Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 218 (2023) 969–978

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Machine Learning and Data Engineering

Impact on Air Quality Index of India Due to Lockdown

Aditya Dubey[a,*], Akhtar Rasool[b]

[a,b]*Department of Computer Science and Engineering, Maulana Azad National Institute of Technology, Bhopal and 462003, India*

Abstract

For the very first time, on 22-March-2020 the Indian government forced the only known method at that time to prevent the outburst of the COVID-19 pandemic which was restricting the social movements, and this led to imposing lockdown for a few days which was further extended for a few months. As the impact of lockdown, the major causes of air pollution were ceased which resulted in cleaner blue skies and hence improving the air quality standards. This paper presents an analysis of air quality particulate matter (PM)2.5, PM10, Nitrogen Dioxide (NO2), and Air quality index (AQI). The analysis indicates that the PM10 AQI value drops impulsively from (40–45%), compared before the lockdown period, followed by NO2 (27–35%), Sulphur Dioxide (SO2) (2–10%), PM2.5 (35-40%), but the Ozone (O3) rises (12–25%). To regulate air quality, many steps were taken at national and regional levels, but no effective outcome was received yet. Such short-duration lockdowns are against economic growth but led to some curative effects on AQI. So, this paper concludes that even a short period lockdown can result in significant improvement in Air quality.

Keywords: Stochastic Forecasting Models; Supervised ML Models; Soft Computing-based Models; Deep Learning Models.

1. Introduction

COVID-19 is deemed one of the greatest disasters that had a detrimental effect on the lives of people worldwide. The infection rapidly spread all through the globe. Numerous nations responded too late to execute preventive measures driving a sudden upsurge in cases around the world. The novel coronavirus plague first broke out in Wuhan in December 2019 and since then has been spreading rapidly. All nations around the globe started being affected by COVID-19 gradually [1, 2]. In Dec-2019, WHO was informed by China about various cases of unfamiliar lung diseases, dominated by pneumonia in Wuhan, a city in Hubei Province [3]. On 7-January-2020 the unfamiliar virus was recognized as SARS-CoV-2. In January, WHO declared public health trauma when it spread widely from one

* Corresponding author. Tel.: +91-750-985-2725.
E-mail address: dubeyaditya65@gmail.com

10.1016/j.procs.2023.01.077

nation to the other. On March 24, the number of COVID-19 cases reached 500 in India, and the complete lockdown was announced in the republic [4]. The essential strategy of abating the contamination rate has been to force strict social distancing controls, where individuals are limited to their homes and all the economic activities were ceased [5].

In many developing countries, specifically India, the increasing figures of motorized automobiles and mechanized activities are believed to be the chief sources of air pollution. The major pollutants for deteriorating air quality comprise PM10, PM2.5, NO2, SO2, carbon monoxide, ammonia, ozone, volatile organic compounds, etc., among which PM2.5 is the most dominant. Few of the major sources of PM 2.5 emissions are vehicular traffic, industries, and construction works [6, 7, 8]. Apart from restrictions, the lockdown had a positive impact, which led to a reduction in the emission of many harmful pollutants, a major cause of a considerable drop in AQI. During the 4-month lockdown period the annual concentrations of NO2, O3, PM10, and PM2.5 dropped significantly [9]. For this exploration day-to-day air pollution data of some of the major cities of India [10], are taken into consideration which is not dominated by the mining belt. The decrement of NO2, O3 PM10, and PM2.5 was more in residential cities than industrial areas [11]. Mutual estimations of the health effects of exposure to toxic volatile substances from quantitative studies could give crucial information for health-related diseases, but estimates for the Indian population are constrained [12, 13].

The rest of the paper is sketched out as follows in section II under the heading Methodology presents a brief explanation of the algorithms used during the study, while the III section presents Experimental Analysis and Result. At last, the paper is concluded in Section IV, which forms the basis for the analysis of air quality and the results of this paper. Following statements includes the research contribution-

- This article investigates how a lockdown affects air quality.
- Using some machine learning algorithms, the difference in the percentage of above-mentioned harmful pollutants before and post lockdown is compared.
- The suggestions of COVID-19 widespread on discussing contamination, environment analysts, and environment in India were examined. To guarantee that the discuss quality amid this lockdown period was moderately cleaner.
- In this ponder, concentrations of six criteria poisons, PM10, PM2.5, NO2 amid 2015-01-01 to 2020-07-01 in a few cities covering diverse locales of India, were dissected.
- The changes in contamination in this lockdown period can explain the achievability of discussing quality enhancement when there are critical confinements in emanations from numerous sources and provides controllers superior plans to control discuss contamination.

2. Literature Survey

Mahato et al. in his research revealed the fact that the air quality dramatically improves during lockdown [14]. PM10 and PM2.5 concentrations have experienced the greatest reduction (N50%) among the chosen pollutants compared to the pure lockdown phase. When compared to the previous year (2019), the reduction in PM10 and PM2.5 during the time mentioned above period is as high as roughly 60% and 39%, respectivelyNO2 (52.68 percent), and CO (30.35 percent) levels have decreased throughout the lockdown period, among other contaminants. After just four days of lockdown, an improvement in the air quality of between 40% and 50% is seen. Based on a difference-indifferences design, this analysis estimates the causal effects of 8 different lockdown measures on changes in a variety of individual pollutants. The findings demonstrate that the NO2 air quality index strength decreases more dramatically (23–37%). The findings indicate that, compared to the pre-lockdown period, the NO2 air quality index value decreases more sharply (23-37%), followed by PM10, SO2, PM2.5, and CO, whereas the O3 value increases 10-27%. PM10, SO2, PM2.5, and CO are all increased during the re-lockdown time, although O3 rises [15]. Additionally, intra- and intercity traffic restrictions are more effective at reducing air pollution [16]. Therefore, this study aims to examine the pattern of air pollution in the various metropolitan areas and their neighboring regions before and after a lockdown. In this study, the non-parametric Maan-Kendall test and Sen's slope estimator have been used to determine air pollution trends before and during the lockdown. The Air Quality Index (AQI) demonstrates improvement during the lockdown as its value dramatically dropped (45 - 68 %). An intriguing finding was that O3 fell during the first week of the lockdown but later increased by 19 to 27% [17]. Traffic restrictions and the temporary closure of factories and industries were the main causes of this drop [18]. However, because thermal power plants were operating throughout lockdown, there was little to no improvement in the air quality in those areas. Overall, throughout the lockdown, there was a noticeable increase in air quality, which improved many seasonal illnesses like asthma and other cardio-respiratory problems in people, better climate conditions, and less pollution.

3. Proposed Work

To ponder the changes in air quality amid the lockdown period, the information from a few cities covering diverse districts of India was analyzed. To evaluate the adequacy of the different emissions diminishments connected with the downfall in human and mechanical activities due to the COVID-19 limitations, emission diminishments related to those restrictions are evaluated, and further study is conducted to analyze the changes in the levels of PM2.5, PM 10, NO2, and O3.

3.1 Data Preprocessing

Initially, the dataset consisted of the data for every city from where the cities mentioned above were extracted using Python libraries yearly from the period 2015 to 2020. Many AQI datasets are incomplete for various reasons, containing some missing data, which are either represented by blanks, NaNs, or other proxies. It is predefined in scikit learn that all the values needed are numerical and hold values but this is incompatible with some datasets. Therefore, before applying the machine learning models, the datasets must be compatible with libraries by eliminating all the missing values. An elementary approach to manipulating incomplete datasets is to abandon a complete row or column that initially contained missing values. However, this can lead to losing some useful information from the datasets. Hence, a much better approach is to ascribe the missing values or gather them from the known part of the data. In this paper, a univariate type of imputation, an algorithm, is used in which the values of any particular feature dimension are imputed by considering only that feature dimension. (impute.SimpleImputer). The SimpleImputer class gives a direct methodology for ascribing missing values by supplanting them with either steady values or utilizing the measurable method like mean, median, or mode of each feature in which the lost values are found.

3.2 Training of Models

In this research paper, the training model is developed on the Python platform with the help of sklearn and xgboost libraries. AQI, PM2.5, NO2, and PM10 data are primarily divided into two parts: training set 80%, and test set 20%. Only 80% of data is taken as the training set to reduce the risk of model over-fitting. Training is carried out by SVR and using XGBRegressor models for each of the components mentioned above. The XGBoost model was trained using historical data observation. The input of 5 consecutive days is taken into consideration to predict the AQI of the current day. To quantitatively assess the prediction precision, this paper uses R2_score and mean_absolute_percentage_error from the metrics module.

3.3 Prediction Model via SVR

The dataset is split by importing the train_test_split function from the sklearn library. To ensure that the model is appropriately trained, the training set to testing set is maintained in the ratio 80:20 by the percentage ratio. SVR implementation is based on the modules of the sklearn library, and the kernel selected in SVR is ‘linear’ [19]. The kernel functions play a vital role in SVM. Their work is to take information as input and change it into any required frame which offers assistance in deciding many important features.

3.4 Extreme Gradient Boosting (XGB)

XGB is a leading model and an open-source library for working with standard data or time series forecasting which is generally present in the tabular format. (E.g., Pandas DataFrames). To achieve accuracy, XGB requires more training data as compared to traditional algorithms such as Random Forest. XGB is implemented using gradient boosted decision trees. An ensemble model is prepared by repeatedly forming new models and calculating errors for each model using some built-in functions. At the mere beginning, the predicted model is highly inaccurate but with further repetitions, the percentage error is minimized. The main intention behind implementing these trees is to achieve higher execution speed and greater model accuracy. The predictions are made using the XGBRegressor which is present as a module of the XGB library supported by python.

3.5 Comparison of Models

"Fig. 1", The actual and predicted values are depicted using SVR. "Fig. 2", The actual and predicted values for the year 2020 during lockdown are plotted using XGBoost. Table I has three columns in which different models and their coefficient of determination and their Mean absolute percentage error are depicted. After comparing the mean absolute percentage error and coefficient of determination for both SVR and Xgboost models, it is observed that the Xgboost technique is more appropriate for the prediction.

Fig. 1. SVR Comparison on Testing data.

Fig. 2. XGBoost Comparison on Testing data.

Table 1. Comparison Table of coefficient of determination and mean absolute percentage error

Models	R2 (Coefficient of Determination)	Mean Absolute Percentage Error
SVR	0.804	0.183
XGBOOST	0.703	0.177

4. Experimental Analysis and Result

4.1 Data Sources and Description

To examine the alterations in the air quality due to lockdown, the data of a few cities from different zones of India were taken from Kaggle [10] and analyzed. Also, the hourly data of all cities is acquired from the CPCB online portal for air quality data description [20]. The studied cities are Delhi, Chennai, Hyderabad, Jaipur, Lucknow, Patna, Visakhapatnam, Jaipur, Gurugram, Kolkata, Bengaluru, Nagpur, Agra. The data was collected daily, which included the concentration of different pollutants between 01 January 2015 to 01 July 2020. The dataset has 16 columns and 29531 rows, of which 12 columns are for the concentration of pollution, 1 for the city name, 1 for the date, and the rest of 2 for AQI (Air Quality Index) description. The different pollutants that are given in the dataset are PM2.5, PM10, NO2, SO2, O3.

4.2 Barometric Changes During Lockdown period

The previous year's seasonal trends were studied to see whether there is any climatical change in the lockdown period where they happen or not. “Fig. 3”, It mainly depicts two pollution level peaks, one during October and the other one during January. And the lowest amount of pollution is around July-September after this period has a tragic increment in this. Similarly, another depreciation is observed is from January to July. This spike in the winters is due to a combination of factors. Here is one point to be noted that majorly the North Indian states have a higher % of pollution. The spike is due to factors including Winter aversion (explicate), valley effect (explicate), periodic aspects related to various seasons for instance dust, storms, crop fires, burning of solid fuels, firework-related pollution, etc

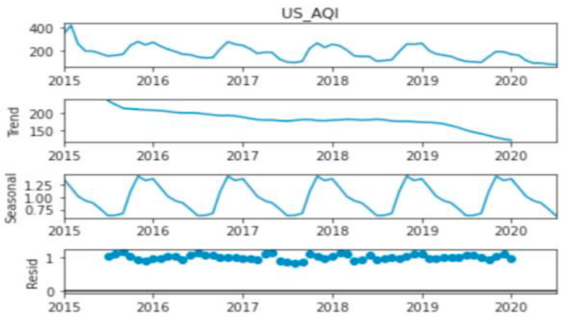

Fig. 3. Seasonal changes in AQI during January 2017 – July 2020

4.3 Changes in Concentration of Pollutants

After considering every point, it observes that all the pollutants' curves and values are decreased during the whole lockdown period, especially in the area that is not under the industrial sector or mining sector. “Fig. 4”, It shows the data of different pollutants according to city wise, in which it is studied that some major cities of India were worst affected by some pollutants before the lockdown happens in India. “Fig. 5”, It can be observed PM2.5 curve bent sharply, 40.58% of the PM2.5 drop observes in the lockdown period. “Fig. 6”, The plot of PM2.5 of different cities before the lockdown from 2015-to 2019 is shown. “Fig. 7”, The plot of actual as compared to the prediction of PM2.5 level of 2020 is presented. By observing the graph, the mean value of PM2.5 level predicted if lockdown did not happen in India was 86.95µg/m3, but the actual mean value of PM2.5 level in lockdown period is 48.09µg/m3. Almost all the cities show a fall in the PM2.5 level, especially the cities which are mostly affected due to pollution through vehicles like Delhi, Bengaluru, etc. because the vehicles are measured cause of PM2.5 pollutant and in the lockdown, all the vehicles’ activities were stopped. While the change in the concentration of PM10 was enormous than the changes in the concentration of PM2.5.due to biogenic changes [21]. Figure 8 shows the PM10 level of different cities of India before the lockdown. “Fig. 9”, The mean value of PM10 all over India before the lockdown period was

110.78µg/m3, which was reduced to 60.84µg/m3 during the lockdown. The graph of the actual and predicted value of PM10 in the lockdown period is outlined.

	City	PM2.5		City	PM10		City	NO2		City	SO2
0	Patna	123.500000	0	Delhi	232.810000	0	Ahmedabad	59.030000	0	Ahmedabad	55.250000
1	Delhi	117.200000	1	Gurugram	191.500000	1	Delhi	50.790000	1	Jorapokhar	33.650000
2	Gurugram	117.100000	2	Talcher	165.770000	2	Kolkata	40.400000	2	Talcher	28.490000
3	Lucknow	109.710000	3	Jorapokhar	149.660000	3	Patna	37.490000	3	Patna	22.130000
4	Ahmedabad	67.850000	4	Patna	126.750000	4	Visakhapatnam	37.190000	4	Kochi	17.600000
5	Kolkata	64.360000	5	Brajrajnagar	124.220000	5	Lucknow	33.240000	5	Delhi	15.900000
6	Jorapokhar	64.230000	6	Jaipur	123.480000	6	Jaipur	32.420000	6	Mumbai	15.200000
7	Brajrajnagar	64.060000	7	Bhopal	119.320000	7	Bhopal	31.350000	7	Guwahati	14.660000
8	Guwahati	63.690000	8	Guwahati	116.600000	8	Coimbatore	28.780000	8	Amaravati	14.260000
9	Talcher	61.410000	9	Kolkata	115.630000	9	Hyderabad	28.390000	9	Bhopal	13.060000

Fig. 4. The mean value of different pollutants before the lockdown

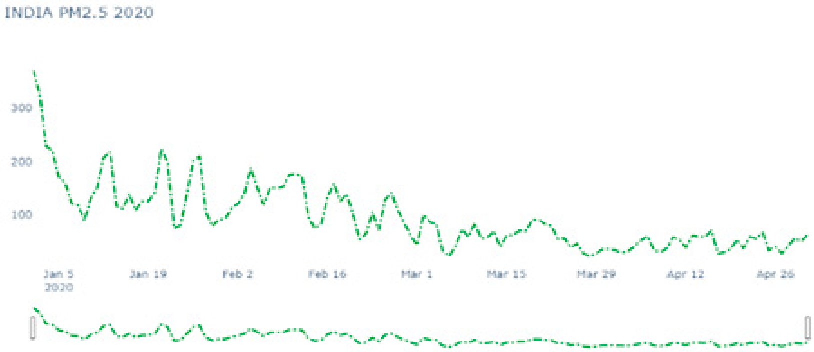

Fig. 5. Shows the trend of PM2.5 pollutants in the lockdown period.

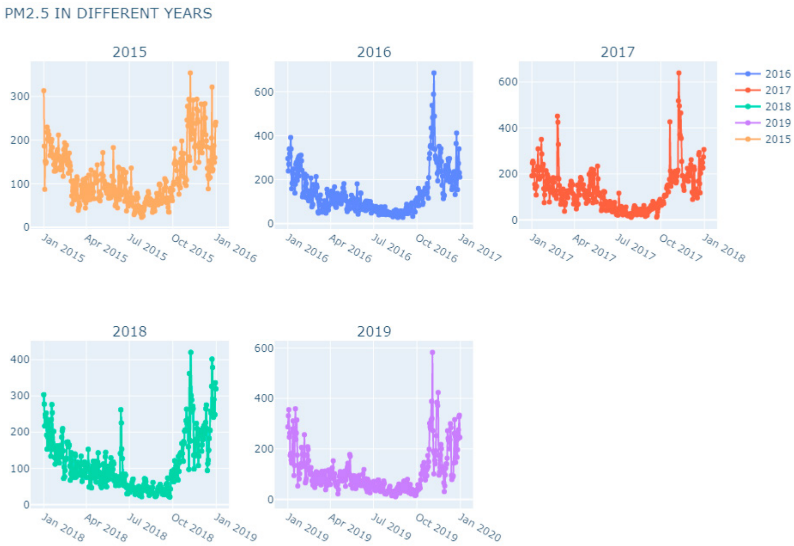

Fig. 6. Plots of PM2.5 level from the year 2015-2019

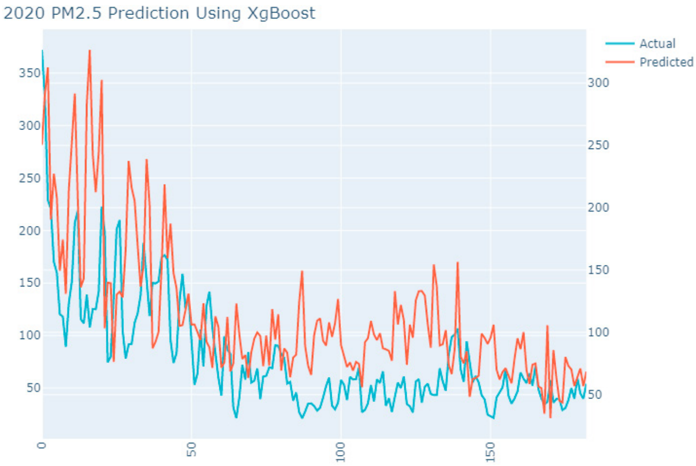

Fig. 7. Graph of Actual vs Predicted value of PM2.5 level during lockdown.

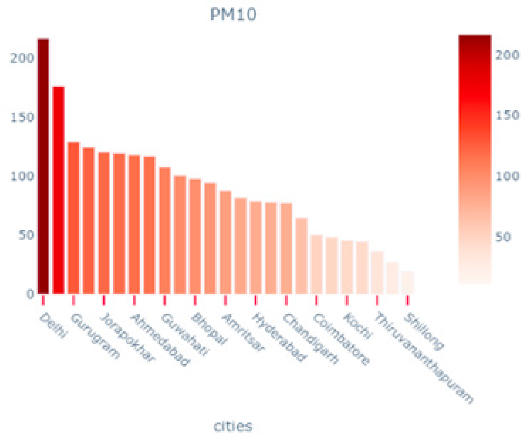

Fig. 8. PM10 level of different cities of India before the lockdown

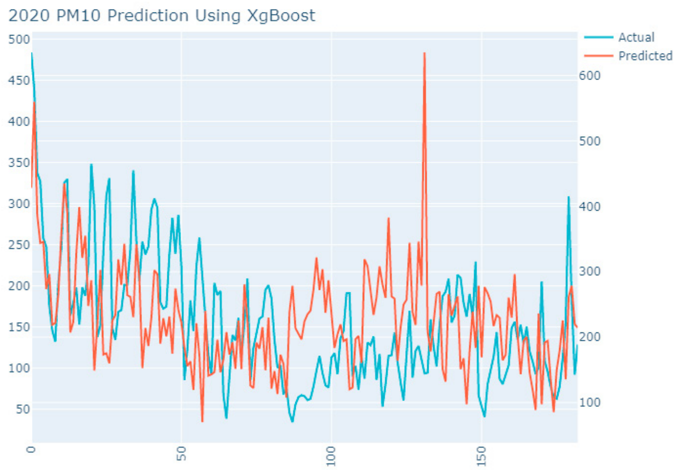

Fig. 9. Graph of Actual vs Predicted PM10 level during the lockdown.

The NO2 level was also dropped during this period, but the cities related to the coal mining belt were not showing significant change during the lockdown because the coal mining was not stopped during the lockdown. “Fig. 10”, The annual concentration of NO2 in the troposphere from the year 2017-to 2020. “Fig. 11”, Another plot shows the actual against predicted NO2 during the lockdown period. The predicted mean value of NO2 during the lockdown period is 35.86µg/m3. But the actual mean value of NO2 is 24.43µg/m3. This shows the 31.87% changes in NO2 concentration during the lockdown

Table II has three columns of Pollutants, the value of Pollutant Before Lockdown and the value of Pollutant After Lockdown in micrograms/cubic meter (µg/m3). It has four rows in values of 4 pollutants PM2.5, PM10, NO2, SO2, and O3 are written. It shows that there is no change in SO2 level, while the O3 (Ozone) level increases which gives a good impact on fulfilling the ozone hole.

4.4 Air Quality Changes before and after lockdown

Previous studies also show the significant changes in AQI during the lockdown [1]. To measure those changes XGBoost technique is used. As the concentration of all the pollutants decreases in the lockdown period, the AQI index also decreases. “Fig. 12”, India's past AQI index graph from 2015-to 2019 for the observation of the AQI index. This graph also shows the drop in the AQI level due to lockdown. 40% of the overall reduction was measured in the AQI of India from the predicted AQI for the year 2020 than the AQI, which was actually in the year 2020 due to the lockdown. “Fig. 13”, The graph of the actual against predicted AQI of India during the lockdown period. It is visible that in the lockdown period, the AQI graphs sharply drop than its predicted value. The mean value of the AQI Index of India before the lockdown was 175 AQI, which was reduced to 95 AQI in the lockdown period. The changes in the AQI were observed due to the stopping of all the industrial activities and also the restriction of human activities.

Table 2. Values of different pollutants of air before the lockdown and after the lockdown

Pollutants	Before Lockdown (µg/m3)	After lockdown (µg/m3)
PM2.5	86.95	48.09
PM10	110.78	60.84
NO2	35.86	24.43
SO2	18.28	16.48
O3	33.57	53.48

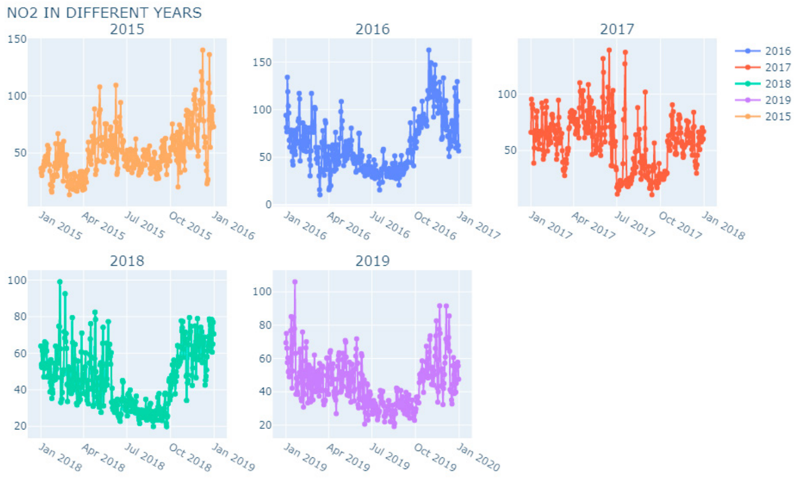

Fig. 10. Graph of NO2 level of India from the year 2015-2019

Fig. 11. Graph of Actual vs Predicted value of NO2 level during the lockdown.

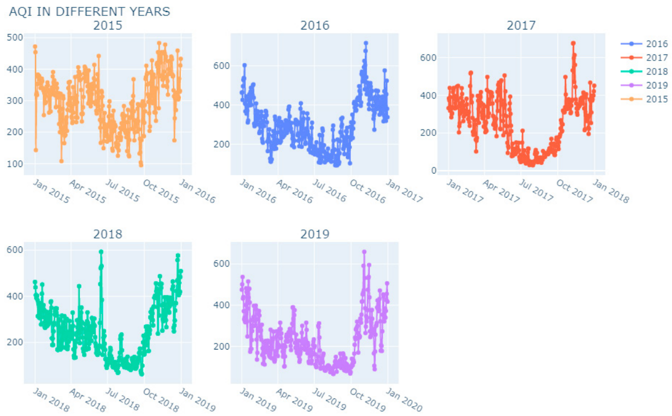

Fig. 12. Graph of AQI level of India from the year 2015-2019

Fig. 13. Actual vs Predicted value of AQI level during the lockdown.

5. CONCLUSION

With the spread of COVID-19 all over the planet, the inescapable and fast state-run administrations' reactions have brought about clearing impacts. Among these, air quality effects may be relied upon to encounter a spectacular enhancement. This paper evaluates the underlying effect of different lockdown measures on the air quality (estimated by PM2.5, SO2, PM10, O3, NO2, and individual AQI) in many major cities of India from 1-January-2020 to 1-July-

2020. This pandemic permits us to inspect the general changes in air quality during ordinary times through a "biggest scale analyze ever" from COVID-19 restrictions. It is observed that different air toxins react diversely to the lockdown measures. This paper emphasizes the changes in the trends of annual concentrations of these pollutants during the lockdown period. The research signifies that the PM10 air quality index value drops impulsively from (40–45%), compared prior to the lockdown period, followed by NO2 (27–35%), SO2 (2–10%), PM2.5 (35-40%), but the O3 rises (12–25%).

References

[1] Feng Liu, Meichang Wang, and Meina Zheng (2021) “Effects of COVID-19 lockdown on global air quality and health.” *Science of the Total Environment* **755**: 1-13.

[2] Manish Kumar Sharma, Prince Kumar, Akhtar Rasool, Aditya Dubey, and Vishal Kumar Mahto (2021) “Classification of Actual and Fake News in Pandemic.” *International Conference on IoT in Social, Mobile, Analytics and Cloud (I-SMAC)*: 1168-1174.

[3] Parth Vyas, Falak Sharma, Akhtar Rasool, and Aditya Dubey, “Supervised Multimodal Emotion Analysis of Violence on Doctors Tweets.” *International Conference on IoT in Social, Mobile, Analytics and Cloud (I-SMAC)*: 962-967.

[4] Anchal Garg, ArvindKumar, N.C. Gupta (2021) “Comprehensive study on impact assessment of lockdown on overall ambient air quality amid COVID-19 in Delhi and its NCR, India.” *Journal of Hazardous Materials Letters* **2**: 1-6

[5] Matthew A. Cole, Robert J R Elliott & Bowen Liu (2020) “The Impact of the Wuhan Covid-19 Lockdown on Air Pollution and Health: A Machine Learning and Augmented Synthetic Control Approach.” *Environmental and Resource Economics* **76(4)**: 553–580.

[6] Shubham Sharma, Mengyuan Zhang, Anshika, Jingsi Gao, Hongliang Zhang, and Sri Harsha Kota (2020) “Effect of restricted emissions during COVID-19 on air quality in India.” *Science of the Total Environment* **728**: 1-8

[7] Justin Beaudoin and C.Y. Cynthia Lin Lawell (2017) “Is Public Transit’s ‘Green’ Reputation Deserved? Evaluating the Effects of Transit Supply on Air Quality.”: 1-46.

[8] Aditya Dubey, and Akhtar Rasool (2021) “Efficient technique of microarray missing data imputation using clustering and weighted nearest neighbour.” *Scientific Reports* **11**: 1-12.

[9] Zander S. Venter, Kristin Aunan, Sourangsu Chowdhury and Jos Lelieveld (2020) “COVID-19 lockdowns cause global air pollution declines.” *Earth Atmospheric and Planetary Sciences* **117 (32)**: 18984-18990.

[10] “Air Quality Data in India (2015 - 2020),” *Kaggle,* Jan. 27, 2022, [Online]. Available: https://www.kaggle.com/rohanrao/air-quality-data-in-india

[11] Ramesh P. Singh and Akshansha Chauhan (2020) “Impact of lockdown on air quality in India during COVID-19 pandemic.” *Air Quality, Atmosphere and Health* **13 (8)**: 921–928.

[12] Maoyong Fan, Guojun He, and Maigeng Zhou (2020) “The winter choke: Coal-Fired heating, air pollution, and mortality in China.” *Journal of Health Economics,* **71**: 1-17.

[13] Feng Lu, Dongqun Xu, Yibin Cheng, Shaoxia Dong, Chao Guo, Xue Jiang, Xiaoying Zheng (2015) “Systematic review and meta-analysis of the adverse health effects of ambient PM2.5 and PM10 pollution in the Chinese population.” *Environ Res*. **136**:196-204

[14] Susanta Mahato, Swades Pal Krishna and Gopal Ghosh (2020) “Effect of lockdown amid COVID-19 pandemic on air quality of the megacity Delhi, India” *Science of The Total Environment* **730**:1-23.

[15] Feng Liu, Meichang Wang, Meina Zheng (2021) “Effects of COVID-19 lockdown on global air quality and health” *Science of The Total Environment* **755:**1-13.

[16] Niladri Das, Subhasish Sutradhar, Ranajit Ghosh, Prolay Mondal (2021) “Asymmetric nexus between air quality index and nationwide lockdown for COVID-19 pandemic in a part of Kolkata metropolitan, India” *Urban Climate* **36**:1-17.

[17] Anchal Garg, Arvind Kumar, N.C. Gupta (2021) “Comprehensive study on impact assessment of lockdown on overall ambient air quality amid COVID-19 in Delhi and its NCR, India” *Journal of Hazardous Materials Letters* **2:**1-6.

[18] C.T. Resmi, T. Nishanth, M.K. Satheesh Kumar, M.G. Manoj,M. Balachandramohan and K.T. Valsaraj (2020) “ Air quality improvement during triple-lockdown in the coastal city of Kannur, Kerala to combat Covid-19 transmission” *peerj :*1-20.

[19] Aditya Dubey, and Akhtar Rasool (2022) “Recent advances and applications of deep learning technique.” *Journal of Theoretical and Applied Information Technology* **100 (13)**: 5000-5009.

[20] “Central Pollution Control Board.”, [Online]. Available: https://app. cpcbccr.com/ccr/#/caaqm-dashboard-all/caaqm-landing.

[21] Zbigniew Klimont, Kaarle Kupiainen, Chris Heyes, Pallav Purohit, Janusz Cofala, Peter Rafaj, Jens Borken-Kleefeld, and Wolfgang Schöpp (2017) “Global anthropogenic emissions of particulate matter including black carbon” *Atmospheric Chemistry and Physics Discussions* **17 (14)**: 8681-8723.