality index in Selangor. The study concluded that the AFRIMA model gave the best prediction, whereas ARIMA failed to accurately forecast

**Table 1:** The description of air pollution models

No	Method	General Equation	Description
1	Univariate time series	$Y_t = \phi_1 Y_{t-1} + \varepsilon_t - \phi_1 e_{t-1}$	The univariate time series is a set continuous observation of one variable with constant time interval [17].
2	Multivariate time series	$(I - \Phi_1 B)Z_t = (I - \Theta_1 B)a_t$	The multivariate time series is defined as observations of two or more variables are often taken simultaneously and describe the interrelationships among the series and using the data to develop a model to forecast the future prediction [18].
3	Artificial neural network (ANN)	Hidden neuron $h_j$ , $h_j = \varphi(z_j),$ The mathematical relations between input and output $f(x)$ and the input $(x_i, y_i)_{i=1}^l$ follows the below equations $f(x) = w_o + \sum_{j=1}^q w_j h_j$	The ANN developed in three elements and the connection to each variable relation characterized by their strength and a linear combiner that combines the weighted inputs signals. Moreover, it has an activation function for limiting the amplitude range of the neuron's output to some finite value [19].
4	Principal component regression (PCR)	The PCA equations as follows $PC_i = l_{1i} X_1 + l_{2i} x_2 + \dots + l_{mi} x_m$ While for MLR is $y = b_o + \sum_{i=1}^p b_i x_i + \varepsilon$	PCR is a combination of principal component analysis (PCA) and multiple linear regressions (MLR). The PCA as input in MLR is intended to reduce the complexity and multicollinearity problems of the model. The three main step in PCR is run the PCA on the table of input variable, then the MLR will run on the components selected by PCA in previous step and computed of the parameter of the model from the correspond to the input variables [20].

the actual value of air quality. In addition, the ARIMA model emerged as the best model for long-term prediction of ozone ( $O_3$ ) concentration, which could be applied in Kemaman, Terengganu. This model was developed based on 144 samples obtained from the Automatic Air Quality Monitoring System applied at the Kemaman station for one year period; 1996-1997 [23].

Meanwhile, Sansuddin *et al.*, [24] concluded that the best time series model at four different locations, which represented industrial (Nilai and Johor Bharu) and residential (Kota Kinabalu and Kuantan), had been the AR. The seasonal ARIMA was also used to predict the ozone precursor at Brazil [25]. Additionally, Rahman *et al.*, [26] revealed that the univariate time series was suitable to be applied at two urban monitoring stations in Johor. In order to predict the  $PM_{10}$  concentrations, the seasonal ARIMA was found to be the most suitable model to be applied in Nilai, Negeri Sembilan [27], in which the study used secondary data gathered from April 2008 until March 2009. Hamid *et al.*, [28], in another study, discovered that the most appropriate model at two locations in Bachang, Melaka and Kuala Terengganu was indeed the ARMA. Meanwhile, Bas *et al.*, [29] found that the seasonal ARIMA had a power explanatory and predictive technique in forecasting air pollutant concentrations.

## 2.2. Multivariate Statistical Modelling

Several methods in multivariate statistical modelling can be used for air pollution prediction. This section summarizes past studies related to multivariate time series, ANN, and PCR. Forecasting using the multivariate time series has limited references, in comparison to univariate time series in air pollution modelling. The multivariate time series has been widely used in economic forecasting, when compared to other disciplines. To date, none has

comprehensively looked into this method for air pollution modelling. In an instance, Hsu [30] used the multivariate time series method to analyze the interdependence among air pollutants in Taipei. The vector autoregressive model (VAR) was chosen to examine the significant relationships between  $NO$ ,  $NO_2$ , and  $O_3$  pollutants. These studies concluded that the simple tools of VAR could be used for assessment of air pollution problem. Other than that, Xi and Lin [31] also utilized this method to analyze the influential factors of the changes in carbon dioxide emission. The VAR analyses revealed that the significant effect on carbon dioxide emissions was due to large-scale population movements and the transformation of the industrial sector.

The mathematical model in analysis inspired by the functional of nervous system is termed as artificial neural network (ANN) model, which is composed by a number of interconnected entities called the artificial neuron [32]. There are many different ways for the neurons to relate depending on the characteristics based on the problem. The procedure for ANN, nonetheless, is more complicated since the model employs structure neuron model to select ANN model. Moreover, in air pollution modelling, the ANN has been widely used to predict future concentration of air pollution. For example, Mok and Tam [33] found that the errors in ANN model ranged between 14.45% and 13.71% for two testing periods in order to predict daily  $SO_2$  concentration in Macau.

The three-layer feed-forward ANNs had been discovered as the best model. In fact, Gennaro *et al.*, [34] also used the ANN method to predict the concentration of  $PM_{10}$  in two urban areas in Spain. The correlation coefficient between real and forecast data had been  $R^2=0.86$  and  $R^2=0.73$ . Moreover, similar methods were used in Portugal and found the Matthews correlation coefficient (MCC) displaying between 0.65 and 0.92 for forecast and real data [35]. Besides, He *et al.*, [36] found a few ANN models with limited accuracy owing to their potential convergence to a local minimum and over fitting when ANN was applied to air quality forecasting in urban areas. Other than that, Pawul and Śliwka [37]

predicted  $PM_{10}$  concentrations in Poland as factors determining the occurrence of smog phenomena by using the ANN method. The three-layer perceptron with back-propagation algorithm of ANN model received a good fit in all monitoring stations. The study on modelling air quality status at monitoring stations located nearest to the industrial areas, which is approximately 7 km from the city, had been conducted and indicated the ANN model as a good model to predict air quality status in Kurichi, India [38]. Similar results were found in Bai *et al.*, [39], where ANN displayed good performance in forecasting air quality. Furthermore, the ANN model of multilayer perceptron (MLP) emerged as the best model that could predict ozone concentrations in a study conducted by Kumar *et al.*, [40]. The study also revealed that the ANN method cannot handle linear patterns equally well as non-linear patterns [41]. Meanwhile, Tealab *et al.*, [42] revealed that the common neural networks were inefficient in recognizing the behavior of non-linear or dynamic time series with moving average terms and hence, low forecasting capability.

Another short-term prediction method is principal component regression (PCR). The PCR model refers to the combination of PCA and multiple linear regressions (MLR), where PCA is used to analyze data to remove the most uncorrelated variable, while the MLR regresses each variable to develop a suitable model. The PCA is indeed a popular multivariate method that can be used to predict air quality [43]. For instance, Lengyel *et al.*, [44] evaluated and predicted the level of ozone concentrations at Miskole, Hungary by using multivariate techniques of PCA, MLR, PCR, and PLS. This study listed the disadvantages of PCR and PLS, which failed to provide physically relevant factors.

Another study by Martin [45] revealed that MLR outperformed PCR in modelling outdoor air pollutants. Meanwhile, Dominick *et al.*, [46] investigated air quality pattern in Malaysia to examine the major sources of air pollution, percentage contribution of pollutant, apart from percentage contribution of each air pollutant. The PCA identified that the predominant source of air pollution was released from motor vehicles and industrial activities, whereas MLR demonstrated that the concentrations of  $PM_{10}$  in the atmosphere within the study areas had been predominantly correlated with CO. The correct identification of the importance of all variables involved in determining indoor pollutants is a major PCR drawback [47].

The methods of PCR and MLR were also applied by Elbayoumi *et al.*, [48], who discovered significant correlation and pointed out that the MLR model was better than the PCR model. Additionally, Awang *et al.*, [49] employed PCA, MLR, and PCR methods to predict the level of ozone during daytime (DT), nighttime (NT), and critical conversion time (CCT) by using the 2010 data recorded at Shah Alam monitoring stations. As a result, this study revealed that the PCR during CCT significantly exerted higher performance during NT and DT, respectively. Meanwhile, Balachandran *et al.*, [50] evaluated the fire weather forecast by using multivariate analysis for  $PM_{2.5}$  pollutant. The study concluded that the  $PM_{2.5}$  showed positive sensitivity towards forecast precipitation and prescribed that burning decisions should be based on the forecast released in the morning of the potential burn since forecast of AM is more stable for PCR treatment and produced less uncertain outcomes. The summary of short term prediction model on air pollutants concentrations study is shown in Table 2.

**Table 2:** Summary of short-term prediction on air pollutants concentrations

Method	Author(s)	Model used
Univariate time series modelling	Sansuddin <i>et al.</i> , 2011	AR
	Hamid <i>et al.</i> , 2017	ARMA
	Ismail <i>et al.</i> , 2001; Rahman <i>et al.</i> , 2016	ARIMA
	Lim <i>et al.</i> , 2008	AFRIMA
	Castaneda <i>et al.</i> 2014; Hamid <i>et al.</i> , 2016; Bas <i>et al.</i> 2017	SARIMA
Multivariate statistical modelling	Hsu, 1992; Xi <i>et al.</i> , 2015	MTS
	Mok and Tam, 1998; Gennaro <i>et al.</i> , 2013; Fontes <i>et al.</i> , 2014; He <i>et al.</i> , 2014; Pawul and Śliwka, 2016; Bai <i>et al.</i> 2016; Park <i>et al.</i> , 2016; Aravind <i>et al.</i> , 2016; Kumar <i>et al.</i> , 2017; Panigrahi <i>et al.</i> , 2017; Tealab <i>et al.</i> , 2017	ANN
	Lengyel <i>et al.</i> , 2004; Martin, 2011; Dominick <i>et al.</i> , 2012; Chatterjee and Hadi, 2013; Elbayoumi <i>et al.</i> , 2014; Awang <i>et al.</i> , 2015; Balachandran <i>et al.</i> , 2017	PCR

### 3. Conclusion

Studies pertaining to air pollution are mostly related to the assessment of pollutant concentration and descriptive statistics in determining the behavior of pollutants. The assessment of air pollution usually focuses on the selected area and the results are valid only at the area of research. The air pollution-forecasting model, hence, is a useful tool that can be embedded as part of an early warning system to provide early indication of future air quality conditions. The short term prediction of air pollution is also useful for future planning and development, especially in determining if the development areas should conduct assessment related to environment so as to reduce risks in the particular areas. Since past short-term prediction of air quality focused on univariate time series, ANN, and PCR, it is worthwhile to develop a multivariate time series model with consideration of meteorological parameters and other pollutants in order to improve the accuracy of prediction.

### Acknowledgement

Authors would like to thank Postgraduate Research Grant (GPPS) (Vote 972), Universiti Tun Hussein Onn Malaysia for the support that allowed us to carry out this research. Also Fundamental Research Grant Scheme (Vote 1523) for publication sponsorships.

### References

- [1] Standers, L.H.. Regulatory Aspects of Air Pollution Control in the United States. Air & Waste Management Association, pp.8-21, (2000).
- [2] Yahaya, A.S., Ramli, N.A. & Hamid, H.A. Review of Fitting Distribution on Air Pollution Modelling. Prosiding Simposium Kebangsaan Sains Matematik ke-XV, Malaysia. (2007).
- [3] Mabahwi, N.A., Leh, O.L.H. and Omar, D. Urban air quality and human health effects in Selangor Malaysia. Procedia-Social and Behavioral Sciences 170, pp.282-291. (2015).
- [4] Chen, R., Kan, H., Chen, B., Huang, W., Bai, Z., Song, G., Pan, G. Association of particulate air pollution with daily mortality the China air pollution and health effects study. Am. J. Epidemiol. 175 (11), pp.1173-1181. (2012).
- [5] Maji, S., Ahmed, S., Siddiqui, W.A., and Ghosh, S. Short term effects of criteria air pollutants on daily mortality in Delhi, India. Atmospheric Environment. 150, pp.210-219. (2017).
- [6] Mabahwi, N.A., Leh, O. L, H and Omar, D. Human Health and Wellbeing: Human health effect of air pollution. Procedia – Social and Behavioral Science 153, pp.221-229. (2014).
- [7] Peng, Z., Liu, C., Kan, H., and Wang, W. Long term exposure to ambient air pollution and mortality in a chinese tuberculosis cohort. Science of the Total Environment. (2016).

- [8] Khaniabadi, Y.O., Hopke, P.K., Goudarzi, G., Daryanoosh, S.M., Jourvand, M. and Basiri, H. Cardiopulmonary mortality and OCPD attributed to ambient ozone. *Environmental Research* 152, pp.336-341. (2017).
- [9] Zhang, D., Aunan, K., Seip, H.M., Larssen, S., Liu, J. and Zhang, D. The assessment of health damage caused by air pollution and its implication for policy making in Taiyuan, Shanxi, China. *Energy Policy* 38, pp.491-502. (2010).
- [10] Mohamad, N.D., Ash'aari, Z.H., & Othman, M. Preliminary Assessment of Air Pollutant Sources Identification at Selected Monitoring Stations in Klang Valley, Malaysia. *International Conference on Environmental Forensics. Procedia Environmental Sciences* 30, pp.121-126. (2015).
- [11] Fang, X., Bi, X., Wu, J., Zhang, Y., and Feng, Y. Source apportionment of ambient PM<sub>10</sub> and PM<sub>2.5</sub> in Haikou, China. *Atmospheric Research* 190, pp.1-9.(2017).
- [12] Kim, K.H., Kabir, E., and Kabir, S. A Review on the Human Health Effect of Airborne Particulate Matter. *Environmental International* pp.136-142. (2015).
- [13] Forsyth, T. Public concerns about transboundary haze: A comparison of Indonesia, Singapore and Malaysia. *Global Environmental Change, Volume 25*, pp.76-86. (2014).
- [14] Ling, Z.H., Zhao, J., Fan, S.J. and Wang, X.M. Source of formaldehyde and their contribution to photochemical O<sub>3</sub> formation at an urban site in the Pearl River Delta, Southern China. *Chemosphere* 168, pp.1293-1301. (2017).
- [15] Li, M., and Zhang, L. (2014). Haze in China: Current and future challenges. *Environmental pollution, Volume 189*, pp.85-86. (2014).
- [16] Ausati, S. and Amanollahi, J. Assessing the accuracy of ANFIS, EEMD-GRNN, PCR and MLR models in predicting PM<sub>2.5</sub>. *Atmospheric Environment* 142, pp.465-474. (2016).
- [17] Brockwell, P. J and Davis, R.A. *Introduction to Time Series and Forecasting*. Springer-Verlag New York, Inc. New York. (2002).
- [18] Chatfield, C. *Time Series Forecasting*. Chapman & Hall/CRC, United State of America. (2002).
- [19] Suleimen, A., Tight, M.R. and Quinn, A.D. Assessment and prediction of the impact of road transport on ambient concentrations of particulate matter PM<sub>10</sub>. *Transportation Research Part D* 49, pp. 301-312. (2016).
- [20] Antanasijević, D.Z., Pocajt, V.V., Povrenovic D.S., Ristić, M.D., Grujić, A.A.P. PM<sub>10</sub> emission forecasting using artificial neural networks and genetic algorithm input variable optimization. *Science of Total Environment* 443, pp.511-519. (2013).
- [21] Wang, Y., Wang, J., Zhao, G., Dong., Y. Application of residual modification approach in seasonal ARIMA for electricity demand forecasting: A case study of China, *Energy Policy* 48, pp.284-294. (2012).
- [22] Lim, Y. S., Lim, Y. C., & Mah, J. W. Arima and Integrated Afrima Models for Forecasting Air Pollution Index in Shah Alam, Selangor. *The Malaysian Journal of Analytical Vol.12 No.1*. (2008).
- [23] Ismail, M., Ibrahim, M. Z., Ibrahim T. A., & Abdullah, A. M. Time Series Analysis of Surface Ozone Monitoring Records in Kemaman, Malaysia. *Sains Malaysiana* 40(5), pp. 411-417. (2011).
- [24] Sansuddin, N., Ramli, N. A., Yahaya, A. S., MD Yusof, N. F. F., Ghazali, N. A., & Al Madhoun, W. A. Statistical Analysis of PM<sub>10</sub> Concentration at different Location in Malaysia. *Environ Monit Assess* (180), pp.573-588. (2012).
- [25] Castaneda, D. M. A., Teixeira, E. C. and Pereira, F. N. Time series analysis of surface ozone and nitrogen oxides concentrations in an urban area at Brazil. *Atmospheric Pollution Research* 5, pp.411-420. (2014).
- [26] Rahman, N. H. A., Lee, M. H., Suhartono. and Latif, M.T. Evaluation performance of Time Series Approach for Forecasting Air Pollution Index in Johor, Malaysia. *Sains Malaysiana* 45(11), pp.1625-1633. (2016).
- [27] Hamid, H.A., Yahaya, A.S., Ramli, N.A, Ul-Saufie, A.Z. and Yasin, M.N.. Short term prediction of PM10 concentrations using seasonal time series analysis. *MATEC Web of Conferences* 47, 05001.(2016).
- [28] Hamid, H.A., Ul-Saufie, A.Z. and Ahmat, H. Characteristic and prediction of carbon monoxide concentrations using time series analysis in selected urban area in Malaysia. *MATEC Web of Conferences* 103, 05001. (2017).
- [29] Bas, M. d. C., Ortiz, J., Ballesteros, L. and Martorell, S. Forecasting 7BE concentrations in surface air using time series analysis. *Atmospheric Environment* 155, pp.154-161. (2017).
- [30] Hsu, K. J. Time series analysis of the interdependence among air pollutants. *Atmospheric Environment, Vol.26B, No. 4*, pp.491-503. (1992).
- [31] Xi, B. and Lin, B. Carbon dioxide emissions reduction in China's transport sector: A dynamic VAR (vector autoregression) approach. *Energy* 83, pp. 486-495. (2015).
- [32] Russo, A., Raischel, F. and Lind, G. Air quality prediction using optimal neural network with stochastic variables. *Atmospheric Environment* 79, pp. 822-830. (2013).
- [33] Mok, K.M. and Tam, S.C. Short-term prediction of SO<sub>2</sub> concentration in Macau with artificial neural networks. *Energy and Buildings* 28, pp.279-286. (1998).
- [34] Gennaro, G. D., Trizio, L., Gilio, A. D., Pey, J., Pérez, N., Cusack, M., Alastuey, A. and Querol, X. Neural network model for the prediction of PM10 daily concentrations in two sites in the Western Mediterranean. *Science of the Total Environment* 463, pp.875-883. (2013).
- [35] Fontes, T., Silva, L.M., Silva, M.P., Barros, N. and Carvalho, A.C. Can artificial neural networks be used to predict the origin of ozone episodes?. *Science of The Total Environment* 488-489, pp.197-207. (2014).
- [36] He, H. D., Lu, W.Z. and Xue, Y. Prediction of particulate matter at street level using artificial neural networks coupling with chaotic swarm optimization algorithm. *Building and Environment* 78, pp.111-117. (2014).
- [37] Pawul, M. and Śliwka, M. Application of artificial neural networks for prediction of air pollution levels in environmental monitoring. *Journal of Ecological Engineering, Volume 17, Issue 4*, pp.190-196. (2016).
- [38] Aravind, T.P.A., Haneef, M., Sulthan, M.A.M., Solomon, M.I. and Jebamani. Monitoring, Analysis and Modelling of Ambient Air Quality Status at Indoshell Mould Ltd., Sidco, Coimbatore using Artificial Neural Network. *International Journal of Scientific Research in Science, Engineering and Technology (IJRSET), Volume 2, Issue 2*, pp.2394-4099. (2016).
- [39] Bai, Y., Li, Y., Wang, X., Xie, J. and Li, C. Air pollutants concentrations forecasting using back propagation neural network based on wavelet decomposition with meteorological conditions. *Atmospheric pollutant research* 7, pp. 557-566. (2016).
- [40] Kumar, N., Middey, A. and Rao, P.S. Prediction and examination of seasonal variation of ozone with meteorological parameter through artificial neural network at NEERI, Nagpur, India. *Urban Climate* 20, pp.148-167. (2017).
- [41] Panigrahi, S. and Behera, H.S. A hybrid ETS-ANN model for time series forecasting. *Engineering Applications of Artificial Intelligence* 66, pp.49-59. (2017).
- [42] Tealab, A., Hefny, H. and Badr, A. Forecasting of non linear time series using artificial neural network. *Future Computing and Informatics Journal*, pp.1-9. (2017).
- [43] Wahab, S.A.A., Bakheit, C.S. and Al-Alawi, S.M. Principal component and multiple regression analysis in modelling of ground-level ozone and factors affecting its concentrations. *Environmental Modelling & Software* 20, pp.1263-1271. (2005).
- [44] Lengyel, A., Héberger, K., Paksy, L., Bánhidí, O. and Rajkó, R. Prediction of ozone concentration in ambient air using multivariate methods. *Chemosphere* 57, pp.889-896. (2004).
- [45] Martin, E., 2011. Comparative Performance of Different Statistical Models for Predicting Ground-level Ozone (O<sub>3</sub>) and Fine Particulate Matter (PM<sub>2.5</sub>) Concentrations in Montréal, Canada. The Department of Building, Civil and Environmental Engineering. Concordia University, Montréal, Canada.
- [46] Dominick, D., Juahir, H., Latif, M.T., Zain, S.M., and Aris, A.Z. Spatial assessment of air quality patterns in Malaysia using multivariate analysis. *Atmospheric Environment* 60, pp.172-181. (2012).
- [47] Chatterjee, S., Hadi, A.S. *Regression analysis by example, fifth ed.* John Wiley & Sons. (2013).
- [48] Elbayoumi, M., Ramli, N.A., Yusof, N.F.F.M., Yahaya, A.S., Madhoun, W.A. and Ul-Saufie, A.Z. Multivariate methods for indoor PM10 and PM2.5 modelling in naturally ventilated schools buildings. *Atmospheric Environment* 94, pp.11-21. (2014).
- [49] Awang, N. R., Ramli, N.A., Yahaya, A.S. and Elbayoumi, M. Multivariate methods to predict ground level ozone during daytime, nighttime and critical conversion time in urban areas. *Atmospheric Pollution Research* 6, pp.726-734. (2015).
- [50] Balachandran, S., Baumann, K., Pachon, J.E., Mulholland, J.A. and Russel, A.G. Evaluation of fire weather forecast using PM2.5 sensitivity analysis. *Atmospheric Environment* 148, pp.128-138. (2017).