iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** VII **Month of publication:** July 2021

DOI: https://doi.org/10.22214/ijraset.2021.37061

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VII July 2021- Available at www.ijraset.com

Comprehensive Multi-Model Framework for establishing correlation between Particulate Matter in Residential and Commercial regions of Lucknow

Shivam Srivastava[1], Dr. AK Shukla[2]
[1]Student, M.Tech Environmental Engineering, IET Lucknow (U.P.) , India
[2]Professor, Department of Civil Engineering, IET Lucknow (U.P.) , India

Abstract: Our Country India is a developing Country and with the onset of heavy production, urbanization and Industrialization in the recent years has imposed a heavy threat on the Environment, The direct impact of which could be seen in the deteriorating Air Quality. Since air pollution continues to rise at an alarming rate, it affects economies and inhabitants quality of life leading to health emergency situations. This paper presents the Air Quality data interpretation and modelling to study for correlation between different Air Quality Parameters specifically Particulate Matter in Residential and Commercial regions of Lucknow- The Second largest city of Uttar Pradesh. The levels of PM2.5 and PM10 was found to be exceeding than 24 Hrs avg as well as Annual Average value (as set by NAAQS) in Commercial Region , while in Residential Region the levels of PM2.5 and PM10 were almost in acceptable range as set by NAAQS . Later on studying the comprehensive correlation between the PM2.5 and PM10 by applying various regression models, As per the levels of PM as obtained in commercial region , Compound Curve Regression Model seems to be of highest significance showing 92.1% relation response with highest Standardized Beta Coefficient having value of 2.611 . While for Residential region as per the levels of PM obtained, Cubic Curve Regression Model gave best suited result depicting 77.9% relation response amongst PM2.5 and PM10 with highest Standardized Beta Coefficient Value of 3.028 .
Keywords: Environment, Air Quality, NAAQS, Regression, Correlation

I. INTRODUCTION

Air quality is a cause for concern in India, particularly in cities and air pollutants including particulate matter (PM), sulfur dioxide (SO_2), nitrogen oxides (NOx), carbon monoxide (CO) often exceed the National Ambient Air Quality Standards (NAAQS). According to the World Health Organization (WHO), 37 cities from India feature in the top 100 world cities with the worst PM10 pollution, and the cities of Delhi, Raipur, Gwalior, and Lucknow are listed in the top 10 (WHO, 2014). A similar assessment by WHO, in 2011, listed 27 cities in the top 100. More than 100 cities under the national ambient monitoring program exceed the WHO guideline As cities are increasing in size and population, there is a steady demand for motorized vehicles in both personal and public transport sectors. This puts substantial pressure on the city's infrastructure and environment, particularly since most Indian citeshave mixed land use. (Sarath A Guttikunda et. al.)

According to the research been carried out by the WHO and Joint Research Centre (JRC) in 2015 on 50 leading countries of the world in order to gain better understanding of the leading causes of urban air pollution , on analysing the data ; the Researchers identified Traffic as being the top contributor to Air Pollution . On India level the study been conducted by Indian Institute of Toxicology Research in Lucknow in 2020 in context to rising air Pollution have concluded that Vehicular traffic has been found to be the main source of air pollution in Lucknow city. The city has become denser with traffic congestion which increases the vehicle emissions and subsequent health impact mainly for drivers, commuters, and individuals living near roadways.

Particles in atmosphere play important roles in various prospects , including epidemiological, degradation of visibility, and climate change globally , owing to their optical properties . Epidemiological studies have depicted statistical associations between health effects on human and ambient concentrations of particulate matter (PM), particularly of finer particles, which have high probability to penetrate into the lungs and are therefore more likely to increase the chances of respiratory and mutagenic diseases (S.C. Barman et.al.) On the basis of epidemiological studies been conducted have reported that , an increase of 10 µg/m³ PM2.5 (as annual average) was associated with a relative risk for total mortality of 1.14 and 1.07, respectively. It was also reported that both short and long term exposures to PM10 and PM2.5 are associated with an increase in mortality, especially in case of elderly adults with cardiopulmonary illness .

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VII July 2021- Available at www.ijraset.com

Air pollution is a growing problem in our state Uttar Pradesh primarily because of the amount of driving that takes place. Motor vehicles are the source of most of our air pollution. On the average, for every kilometers driven, a huge load of pollution is emitted into the air. Thus, through our daily transportation choices, each of us can be a part of the pollution problem or solution. Most of us are aware that driving causes air pollution, and some are even beginning to explore alternative modes of transportation. Nevertheless, many of us are entrenched in our driving habits, and rarely consider carpooling, taking the bus, or riding a bike. Because breathing clean air is one of our most basic needs, and because the cause and effect relationship between transportation and pollution is so closely linked to our everyday lives ,Understanding these connections can empower us to make a contribution toward solving this important environmental problem and make a difference in improving our own future.

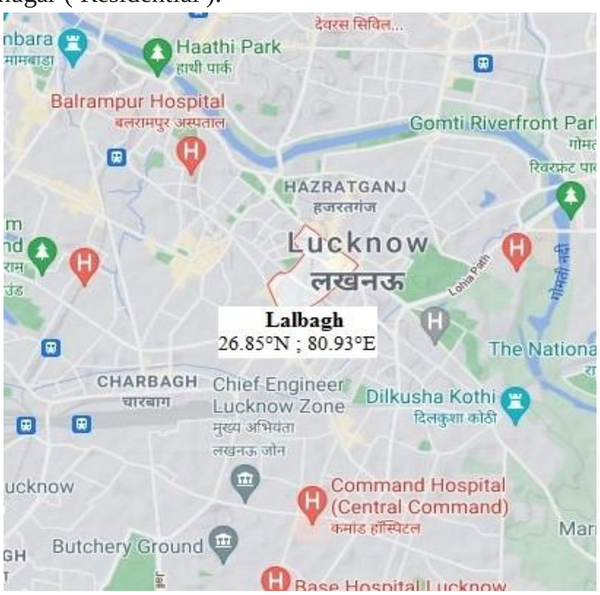
II. METHODOLOGY

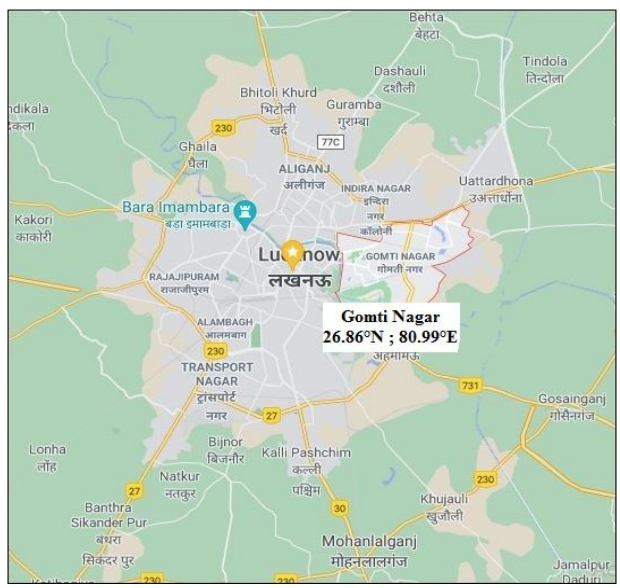
A. Study Area

The selected study area for this work is Lucknow city, the capital city of Uttar Pradesh. The city lies at 80º 57’ E longitude and 26º 51’ N latitude being 123 m above sea level. Having total area of approximately 610 sq Km. Having good infrastructure, possessing various other amenities, well feasible connected routes of transportation makes the city densely populated with a population of 28.15 lakhs (2011 census) and projected population of 37.65 lakhs in 2021 . Being the capital, Lucknow has a well connected system of roads within itself and between other cities. Due to large population of the city and presence of number of roads, numerous vehicles run on the roads of the city ranging from public transport vehicles, heavy diesel vehicles, personal owned vehicles etc. The total vehicular population in Lucknow city as on 31st march, 2020 was 2407190, which showed 9.70% growth in its number in comparison to previous year (source: RTO, Lucknow). In this phase of the work, to establish and understand good statistical coorelation amongst the various air quality parameters two sites in Commercial and Residential locations were been selected .

B. Selected Locations

Two locations were selected in the Commercial and Residential regions of Lucknow namely Lalbagh (Commercial) and Gomtinagar (Residential).





(Fig1 : Selected Locations on Map)

C. Data Monitoring and Analysis Tool

The real time Air Quality Parameters Monitoring data for Commercial Region was taken from Central Pollution Control Board , Lalbagh (near Dayanidhan Park) in the month of March 2021 and for Residential Region the real time Monitoring data was taken from Central Pollution Control Board ,Central School ,GomtiNagar (near Pickup Building , Vibhuti Khand) in the month of May 2021 .For better understanding of results through statistical model the 4-Hour average data was taken in accountabaility.

For the statistical analysis SPSS Software was been used. The SPSS software is the platform offering advanced statistical analysis, and a vast library of machine learning algorithms , analysis of texts , open source extensibility, big data integration and seamless deployment into applications. Its ease of use, flexibility and scalability make SPSS accessible to users of all skill levels. It ts suitable for projects of all sizes and levels of complexity.

Within the SPSS software family of products, SPSS Statistics supports a top-down, hypothesis testing approach to the data while SPSS Modeler exposes patterns and models hidden in data through a bottom-up, hypothesis generation approach.

Keeping in account the pros of SPSS ; the data analysis been conducted was done on SPSS Software employing various techniques which include Analysis of Variance (ANOVA) on Linear Regression Model , Logarithmic Regression Model, Inverse Regression Model, Quadratic Regression Model, Cubic Regression Model, Compound Regression Model, Power Regression Model , S Regression Model , Growth Regression Model , Exponential Regression Model and Logarithmic Regression Model .

All these models were been applied one by one and depending on R square values , ANOVA F Value and Standardized Beta Coefficients the best fit model was been predicted .

D. Data Analysis and Interpretation

Regression analysis is a set of statistical processes for estimating the relationships between a dependent variable (often called the 'outcome' or 'response' variable) and one or more independent variables (often called 'predictors', 'covariates', 'explanatory variables' or 'features').

The most common form of regression analysis is linear regression, in which one finds the line (or a more complex linear combination) that most closely fits the data according to a specific mathematical criterion. For example, the method of ordinary least squares computes the unique line (or hyperplane) that minimizes the sum of squared differences between the true data and that line (or hyperplane).

For specific mathematical reasons (see linear regression), this allows the researcher to estimate the conditional expectation (or population average value) of the dependent variable when the independent variables take on a given set of values. Less common forms of regression use slightly different procedures to estimate alternative location parameters (e.g., quantile regression or Necessary Condition Analysis) or estimate the conditional expectation across a broader collection of non-linear models (e.g., nonparametric regression).

Regression analysis is primarily used for two conceptually distinct purposes. First, regression analysis is widely used for prediction and forecasting, where its use has substantial overlap with the field of machine learning. Second, in some situations regression analysis can be used to infer causal relationships between the independent and dependent variables. Importantly, regressions by themselves only reveal relationships between a dependent variable and a collection of independent variables in a fixed dataset

Correlation is a Bivariate analysis that measures the strengths of association between two variables. In statistics, the value of the correlation coefficient varies between +1 and -1. When the value of the correlation coefficient lies around ± 1, then it is said to be a perfect degree of association between the two variables. As the correlation coefficient value goes towards 0, the relationship between the two variables will be weaker. Usually, in statistics, we measure three types of correlations: Pearson correlation, Kendall rank correlation and Spearman correlation.

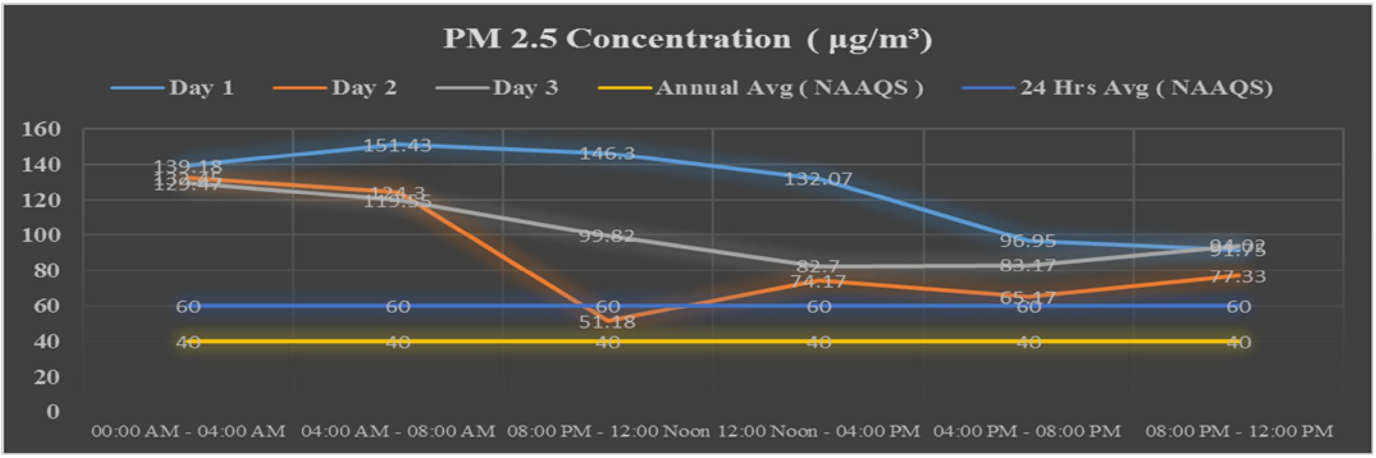
Pearson r correlation: Pearson correlation is widely used in statistics to measure the degree of the relationship between linear related variables.

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

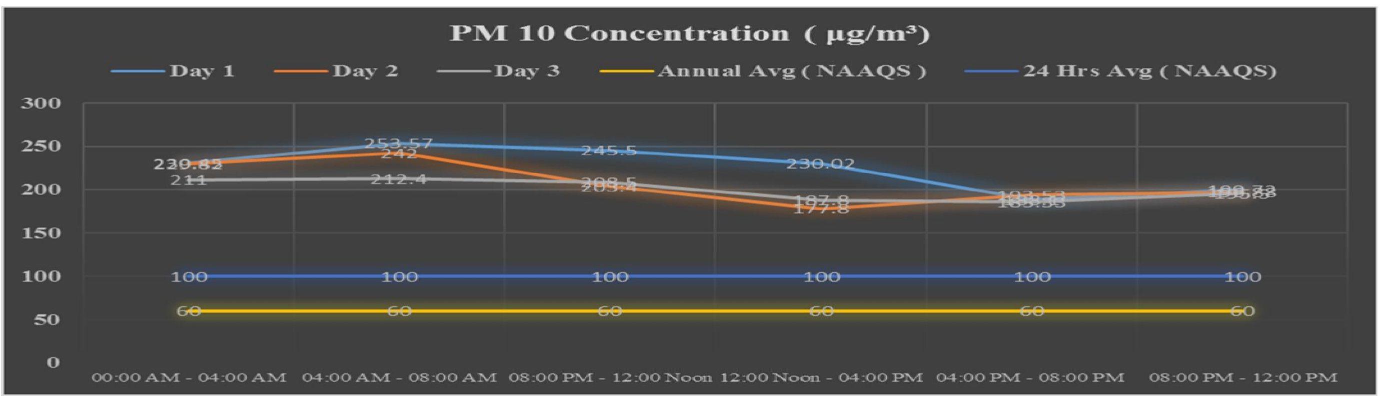
Where $\bar{x}$ is the mean of variable x values, and $\bar{y}$ is the mean of variable y values.

III. RESULTS

The Four Hourly variation data is been taken into account on three consecutive days in months of March and corresponding variation of PM 2.5 and PM 10 Concentration in Lalbagh region is plotted in Fig.2 and Fig.3 respectively.



(Fig. 2 : Variation in concentration of PM2.5 in Lalbagh Region)



(Fig. 3: Variation in concentration of PM10 in Lalbagh Region)

After applying various regression models taking PM10 as independent and PM 2.5 as dependent variable the following best suited results were been derived ;

TABLE I
Statistical Multimodel Analysis Results
(For Lalbagh Region)

Model	R	R Sqaure	Std. error of estimate	ANOVA F- Value	Standardized Beta Coefficient
Linear	0.968	0.937	7.180	59.734	0.968
Logarithmic	0.968	0.936	7.243	58.641	0.968
Inverse	0.965	0.932	7.500	54.407	-0.965
Quadratic	0.968	0.937	8.275	22.491	1.363
Cubic	0.967	0.938	8.262	22.566	1.232
Compound	0.960	0.921	0.068	46.665	2.611
Power	0.960	0.922	0.068	47.360	0.960
S	0.959	0.920	0.069	45.702	-0.959
Growth	0.960	0.921	0.068	46.665	0.960
Exponential	0.960	0.921	0.068	46.665	0.960
Logistic	0.961	0.932	0.068	46.668	0.383

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VII July 2021- Available at www.ijraset.com

Compound

Model Summary

R	R Square	Adjusted R Square	Std. Error of the Estimate
.960	.921	.901	.068

The independent variable is PM10.

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Regression	.217	1	.217	46.665	.002
Residual	.019	4	.005		
Total	.236	5			

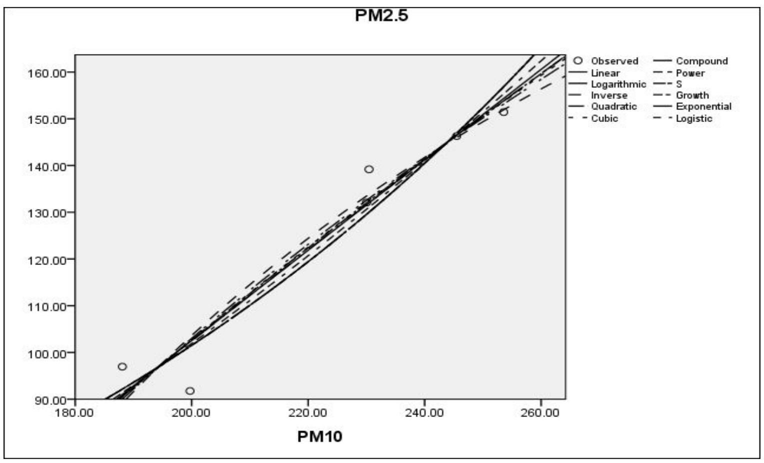
The independent variable is PM10.

Coefficients

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
PM10	1.008	.001	2.611	840.739	.000
(Constant)	19.987	5.368		3.724	.020

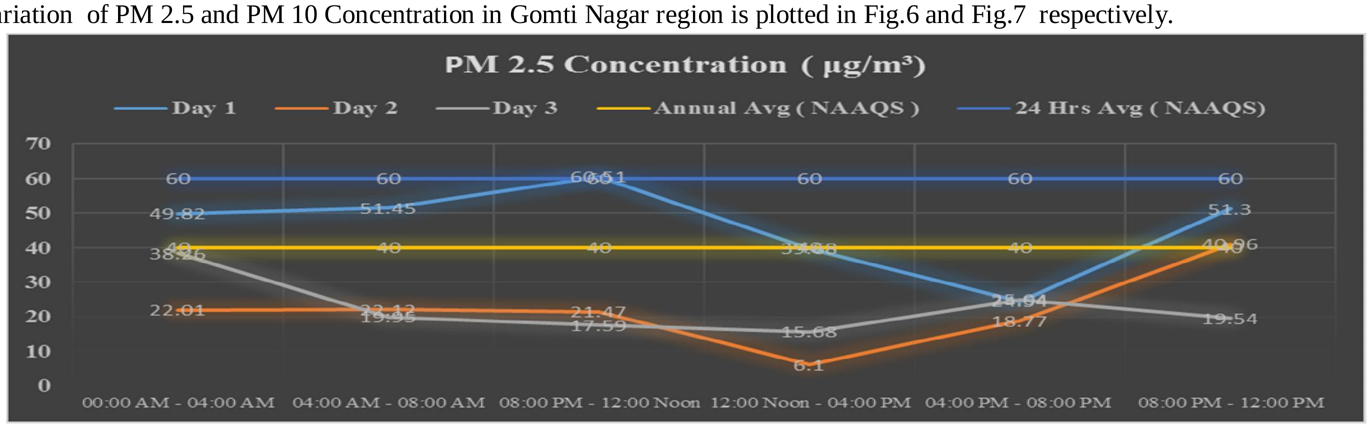
The dependent variable is ln(PM2.5).

(Fig 4 : Best Suited Model for Correlation analysis between Particulate Matter in Lalbagh Region)

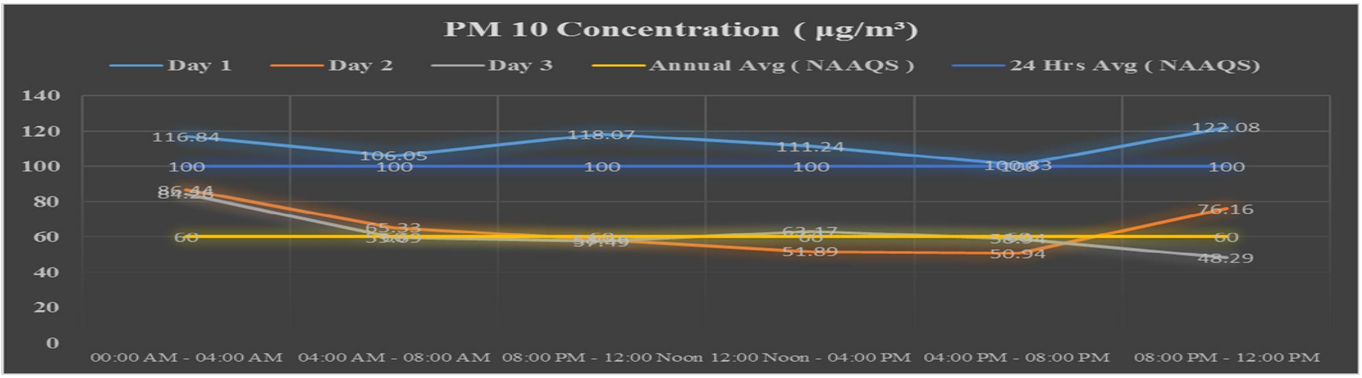


(Fig 5 : Plot of results of Various models for correlation analysis between Particulate Matter in Lalbagh Region)

The Four Hourly variation data is been taken into account on three consecutive days in months of March and corresponding variation of PM 2.5 and PM 10 Concentration in Gomti Nagar region is plotted in Fig.6 and Fig.7 respectively.



(Fig. 6 : Variation in concentration of PM2.5 in Gomti Nagar Region)



(Fig. 7: Variation in concentration of PM10 in Gomti Nagar Region)

After applying various regression models taking PM10 as independent and PM 2.5 as dependent variable the following best suited results were been derived ;

Table II
Statistical Multimodel Analysis Results
(FOR GOMTI NAGAR REGION)

Model	R	R Sqaure	Std. error of estimate	ANOVA F-Value	Standardized Beta Coefficient
Linear	0.877	0.768	7.782	56.382	0.877
Logarithmic	0.858	0.737	8.297	47.549	0.858
Inverse	0.812	0.660	9.427	33.003	-0.812
Quadratic	0.879	0.773	7.945	27.204	0.346
Cubic	0.883	0.779	8.087	17.649	3.028
Compound	0.830	0.689	0.343	37.657	2.293
Power	0.838	0.703	0.335	40.195	0.838
S	0.822	0.676	0.350	35.514	-0.822
Growth	0.830	0.689	0.343	37.657	0.830
Exponential	0.830	0.689	0.343	36.540	0.825
Logistic	0.831	0.690	0.345	37.658	0.436

Cubic

Model Summary

R	R Square	Adjusted R Square	Std. Error of the Estimate
.883	.779	.735	8.087

The independent variable is PM10.

ANOVA

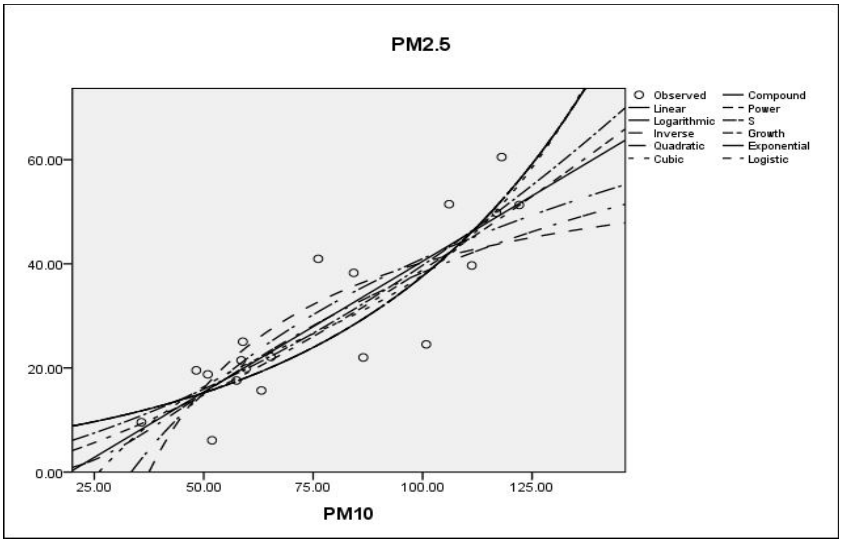
	Sum of Squares	df	Mean Square	F	Sig.
Regression	3462.908	3	1154.303	17.649	.000
Residual	981.055	15	65.404		
Total	4443.963	18			

The independent variable is PM10.

Coefficients

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
PM10	1.735	2.382	3.028	.728	.478
PM10 ** 2	-.019	.032	-5.498	-.601	.557
PM10 ** 3	8.836E-5	.000	3.390	.664	.517
(Constant)	-33.857	57.162		-.592	.562

(Fig 8 : Best Suited Model for Correlation analysis between Particulate Matter in Gomti Nagar Region)



(Fig 9 : Plot of results of Various models for correlation analysis between Particulate Matter in Gomti Nagar Region)

IV. CONCLUSIONS

Atmospheric particulate matter is a central component of the atmospheric chemical and climate system, a major air pollutant harmful to human health, and a component of biological systems and global biogeochemical cycles , Hence it is a matter of primary concern.

Based on the study done in Lalbagh and Gomtinagar Region it could be clearly seen that In Lalbagh Region , PM2.5 and PM10 both exceed the 24 Hrs Average as well as Annual Average standards as set by National Ambient Air Quality Standards (NAAQS) making it a great cause of concern for peoples living in that area , the reason for this could be attributed to its geographical location in Lucknow , also this region being a commercial area with high population density and High Traffic Density as well .

On the contrary Gomti Nagar being Residential Area with less Population Density and Less Traffic Density was been found with almost acceptable concentrations of both PM2.5 and PM10.

PM2.5 in Gomti Nagar exceeded Annual Average standard on a few days but it was below 24 Hrs average on almost each day ; while PM10 exceeded 24 Hrs average on a day but on other days it was less than 24 Hrs avg as well as Annual average range as set by NAAQS.

On application of Multimodel Regression Analysis for finding correlation between PM2.5 and PM10 for Lalbagh Region for selected days , Compound Curve Regression Model seems to be of highest significance showing 92.1% relation response with highest Standardized Beta Coefficient having value of 2.611 , though Cubic curve regression model was showing 93.8% relation response but it could not be considered as best fit model due to its low value of standardized beta coefficient value of 1.232 .

While for Gomti Nagar region on selected days , Cubic Curve Regression Model gave best suited result depicting 77.9% relation response amongst PM2.5 and PM10 with highest Standardized Beta Coefficient Value of 3.028 .

Hence from both the regions it could be predicted that both PM2.5 and PM10 show high degree of relativeness in their variation of concentration .

REFERENCES

[1] Pandey, P., Khan, A.H., Verma, A.K., Singh, K.A., Mathur, N., Kisku, G.C. and Barman, S.C., 2012. Seasonal trends of PM 2.5 and PM 10 in ambient air and their correlation in ambient air of Lucknow City, India. Bulletin of Environmental Contamination and Toxicology, 88(2), pp.265-270.

[2] Guttikunda, S.K., Goel, R. and Pant, P., 2014. Nature of air pollution, emission sources, and management in the Indian cities. Atmospheric environment, 95, pp.501-510.

[3] Barman, S.C., Singh, R., Negi, M.P.S. and Bhargava, S.K., 2008. Fine particles (PM2. 5) in residential areas of Lucknow city and factors influencing the concentration. Clean–Soil, Air, Water, 36(1), pp.111-117.

[4] Pandey, P., Patel, D.K., Khan, A.H., Barman, S.C., Murthy, R.C. and Kisku, G.C., 2013. Temporal distribution of fine particulates (PM2. 5, PM10), potentially toxic metals, PAHs and Metal-bound carcinogenic risk in the population of Lucknow City, India. Journal of Environmental Science and Health, Part A, 48(7), pp.730-745.

[5] Gupta, V. and Arya, A.K., 2016. Demonstrating urban pollution using heavy metals in road dusts from Lucknow City, Uttar Pradesh, India. Environment Conservation Journal, 17(1&2), pp.137-146.

[6] Behera, S.N., Sharma, M., Nayak, P., Shukla, S.P. and Gargava, P., 2014. An approach for evaluation of proposed air pollution control strategy to reduce levels of nitrogen oxides in an urban environment. Journal of Environmental Planning and Management, 57(4), pp.467-494.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VII July 2021- Available at www.ijraset.com

[7] Ramachandra, T.V., 2009. Emissions from India's transport sector: statewise synthesis. Atmospheric Environment, 43(34), pp.5510-5517.
[8] Muralikrishnan, R., Swarnalakshmi, M. and Nakkeeran, E., 2014. Nanoparticle-membrane filtration of vehicular exhaust to reduce air pollution–a review. Int Res J Environ Sci, 3(4), pp.82-86.
[9] Kumar, A., Patil, R.S., Dikshit, A.K. and Kumar, R., 2016. Comparison of predicted vehicular pollution concentration with air quality standards for different time periods. Clean Technologies and Environmental Policy, 18(7), pp.2293-2303.
[10] Sharma, A.R., Kharol, S.K. and Badarinath, K.V.S., 2010. Influence of vehicular traffic on urban air quality–A case study of Hyderabad, India. Transportation Research Part D: Transport and Environment, 15(3), pp.154-159.
[11] Gulia, S., Tiwari, R., Mendiratta, S., Kaur, S., Goyal, S.K. and Kumar, R., 2020. Review of scientific technology-based solutions for vehicular pollution control. Clean Technologies and Environmental Policy, pp.1-12.