

Modelling of Tropospheric Ozone Concentration in Urban Environment

Nadiyahatur Rahmatikal Wasi'ah^{1*} and Driejana^{2#}

Environmental Engineering Graduate Program¹, Air and Waste Management Research Group²

Faculty of Civil and Environmental Engineering, Institut Teknologi Bandung

Jl. Ganesha 10 Bandung 40132

[#]Corresponding author: driejana@tl.itb.ac.id

Abstract— Tropospheric ozone is harmful to human health and plants. It is resulted from photochemical processes involving NO_x and VOCs from reactions of motor vehicle emissions and solar radiation in polluted urban environment. Historical data in Jakarta indicated that ozone concentrations often exceeded ambient standard threshold. To minimize its impact to human health it is important to predict its concentration. This paper reports the use of multivariate statistical method to predict ozone concentration, using precursor concentration and meteorological parameters. CH₄, CO, NMHC, NO, NO₂, THC data concentration, wind direction and speed, temperature, solar radiation and relative humidity during 2011 - 2012 were used to build the model. Multiple linear regressions were applied to predict ozone concentration at Thamrin Station, Jakarta. These data were used as predictors at time (t) to estimate the ozone concentration at time (t +1). Meteorological conditions were found to strongly affect the concentration of ozone. The strongest relationship was found between ozone and temperature (0.513, p = 0.000). Weaker but significant positive correlations were found for solar radiation and NO₂ (r = 0.242, p= 0.000),. NMHC and NO correlation (r= 0.353, p= 0.000). Both NO and NMHC are freshly emitted from exhaust gas. Correlations between humidity, wind speed and direction were negative. Methana, NMHC, were negatively correlated with ozone due to their roles for producing NO₂ as the main precursor, while NO was for its scavenging reaction with O₃. Based on Adjusted R² value, all predictors could explain variation in ozone concentration of approximately 46.32%. These findings will be useful as input in urban transportation planning and management in cities with tropical climate like Indonesia, as all precursors are emitted from vehicle combustion.

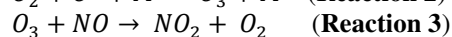
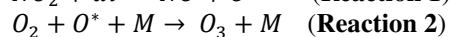
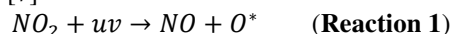
Keywords—urban air pollution, sustainable city, tropical climate, urban planning, air quality modeling

I. INTRODUCTION

Ozone is a secondary pollutant formed by photochemical reaction among precursors such as volatile organic compounds (VOCs) and NO_x with short-wave solar radiation ([1]. Ozone is a powerful oxidant and harmful to human and plant [2]. Based on WHO (2000), high levels of ozone exposure cause impairment of pulmonary function and inflammatory response of lung. Various studies and air quality monitoring data reported that ozone concentration often exceeded the standard. . One-hour mean ozone concentration reach 350 µg/m³ in certain parts of Europe meanwhile in USA, its 1 hour-concentration often exceed 400 µg/m³ [3]. In 1990-1992, Ozone exceeded the quality standard on average per day by 75% the day of the study in Mexico [4]. While research of the ozone concentration that was conducted by Driejana in Bandung in 2001 showed the average concentration of ozone per month in Bandung was greater than 200 ppb with the highest concentration of 287 ppb found in June [5]. These factors emphasizes the importance of ozone concentration in monitoring ambient air .

Tropospheric ozone's formation involves precursors and the presence of solar radiation. At first, NO will be oxidized to NO₂ by organic radical. The photolysis of NO₂ (λ = 420 nm) yields NO and atomic oxygen (O*)

(**Reaction 1**). The radical oxygen atoms react with oxygen to form ozone (**Reaction 2**). In the normal state, formed ozone will react with NO to form NO₂ and oxygen (**Reaction 3**). This cycle continues therefore the tropospheric ozone concentration remains stable [6]. Volatile organic compounds (VOCs) contribute to the ozone production as radical organic reservoir which formed by the photolysis of VOCs or reaction with OH* radical [7]



In previous studies, statistical methods predicted well tropospheric ozone concentrations in some urban area such as Dilovasi, Turkey [8] Mexico City [9], Spain [10] and Kuwait [11]. These studies suggested that this approach is suitable to be applied in varied geographical locations and meteorological conditions. It should be noted however, that the complexity of tropospheric ozone formation reaction, meteorological conditions in urban areas and uncertainties in the measurement of all parameters cause difficulty of ozone modeling [8]. Moreover, the relationship between ozone and its precursor is complex because meteorological conditions and the rate of chemical reactions change very rapidly to very slowly [12]. Therefore, to simplify the complex processes of ozone formation, multivariate statistical methods is applied to model the air quality monitoring data in urban area, Jakarta, Indonesia.

*Presenting author: wasiah@tl.itb.ac.id

#Corresponding author: driejana@tl.itb.ac.id

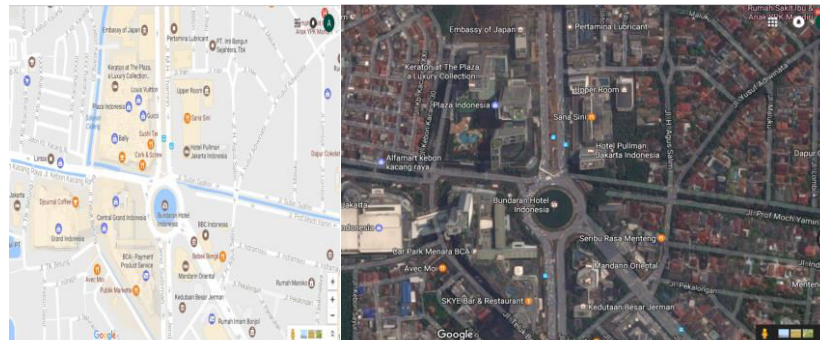


Figure 1. Air Quality Monitoring Station at Thamrin Site[24]

The applied approach is a multivariate statistical technique using Multiple Linear Regression (MLR) modeling. MLR is a statistical method that describes the relationship between one dependent variable with multiple independent variables [12], [13]. In this study, MLR is useful to model the relationship between ozone concentration (the dependent variable) with the precursors concentrations and meteorological conditions (the independent variables). Therefore MLR predicts ozone concentration at time (t +1) using previous monitored data of at time t.

II. METHODOLOGY

II.1. Location

The research was conducted in urban area of Jakarta. Air quality monitoring station is located in Thamrin, Central Jakarta ($6^{\circ}11'40.0''S$ $106^{\circ}49'24.3''E$). Central Jakarta represents community center (urban area) where offices, department stores, hotels and tourism spots are located there. This area is very crowded throughout the day. The sampling locations can be seen in **Figure 1**. This location is very important due to its function as a representative of near-road side concentration in the central of human activities. Office workers, police officers and pedestrians who passed this area are potentially exposed to air pollutants with their health impact consequences.

II.2. Data Collection and Analyses

Two-year data of 2011-2012 were gathered from Thamrin air quality monitoring station in Jakarta, Indonesia. The monitoring station belongs to Environmental Protection Agency of DKI Jakarta Province. Pollutant concentration and meteorology affect the formation and decomposition processes of ozone in ambient air. Data consisted of methane (CH_4), non-methane Hydrocarbon (NMHC), total hydrocarbon (THC), nitrogen monoxide (NO), nitrogen dioxide (NO_2), and nitrogen oxides (NO_x) concentration. In addition, meteorological data of temperature, solar radiation, relative humidity, wind speed and direction were also collected. These were 30-minutes average data that were monitored continuously. Location of Thamrin monitoring site can be seen in **Figure 1**.

Empty or missing data will affect the statistical analysis and cause bias to the model. In data screening process, missing data is filled by mean value. The missing data were replaced by average value of the same time of the year. As the requirement, the total missing data in one day-measurement should not be more than 25%. If the total missing data is more than 25 %, the data of one day-measurement is invalid and was excluded in statistical analysis.

Descriptive statistics explained the characteristics of the data group. Bivariate correlation analysis investigated the relationship among the parameters and determined the linear correlation between the parameters. Before using MLR, time lag was used in this model to assume that ozone at t time was formed by precursor concentration before t time. Ozone formation process also takes time depending on meteorological factors and precursors concentration. The next step, MLR modeling is applied with 3 variations of conditions to know the impact of different independent variable (x variable/predictor variable) to dependent variable (y variable/ozone concentration) namely, 1st variation (the meteorological parameters and ozone), 2nd variation (precursors and ozone), and 3rd variation (meteorological parameters, precursors and ozone). Multi-collinearity test was used to investigate collinearity among independent variables (meteorological parameters and precursor concentrations).

III. RESULTS AND DISCUSSION

III.1. Diurnal variation of meteorological parameters and pollutant concentration

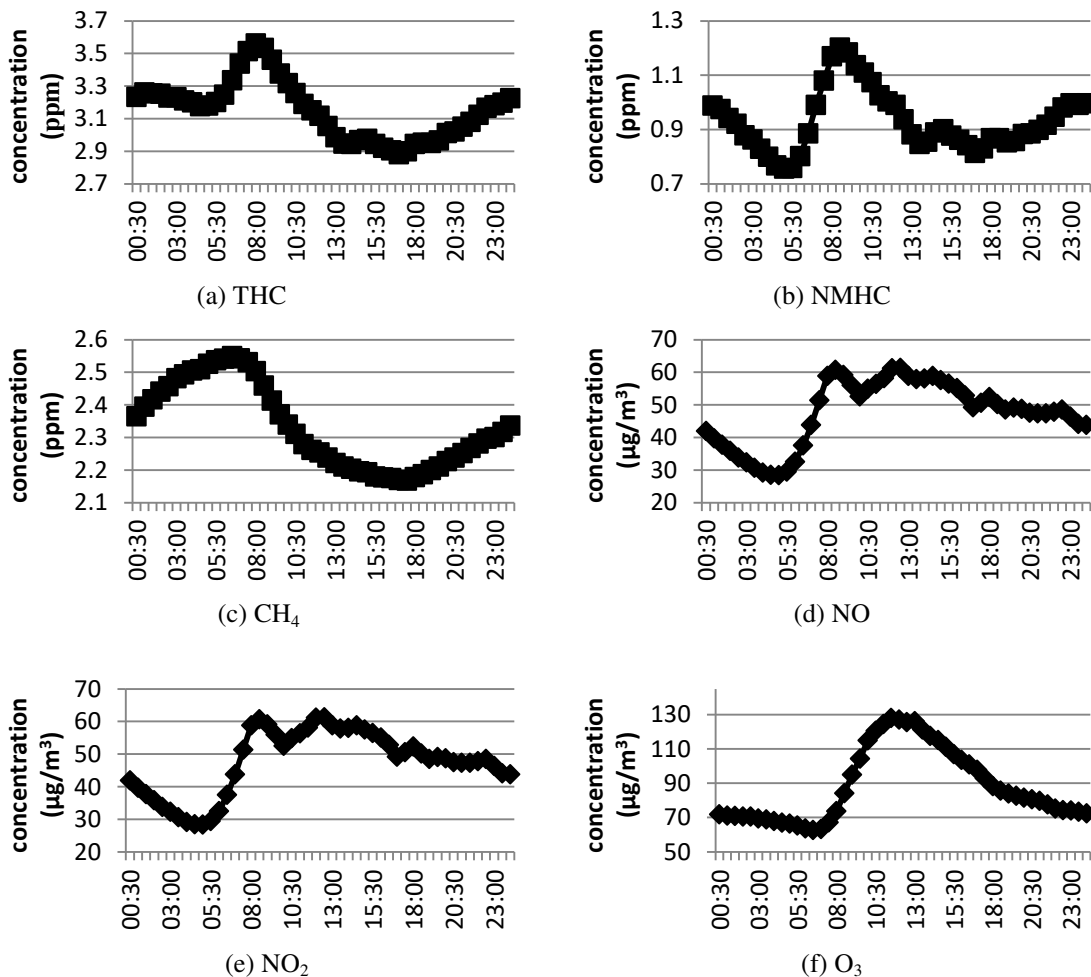
Thamrin air quality monitoring station recorded meteorological conditions such as the solar radiation, wind speed and direction, temperature and humidity as well as the pollutant concentration such as ozone, nitrogen dioxide and nitrogen oxides. In addition, the Thamrin station recorded hydrocarbon group such as methane (CH_4) and non-methane hydrocarbons (NMHC) that made up for total hydrocarbons (THC). Diurnal variation illustrates the fluctuations of measured meteorological conditions, ozone and precursor concentrations at the monitoring station. It shows the mean-hour distribution of the measurement data in two years. There are two pattern of diurnal variation of pollutant concentration. Firstly, diurnal variation with one peak is achieved by NO_2 and secondary pollutant O_3 .

due to photochemical reactions that only occur with the maximum sunlight during the day. Secondly, diurnal variation with double peak of concentrations is associated with primary pollutants such as NMHC, NO, THC, CO, CH₄ concentrations where their values were influenced by anthropogenic sources. Diurnal variations of data for two years were indicated by **Figure 2** and **Figure 3**.

Fluctuations of concentration are influenced by various factors. Motor vehicle contribute mostly primary pollutants concentration such as NMHC, NO, THC, CH₄. Combustion of fuel that contained sulfur compound is source of sulfur dioxide. Based on **Figure 2 (e) and (f)**, while CO is emitted by incomplete combustion of fuel. It usually formed when traffic jam occurred. It seemed that the high CO concentration tends to occur at peak hours in the morning or afternoon. Ozone formed several hours

after the peak concentration of precursor occurred. It indicates that the photochemical processes between the precursors and sunlight produce ozone are quite complex and takes time to react.

When the mixing process occurred and supported by favorable weather conditions, these factors can increase the secondary pollutant concentration (ozone). At night, because human activity tends to decrease, vertical mixing and photochemical reactions in the troposphere is less effective, ozone concentration also decreases [14]. For meteorological factors profile, range of wind speed is 0.8 - 1.3 m/s. Solar radiation and temperature have the same patterns when they tend to increase after sunrise and plummet after sunset. Their patterns reverse humidity that is greater at night and dawn. When it is noon, humidity reaches the lowest point.



III.2. Bivariate Analysis

To know relationship between parameters, bivariate correlation analysis (Pearson coefficient, $\alpha = 0.05$) was applied to analyze the data. **Table 1** shows that p-values between ozone concentrations with all independent variables are < 0.001 , indicating significant linear correlation between variables.

At the Thamrin station, correlation coefficient between ozone and its precursors (CH_4 , NO , NMHC , THC), wind speed and humidity are negative. These results indicate that these pollutants are playing role as precursors to ozone photochemical reaction, so that their concentration as the reactants decreased when ozone was produced.

Precursor concentration especially VOC decreased due to the photochemical reaction producing ozone [15]. The role of NO_2 as a precursor should be indicated by a negative relationship. But in the contrast, Thamrin showed positive relationship between ozone and NO_2 . Thamrin area is very crowded throughout the day with considerably high NMHC emission, therefore it might be that NO_2 was produced in a much faster rate than its consumption due to ozone production. Another reason due to thermal inversion that relates positive association between ozone and NO_2 [14].

Meteorological factors influence air pollution phenomenon [16], [17] and relate to pollutant concentration [18]. Negative correlation coefficient happens between ozone concentration with humidity, wind direction and speed. This correlation does not indicate that wind direction and have important role in ozone formation. Wind helps mechanical mixing and displacement of primary pollutants from anthropogenic sources (motor vehicle) to another place. However the negative relationship occurs between humidity and ozone concentration describes the positive correlation between temperature, solar radiation to ozone concentration. The high temperature is affected by solar radiation and lower humidity causes the tropospheric ozone formation reacts rapidly. This study support that air temperature and solar radiation correlate linearly with ozone concentration [19],[1],[20]. Besides, solar radiation is a major factor in ozone formation by initiating the photolysis of NO_2 and VOC as the main precursors of ozone.

III.3. Statistical Modeling

Ozone is a secondary pollutant and its concentration in ambient area largely depends on two independent factors, namely the emissions rate of primary pollutants and meteorological conditions. The multiple linear regression (MLR) modeling estimate ozone concentration as a function of various parameters [13],[21]. These parameters had been measured at the monitoring station. Independent variables consisted of CH_4 , CO , NMHC , NO , NO_2 , THC , wind speed, temperature, solar radiation and humidity, while ozone concentration was the dependent parameter. Three scenarios of predictor variations were applied in MLR. **Table 2.** shows the MLR results with its significance value.

The MLR equations show that the relationship between ozone and its precursor parameters can not be

described by a linear relationship only because it has small adjusted R^2 value on the first variation (0.1548 of Thamrin). Complex relationship between ozone and precursors of chemical reactions is difficult to be captured by a linear relationship. In the contrast, the linear relationship of different meteorological parameters with ozone or all parameters with ozone shows higher value of R^2 . In the second variation of MLR, adjusted R^2 value is 0.4284 and the third variation of adjusted R^2 is 0.4632. This value shows that independent variables could reasonably predict ozone concentration and can explained 46.32% of its concentration variation.

Other studies with the same modeling method showed that R^2 value were relatively higher than this study such as Turkey during warming seasons (0.92[8]), Taiwan (0.86[12]), Kuwait(0.69[11]), Greece (0.653[22]), Croatia during summer season (0.8 [23]). Many factors could cause the differences. Firstly, their sampling locations are located in high places [23] and no tall building near the sampling station [11]. In this paper, the sampling location was very near road side with heavy traffic and most calm winds were detected, therefore pollutants tend to accumulate. When ozone is formed, precursors from motor vehicle emissions were still high. Secondly, the scavenging process of ozone and nitrogen oxide (NO) might be altered due high production of NO_2 reacts with VOCs then produce more ozone. Photochemistry and meteorological factors variously affect ozone, depend on time and location. Their concentration profiles based on MLR model were illustrated in **Figure 4.**

In this modeling with many independent variables, multi-collinearity test was needed. This test helps to investigate high correlation between independent variables in the regression model. Multi-collinearity occurs when most of interrelated variables are used as independent variable of regression model. Multi-collinearity may cause problems such as regression coefficient estimates cannot be determined and/or the error variance and standard are infinite. In other conditions, statistics can estimate regression coefficients, but the error variance and standard value are large. Multi-collinearity is calculated by Variation Inflation Factor (VIF) value[25]. In all three models, temperature's VIF value is always higher than 5 but lower than 10. The highest VIF value in third variation is 6.54. This value indicates that temperature is correlated to other variables.

Based on the model diurnal variation, the second scenario predicted ozone diurnal pattern closely to the measured ozone. Meanwhile, the other variations tend to underestimate the predicted ozone concentration. The second model did not include any of precursor concentrations, suggesting the sole influence of meteorology on ozone concentration variation, with $R^2 = 0.43$. In the second model, temperature was highly correlated with solar radiation (0.755), while temperature and solar radiation correlations to ozone were 0.513 and 0.397, respectively. It seemed that temperature is a better predictor of ozone than solar radiation.

Table 2. Multiple Linear Regression Results Thamrin Station

Thamrin	predictor	coef. Estimate	Std. error	t value	P-value	multiple R ²	Adjusted R ²
1st Variation	(Intercept)	30.67865	1.967	15.597	<2x10 ⁻¹⁶	0.1548	0.1546
	CH ₄	16.96437	0.88337	19.204	<2x10 ⁻¹⁶		
	NMHC	-16.47764	0.84279	-19.551	<2x10 ⁻¹⁶		
	NO	-0.09094	0.01011	-8.998	<2x10 ⁻¹⁶		
	NO ₂	0.72315	0.01004	72.013	<2x10 ⁻¹⁶		
	THC	0.66405	0.70188	0.946	0.171		
2nd variation	(Intercept)	-565.6	7.543	-74.99	<2x10 ⁻¹⁶	0.4285	0.4284
	Wind Speed	-16.07	0.248	-64.72	<2x10 ⁻¹⁶		
	Radiation	0.026	0.002	15.71	<2x10 ⁻¹⁶		
	Humidity	2.334	0.036	65.21	<2x10 ⁻¹⁶		
	Temperature	17.24	0.186	92.69	<2x10 ⁻¹⁶		
3rd variation	(Intercept)	-568.6	7.686	-73.985	<2x10 ⁻¹⁶	0.4634	0.4632
	CH ₄	6.537	0.733	8.917	<2x10 ⁻¹⁶		
	Wind Speed	-13.42	0.272	-49.288	<2x10 ⁻¹⁶		
	Radiation	0.04	0.002	23.781	<2x10 ⁻¹⁶		
	Humidity	2.576	0.036	72.505	<2x10 ⁻¹⁶		
	NMHC	-11.55	0.679	-17.009	<2x10 ⁻¹⁶		
	NO	-0.011	0.008	-1.354	0.176		
	NO ₂	0.343	0.009	36.746	<2x10 ⁻¹⁶		
Temperature	16.57	0.186	89.037	<2x10 ⁻¹⁶			
THC	-6.268	0.572	-10.968	<2x10 ⁻¹⁶			

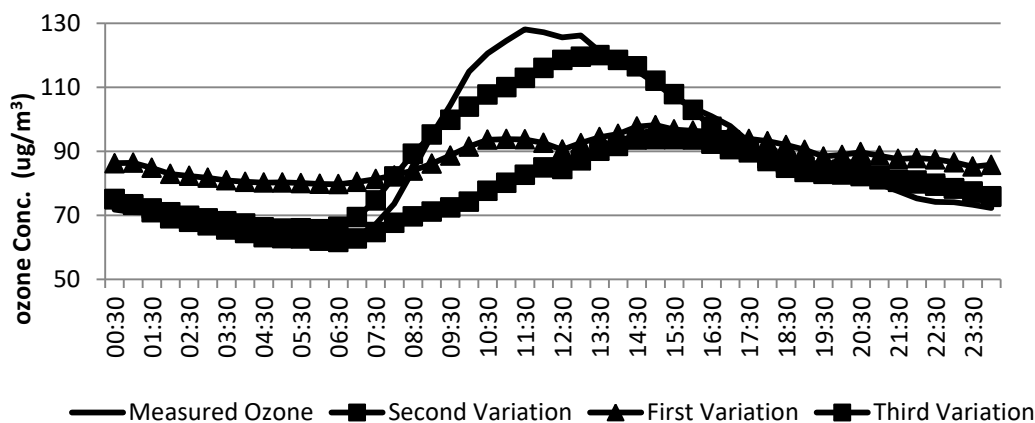


Figure 4. Ozone Concentration Measurement Profile and Multiple Linear Regression in Thamrin

IV. CONCLUSION

Multiple linear regression (MLR) analysis was used to develop ozone prediction model. Based on MLR result, meteorological parameters and precursors could model ozone with R² is 0.4634 in Thamrin station. The diurnal pattern was found to be affected strongly by meteorology, particularly temperature and solar radiation. Further work to improve the model can be done with more advance statistical methods.

ACKNOWLEDGEMENT

This research was part of Riset Desentralisasi DIKTI 2014 entitled "Prediction model for method development of air-pollution public warning system" funded by the Ministry of Research and Higher Education of the Republic of Indonesia (*Menristekdikti*). The authors wish to thank Environmental Protection Agency of DKI Jakarta Province for providing the data in this study.

V. REFERENCES

- [1] A. Akdemir, U. Ozel, and O. N. Ergun, "Multivariate Regression Analysis for Ground-Level," vol. 89, pp. 84–89, 2013.
- [2] K. Wark and C. F. Warner, *Air Pollution: Its Origin and Control*, 2nd ed. New York: Harper & Row, 1981.
- [3] W. H. O. WHO, "Ozone and other photochemical oxidants," 2000.
- [4] B.-A. H., P. Victor, L. Dana, B. S. I., S. C. M., and R. A. Rascon-Pachecho, "Ozone, Suspended Particulates and Daily Mortality in Mexico City," *Am. J. Epidemiol.*, vol. 145, no. 3, 1997.
- [5] R. Driejana, "Initial Assessment Of Ambient Ozone Concentration In Bandung, Indonesia," in *The regional workshop of Better Air Quality in Asia and Pacific Rim Cities Conference*, Manila, December 2003.
- [6] J. H. Seinfeld and J. Wiley, *ATMOSPHERIC From Air Pollution to Climate Change, Second Ed.* Canada: John Wiley & Sons, 2006.
- [7] M. E. Jenkin and K. C. Clemitshaw, "Chapter 11 Ozone and other secondary photochemical pollutants: chemical processes governing their formation in the planetary boundary layer," *Dev. Environ. Sci.*, vol. 1, pp. 285–338, 2000.
- [8] B. Özbay, G. A. Keskin, and S. Ayberk, Şenay Çetin Doğruparmak, "Multivariate methods for ground-level ozone

- modeling,” vol. 102, pp. 57–65, 2011.
- [9] A. Nebot, V. Mugica, and A. Escobet, “Ozone prediction based on meteorological variables: a fuzzy inductive reasoning approach,” 2008.
- [10] S. Samadianfard, R. Delirhasannia, and E. Agirre-basaruko, “Comparative analysis of ozone level prediction models using gene expression programming and multiple linear regression,” vol. 30, 2013.
- [11] S. A. Abdul-wahab, C. S. Bakheit, and S. M. Al-alawi, “Principal component and multiple regression analysis in modelling of ground-level ozone and factors affecting its concentrations,” vol. 20, pp. 1263–1271, 2005.
- [12] T. Pai, P. Sung, C. Lin, H. Leu, and Y. Shieh, “Predicting hourly ozone concentration in Dali area of Taichung County based on multiple linear regression method,” pp. 127–132, 2010.
- [13] A. Lengyel, K. He, and R. Rajko, “Prediction of ozone concentration in ambient air using multivariate methods,” vol. 57, pp. 889–896, 2004.
- [14] S. Abdul-Wahab, W. Bouhamra, Lh. Ettouney, B. Sowerby, and B. D. Crittenden, “Predicting Ozone Levels: A Statistical Model for Predicting Ozone Levels in the Shuaiba Industrial Area, Kuwait,” *Environ. Sci. Pollut. Res.*, vol. 3, no. 4, pp. 195–204, 1996.
- [15] W. L. Chameides, • F. Fehsenfeld, • J. Martinez M. O. Rodgers, • C. Cardelino, • D. Parrish, J. G. W. Lonneman, A. R. Lawson, R. A. Rasmussen, P. Zimmerman, and A. T. W. P. Middleton, “Ozone Precursor Relationships in the Ambient Atmosphere,” *J. Geophys. Res.*, vol. 97, no. 91, pp. 6037–6055, 1992.
- [16] E. W. Hewson, “Meteorological Factors Affecting Causes and Controls of Air Pollution Meteorological Factors Affecting Causes and Controls of Air Pollution *,” *J. air Pollut. Control Assoc.*, vol. 5, no. April 2014, pp. 37–41, 1956.
- [17] F. Hosseinibalam and A. Hejazi, “Influence of Meteorological Parameters on Air Pollution in Isfahan,” vol. 46, 2012.
- [18] M. B. Celik and İ. Kadi, “The Relation Between Meteorological Factors and Pollutants Concentrations in Karabük City,” *J. Sci.*, vol. 20, no. 4, pp. 87–95, 2007.
- [19] E. Stathopoulou, G. Mihalakakou, M. Santamouris, and H. S. Bagiorgas, “On the impact of temperature on tropospheric ozone concentration levels in urban environments,” vol. 2, no. 2, pp. 227–236, 2008.
- [20] D. Pudasainee, B. Sapkota, and M. Lal, “Ground level ozone concentrations and its association with NO_x and meteorological parameters in Kathmandu valley, Nepal,” vol. 40, no. x, pp. 8081–8087, 2006.
- [21] S. I. V Sousa, F. G. Martins, and M. C. Pereira, “Prediction of ozone concentrations in Oporto city with statistical approaches,” vol. 64, pp. 1141–1149, 2006.
- [22] K. P. Moustris, P. T. Nastos, I. K. Larissi, and A. G. Paliatsos, “Application of Multiple Linear Regression Models and Artificial Neural Networks on the Surface Ozone Forecast in the Greater Athens Area, Greece,” *Adv. Meteorol.*, vol. 2012, p. 8, 2012.
- [23] E. Kova, J. Brana, and V. Gvozdi, “Ecological Informatics Impact of meteorological factors on ozone concentrations modelled by time series analysis and multivariate statistical methods,” vol. 4, pp. 117–122, 2009.
- [24] Google Maps., 2017. Accessed 1 May 2017 .
- [25] B.G. Tabachnick and L.S.Fidel, *Using Multivariate Statistics*, 6th.Ed., Pearson, Boston, 2013.