

Temporal and Correlational Analysis of Air Pollution and Covid-19 Across Major Metropolis in India



Anurag K S V, Bhaskara Rao B

Abstract: This study aimed to check the temporal variations in Air Pollution levels using Air Quality Index in major cities spanning geographically across India and understand the correlation between the severity of Covid-19 and the concentrations of the six key air pollutants. The study considered a tenure between May-June 2020 when there was a drastic shift in human behavior due to Unlock period in India. It employed Spearman's Rank Correlation Coefficient, which better understands monotonic function association between two variables for correlation. The study successfully established statistical evidence of improvement in the air quality of 29% (average) in May and June of 2020 compared to 2019 in the cities under study. A significant correlation was also established between air pollution and covid-19 across various cities in India, with Delhi and Mumbai exhibiting a weak negative correlation. At the same time, Hyderabad and Bengaluru showed a moderate negative correlation. On the other hand, the city of Chennai implied a strong positive correlation between air pollution and Covid-19.

Keywords: Correlational Analysis, Air Pollution, Covid-19, Temporal Variations.

I. INTRODUCTION

To contain the Covid-19 spread, India implemented a phase-wise lockdown in the country from 25 March 2020 to 31 May 2020. This lockdown was conducted in four phases with incremental liberation of essential facilities and anthropogenic movement. Starting in June 2020, a phase-wise unlock action plan was implemented until the second wave hit India in February 2021[1]. Studies conducted in the Northern parts of India suggested that the lack of anthropogenic movement decreased air pollution levels all over the country's northern regions[2]. The strong effects of air pollution on humans have a long-standing impact that is often ignored but has a staggering impact over time, which is brought to light due to the spread of the disease extensively in areas typically with worse pollution levels than others[3]. A stage for significant debate recently was how the lockdown impacted the air pollution levels of the country overall? This study attempted to undertake a comparative

analysis of the influencing factors over a span of two months, starting from May 2020 to June 2020. This period included a period of complete lockdown and a brief unlock period of public movement restrictions. This study also covered the correlation between air pollution and the number of Covid-19 cases across the major metropolis in India during these periods. The geographical diversity of these metropolises and the population density in these cities enables us to get a holistic understanding of the correlation.

II. LITERATURE REVIEW

An extensive amount of research was conducted in Computational and Environmental Science to understand the spread of Covid-19 and its relation to various environmental factors such as temperature and pollution. Studies in the U.S used land-based sensor observation data to determine a 25.5% decrease in NO₂ levels compared to pre-covid pollution levels[4]. At the same time, Indian studies across major cities show the reasons for the decline in pollution levels: a lack of anthropogenic movement in India, where we could find a dip up to 62% in pollution levels[5]. Satellite observation over Europe via Sentinel-5P satellite [United Space in Europe (2020)] which could track real-time pollution levels, suggested a decline in air pollution levels during the lockdown period[6]. Likewise, Space-borne observations were taken into consideration extensively, where Moderate Resolution Imaging Spectrometer (MODIS) was used to estimate the global distribution of aerosols[7]. Space-borne observations using satellite imagery and MEERA-2 to monitor air pollution parameters over 20 countries were done[8], which showed an overall drop in PM_{2.5} in India during April 2020. Research in China confirmed a mild correlation between primary air pollutants with inner-city activities in Wuhan and Beijing in China[9].

III. OBJECTIVE

Studies so far considered the first phase of the lockdown period of Covid-19 (25th March 2020 to 14th April 2020) to determine the decrease in air pollution levels across various states in India. This study is an attempt to answer some of the questions, which include: what happens when you consider the subsequent phases of the Covid-19 lockdown when restricted movement was permitted? How did the first unlock phase, which eased significant restrictions imposed by the lockdown, impact air pollution, and Covid-19? Hence, this study aimed to understand and find the correlation between the severity of Covid-19

Manuscript received on 31 August 2022 | Revised Manuscript received on 04 September 2022 | Manuscript Accepted on 15 October 2022 | Manuscript published on 30 October 2022.

*Correspondence Author (s)

Anurag K S V, School of Technology, Gandhi Institute of Technology and Management, Visakhapatnam (A.P.), India. Email: anurag.krovvidi@gmail.com

Dr. Bhaskara Rao B*, Associate Professor, School of Technology, Gandhi Institute of Technology and Management, Visakhapatnam (A.P.), India. Email: bboddu@gitam.edu

©The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC-BY-NC-ND license: <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Temporal and Correlational Analysis of Air Pollution and Covid-19 Across Major Metropolis in India

(via the number of cases increased per day) and the concentrations of the six classical air pollutants in six major metropolises Bengaluru, Chennai, Delhi, Hyderabad,

Kolkata, and Mumbai; spanning geographically across India. In extension, checked the temporal variations in Air Pollution levels using Air Quality Index in these cities.

IV. DATA

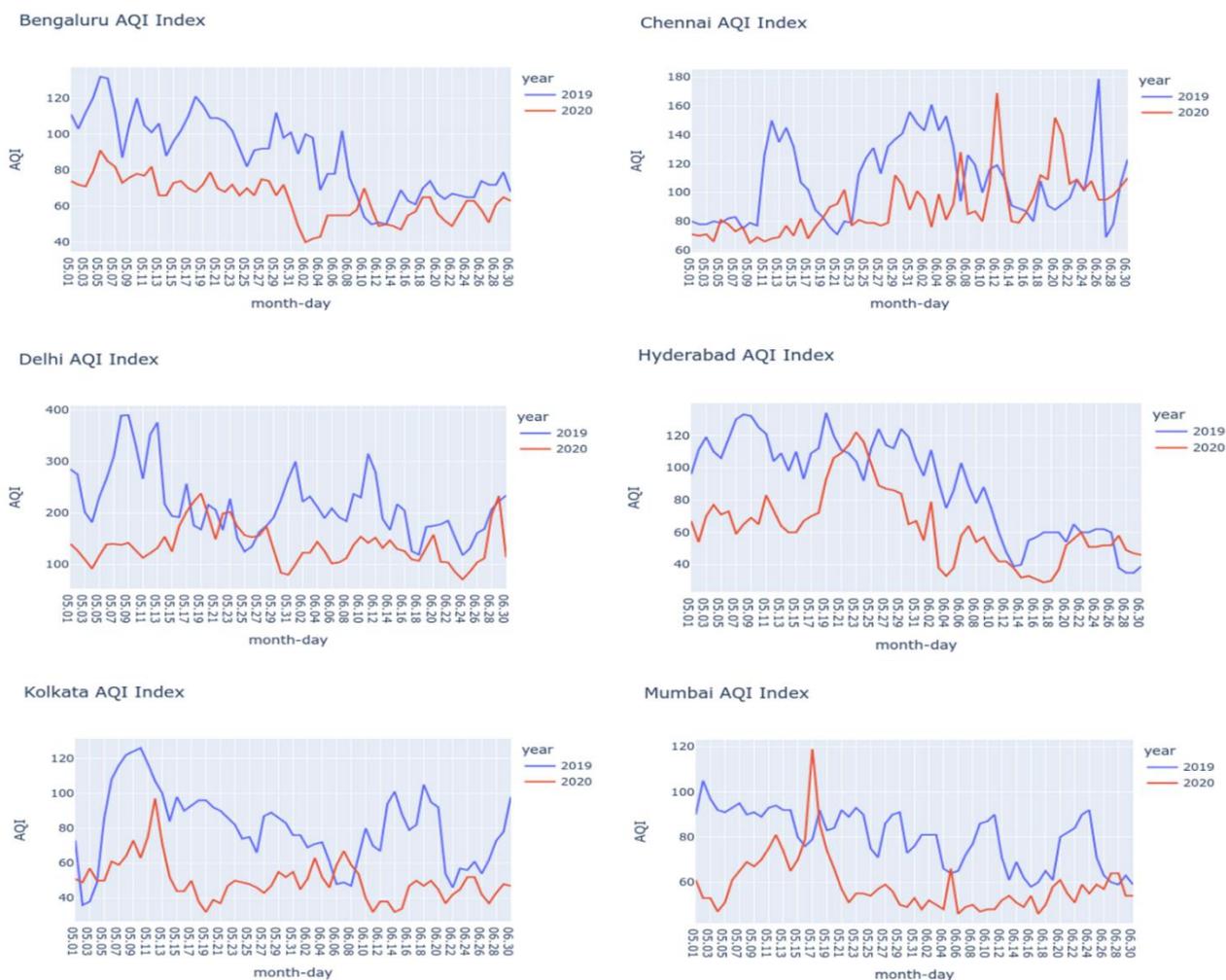


Fig. 1. Time Series Analysis of AQI across the major metropolis in India.

Table-I: Mean difference between AQI in 2019 and 2020

City	Mean AQI 2019	Mean AQI 2020	Mean AQI % Difference
Bengaluru	88.25 (Satisfactory)	64.31 (Satisfactory)	27.12%
Chennai	108.07 (Moderately Polluted)	90.67 (Satisfactory)	16.10%
Delhi	217.11 (Poor)	137.38 (Moderately Polluted)	36.73%
Hyderabad	89.44 (Satisfactory)	63.48 (Satisfactory)	29.03%
Kolkata	80.23 (Satisfactory)	49.87 (Good)	37.84%
Mumbai	80.07 (Satisfactory)	58.70 (Satisfactory)	26.68%

The secondary data was obtained from the air pollution data and Covid 19 of India published for the months of May and June 2019 from the following sources: A publicly available dataset[10] sourced by the Central Pollution Control Board (CPCB)[11] for Air Pollution, an open-source API[12] for Covid-19. The data for the months of May and June during 2019 and 2020 was considered for this study since:

1. Air Pollution data was used as a standard for comparison between 2019 and 2020.
2. The data from May 2020 and June 2020 accounts for the various changes in human behavior due to restrictions that better depict the real-life case of human activity.

The data points where data was not available were assumed to be of Zero Value to enable data continuity. The consolidated data of Telangana state was used as the data for Covid-19 cases in the city of Hyderabad. This particular methodology to handle missing data was viable for the study as the statistical methods used were resilient in handling them.

V. METHODOLOGY

A. Air Pollution Analysis

Temporal analysis of the data was conducted for various cities by exploring the Air Quality Index (AQI) of these cities over a time-series line plot against each day over May and June of 2019 and 2020, as shown in “Fig.1”. AQI gave us a consolidated picture of the Air Quality over that area by averaging the air pollutants' concentrations and their respective permissible standards [13]. CPCB summarized that the lower the AQI better the air quality in the area of

study [13]. Therefore, observations made on the trend of AQI across these cities suggested a consistent dip in the AQI level inferring better air quality during 2020. A visual representation of the same is shown in “Fig.1”.

Further statistical analysis quantified the significant drop between AQI averages in 2019 and 2020, demonstrated in “Table. I”. Observations suggested a change in pollution tiers in a few cities, such as Chennai from Moderately Polluted to Satisfactory, Delhi from Poor to Moderately Polluted, and Kolkata from Satisfactory to Good. The other three cities, Bengaluru, Hyderabad, and Mumbai, showed a significant increase in Air Quality of 27.12%, 29.03%, and 26.68%, respectively, on a comparative basis between 2019 and 2020.

This initial analysis paved the way to explore more about the variations in major air pollutants (PM2.5, PM10, O3, NO2, SO2, and CO) and their association with Covid-19 across these cities, as shown in “Table. II”.

Table-II: Average of various air pollutants across all the cities under observation during May and June of 2020

City	Air Pollution													
	PM2.5		PM10		O3		SO2		NO2		CO		AQI	
	May	June	May	June	May	June	May	June	May	June	May	June	May	June
Bengaluru	21.3	18	59.8	41.1	32.4	16.2	6.48	7.1	15.4	13.7	0.78	0.63	73.2	55.2
Chennai	18.9	29.5	22.8	52.2	46.7	58.9	5.1	6.24	7.93	11.8	0.86	1.08	78.7	103
Delhi	55.5	46.7	138	121	56.8	39.6	15.4	12.4	24.9	23.7	0.91	0.92	149	126
Hyderabad	30.1	18.6	80.5	45.1	31.2	18.5	6.14	6.52	26.2	21.7	0.46	0.31	79.4	47.1
Kolkata	16.1	13.6	37.9	34.8	33.6	26	5.64	6.51	7.38	10.4	0.33	0.36	53.2	46.5
Mumbai	14.7	11.2	53.7	35.2	18.6	12.3	16.2	10.3	7.78	11.7	0.29	0.37	63.9	53.3
Mean	26.1	22.9	65.4	54.9	36.5	28.6	9.16	8.17	14.9	15.5	0.61	0.61	82.8	71.8

B. Covid-19 Analysis

The first two phases of lockdown implemented in India proved to be effective in controlling the spread of Covid-19 in the country. However, during the third and fourth phases of lockdown in May 2020, majority of the strict norms were eased, leading to extensive human movement. The month of June marked the first unlock period for the country after 68

days of phase-wise lockdown [1]. This period showed the peak anthropogenic movement in the country after a long time. Due to this, a steady spike of Covid-19 cases is observed in the rush cities of India, as shown in “Fig.2”. The average increase in the number of cases recovered and deaths registered is also observed to increase between May and June, as shown in “Table. III”.

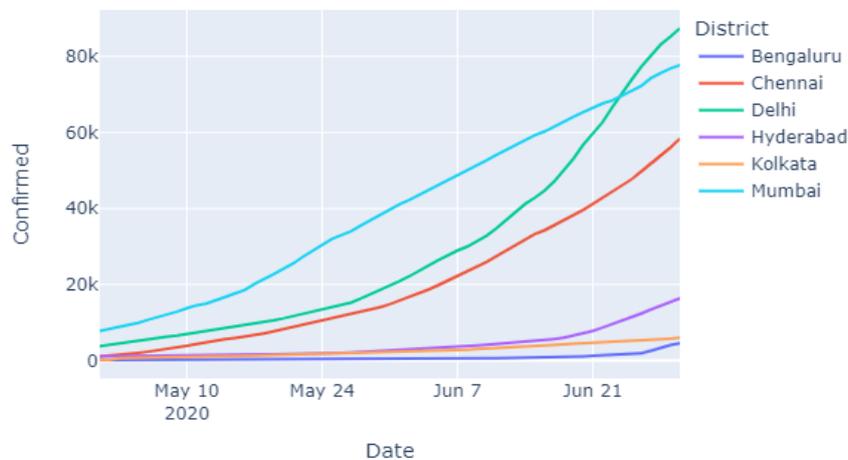


Fig. 2. Covid-19 time-series for May-June 2020.

Table-III: Average number of Covid-19 cases across all cities under observation during May and June of 2020

City	Covid-19					
	Confirmed		Recovered		Deceased	
	May	June	May	June	May	June
Bengaluru	7	140	5	10	0	3
Chennai	448	1451	248	898	4	25
Delhi	527	2251	238	1662	13	76
Hyderabad	54	455	32	196	2	6
Kolkata	63	129	29	95	7	6
Mumbai	1052	1266	506	913	32	109
Mean	358	948	176	629	9.61	37.4

Delhi, Mumbai, and Chennai were prone to more cases with a faster spread of Covid-19 than Bengaluru, Kolkata, and Hyderabad. The day-wise rise in Covid-19 cases across these cities was observed, giving us a better measure of day-to-day change in the severity of Covid-19, as shown in “Fig.3”. This metric was used as a variable for calculating the Correlation between Covid-19 and Air-Pollution metrics across these cities.

The day-wise trend for these cities was different compared to the timeseries as Delhi recorded the highest number of cases per day with a steep spike and drop, Chennai showed an approximately linear increase in cases, and Mumbai depicted a constant rise in the number of cases per day despite its high cumulative numbers. Hyderabad showed a spike in the number of cases in mid of June 2020, whereas Bengaluru cases spiked towards the end of June 2020. On the other hand, Kolkata showed a tightly controlled tread as its cases never reached the 1000 mark.



Fig. 3. Covid-19 day-wise confirmed cases.

C. Correlation Analysis

In order to further understand the inter-relationship between the two major factors identified, the following hypothesis was postulated:

H₀=There is no correlation between Air Pollution and Covid-19 for a given population set.

H₁ = There is a significant correlation between Air Pollution and Covid-19 for a given population set.

These hypotheses were significant to the study as we proved their statistical significance by calculating the p-value (probability) obtained from the data. If the p-value is less than the confidence threshold ($\alpha = 0.05$), we rejected the null hypothesis (H₀) and considered the alternate hypothesis (H₁) to be true.

Since the dataset doesn’t satisfy the normal distribution and the Q-Q plots reveal monotonicity of data, Spearman’s Rank Correlation measure was used to find the correlation between the classical air pollutants and the day-wise Covid-19 cases.

$$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \tag{1}$$

Where,

ρ = Spearman’s Rank Correlation Coefficient.

d_i = Difference between the rank of pollution value and the number of corresponding cases.

n = The total number of days considered.

Equation (1) shows Spearman’s Rank Correlation Coefficient, ‘ ρ -value’ varies between (-1,1) where -1 depicts a very strong negative correlation, +1 indicates a very strong positive correlation, and 0 shows no correlation between the bivariate under study.

Various studies classified the value of correlation based on their absolute magnitude, one of which showed that the coefficients of magnitude between (0.00 - 0.10) exhibit negligible correlation, (0.10 - 0.39) exhibit weak correlation, (0.4 - 0.69) exhibit moderate correlation, (0.70 - 0.89) exhibit strong correlation, and (0.90 - 1.00) exhibit very strong correlation between the variables [14].

VI. RESULTS AND DISCUSSION

The key air pollutants that are considered for the calculation of Air Quality Index (AQI) include Particulate Matter 2.5 (PM2.5), which refers to particles that are 2.5 microns or less in diameter, Particulate Matter 10 (PM10), which refers to particles that are ten microns or less in diameter, ground-level Ozone (O3), Sulphur-Dioxide (SO2), Nitrogen-Dioxide (NO2), and Carbon-Monoxide (CO) [15].

The averages of the above-mentioned air pollutants for May and June of 2020 are consolidated for various cities across India in “Table. III”, which indicates an overall decrease in pollutant concentration from May 2020 to June 2020 and an increase in Air Quality depicted by AQI.

Likewise, a consolidated table for averages of the number of Covid-19 cases confirmed, recovered, and deceased was recorded for May and June of 2020 for these cities detailed in “Table. III”, from which we may infer that June 2020 had a significant increase in the number of Covid-19 cases across all the cities under study.



Based on these observations, the Spearman's Rank Correlation Coefficient 'ρ-value' was calculated towards the inter-relationship between air pollutants and Covid-19; the following observations were made:

PM2.5 and NO2 showed a weak negative correlation, whereas SO2 exhibited a moderate positive correlation with Covid-19 in the city of Bengaluru. PM10 and CO demonstrated a moderate negative correlation; reasons for this may be attributed to restrictive human movement as these pollutants' primary source is vehicular emissions. O3 showed a strong negative correlation with Covid-19 of -0.80, the highest among the observed cities under study.

All the classical air pollutants were positively correlated with Covid-19 in the city of Chennai, where particulate matter showed a moderate positive correlation with Covid-19. The city's other non-particulate/gaseous pollutants show a moderate to high correlation with Covid-19. The classical gaseous air pollutants have their primary sources from vehicular combustion and industrial combustion, which show a significant positive correlation with the rise in Covid-19 in the city, with SO2 at 0.42, NO2 at 0.64, CO at 0.67, and O3 at 0.74. The above-mentioned reasons may be further explored in-depth to understand the root causes behind the strong positive correlation between the two factors in this city. For the city of Delhi, the pollutants PM10, NO2, and CO were at 0.17, 0.70, and 0.35, respectively. In contrast, PM2.5 exhibited a weak correlation. Delhi's atmospheric condition is different from other cities as it suffers from an inversion layer trapping all pollutants to ground level [16]. Due to this, the city suffers from abnormally high pollution levels irrespective of the pollutant mentioned above level changes over a period of time. However, our study successfully established a moderate negative correlation between SO2 and Covid-19 at -0.58. A strong negative correlation between ground-level O3 levels and Covid-19 was found as well at a Spearman's Rank Correlation Coefficient of -0.79. Hyderabad, on the other hand, showed a weak negative correlation between NO2 and Covid-19 at -0.35. Whereas particulate matter (PM2.5, PM10), O3, and CO showed a moderate negative correlation with the spread of Covid-19, a mildly controlled increase comparatively hinted by the daily increase in several cases, as in "Fig.3".

PM2.5 and O3 showed a weak negative correlation with Covid-19 in the city of Kolkata. CO and NO2, both byproducts of vehicular pollution [17], [18], showed a weak and moderate positive correlation at 0.36 and 0.51, respectively. It may be speculated that this positive correlation was a result of increased anthropogenic movement in the city, which requires further study for conclusive evidence.

Mumbai showed a moderate positive correlation with NO2, hinting that vehicular pollution might put the public at a higher risk of contracting the disease and aiding its spread [17]. SO2 exhibited a moderate negative correlation of -0.40. One might infer that the lockdown period curtailed industrialization involving thermal and combustion of fossil fuels as sources for SO2[19], which led to a decrease in its emissions. According to the study, plateau cities located in the southern part of the country, Bengaluru and Hyderabad, showed similar patterns with an overall moderate negative correlation with all the classical air pollutants except for SO2.

Chennai, Mumbai, and Kolkata, which are under the active influence of water bodies, exhibited a positive correlation w.r.to NO2. Delhi, on the other hand, located in Northern India, due to its peculiar atmospheric conditions, showed a significant negative correlation with O3 and SO2. Chennai was the only city with a consistent positive correlation between all observed air pollutants and Covid-19.

VII. CONCLUSION

This study was critical in examining the correlation between two key factors of air pollution and Covid-19 across time periods of 2019 and 2020 in the major metropolis of India. Evidence was established that in May and June of 2020, during lockdown 3.0, lockdown 4.0, and Unlock 1.0 phases. There was a 29% improvement in Air Quality on average, across all the cities. A more detailed overview is listed in "Table. I".

To understand the overall influence of air pollution on Covid-19, Spearman's Rank Correlation Coefficient was calculated between the AQI of these cities to the number of Covid-19 cases confirmed. The null hypothesis (H0) was accepted for Kolkata as the p-value of 0.10 is higher than the confidence threshold; the other cities rejected the null hypothesis (H0) in favor of the alternate hypothesis (H1). Delhi and Mumbai exhibited a weak negative correlation of -0.28 and -0.31, respectively. Hyderabad and Bengaluru showed a moderate negative correlation of -0.63 and -0.65, respectively. The only positive correlation was observed in the city of Chennai with 0.77, which implied a strong correlation between air pollution and Covid-19 in that city. A detailed heatmap of Spearman's Rank Correlation Coefficient between various cities and key air pollutants, AQI, is shown in "Fig.4".

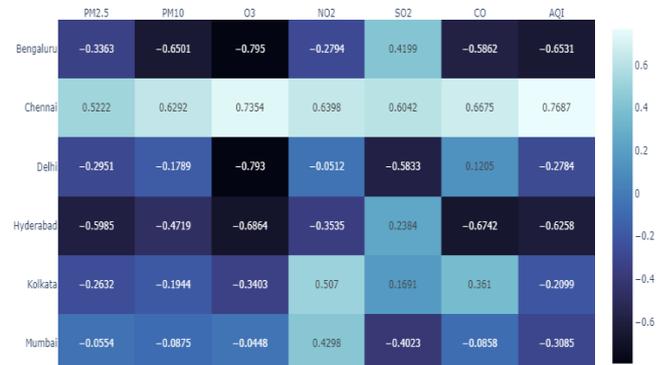


Fig. 4. Spearman's Correlation Coefficient between classical Air Pollutants and Covid-19 cases across the major metropolis in India.

The results and analysis of the data in this study indicated the requirement of further evaluation of the interrelationship between the two variables considered, hinting at expanding the scope for understanding the possibility of other factors which could impact the variations in AQI, Covid-19 and its impact on activities of mankind on the Earth, and its implications on the air quality, which can help us understand the possible ways of protecting the air quality during the post-Covid times.



REFERENCES

1. COVID-19 lockdown in India [Internet]. Wikipedia. 2021. Available from: https://en.wikipedia.org/wiki/COVID-19_lockdown_in_India
2. Gautam AS, Dilwaliya NK, Srivastava A, Kumar S, Bauddh K, Singh D, et al. Temporary reduction in air pollution due to anthropogenic activity switch-off during COVID-19 lockdown in northern parts of India. Environment, Development and Sustainability. 2020 Sep 24; Available from: <https://doi.org/10.1007/s10668-020-00994-6> [CrossRef]
3. Vadrevu KP, Eaturu A, Biswas S, Lasko K, Sahu S, Garg JK, et al. Spatial and temporal variations of air pollution over 41 cities of India during the COVID-19 lockdown period. Scientific Reports. 2020 Oct 6;10(1). Available from: <https://doi.org/10.1038/s41598-020-72271-5> [CrossRef]
4. Berman JD, Ebisu K. Changes in U.S. air pollution during the COVID-19 pandemic. Science of The Total Environment [Internet]. 2020 Oct 15 [cited 2020 Jun 8]; 739:139864. Available from: <https://www.sciencedirect.com/science/article/pii/S0048969720333842#0005> [CrossRef]
5. Ravindra K, Singh T, Biswal A, Singh V, Mor S. Impact of COVID-19 lockdown on ambient air quality in megacities of India and implication for air pollution control strategies. Environmental Science and Pollution Research. 2021 Jan 7; Available from: <https://doi.org/10.1007/s11356-020-11808-7> [CrossRef]
6. Messaoud TA, Smiti A. The COVID-19 Pandemic: What About Air Pollution? 2020 4th International Conference on Advanced Systems and Emergent Technologies (IC_ASET). 2020 Dec 15; Available from: https://doi.org/10.1109/ic_aset49463.2020.9318299 [CrossRef]
7. Ramanathan V. Aerosols, Climate, and the Hydrological Cycle. Science. 2001 Dec 7;294(5549):2119–24. Available from: <https://doi.org/10.1126/science.1064034> [CrossRef]
8. R. D. Das, S. Bandopadhyay, M. Das and M. Chowdhury. Global Air Quality Change Detection During Covid-19 Pandemic Using Space-Borne Remote Sensing and Global Atmospheric Reanalysis. 2020 IEEE India Geoscience and Remote Sensing Symposium (InGARSS), 2020, pp. 158-161. Available from: <https://doi.org/10.1109/InGARSS48198.2020.9358918> [CrossRef]
9. Zhan C, Jiang W, Li J, Xu H, Sha W. Impact of COVID-19 Lockdown on Human Activity and Air Quality in China. 2020 IEEE International Symposium on Product Compliance Engineering-Asia (ISPCE-CN). 2020 Nov 6; Available from: <https://doi.org/10.1109/ISPCE-CN51288.2020.9321817> [CrossRef]
10. Rao R. Air Quality Data in India (2015 - 2020) [Internet]. kaggle.com. Available from: <https://www.kaggle.com/rohanrao/air-quality-data-in-india>
11. CPCB/Central Pollution Control Board [Internet]. Cpcb.nic.in. 2019. Available from: <https://cpcb.nic.in/>
12. COVID19-India API [Internet]. Available from: <https://data.incovid19.org/>
13. National Air Quality Index [Internet]. https://app.cpcbccr.com/ccr_docs/FINAL-REPORT_AQI.pdf. p. 35, 36. Available from: https://app.cpcbccr.com/ccr_docs/FINAL-REPORT_AQI.pdf
14. Schober P, Boer C, Schwarte LA. Correlation Coefficients: Appropriate use and interpretation. Anesthesia & Analgesia [Internet]. 2018 May;126(5):1763–8. Available from: https://journals.lww.com/anesthesia-analgesia/fulltext/2018/05000/correlation_coefficients_appropriate_use_and.50.aspx [CrossRef]
15. WHO. New WHO Global Air Quality Guidelines aim to save millions of lives from air pollution [Internet]. www.who.int. 2021. Available from: <https://www.who.int/news/item/22-09-2021-new-who-global-air-quality-guidelines-aim-to-save-millions-of-lives-from-air-pollution>
16. November 02, Jaiswal A. Return of the Delhi Smog [Internet]. NRDC. 2018 [cited 2022 Jan 9]. Available from: <https://www.nrdc.org/experts/anjali-jaiswal/return-delhi-smog>
17. Chakraborty P, Jayachandran S, Padalkar P, Sitlhou L, Chakraborty S, Kar R, et al. Exposure to Nitrogen Dioxide (NO₂) from Vehicular Emission Could Increase the COVID-19 Pandemic Fatality in India: A Perspective. Bulletin of Environmental Contamination and Toxicology. 2020 Jul 15;105(2):198–204. Available from: <https://doi.org/10.1007/s00128-020-02937-3> [CrossRef]
18. Senaul H, Singh R. B. Air Pollution and Human Health in Kolkata, India: A Case Study. Climate. 2017 Oct 12;5(4):77. Available from: <https://doi.org/10.3390/cli5040077> [CrossRef]
19. Tricker R, Tricker S. Pollutants and contaminants. Environmental Requirements for Electromechanical and Electronic Equipment. 1999;158–94. Available from: <https://doi.org/10.1016/b978-075063902-6.50010-3> [CrossRef]

AUTHORS PROFILE



Anurag K S V, completed his Bachelor of Technology in Computer Science Engineering from Gandhi Institute of Technology and Management, Visakhapatnam (Andhra Pradesh), India. He worked on temporal and bivariate correlational data analysis using python along with surveying the field of Quantum Computing using Qiskit. He worked as a part-time Business Data Analyst for Hexawel Healthcare, looking after the firm's data. His research interests include Computational Science and Quantum Computing. He is currently enrolled in the Defence Institute of Advanced Technology, Pune (Maharashtra), India, pursuing his Master of Technology in Quantum Computing with a full scholarship from AICTE.



Dr. Bhaskara Rao B, completed his Master of Technology in Computer Science Engineering and Doctorate of Philosophy (Ph.D.) from Andhra University, Visakhapatnam (Andhra Pradesh), India, focusing on personalized search engines based on the semantic web for analysis. He is currently working as an Associate Professor at Gandhi Institute of Technology and Management, Visakhapatnam (Andhra Pradesh), India, with 18+ years of research experience in various fields of study such as semantic web, information retrieval, data mining, and teaching experience in basics of programming, including Object Oriented Programming using C++ and Java along with Data Structures.