Journal of Civil Engineering and Environmental Technology
p-ISSN: 2349-8404; e-ISSN: 2349-879X; Volume 7, Issue 1; January-March 2020, pp. 16-21

http://www.krishisanskriti.org/Publication.html

Use of Multiple Linear Regression Technique to Validate Monitored Data of PM2.5 and Ozone

Udit Mamoriya[a], Athar Hussain[b], Mamta Prakash[c], Garima[b] and Mridul Sharma[b]

[a]*Department of Civil Engineering, School of Engineering, Gautam Buddha University, Greater Noida, U.P. India*
[b] *Department of Civil Engineering, Ch. Brahm Prakash Government Engineering College, Jaffarpur, New Delhi 110073, India*
[c]*Department of Environmental Engineering, National Environmental Engineering Research Institute, Naraina, New Delhi, India*

Abstract—*Increasing air pollution across the globe is of a serious concern for most of the countries worldwide. It is difficult to control the rate of pollution in developing countries like India. In present study major pollutant PM2.5 and ozone as secondary pollutant has been monitored in ambient air and concentration for the same was being determined in the laboratory. The obtained results has been used to validated using the multiple linear regression model. The pollutants responsible for production of selected pollutants has also been considered in the study. This model was made by using all the real time monitoring pollutants and meteorology data. Real time monitoring data from CPCB (Central pollution control board) was taken of 6 months and the study area was Shadipur. This data includes SO_2, NO_2, CO PM2.5, Ozone and including meteorology data of wind speed, temperature, and relative humidity, solar radiation was also collected. All the pollutants were analyzed manually too, and the study area included Narayana industrial area.*

Introduction

Outdoor air pollution is a major environmental health problem. Exposure to air pollution leads to increased risk of respiratory diseases like acute respiratory infections and chronic obstructive pulmonary diseases; and cardiovascular diseases, such as stroke and ischemic heart disease. According to World Health Organization, one in eight total global deaths occurs as a result of exposure to air pollution. Over 3.5 million people die each year from outdoor air pollution. Low- and middle-income countries, especially the Western Pacific and South-East Asian countries account for about 88% of those premature deaths. According to a recent study by the Organization for Economic Co-Operation and Development (OECD), in India, the cost of air pollution to society in 2010 was estimated at US$ 0.5 trillion. Ozone is one of the air pollutants of major concern globally. Higher levels of ozone in the air can affect human health, leading to breathing problems, asthma exacerbation and reduced lung function. Several studies have also shown that daily mortality and heart diseases increase with exposure to high levels of ozone. Thestudy portrays that ozone level is an important air pollutant causing cardiovascular and respiratory diseases. Its levels are above the recommended threshold at various stations in India. Therefore, there is a need to monitor it and take steps to reduce its levels.

Most of the evidence on the health impacts of ground-level ozone comes from animal studies and controlled clinical studies of humans focusing on short-term acute exposure. Clinical studies have documented an association between short-term exposure to ground-level ozone at concentrations of 200–500 $\mu g/m^3$and mild temporary eye and respiratory irritation as indicated by symptoms such as cough, throat dryness, eye and chest discomfort, thoracic pain, and headache (WHO 1979, 1987). Ozone (O_3) is a reactive oxidant gas produced naturally in trace amounts in the Earth's atmosphere. Most of the Earth's atmospheric ozone (about 90%) is found in the stratosphere where it plays a critical role in absorbing ultraviolet radiation emitted by the Sun. Figure 1. shows the stratospheric ozone at 35°N in September 1996.

The peak in ozone molecular number density (concentration) occurs in the region of 20-30 km. The so-called stratospheric ozone layer absorbs virtually all of the solar ultraviolet radiation of wavelengths between 240 and 290 nm.

Such radiation is harmful to unicellular organisms and to surface cells of higher plants and animals. In addition, ultraviolet radiation in the wavelength range 290-320 nm, so- called UV-B, is biologically active. A reduction in stratospheric ozone leads to increased levels of UV-B at the ground, which can lead to increased incidence of skin cancer in susceptible individuals. (An approximate rule of thumb is that a 1% decrease in stratospheric ozone leads to a 2% increase in UV-B.)

A significant amount of naturally occurring ozone, about 10-15% of the atmospheric total, is found in the troposphere (Fishman et al. 1990). The total amount of O3 in the atmosphere, stratosphere and troposphere combined, is extremely small. Atmospheric ozone made at Montsouris, near Paris from 1876 to 1910 have been reanalyzed by Volz and Kley (1988), who recalibrated the original measurement technique. Their analysis showed that surface ozone mixing ratios near Paris over 100 years ago averaged about 10 ppb; current mixing ratios in the most unpolluted parts of Europe average between 20 and 45 ppb (Volz and Kley 1988; Bojkov 1988; Crutzen 1988; Staehelin and Schmid 1991; Oltmans and Levy 1994; Janach 1989).

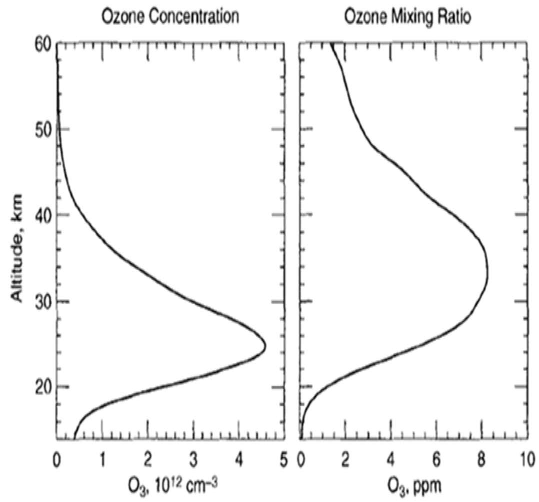

Figure 1: Stratospheric ozone profile over Northern Hemisphere mid-latitude (35°N) in September 1996 as measured by satellite with the Jet Propulsion Laboratory FTIR (Fourier transform infrared) spectrometer

Methodology

Delhi city had been identified within 24 critically polluted areas in India as per Central Pollution Control Board (CPCB).Air quality monitoring was carried out for a period of 6 months on the terrace of NEERI Zonal Laboratory, Naraina industrial area Phase-1 in Delhi. 24 hourly averaged real time air quality data of PM 2.5, Sulphur dioxide (SO_2), nitrogen dioxide (NO_2) ozone (O_3), carbon monoxide (CO) for a period 6 months from October 2015- March 2016 at Shadipur, obtained from Central Pollution Control Board (CPCB), Delhi has been used in this study. Shadipur has been chosen as it is one of the continuous air quality monitoring station of CPCB.

APM-460 NL respirable dust samplers (RDS) with provision for gaseous sampling APM-411 (Make: Envirotech) were used for measuring the concentrations of NO_2, O_3. SO_2 andAPM550 PM2.5 Fine particulate sampler (Make: Envirotech) was used for measuring PM2.5 concentration in the ambient air. The RDS APM-460 and APM550 were placed at the terrace of the building to measure the concentration of gaseous air pollutant (SO_2, NO_2, Ozone,) and Particulate matter (PM2.5).

Results and Discussion

The wind direction at shadipur was analyzed using a wind rose diagram, the objective was to find the prevailing wind direction at shadipur. The wind rose diagram showed the wind direction and the relative frequency of the wind, a color scale is used to show the wind speed in m/s. The wind rose diagram was generated using a software named Hydrognomon.

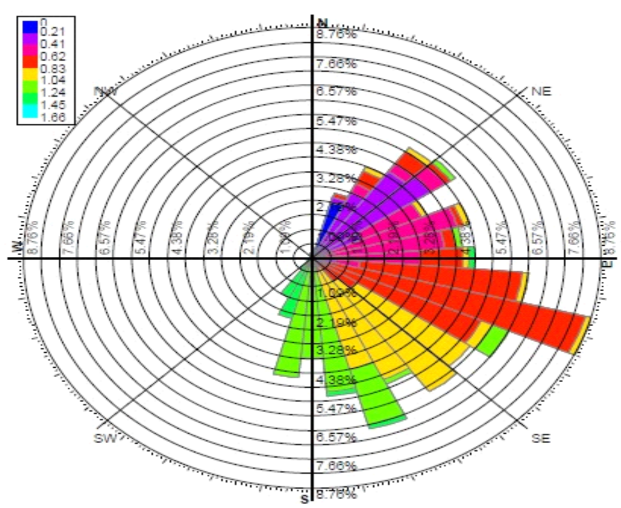

Figure 2: Wind rose diagram of the study period of the present study area

From the Figure.2 it can be concluded that during the six months period the prevailing wind direction at the site shadipur was from (SE) direction. Noticeably there was also wind coming from the (NE), and (S) direction. But (SE) appears to be dominating direction however.

Firstly, CPCB Shadipur data is selected to make a base model for validation of NEERI Zonal Lab, Nariana industrial area. Using the multiple regression model all the variables including Ozone, Pm2.5, SO_2, NO_2, RH, Temperature, Solar radiation, wind speed and CO were used. Considering ozone as dependent variable and all the other parameters as the independent variables the regression analysis was being carried out. The data summarized in Table.1 indicates that all the variables are correlated with each other and the relation coming out of the used values has been summarized. It has also been observed that the temperature values as a variable shows a high correlation of 0.879 with ozone. It can also be seen that Ozone is having a negative correlation with NO_2 and is found to be averagely significant.

Table 1: Correlation matrix for ozone evaluation used in present study

Pearson Correlation OZONE	**OZONE**	**CO**	**NO_2**	**SO_2**	**SR**	**RH**	**WS**	**TEMP**	**PM2.5**
	1.000	-0.274	-0.343	0.126	0.131	0.165	-0.254	0.879	0.262
CO	-0.274	1.000	0.529	0.018	-0.053	0.159	-0.055	-0.242	0.004
NO_2	-0.343	0.529	1.000	-0.206	-0.208	-0.047	0.165	-0.333	-0.133
SO_2	0.126	0.018	-0.206	1.000	0.009	0.177	-0.340	0.103	0.364
SR	0.131	-0.053	-0.208	0.009	1.000	0.017	-0.133	0.124	-0.219
RH	0.165	0.159	-0.047	0.177	0.017	1.000	-0.320	0.175	0.296
WS	-0.254	-0.055	0.165	-0.340	-0.133	-0.320	1.000	-0.284	-0.650
TEMP	0.879	-0.242	-0.333	0.103	0.124	0.175	-0.284	1.000	0.265
PM2.5	0.262	0.004	-0.133	0.364	-0.219	0.296	-0.650	0.265	1.000

The CO as a variable possess the second highest correlation of -0.274 with ozone which is also not very significant. Therefore a stepwise method is being adapted using multiple linear model to eliminate the unwanted correlation variables. Therefore, keeping in view the related variables using the existing model includes parameters CO and Temperature and the correlated values has been assessed. The adjusted correlation values for Ozone with Temperature and CO has been found to be 0.771 and 0.75 respectively. Therefore the present multiple linear model proved to be authentic in satisfying the derived results.

Table 2: Parameters of model validation used for ozone predicting present study

MODEL	**R**	**R Square**	**Adjusted R Square**	**Std. Error of the estimate**
1	0.881^b	0.776	0.775	0.09757

The standard coefficients used for creating a multiple linear regression equation for Ozone is summarized in Table.3.Using these standardized coefficients following equation was derived (Eq.1)

$$O_3=(0.863)(\text{temperature})+(-0.065)(CO)(1)$$

Using the above as model equation for CPCB data, validation of NEERI data was being carried out. Using the model equation the simulated data was being produced the study period of both study areas has been compared.

Table 3: Standard coefficient used for creating a multiple linear regression equation for Ozone

MODEL	**Unstandardized coefficients**		**standardized coefficient**	**t**	**Sig.**	**Correlations**		
	B	**Std. Error**	**Beta**			**Zero -order**	**Partial**	**Part**
(CONSTANT)	0.072	0.013	-	5.528	0	-	-	-
TEMP	0.849	0.023	0.863	36.331	0	0.879	0.870	0.837
CO	-0.111	0.041	-0.065	-2.724	0.007	-0.274	-0.131	-0.063

The variation in simulated values obtained using CPCB and NEERI data in present study is depicted through Figure3. Perusal of data prescribed in Figure 3 indicates that there is no relationship between the data obtained by simulation using Ozone CPCB and NEERI data.

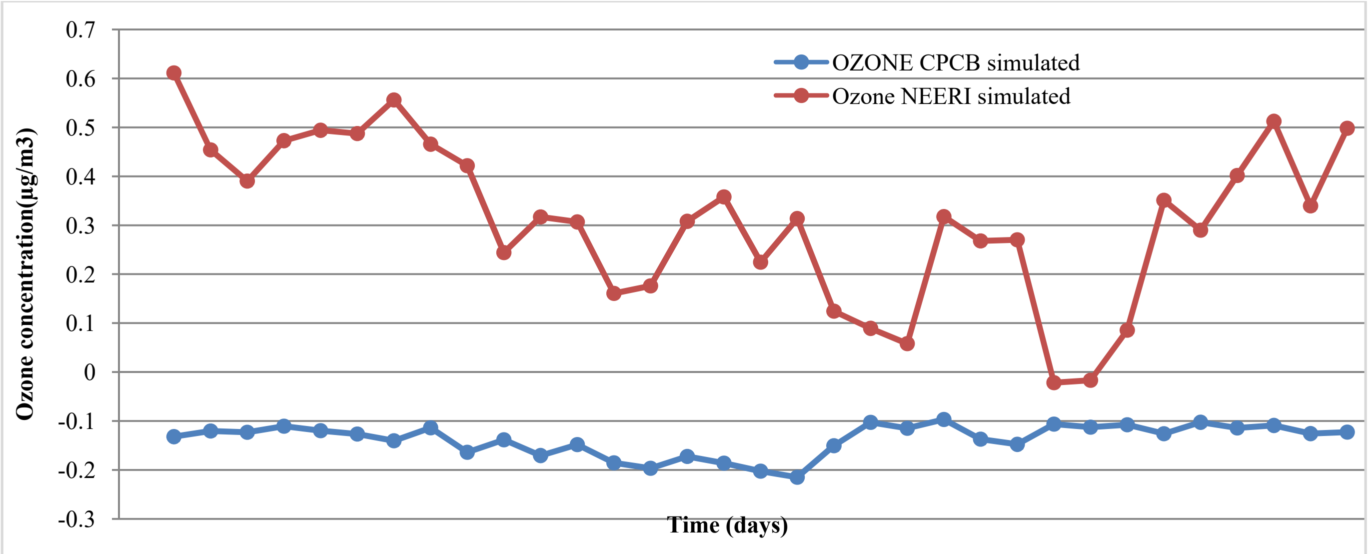

Figure 3:1 Variation in CPCB and NEERI simulated data

PM2.5

The multiple regression model has been used with all the variables including Ozone, PM2.5, SO_2, NO_2, Rh, Temperature, Solar radiation, wind speed and CO. However, PM2.5 was being considered as the dependent variable with all the other parameters as independent variables. The obtained values as summarized in Table.4 indicates all the variables were found to be correlated with each other as the relationship with some appropriate values has been observed. From the table it can be seen that the CO and Solar radiation as a variable possess a correlation with 0.329 and -0.350 against PM2.5. It can also be seen that Ozone is having a negative correlation with NO_2 which not very significant average relationship. However, the third highest correlation of -0.266 has been observed with temperature. Overall it can be seen that there was no such correlation between any of the variables.

Table 4: Correlation matrix for PM2.5 for model used in present study

Pearson Coefficient PM2.5	**PM2.5**	**CO**	**NO_2**	**SO_2**	**SR**	**RH**	**WS**	**TEMP**	**OZONE**
	1.000	0.329	0.256	0.155	-0.350	0.112	-0.095	-0.266	-0.236
CO	0.329	1.000	0.508	-0.153	-0.021	0.154	-0.519	-0.094	-0.156
NO_2	0.256	0.508	1.000	-0.225	-0.016	0.242	-0.349	-0.144	-0.147
SO_2	0.155	-0.153	-0.225	1.000	-0.156	-0.198	0.215	-0.224	-0.042
SR	-0.350	-0.021	-0.016	-0.156	1.000	-0.053	-0.014	0.159	0.129
RH	0.112	0.154	0.242	-0.198	-0.053	1.000	-0.212	-0.578	-0.571
WS	-0.095	-0.519	-0.349	0.215	-0.014	-0.212	1.000	0.197	0.174
TEMP	-0.266	-0.094	-0.144	-0.224	0.159	-0.578	0.197	1.000	0.620
OZONE	-0.236	-0.156	-0.147	-0.042	0.129	-0.571	0.174	0.620	1.000

However, stepwise method using multiple linear model eliminated the unwanted correlation variables. Keeping the related variable the prepared model includes solar radiation, CO and Temperature and overall adjusted correlation value for Ozone with solar radiation, CO and Temperature has been found to be 0.117, 0.217, and 0.247 respectively. The obtained values does not show a very satisfying model using multiple linear model.

Table 5: Model summary for PM2.5 assessed in present study

MODEL	R	R Square	Adjusted R Square	Std. Error of the estimate
1	0.350[a]	0.123	0.117	0.19686
2	0.476[b]	0.226	0.217	0.18542
3	0.510[c]	0.260	0.247	0.18185

Table 6: Standard coefficients of PM2.5 with existing model

MODEL	Unstandardized coefficients		standardized coefficient	t	Sig.	Correlations		
	B	Std. Error	Beta			Zero - order	Partial	Part
(Constant)	0.565	0.027	-	20.800	0.000	-	-	-
SR	-0.593	0.123	-0.350	-4.830	0.000	-0.350	-0.350	-0.350
(Constant)	0.467	0.033	-	14.112	0.000	-	-	-
SR	-0.582	0.116	-0.343	-5.027	0.000	-0.350	-0.364	-0.343
CO	0.346	0.073	0.322	4.717	0.000	0.329	0.344	0.322
(Constant)	0.550	0.044	-	12.396	0.000	-	-	-
SR	-0.532	0.115	-0.314	-4.626	0.000	-0.350	-0.339	-0.310
CO	0.328	0.072	0.305	4.537	0.000	0.329	0.333	0.304
TEMP	-0.159	0.58	-0.187	-2.752	0.007	-0.266	-0.209	-0.184

Using the standardized coefficients following equation was derived (Eq.2).

$$PM2.5 = (-0.314)(\text{solar radiation}) + (0.305)(CO) + (\text{temperature})(-0.187) \quad (2)$$

Using the equation as the model equation for CPCB data, validation of NEERI data was done. Firstly, data of variables of the same dates for both study area CPCB and NEERI was taken. And using the model equation simulated data was produced and both study areas has been compared and assessed.

However using Eq.2 for simulation in both the study period values a similar kind of pattern throughout the study period has been observed thus indicating a good validation model. However despite the fact that correlation values were not much unsatisfying but simulated values were found to be significantly correlated and satisfying relationship.

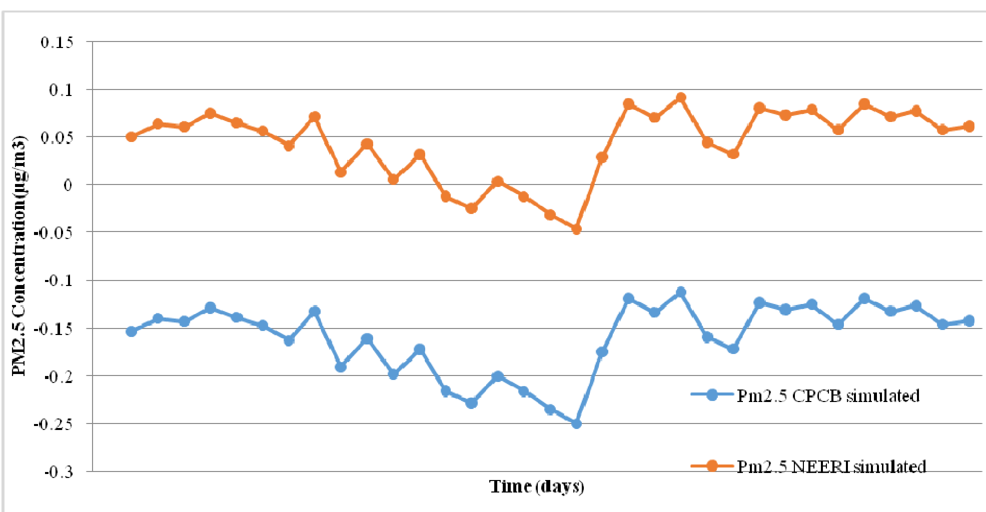

Figure 6: Variation of PM2.5 concentration in data obtained from CPCB and NEERI

Conclusion

From the present study it can be concluded that using multiple regression model the temperature values as a variable shows a high correlation of 0.879 with ozone. It can also be seen that Ozone is having a negative correlation with NO_2 and is found to be averagely significant. Also, the CO and Solar radiation as a variable possess a correlation with 0.329 and -0.350 against PM2.5. Considering SO2 as a dependent variable and correlating it with all other parameters it can be seen that SO2 is having highest positive correlation with PM2.5 and highest negative correlation with wind speed. Considering SR as a dependent variable it can be seen that SR is having maximum positive correlation with ozone and maximum negative correlation with PM2.5. However, the obtained simulation equation in present study shows a good validation model and the obtained simulated values showed significant relationship.

References

[1] Fishman, J., C. E. Watson, J. C. Larsen, and J. A. Logan, “Distribution of tropospheric ozone determined from satellite data”, J. Geophys. Res., 95, 1990, 3599 – 3617.

[2] Andreas Volz and Dieter Kley,“Evaluation of the Montsouris series of ozone measurements made in the nineteenth century”, Nature, Volume 332, 1988, Issue 6161, pp. 240-242.

[3] Rumen D. Bojkov, Carlton L. Mateer, Anne L. Hansson, “Comparison of ground-based and total ozone mapping spectrometer measurements used in assessing the performance of the global ozone observing system”, Journal of Geophysical Research Atmospheres 93(D8):9525-9533, August 1988.

[4] Crutzen P J, “Tropospheric ozone: an overviewTropospheric Ozone: Regional and Global Scale Interaction”, ISA Isdksen Ed. (Dordrecht: Reidel),1988, pp. 3–32.

[5] Staehelin, J., and W. Schmid, “Trend Analysis of Tropospheric Ozone Concentrations Utilizing the 20-Year Data Set of Ozone Balloon Soundings over Payerne, (Switzerland)” Atmospheric Environment, 1991, 25A, pp. 1739-1749.

[6] Oltmans, S.J., and H. Levy H, “Surface ozone measurements from a global network”, Atmos. Environ., 1994, 28:9-24.

[7] Janach, W. E., “Trend details, seasonal variations, and interpretation”, J. Geophys. Res., 1989, 94(D15), 18289-1 8295.