International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-3, January 2020

# Modeling Air Pollution and Temperature Components to Identify Their Effects on India's Capital using Python

Sandeep Mathur, Satvik Sharma

***Abstract***— ***Various past research works have shown that temperature can alter the effect of ambient fine particles category PM 2.5 which causes the high mortality risk. In surveying air contamination impacts, temperature is generally considered as a confounder. In any case, encompassing temperature can change individuals' physiological reaction to air contamination and might alter the effect of air contamination on wellbeing results. This study investigates the interaction between monthly values of PM2.5 and monthly average temperature values in Delhi, India using data for the period 2010–2018. The computer language Python is used to analysis facts and produce the outcomes which can shape the future research work and policies to overcome both global issues- pollution and Global warming.***

***Keywords—-Data analytics, Air pollution, Metrological factors, Python, CPCB***

## I. INTRODUCTION

In recent times, air pollution and rise in temperature are two biggest issue in most of the developed and developing countries. PM2.5 could be characterized as the expansion of destructive particles to the climate which is massing in a gigantic sum causing a threatening situation for the individuals on earth. Previously various research works have been carried out to find out the impact of metrological factors on pollution contents. In [1] the factors affecting the air quality and pollution components have been analyzed for the urban city Shanghai, China. The research work stated that population and building density with energy consumptions are the major factors for the increase in air pollution in Shanghai. The performance of the pollution components has been studied and analyzed from 1994 to 2013 for the Metropolitan Area of Sao Paulo, Brazil. This research work indicated the impact of metrological factors such as ventilation and humidity on the concentration of the aerosol component [2]. The research work has been carried on the air pollution and Metrological factors data of Trabzon, Turkey city in [3]. This research work indicated the relationship between supervised outdoor air quality data and meteorological factors, such as wind speed, relative humidity ratio and temperature, is statistically analyzed, using the code SPSS.

It performs a regression analysis on the $SO_2$ data and wind speed values. As Delhi is among the list of high polluted city in the world in India similar research work have been carried extensively. In [4, 9] Air quality of Delhi, India has been monitored during XIX commonwealth games. The study showed that during the commonwealth games wind speed blown the unexpecting dust particles at the alarming level and emission rate is beyond from permittable World Health Organization limit. Concentration of air pollution is high in urban areas as compared to rural areas because the urban areas are nearer to industries [5].

In the heaviest pollution region in china, an important Convergent cross mapping research method has been deployed to analysis the relationship of metrological factors and pollution component pm 2.5. It reflects the impacts of meteorological factors such as temperature and wind speed on the PM 2.5 concentration in the local geographical regions in China. This research work proves the influence of metrological factors on the pollution component pm 2.5 [6]. The extreme value analysis has been done on the sixteen year hourly based pollution data of the Metropolitan Area of Sao Paulo, Brazil and seven-year data on the pollution data of the Metropolitan Area of Rio de Janeiro [7]. Research work has been carried out in Utah's urbanized Salt Lake Valley based on forty years of data. The interesting findings of this research work suggested that in Salt Lake valley PM2.5 is closely related to integrated atmospheric stability. Secondly, PM 2.5 is extremely above the allowed upper thrust limit in the winter seasons of the Salt Lake valley. The various pollution components such as nitrogen dioxide, carbon monoxide, Sulphur dioxide and other related components emitted from vehicles have been studied to show their hazardous effects on Delhi' citizens [10, 11]

## II. DATABASE AND METHODOLOGY

For the present study, nine successive years data from 2010-2018 have been taken into consideration which is obtained from Central Pollution Control Board (CPCB), busiest ITO station for desired PM 2.5 data. Figure 4 represents the collected data set. This dataset has been extracted from the stored data repository of CPCB [12].

**Revised Manuscript Received on January 30, 2020.**
* Correspondence Author
**Sandeep Mathur*,** Amity Institute of Information Technology Amity University, Uttar Pradesh, India. sandeep2809@gmail.com
**Satvik Sharma,** Amity Institute of Information Technology Amity University, Uttar Pradesh, India. satviks84@gmail.com

It should be in consideration that in case of not getting exact data from reliable resources average annual data value or whenever no annual average data is available Nil value has been used in analysis.

The temperature data is taken from National enters for Environmental Information (NOAA) and National Weather Service. Figure 6 signifies extracted Monthly average temperature data set for the nine years from 2010 to 2018 [13]. In any case, what was seen in 2016 was route past whatever a temperature drop could sensibly clarify [14, 15]. Accordingly started, late on fourth November, and proceeding with well into the late long periods of sixth November, a timeframe when AQI of 500 was continually recorded.

As an aside, 500 is the greatest incentive for the AQI. Regardless of whether the centralization of individual contaminations surpasses the qualities required to record this worth, the estimation of AQI will stay consistent at 500. Another fascinating perception that is seen is that the AQI esteemed in 2015 was not near 500 anytime. Be that as it may, what was more vexing was the degrees of PM 2.5(a significant toxin whose focus in PPM tracks the AQI) in 2016 [16].

They were following a pattern near that of 2015, until late on the fourth of November. By then, they relentlessly shot up to and kept up a level higher than those saw during Diwali.

This began, as indicated by the gadgets introduced by the CPCB, in places in Haryana close to the NCR area, and consistently advanced towards the NCR district.

A peculiar ascent in the degrees of PM2.5 was watched late on fourth November, and it continued towards Delhi, and during the early long periods of fifth November, the PM2.5 levels in many places in Delhi had started to mirror this spike.

This step by step developed, until the night of sixth November. This is a distinct difference to the $PM_{2.5}$ levels in the earlier year. Regardless of the praiseworthy estimates taken by the administration, it is conceivable that there probably won't be a significant change in the measure of contamination Delhi is confronting, for the straightforward explanation that it isn't the significant supporter of the contamination levels seen for the current year [17, 18].

## III. RESULT AND DISCUSSION

Python is the current generation powerful computer programming language for data analytics. It provides flexibility and ease of programming. Here python framework has been used for analyzing and interpreting air pollution data collected from Center of pollution control board (CPCB) ITO, Delhi center. Following python block of codes have been developed and used for data analysis.

==============================

**To get Pearson correlation month wise**

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
da=pd.read_excel("C:\\Users\\91963\\Desktop\\sandeep sir\\PMCO.xlsx")
df=pd.read_csv("C:\\Users\\91963\\Desktop\\sandeep sir\\temp.csv")
frames=[df,da]
corr=pd.concat(frames,axis=1)
corr=corr.corr()
pear=corr.corr(method='pearson')
plt.plot(pear.X,pear.Y,'-og')
plt.plot(corr.X,corr.Y,'-ob')
plt.show()
```

**to get correlation month wise:**

```
corr=pd.read_csv("C:\\Users\\91963\\Desktop\\sandeep sir\\correlation.csv")
plt.bar(corr.X,corr.Y)
plt.xlabel("Month")
plt.ylabel("correlation Value")
plt.title('Correlation')
```

**to compare correlation(month wise)**

```
plt.plot(pear.X,pear.Y,'-go',label='pearson corr.')
plt.plot(corr.X,corr.Y,'-ob',label='correlation')
plt.legend()
plt.xlabel('month')
plt.title('Monthwise')
plt.show()
```

**to get correlation and Pearson correlation year wise:**

```
df=pd.read_csv("C:\\Users\\91963\\Desktop\\sandeep sir\\temp.csv")
da=pd.read_excel("C:\\Users\\91963\\Desktop\\sandeep sir\\PMCO.xlsx")
frames=[df,da]
corr=pd.concat(frames,axis=1)
corr=corr.corr()
corr.to_csv('C:\\Users\\91963\\Desktop\\sandeep sir\\corr(year).csv')
corr=corr.corr(method='pearson')
corr.to_csv('C:\\Users\\91963\\Desktop\\sandeep sir\\pearcorr(year).csv')
da=pd.read_csv('C:\\Users\\91963\\Desktop\\sandeep sir\\corr(year).csv')
df=pd.read_csv('C:\\Users\\91963\\Desktop\\sandeep sir\\pearcorr(year).csv')
```

**to plot correlation year wise:**

```
plt.plot(da.X,da.Y,'-or')
plt.xlabel('year')
plt.title('correlation(year)')
```

**to plot Pearson correlation year wise**:

```
plt.plot(df.X,df.Y,'-oy')
plt.xlabel('year')
plt.title('pearson correlation(year)')
```

**to compare correlation and Pearson correlation year wise:**

```
plt.plot(da.X,da.Y,'-or',label='pearson corr')
plt.plot(df.X,df.Y,'-oy',label='correlation')
plt.xlabel('year')
plt.title('comapre(year)')
plt.legend()
plt.show()
```

==============================

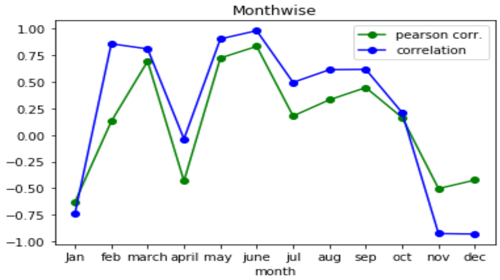


**Figure 1 Monthly Comparison (2010-2018 ) dataset**

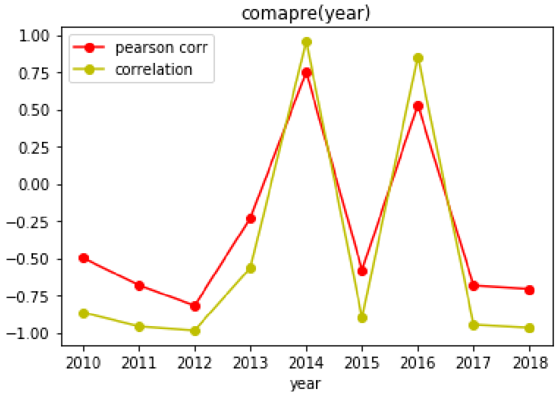


**Figure 2 Yearly Comparison (2010-2018) dataset**

Figure 1 and Figure 2 represents monthly and yearly comparisons of (2010-2018) dataset respectively for air pollution and rise in temperature.

| | Y | X |
|---|---|---|
| 2010(T) | -0.49929 | 2010 |
| 2011(T) | -0.68108 | 2011 |
| 2012(T) | -0.81881 | 2012 |
| 2013(T) | -0.22882 | 2013 |
| 2014(T) | 0.749114 | 2014 |
| 2015(T) | -0.57838 | 2015 |
| 2016(T) | 0.528909 | 2016 |
| 2017(T) | -0.68318 | 2017 |
| 2018(T) | -0.70635 | 2018 |

**Figure 3 Yearly Correlation (2010-2018) dataset**

| | Y | X |
|---|---|---|
| JAN(T) | -0.7417 | Jan |
| FEB(T) | 0.858897 | feb |
| MAR(T) | 0.810677 | march |
| APR(T) | -0.03732 | april |
| MAY(T) | 0.904735 | may |
| JUN(T) | 0.982178 | june |
| JUL(T) | 0.495394 | jul |
| AUG(T) | 0.614784 | aug |
| SEP(T) | 0.618274 | sep |
| OCT(T) | 0.210432 | oct |
| NOV(T) | -0.92685 | nov |
| DEC(T) | -0.93151 | dec |

**Figure 4 Monthly Correlation (Jan-Dec) dataset**

Figure 3 and 4 represents monthly and yearly correlations (Jan- Dec) and (2010-2018) dataset respectively for rise in air temperature.

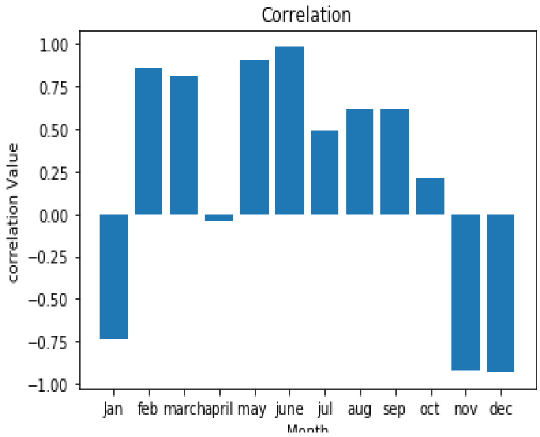


**Figure 5 Monthly Representations of Correlation Through Bar Graph (Jan-Dec)**

| MONTHLY AVERAGE PM 2.5 VALUES ITO center | | | | | | | 216 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | | | CORREL2014 |
| jan | 0 | 213 | 189 | 125 | 0 | 190 | 190 | 171 | 206 | | 2014 | 0.901343796 |
| feb | 0 | 160 | 118 | 153 | 0 | 153 | 153 | 103 | 117 | | 2015 | -0.578381069 |
| mar | 152 | 123 | 111 | 193 | 0 | 0 | 0 | 87 | 88 | | 2016 | |
| apr | 110 | 130 | 76 | 171 | 112 | 0 | 0 | 77 | 81 | | 2017 | |
| may | 114 | 93 | 85 | 0 | 150 | 0 | 0 | 105 | 76 | | 2018 | -0.706351427 |
| june | 124 | 69 | 0 | 0 | 147 | 0 | 0 | 116 | 159 | | | |
| jul | 59 | 72 | 0 | 0 | 148 | 0 | 0 | 67 | 53 | | | |
| aug | 54 | 52 | 0 | 109 | 88 | 0 | 0 | 65 | 49 | | | |
| sept | 65 | 53 | 0 | 134 | 107 | 0 | 0 | 87 | 38 | | | |
| oct | 187 | 153 | 70 | 171 | 44 | 0 | 0 | 25 | 129 | | | |
| nov | 290 | 23 | 115 | 216 | 0 | 0 | 200 | 220 | 202 | | | |
| dec | 265 | 266 | 113 | 0 | 0 | 0 | 228 | 188 | 229 | | 2011 | -0.681077261 |

MONTHLY AVERAGE Temprature VALUES

| | JAN | FEB | MAR | APR | MAY | JUN | JUL | AUG | SEP | OCT | NOV | DEC |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2010 | 17 | 20 | 26 | 29 | 30 | 29 | 27 | 27 | 26 | 25 | 22 | 17 |
| 2011 | 16 | 19 | 24 | 27 | 30 | 29 | 27 | 26 | 26 | 25 | 22 | 18 |
| 2012 | 16 | 19 | 24 | 28 | 30 | 30 | 27 | 26 | 26 | 24 | 21 | 18 |
| 2013 | 17 | 19 | 24 | 27 | 30 | 28 | 26 | 26 | 26 | 25 | 20 | 18 |
| 2014 | 17 | 19 | 23 | 27 | 29 | 30 | 27 | 27 | 26 | 25 | 21 | 17 |
| 2015 | 16 | 20 | 23 | 27 | 20 | 20 | 27 | 27 | 27 | 25 | 22 | 18 |
| 2016 | 15 | 18 | 25 | 26 | 33 | 33 | 30 | 30 | 30 | 26 | 20 | 16 |
| 2017 | 14 | 18 | 23 | 30 | 32 | 32 | 30 | 30 | 29 | 26 | 18 | 15 |
| 2018 | 13 | 18 | 24 | 30 | 32 | 34 | 31 | 30 | 28 | 28 | 19 | 15 |

**Figure 6 Monthly Average Values of PM and Monthly Average Temperature Values**

Figure 5 and 6 represents monthly recorded /average values of $PM_{2.5}$ at CPCB ITO, Delhi center and average temperature values.

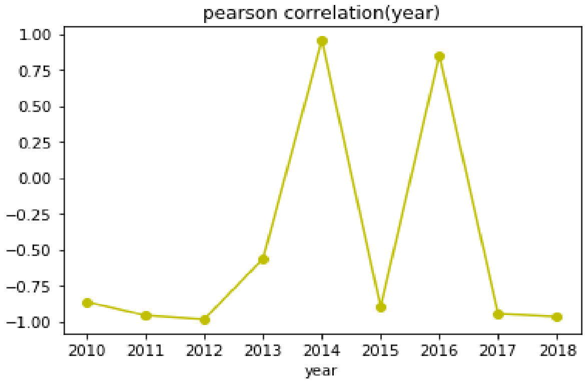


**Figure 7 Pearson's Yearly Correlation (2010-2018) dataset**

| | Y | X |
|---|---|---|
| JAN(T) | -0.6308 | Jan |
| FEB(T) | 0.127909 | feb |
| MAR(T) | 0.694828 | march |
| APR(T) | -0.43306 | april |
| MAY(T) | 0.724753 | may |
| JUN(T) | 0.834575 | june |
| JUL(T) | 0.179218 | jul |
| AUG(T) | 0.331619 | aug |
| SEP(T) | 0.446754 | sep |
| OCT(T) | 0.164644 | oct |
| NOV(T) | -0.50372 | nov |
| DEC(T) | -0.42384 | dec |

**Figure 8 Pearson's Monthly Correlation (2010-2018)**

Figure 7 and 8 represents Pearson's yearly and monthly correlation (2010-2018) and (Jan-Dec) dataset. Figure 9 solely depicts Pearson's Correlation Yearly Bar Graph.

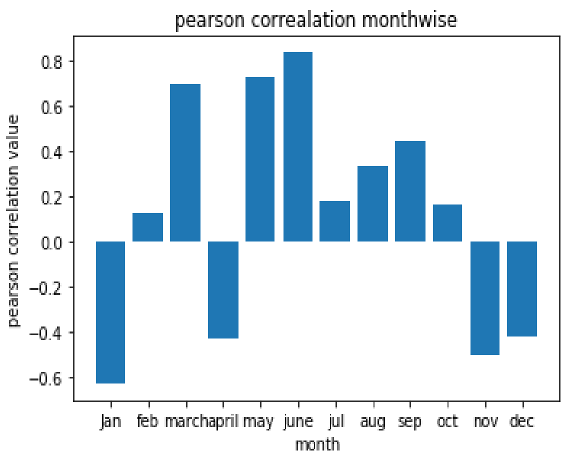


**Figure 9 Pearson's Correlation Bar Graph (2010-2018) dataset**

## IV. CONCLUSION

As it is clear from the figures and charts mentioned above, the most noticeable spike in 2016 begun, at which point in 2015, the AQI was starting to balance out. While there was a spike on the seventeenth which remained till the 21st, it was effectively owing to the unexpected 2° C temperature drop from the 22° C on the sixteenth to 20° C on the seventeenth and past. According to the research outcomes derived from this research and experiments there is little, or no relationship exists between $PM_{2.5}$ and temperature in the yearly comparison. Unlike monthly comparisons that exhibits their partial dependency. From last many years Winter season in Delhi/NCR region has been started with a smog atmosphere in which whole area seems to be a Gas Chamber. Initially it gives the impression to have a connection in change in temperature with air pollution components. This research work is useful to carry further research work on the reasons of this Smog in Delhi/NCR region in a starting winter temperature and initiate smog analysis.

## FUTURE WORK

$PM_{2.5}$ fine particulate issue noticeable all around that are two and one-half microns or less in width will be at the 500+ level and $PM_{10}$ poisons will remain at 446 [19, 20, 21]. Both will break the "serious" mark on the record, provoking metropolitan enterprise experts in the city to arrange streets sprinkled with water. In this research, the measurable techniques that is, Correlation Function have been utilized to discover a connection between the air contaminations and the temperature esteem. Additionally, Pearson's Correlation Method have been utilized to determine the relationship among them. For further research, Sophisticated software techniques could be utilized other than measurable strategies on the bigger dataset of air pollution components from Delhi/NCR regions for performing information investigation and to make the procedure streamlined.

## REFERENCES

1. Chao, Z., *''urban climate and air pollution in Shanghai''*, Energy and Buildings, 16 (1-2): pp. 647-656, (1991).
2. Sánchez-Ccoyllo, O. R., and M. de Fatima Andrade. "The influence of meteorological conditions on the behavior of pollutants concentrations in São Paulo, Brazil." *Environmental Pollution* 116, no. 2 (2002): 257-263.
3. Cuhadaroglu, B. and Demirci, E., ''Influence *of some meteorological factors on air pollution in Trabzon city''*, Energy and Buildings, 25: pp. 179-184, (1997).
4. Sahu, S.K., Beig, G. and Parkhi, N.S. (2011). PM2.5 and PM10 in Delhi during Commonwealth Games 2010. *Atmos. Environ.* 45: 6180–6190.
5. Chen, Ziyue, Xiaoming Xie, Jun Cai, Danlu Chen, Bingbo Gao, Bin He, Nianliang Cheng, and Bing Xu. "Understanding meteorological influences on PM 2.5 concentrations across China: a temporal and spatial perspective." *Atmospheric Chemistry and Physics* 18, no. 8 (2018): 5343-5358.
6. Verma, Kritika, Sandeep Mathur, and S. K. Khatri. "Study on Temperature Variation Pattern Based on Data Analytics." In *2019 Amity International Conference on Artificial Intelligence (AICAI)*, pp. 106-109. IEEE, 2019.
7. Martins, Leila Droprinchinski, Caroline Fernanda Hei Wikuats, Mauricio Nonato Capucim, Daniela S. de Almeida, Silvano Cesar da Costa, Taciana Albuquerque, Vanessa Silveira Barreto Carvalho, Edmilson Dias de Freitas, Maria de Fátima Andrade, and Jorge Alberto Martins. "Extreme value analysis of air pollution data and their comparison between two large urban regions of South America." *Weather and climate extremes* 18 (2017): 44-54.
8. Whiteman, C. David, Sebastian W. Hoch, John D. Horel, and Allison Charland. "Relationship between particulate air pollution and meteorological variables in Utah's Salt Lake Valley." *Atmospheric Environment* 94 (2014): 742-753.

9. Beig, Gufran, Dilip M. Chate, Sachin D. Ghude, A. S. Mahajan, R. Srinivas, K. Ali, S. K. Sahu, N. Parkhi, D. Surendran, and H. R. Trimbake. "Quantifying the effect of air quality control measures during the 2010 Commonwealth Games at Delhi, India." *Atmospheric environment* 80 (2013): 455-463.
10. Marrapu, P., Y. Cheng, Gufran Beig, S. Sahu, R. Srinivas, and G. R. Carmichael. "Air quality in Delhi during the Commonwealth Games." *Atmospheric chemistry and physics* 14, no. 19 (2014): 10619-10630.
11. Nagpure, Ajay S., Ketki Sharma, and Bhola R. Gurjar. "Traffic induced emission estimates and trends (2000–2005) in megacity Delhi." *Urban Climate* 4 (2013): 61-73.
12. Beryland, M.Y., 1975, Contemporary problems of atmospheric diffusion and pollution of the atmosphere. Gidrometezdat, Leningrad, translated into English by NERC, US EPA.
13. Boubel et al. (1994) Fundamentals of Air Pollution, 3rd edition. Academic Press. Briggs, G.A., 1965, A plume rise model compared with observations J.Air Poll. Control Association 15:433.
14. Bosanquet, C.H. (1936) The Spread of Smoke and Gas from Chimmneys. Trans. Faraday Soc. 32:1249.
15. Carmichael, G.R., and Peters, L.K., 1979, Numerical simulation of the regional transport of SO2 and sulfate in the eastern United States, Proc. 4 th Symp. on turbulence, diffusion and air pollution, AMS 337.
16. Huang, G., Lee, D., & Scott, E. M. (2018). Multivariate space-time modelling of multiple air pollutants and their health effects accounting for exposure uncertainty. Statistics in medicine, 37(7), 1134-1148.Deardorff, J.W., and Willis, G.E., 1975, A parameterization of diffusion into the mixed layer J. Appl. Met 14:1451.
17. Deardorff, J. W. (1970). Convective velocity and temperature scales for the unstable planetary boundary layer and for Rayleigh convection. Journal of the atmospheric sciences, 27(8), 1211-1213.
18. J. Atm. Sci. 27, 1211-1213. Dyun, D.W. and J.K.S. Ching, 1999, Science algorithm of the EPA Models-3 Community Multiscale Air Quality (CMAQ) Modeling System., EPA-Dep./600/R-99/030.
19. Egan, B. A., Rao, K. S., & Bass, A. (1976). A three-dimensional advective-diffusive model for long range sulfate transport and transformation. In NATO Comm. on the Challenges of Mod. Soc. Proc. of the 7th Intern. Tech. Meeting on Air Pollution Modeling and its Appl. p 697-714(SEE N 77-27569 18-45).
20. Eliassen, A., and Saltbones, J., 1975, Decay and transformation rates of SO2 as estimated from emission data, trajectories and measured air concentrations Atm. Env. 9:425.
21. Eschenroeder, A.Q. and J.R. Martinez, 1970, "Mathematical Modeling of Photochemical Smog", American Institute Aeronautics and Astronautics, Eight Aerospace Sciences Meeting, New York, Jan 19-21.


## AUTHORS PROFILE

**Dr. Sandeep Mathur,**
PhD CSE Data Analytics, Data Modelling
MITACSIT, SPOC, MIAENG, MIACSIT, MCSTA, MISOC MSCIEI

**Satvik Sharma,** Amity Institute of Information Technology Amity University, Uttar Pradesh
satviks84@gmail.com