Nature Environment and Pollution Technology
An International Quarterly Scientific Journal

p-ISSN: 0972-6268
(Print copies up to 2016)
e-ISSN: 2395-3454

Vol. 22 | No. 1 | pp. 199-210 | 2023

Original Research Paper

https://doi.org/10.46488/NEPT.2023.v22i01.018

Open Access Journal

Assessment and Prediction of Air Quality Level Using ARIMA Model: A Case Study of Surat City, Gujarat State, India

H. N. Mahendra*†, S. Mallikarjunaswamy*, D. Mahesh Kumar*, Shilpi Kumari*, Shubhali Kashyap*, Sapna Fulwani* and Aishee Chatterjee*

*Department of Electronics and Communication Engineering, JSS Academy of Technical Education, Bengaluru-560060, Karnataka, India

†Corresponding author: H. N. Mahendra; mahendrahn@jssateb.ac.in

Nat. Env. & Poll. Tech.
Website: www.neptjournal.com

Received: 26-07-2022
Revised: 06-10-2022
Accepted: 17-10-2022

Key Words:
Air quality
ARIMA model
Monitoring of air
Prediction system

ABSTRACT

Air quality has recently been a huge concern as it directly affects people's lives. An air quality level assessment and prediction system is essential to keep track of air quality. Therefore, developing an efficient air quality assessment and prediction system has become one of the most important concerns. In the present work air quality level of Surat city, India is assessed and predicted for the period from 2020 to 2023 using the Autoregressive integrated moving average (ARIMA) model. Experimental results show that the ARIMA model outperforms the other models. According to the findings, the maximum quantity of SO_2 and NO_2 present in the air in 2020 is 37 mm and 18 mm, respectively, with a maximum of 27 mm and 31 mm in 2021. Thus, we can observe that even though SO_2 has reduced a bit, the amount of NO_2 has increased, thus degrading the quality of air.

INTRODUCTION

Today, many industrial and daily activities have produced significant air pollution, particularly in emerging metropolises like China and India. Exposure to outdoor air pollution can have adverse health effects that can prove to be complicated results of pollutant compositions and concentrations (Badas et al. 2019). Ozone (O_3), nitrogen oxide, volatile organic compounds, metals, particulate matter, and metals are among the significant causes of air pollutants in cities (Arsov et al. 2020, Bhalgat et al. 2019). Increased air pollutants (such as SO_2, O_3, and PM) have been linked to higher death and morbidity rates. The urban environment's air quality (PM2.5 concentrations) has severely harmed people's work and lives in major cities (Al-Awadi 2018). The major sources of outdoor pollution are automobiles, factories, and industries, while smoke, toxins, and pollutants in the home cause indoor air pollution. Primary pollutants and secondary pollutants are the two categories of pollutants that cause air pollution (Baran et al. 2018). Primary pollutants are distinguishable from secondary pollutants as primary pollutants are straight discharged into the air from the source. Secondary pollutants, on the other hand, cause concentrations to rise when they react with primary pollutants or other air particles.

Several toxins are responsible for air pollution, but PM2.5 is the most notable, according to the author and their research (Angelin et al. 2019). The concentrations of PM2.5 can be calculated using logistic regression and autoregression (Zhu et al. 2018; Qingping et al. 2014). Various writers removed the day-by-day predictions of pollution levels by predicting hourly data using algorithms (Kumar & Pande 2022, Cosma & Simha 2018, Samal et al. 2019). The first stage is rating the air quality of an urban setting to actively collect sample air particles in every area of the city. In most nations, the current technique of assessment of air quality is via static air pollution monitoring stations (Li & He 2017, Santos et al. 2020). These reference stations can give very high precise readings from a small count of well-chosen areas that should reflect a variety of distinct environments (Amado & Dela Cruz 2018). Air pollution affects humans and plants negatively (Marjovi et al. 2015, San José et al. 2019). It causes ailments that are not life-threatening, such as throat and nasal discomfort (Daisey et al. 2003). Headaches can progress to more serious illnesses such as lung cancer, respiratory problems, brain disease, renal disease, shortness

ORCID details of the authors:

H. N. Mahendra
https://orcid.org/0000-0003-3854-5500

of breath, and even death. Some masks defend us from rising air pollution, and numerous laws regulate air pollution (Jain & Mandowara 2019, Savita et al. 2018). It is also very important to raise public awareness about the dangers created by air pollution (David et al. 2019). The air quality should be evaluated very accurately, and an air quality prediction system is essential to take the appropriate action to minimize air pollution (Chang et al. 2020).

There are a variety of traditional techniques for measuring it, but the results are not always exact, and it requires a big amount of mathematical computation (Ferrari et al. 2017, Huang et al. 2018). Machine learning is a form of artificial intelligence that requires many mathematical operations (Gaganjot et al. 2018, Zhu et al. 2018). Artificial intelligence, known as machine learning, plays a key role in air quality assessment and prediction (Kumar et al. 2021, Pasupuleti et al. 2020). Machine learning algorithms are being used in several studies to measure the Air Quality Index (Nandini & Fathima 2019, Kostandina & Dimoski 2018, Madan et al. 2020). The first step in reducing air pollution is correctly assessing the air quality level. Machine learning methods are essential for measuring the air quality level (Xiang et al. 2016). In the present study, several algorithms are compared to several conditions in various sectors.

The motivation for doing this study was primarily an interest in focusing on the impact of air quality on human health. Also, people's lifestyles have become heavily interested in information about air quality. Real-time air quality assessment of data is obtained using current air quality monitoring devices, stations, and satellite meteorological data. However, this is insufficient, and it is essential to calculate the expected air pollution trend. Therefore, in the present study, the "Autoregressive integrated moving average (ARIMA) model" is used to assess and predict the air quality of the study area.

RELATED STUDIES

Huang et al. (2018) developed an algorithm to determine the air exchange condition of the car, which extract the pollutant concentration from the state that the concentration trend converges when we open the window, and eventually, the measured convergence. The measured values are specified as equivalent environmental air quality levels. The algorithm gives fast results, but prediction accuracy is low.

Xiang et al. (2016) presented a predictive method for the quality of air that depends on "Spatiotemporal deep learning (STDL)" that essentially considers spatial and temporal correlations. When comparing it with the existing air quality prediction model, this model simultaneously predicts air quality for all centers and shows adequate stability. Still, the prediction accuracy is too low and requires more computation time.

Qingping et al. (2014) developed a hybrid "Empirical mode decomposition general regression neural network (EEMDGRNN) model," which depends on preprocessing of data and analysis for 1-day prediction of PM2.5 concentrations. It gives accurate results but with tremendous time complexity.

Cosma & Simha (2019) presented an intelligent control method that depends on a Support Vector Machine (SVM) classifier (using a human-centric HVAC control system) in their study. Skin temperature is the only input given to the method and has shown good predictive power in acknowledging fixed heat needs. Although one skin temperature is used to predict 80% of heat demand accurately, it does not apply to time-series data analysis and has poor accuracy.

Zhu et al. (2018) formalized 24-hour forecasting as multi-task learning (MTL) by proposing in a study an improved model for predicting hourly air pollution concentrations based on previous meteorological data. It is based on standard regression models (linear or nonlinear) in machine learning. This helps speed up the learning process on big data, but with less accuracy.

Bhalgat et al. (2019) presented an integrated model to predict air pollution levels using artificial neural networks and kriging in their study. This model uses a linear regression protocol and a multilayer perceptron (ANN) to predict the next day. AR and ARIMA models successfully predict SO_2 values, but more studies are needed to predict PM2.5 and calculate AQI.

Chang et al. (2020) proposed a mixed model and framework to enhance reliability in predicting air pollution compared to traditional methods. This approach integrates various predictive models to educate the ensemble, together with the Pearson correlation coefficient to evaluate the dependence of different models on each other. For physical integration of this model, demonstrate predicted air pollution of 1 to 9 hours. The results suggested that the proposed mixed model exceeds one conventional method that teaches the machine in terms of MAE and RMSER. Still, it is difficult to determine the concentration of pollutants.

Pandey et al. (2018) use a wide range of classifiers like neural networks, Bayesian networks, SVMs, and decision trees in their research to determine PM1.0 levels along with UFP levels. Set traffic data and environment variables. The forecast is reasonably accurate for the PM1.0 level but has data limitations. Mahendra et al (2022) have use the SVM to analyse the performance in remote sensing image classification. In another study, Mahendra et al (2019) have analyse supervised classification in remote sensing image classification.

The comparative analysis of different air quality prediction models is shown in Table 1. The comparative study shows that the ARIMA model performs better than other machine learning algorithms.

OBJECTIVES OF THE STUDY

This study aims to assess and predict the air quality of Surat city for the years 2020 to 2023. The main objectives of this study are:

1. To assess pollutant concentrations of the study area by utilizing historical air quality and meteorological data.
2. To predict and forecast pollutant concentrations based on the assessment study of current and previous pollutant concentrations.

MATERIALS AND METHODS

Study Area and Data Set

The study area considered in this work is Surat city, located in Gujarat state, India. Gujarat is a region on the west coast of India with a shoreline of about 1,600 miles, a maximum of that on the Kathiawar peninsula. The capital of Gujarat is Gandhinagar, with a population of about 60.4 million. Surat is one of the most populated cities in Gujarat, with a population of 7,784,276. Surat city has a latitude of 21.1702° and a longitude of 72.8311°.

With its large population, Surat city's air quality is paramount. Air quality directly affects the occurrence of diseases and lowers the quality of life. Proper decision-making at the right time depends on measuring and analyzing air parameters, creating the need for real-time improved air quality monitoring. The use of multi-parameter air quality monitoring systems makes it possible to perform a detailed analysis of the level of significant pollution and its sources. Monitoring systems are integral to many smart city projects to monitor air quality and control the concentration of key pollutants in urban areas.

In the present work, three years of air quality data in 2020, 2021, and 2022 of the study area of Surat city is collected from Gujarat's Environmental Protection Administration (EPA) to build multiple speculative models. Air quality monitoring parameters such as NO_2, SO_2, PM2.5, PM10, CO, and O_3 are obtained from the data. The collected parameters are used to assess the present and predict future air quality of the study area.

The methodology followed in the present work is divided into four distinct phases. They are Data collection, Assessment, Prediction, and Forecast.

Identification

The first stage is initiated by identifying the various contents of the pollutants present in the air. Then, we assemble the previous data set by giving input to the system. It is followed by comparing the latest data to the assembled previous data set.

Table 1: Comparative analysis of air quality prediction model.

Ref.	Prediction Model	Prediction Performance	Pollutants	Study Areas
Korunoski et al. (2019)	Internet of the things-based algorithm.	Accuracy: 92.4%	CO_2, PM2.5, TVOC	Skopj, Methodius
Xiang et al. (2016)	STDL Model	Accuracy: 82.66%	PM2.5	China
Qingping et al. (2014)	Hybrid EEMD-GRNN Model	(Calculated using formula)	SO_2, NO_2, PM10, PM2.5, CO, O_3	Republic of Macedonia
Cosma and Simha (2018)	The method based on SVM Classifier	Accuracy: 80.9%	-	General
Zhu et al. (2018)	Standard regression models (linear or nonlinear)	-	SO_2, PM2.5, O_3	Chicago area
Chang et al. (2020)	Model using ANN and Kriging	-	SO_2	Mumbai, India
Pandey et al. (2018)	Hybrid model (LSTM, SVR, GBT, LSTM2)	High accuracy compared to the traditional method	CO, NO, NO_2, O_3, PM10, PM2.5, SO_2	EPA of Taiwan
Kostandina and Dimoski (2018)	Classifiers (Neural Network, Bayesian Network, SVM, and Decision Trees)	-	PM1.0	Hangzhou, China
Kumar et al. (2021)	ARIMA model	High accuracy compared to machine learning algorithms	CO, NO, NO_2, O_3, PM10, PM2.5, SO_2	Assam, India

Assessment

In the second phase, the concentration of pollutants is estimated by prevailing air standard monitoring devices, stations, and satellite weather specifics for 2020, 2021, and 2022. It can issue information about real-time air status monitoring.

Prediction

By using existing historical air data for weather conditions, prediction is carried out with the help of machine learning for the year 2023. The ARIMA model helps in the processing of high-dimensional large-scale data. We implement the de-exploration of multidimensional time-space using a sliding window mechanism. With access to the concentration of contaminants, we evaluate their impact on health and vegetation.

Forecast

Finally, in the last phase, we forecast estimated pollutants concentration for the year 2020, 2021, and 2022 and predicted pollutants concentration for the year 2023. The methodology followed in the present work is shown in Fig. 1.

ARIMA Model

Box and Jenkins introduced the ARIMA model, which can be termed a time series forecast model, in the early 1970s (Stellwagen & Tashman 2013)). The present value of a time series is linearly characterized by its primary importance in an ARMA model. The ARIMA model incorporates the concept of integration and is an extension of the Autoregressive Move Average. The values are forecasted using the previous stage data by an auto-regression model. Equation (1) describes a satisfactorily central principle that may lead to correct predictions of a chain of problems (1).

$$Y_{hat} = b_0 + b_1 * X_I \quad \ldots(1)$$

The value of X_I is an input for prediction, while the values b0 and b1 are model coefficients.

This approach can be used for time series when lag variables are included in past observations. Equation (2) describes the regression model's expression.

$$X_{(T+1)} = b_0 + b_1 * X_{(T-1)} * b2 + X_{(T-2)} \quad \ldots(2)$$

Based on the results of the previous two phases, the value of the (T+1) can be calculated (T-1 and T-2). Autoregressions are regression models that utilize the same input variable throughout the forecasting process.

RESULTS AND DISCUSSION

In the present work, MATLAB 9.6 R2019a has been used to assess and predict the air quality level from the data set of the study area. Table 1 shows the air quality index value

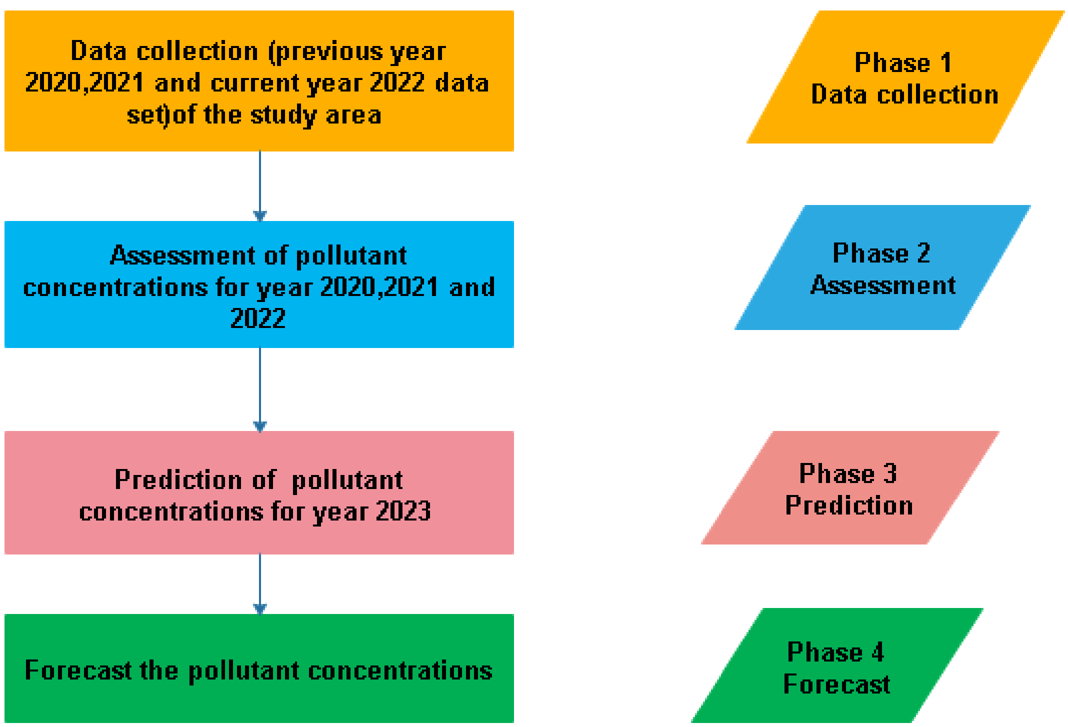

Fig. 1: Methodology followed in the present work.

(AQI). An AQI value of 50-0 indicates the air quality is good, and there is a limited possibility of affecting the environment and human health. Levels between 50-100 are acceptable, but if values exceed 100, they can be harmful to human health. If the value is between 151-200, people might experience more severe health effects. If the value ranges from 201 to 300, it is very unhealthy for the environment. On the other hand, an AQI value of 300 or higher indicates precarious air quality that is likely to harm the environment and public health. Table 2 shows the air quality index value of the different categories.

Air Quality Level Assessment

Table 3 shows the pollutant concentrations assessment for the year 2020. The results show that in 2020, the minimum amount of nitrogen and Sulfur dioxide in the air was 12 mm and 6 mm. The maximum amount of nitrogen and Sulphur dioxide in the air is 37 mm and 18 mm. Nitrogen dioxide and Sulphur dioxide mean values are 20.22 and 13.03, respectively.

Table 4 shows the pollutant concentrations assessment for the year 2021. The results show that in the year 2021, the minimum amount of Nitrogen dioxide and Sulphur dioxide present in the air are 15mm and 6mm, and the maximum amount of Nitrogen dioxide and Sulphur dioxide present in the air are 27mm and 31mm, the mean value of nitrogen dioxide and Sulphur dioxide are 20.36 and 13.57.

Table 5 shows the pollutant concentrations assessment for the year 2022. The results show that in the year 2022, the minimum amount of Nitrogen dioxide and Sulphur dioxide present in the air are 18 mm and 12 mm, and the maximum amount of Nitrogen dioxide and Sulphur dioxide present in the air is 29 mm and 22 mm, the mean value of nitrogen dioxide and sulphur dioxide are 21.14 and 14.61.

Fig. 2 shows the minimum values of the pollutants SO_2, NO_2, PM10, PM2.5, CO, and O_3, in the years 2020, 2021, and 2022 respectively. The results show that the concentration of most pollutants was minimal in 2020 and then gradually increased.

Fig. 3: Shows the maximum values of the pollutants SO_2, NO_2, PM10, PM2.5, CO, and O_3, in the years 2020, 2021, and 2022 respectively. The results show that the concentration of PM2.5 is the maximum in the year 2022,

Table 2: Air Quality Index value.

Air Quality Index Category	Air Quality Index Value	Description Of Air Quality
Good	0-50	Minimal Impact
Moderately Polluted	51-100	Acceptable quality of air
Unhealthy for responsive groups	101-150	Members of responsive groups may undergo health issues.
Harmful	151-200	Common people may experience health effects, while sensitive people might experience more serious health effects
Injurious	201-300	Health Alert
Hazardous	>300	Health warning of emergency conditions

Table 3: Pollutant concentrations assessment for the year 2020.

	SO_2 [m]	NO_2 [mm]	PM10 [mm]	PM 2.5 [mm]	CO [mm]	O_3 [mm]
Min	6	12	53	43	6	53
Max	18	37	179	9	18	179
Mean	13.03	20.22	85.47	28.78	12.81	85.39

Table 4: Pollutant concentrations assessment for the year 2021.

	SO_2 [mm]	NO_2 [mm]	PM10 [mm]	PM 2.5 [mm]	CO [mm]	O_3 [mm]
Min	6	15	63	12	11	63
Max	31	27	180	48	31	180
Mean	13.57	20.36	87.88	30.88	13.96	88.38

Table 5: Pollutant concentrations assessment for the year 2022.

	SO_2 [mm]	NO_2 [mm]	PM10 [mm]	PM 2.5 [mm]	CO [mm]	O_3 [mm]
Min	12	18	66	13	15	58
Max	22	29	133	133	15	133
Mean	14.61	21.14	89.17	31.46	15	88.40

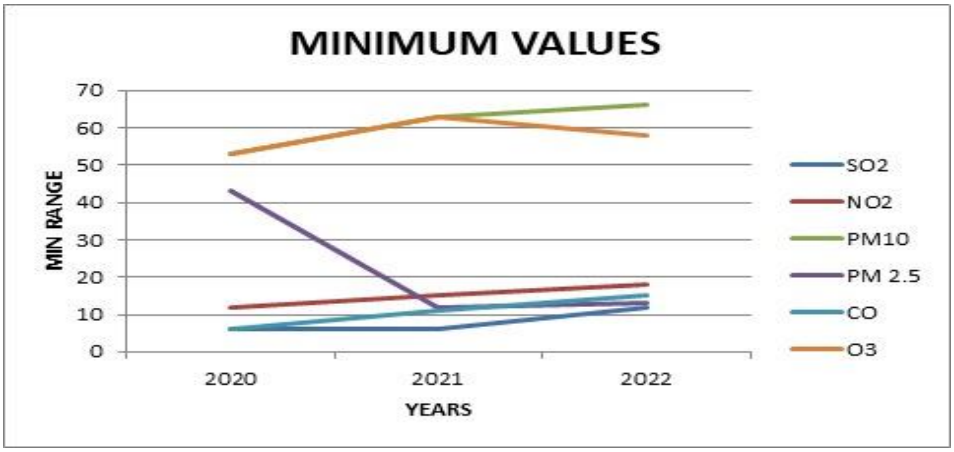

Fig. 2: Minimum values of the pollutants.

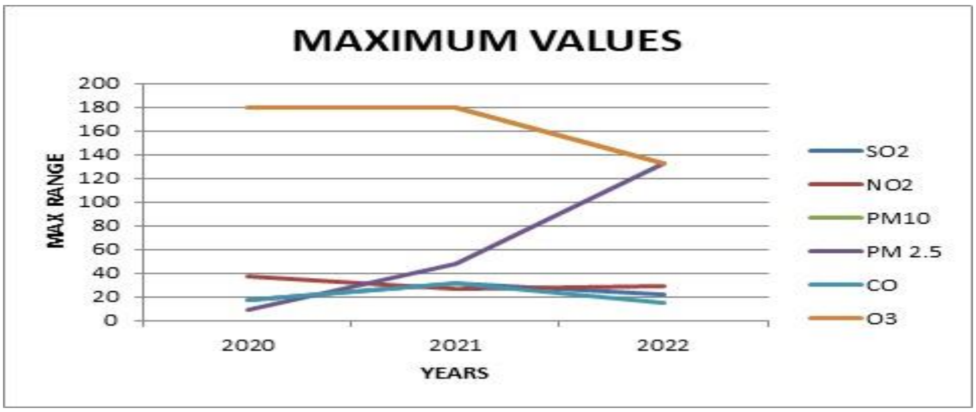

Fig. 3: Maximum values of the major pollutants.

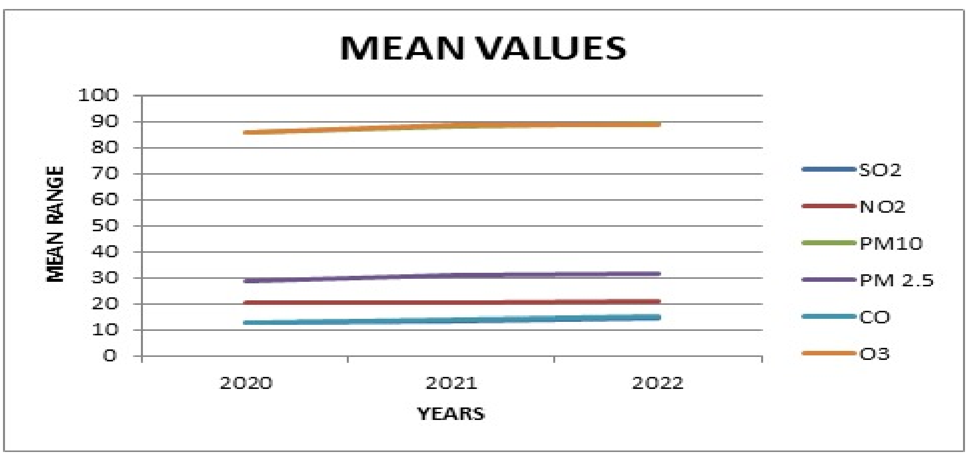

Fig. 4: Calculated mean values.

which means it is increasing day by day. The corresponding calculated mean value for the 3 years is shown in Fig. 4.

Air Quality Level Prediction

The air quality monitoring parameters detected are SO_2 (Sulphur Dioxide in mm), NO_2 (Nitric Dioxide in mm), PM10 (Particulate Matter 10-micrometer in mm), PM2.5 (Particulate Matter 2.5-micrometer in mm), CO, and O_3, which are also used to predict future air quality data. All the above six parameters contain 574 datasets (January 2020- until April 2023). The observed and prediction of all six parameters, i.e., SO_2, NO_2, PM10, PM2.5, CO, and O_3, are shown in Figs. 5, 6, 7, 8, 9, and 10, respectively.

The prediction of Sulphur dioxide (SO_2) levels is shown in Fig. 5. The values of SO_2 are observed from January 2020 to January 2022. By using the ARIMA model, it is forecasted up to August 2023. SO_2 is composed of Sulphur and oxygen and has a strong odor. It is a colorless, reactive air pollutant and contributes to air pollution. Hence, it can be harmful to human, animal, and plant lives. The main sources of Sulphur dioxide emissions are fossil fuel combustion and natural volcanic activity. Due to the COVID-19 pandemic prevention, commercial ships and incineration activities were

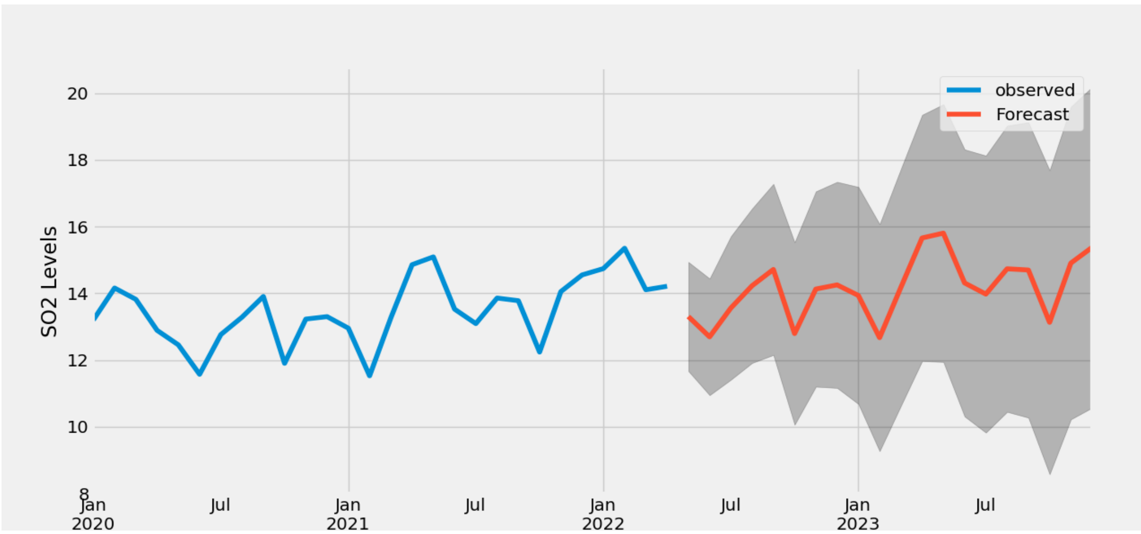

Fig. 5: Prediction of SO_2 levels.

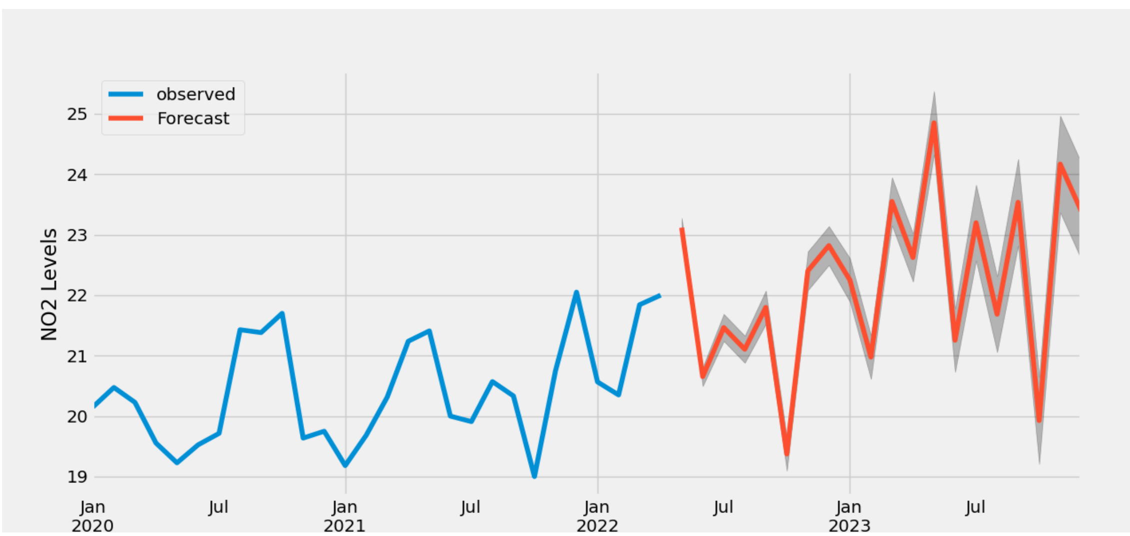

Fig. 6: Prediction of NO_2 levels.

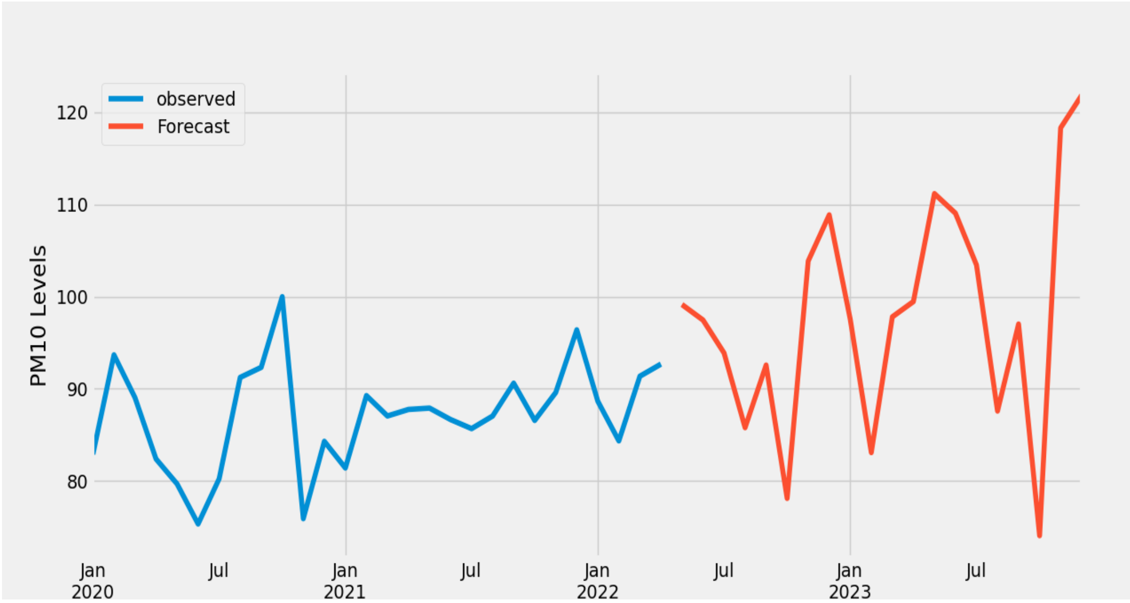

Fig. 7: Prediction of PM10 levels.

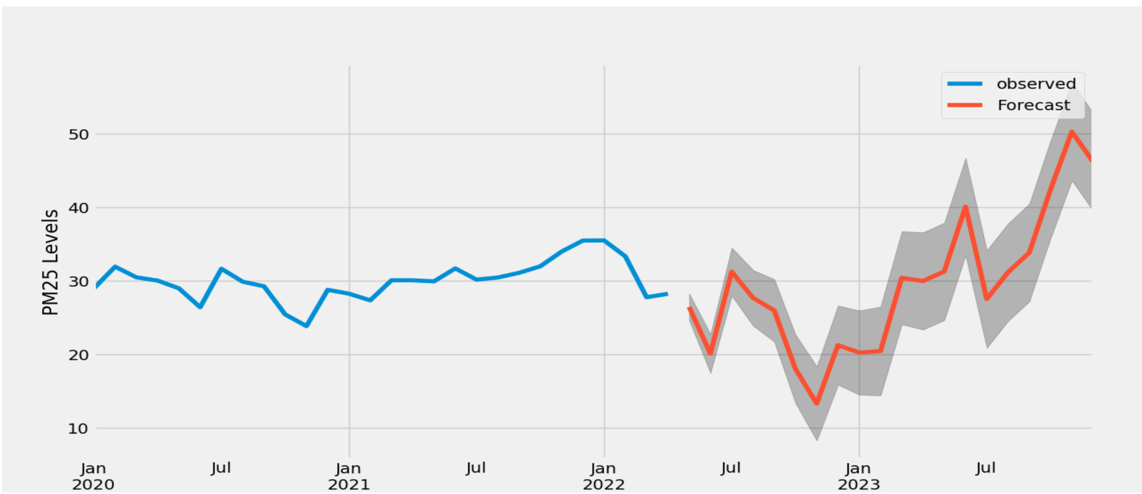

Fig. 8: Prediction of PM2.5 levels.

closed, which decreased the SO_2 emissions in the year 2020. In the short-term period, there was a downfall in the values of SO_2 emissions, but in the long-term period, an improvement in its values is inevitable. Hence, it is important to consider measures to reduce SO_2 levels in the environment.

The prediction of nitrogen dioxide (NO_2) levels is shown in Fig. 6. The values of NO_2 are observed from the duration of January 2020 to January 2022. By using the ARIMA model, it is forecasted up to August 2023. NO_2 is composed of nitrogen and oxygen and arises from high-temperature combustion processes of fossil fuels, like vehicle exhaust and transportation. During the year 2020 (January to July), there was a significant decrease in the values of NO_2 concentrations. The COVID-19 protective measures to terminate commercial activities helped in a limitation in NO_2 emissions from production and automobile exhaust.

The prediction of PM10 levels is shown in Fig. 7. The values of PM10 are observed from the duration of January 2020 to January 2022. By using the ARIMA model, it is forecasted up to August 2023. PM10s are very small particles with a diameter of 10 microns or less, found in dust and smoke. The pattern of PM10 primarily results from roadway

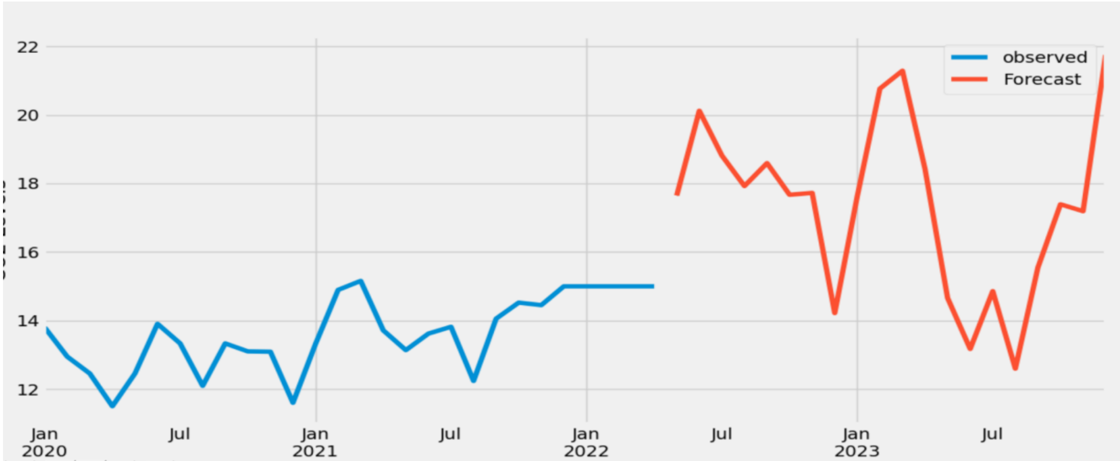

Fig. 9: Prediction of CO levels.

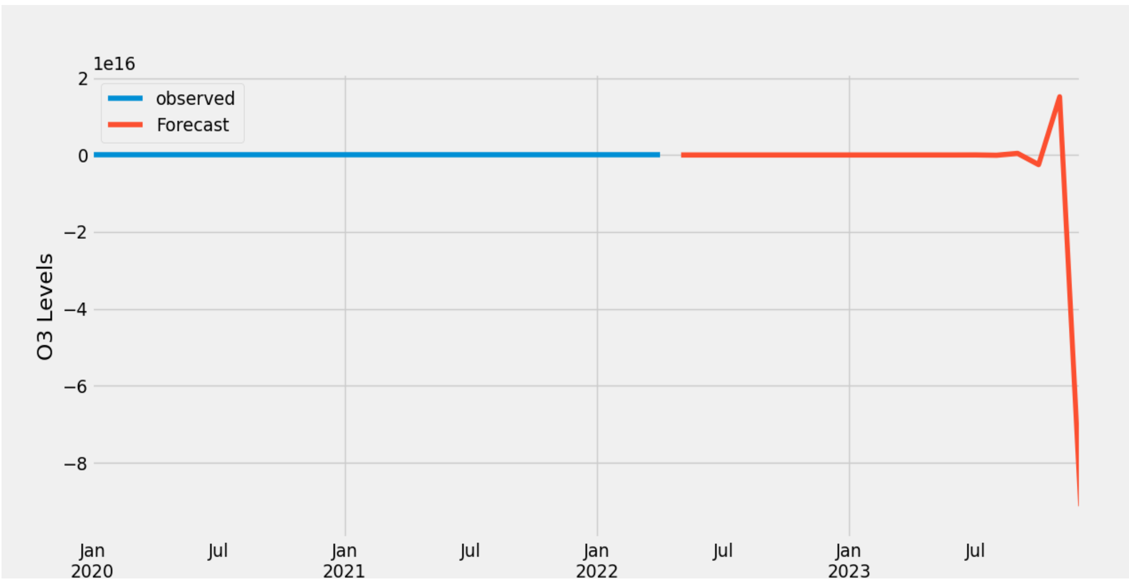

Fig. 10: Prediction of O_3 levels.

traffic, architecture activities, and dust entertainment. Variations in PM10 concentrations can be observed during the period 2020-2021. The consequences of reduced city pollution had a limited effect on PM10 concentrations.

The prediction of PM2.5 levels is shown in Fig. 8. The values of PM2.5 are observed from the duration of January 2020 to January 2022. By using the ARIMA model, it is forecasted up to August 2023. PM2.5 is an air pollutant with a size of 2.5 microns or smaller. Fine particulate matter (PM2.5) is a pollutant in the atmosphere that threatens people's health when its levels are high in the atmosphere. The observed PM2.5 levels were constant for the period 2020 to 2022. The predicted PM2.5 levels indicate an increase in its values for the year 2023.

The prediction of carbon monoxide (CO) levels is shown in Fig. 9. CO values are observed from January 2020 to January 2022. By using the ARIMA model, it is forecasted up to August 2023. Carbon monoxide is a clear, odorless gas that is a common air pollutant. The main sources of CO in the environment are automobiles and industrial emissions. The significant change in the observed and predicted values of CO can be noticed in Fig. 9.

The Ozone (O_3) level prediction is shown in Fig. 10. The values of O_3 are observed from January 2020 to

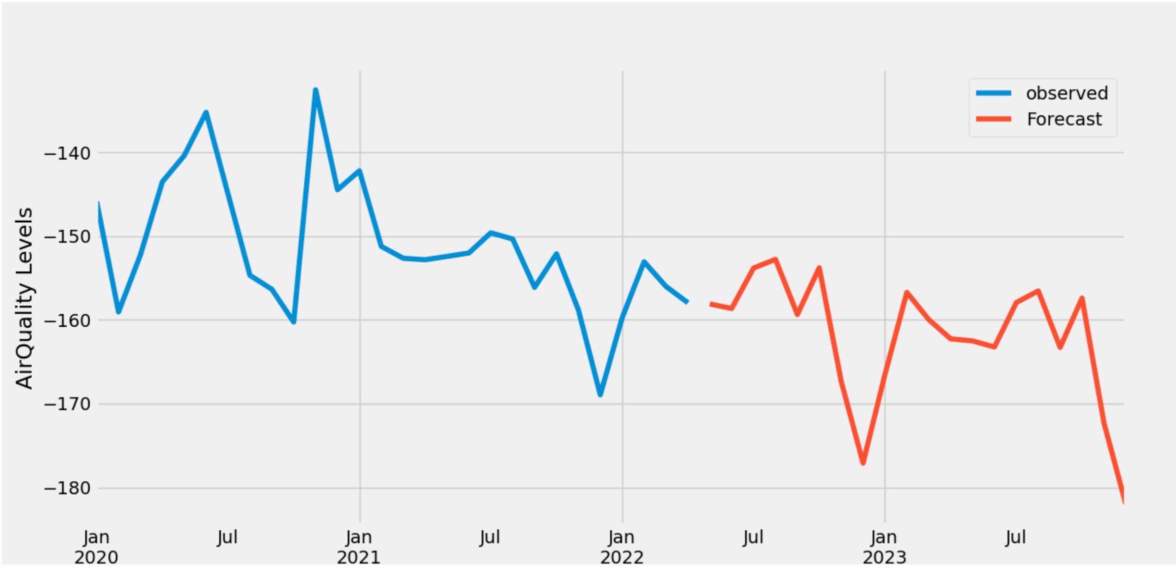

Fig. 11: Prediction of air quality levels.

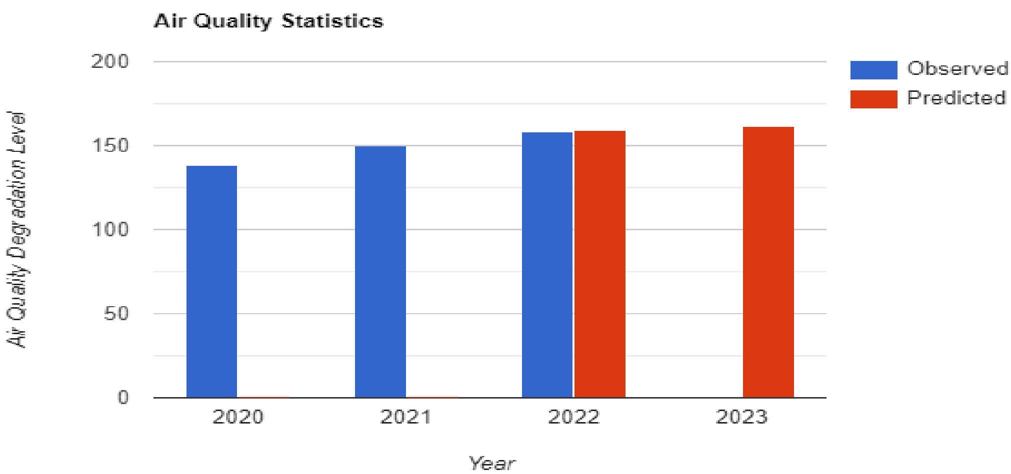

Fig. 12: Statistical representation of air quality level from 2020 to 2023.

January 2022. By using the ARIMA model, it is forecasted up to August 2023. O_3 gas is present in two layers of the atmosphere. High up in the atmosphere, O_3 forms a layer that protects and shields the Earth from harmful ultraviolet rays. The values of O_3 are expected to be uniform till July 2023, but they can be reduced strictly after July 2023.

The prediction of air quality levels is shown in Fig. 11. The blue color curve in the graph represents the observed values of air quality levels, while the red color curve represents the forecasted values. The dataset is observed from the duration of January 2020 to January 2022. The ARIMA model forecasts air quality levels up to August 2023. As we can see in the observed curve in 2020, quality improved linearly from January to July 2020. Again, air quality started degrading from July 2020 to December 2020. The sudden improvement and degradation in the year 2020 were due to the lift of the lockdown during the COVID-19 pandemic. In the year 2021, the quality of air started degrading exponentially. This was the year after the lockdown.

The cause of change in air quality in the year 2021 were conveyance (36.0 %), production (33.0%), and wind-blown dust (21.0%). The major cause of the degradation in air quality levels was the expansion of industries. Total estimated emissions for the year 2021, PM2.5, were almost

6,500 tons from the road, rail, aviation, and shipping transportation modes. Add to this the residential emissions, industrial emissions, dust, and open burning pollutants, emissions from diesel generators, and emissions from brick kilns, and the annual figure rises to a substantial 61,000 tons. Other pollutants measured included PM10, Sulphur dioxide (SO_2), nitric oxide (NOx), and carbon monoxide (CO). As we can see from the red curve in the predicted data graph, the air quality is degrading. The major contributors are industries and vehicles. The leading growth of urbanization and industrialization are making cities like Ahmedabad, Surat, Vadodara, and Rajkot the primary target for air pollution. According to reports, Surat has more than 2000 industries contributing to air pollution as of May 2012. The report states that the number of vehicles has doubled in the past 10 years.

The statistical representation of the air quality level is displayed in Fig. 12. The blue-colored histogram indicates the observed values, and the red-colored histogram shows the predicted values for the years 2022 and 2023. The accuracy of the predicted values can be observed from the bar graph. The observed values of degrading air quality levels were: 135 in July 2020 and 150 in July 2021,155 in July 2022. The predicted values of degrading air quality levels were 162 in July 2022 and 165 in July 2023.

The approach presented in our work has the advantage of following date columns and not exhibiting data fluctuation during the event. It is capable of overcoming unpredictability in data and generating effective results. The large data fluctuations if the data is collected around festivals such as Diwali, a major cause of increasing pollution due to big fireworks displays. It was also observed that the large data fluctuations and not following the date column in sequence might lead to incorrect results.

CONCLUSION

In this present study, we predict and forecast the air quality level using the ARIMA model. The various baseline models fail to forecast air quality due to frequent fluctuations in the dataset. This model enables us to deeply mine and explore the time-series data concept. The experiment results show that in the year 2020, the minimum amount of Nitrogen dioxide and Sulphur dioxide present in the air were 12 mm and 6 mm, and the maximum amount of Nitrogen dioxide and Sulphur dioxide present in the air was 37 mm and 18 mm. The mean value of Nitrogen dioxide and Sulphur dioxide are 20.22 and 13.03. Also, the predicted air quality level for the year 2023, the minimum amount of Nitrogen dioxide and Sulphur dioxide present in the air would be 18 mm and 12 mm, and the maximum amount of Nitrogen dioxide and Sulphur dioxide present in the air would be 29 mm and 22 mm. The mean value of Nitrogen dioxide and Sulphur dioxide would be 21.14 and 14.61, respectively. The predicted values from the system are highly accurate for analyzing time-series data. Hence, it can effectively turn on automatic air quality forecasting in the future.

ACKNOWLEDGEMENT

We would like to thank the JSS Academy of Technical Education Bengaluru, Visvesvaraya Technological University, Belagavi, and Vision Group on Science and Technology (VGST) Karnataka for all the support and encouragement provided by them to take up this research work.

REFERENCES

Al-Awadi, L.J. 2018. Assessment of indoor levels of volatile organic compounds and carbon dioxide in schools in Kuwait. J. Air Waste Manag. Assoc., 68(1): 54-72.

Amado, T.M. and Dela Cruz, J.C. 2018. Development of Machine Learning- based Predictive Models for Air Quality Monitoring and Characterization, TENCON IEEE Region Conference, 16-19 Nov 2020, Osaka, Japan, IEEE, NY, pp. 0668-0672.

Angelin, J. and Sasi Kumar, A. 2019. PM2.5 prediction using machine learning hybrid model for smart health. Int. J. Eng, Adv. Technol., 9(1): 6500-6504.

Arsov, M., Zdravevski, E., Lameski, P., Corizzo, R., Koteli, N., Mitreski, K. and Trajkovik, V. 2020. Short-term air pollution forecasting based on environmental factors and deep learning models. IEEE Conference on Computer Science and Information Systems, pp. 15-22.

Badas, M.G., Salvadori, L., Garau, M., Querzoli, G. and Ferrari, S. 2019. Urban areas parameterization for CFD simulation and cities air quality analysis, Int. J. Environ. Pollut., 66(1/2/3): 5-18.

Baran, B. 2019. Prediction of Air Quality Index by Extreme Learning Machines. 2019 International Conference on Artificial Intelligence and Data Processing (IDAP), Malatya, Turkey, Piscataway, IEEE, pp. 123-139.

Bhalgat, P., Bhoite, S. and Pitare, S. 2019. Air quality prediction using machine learning algorithms. Int. J. Comp. Appl. Technol. Res., 8(9): 367-390.

Chang, Y.S., Abimannan, S., Chiao, H.T., Lin, C.Y. and Huang, Y.P. 2020. An ensemble learning-based hybrid model and framework for air pollution forecasting. Environ. Sci. Pollut. Res., 27(30): 38155-38168.

Cosma, A. and Simha, R. 2018. Machine learning method for real-time non-invasive prediction of individual thermal preference in transient conditions. Build. Environ. 148: 372-383.

Daisey, J.M., Angell, W.J. and Apte, M.G. 2003. Indoor air quality, ventilation, and health symptoms in schools: An analysis of existing information. Indoor Air, 13(1): 53-64.

David, C., Amy, S., Daniel, C., Jo, D., Rod, J., Ian, L., Olalekan, A.M., Popoola, K. and Martin, S. 2019. Urban emission inventory optimisation using sensor data, an urban air quality model, and inversion techniques. Int. J. Environ, Pollut., 66(4): 252-266.

Ferrari, S., Badas, M.G., Garau, M., Querzoli, G. and Seoni, A. 2017. The air quality in narrow two-dimensional urban canyons with pitched and flat roof buildings. Int. J. Environ. Pollut., 62(1): 347-368.

Gaganjot, K.K., Jerry, Z., Gao, S.C., Shengqiang, L. and Gang, X. 2018. Air quality prediction: Big data and machine learning approaches. Int. J. Environ. Sci. Dev., 9(1): 8-16.

Huang, J., Duan, N., Ji, P., Ma, C., Hu, F., Ding, Y. and Yu, Y. 2018. A crowdsource-based sensing system for monitoring fine-grained air quality in urban environments. IEEE IoT J., 6(2): 142.

Jain, S. and Mandowara, V. 2019. Study on particulate matter pollution in Jaipur city. Int. J. Appl. Eng. Res., 14(3): 637-645.

Korunoski, M., Risteska, S., Biljana, M. and Trivodaliev, K. 2019. Internet of Things Solution for Intelligent Air Pollution Prediction and Visualization. IEEE EUROCON 2019 - 18th International Conference on Smart Technologies, 8-11 July, 2019, Novi Sad, Serbia, IEEE, NY, pp. 1-6.

Kostandina, V. and Dimoski, A. 2018. Air quality index prediction using simple machine learning algorithm., Int. J. Emerg, Trends Technol. Comp. Sci., 7(1): 25-30.

Kumar, K. and Pande, B.P. 2022. Air pollution prediction with machine learning: A case study of Indian Cities. Int. J. Environ. Sci. Technol., 51: 28-32

Kumar, T.S., Das, H.S., Choudhary, U., Dutta, P.E., Guha, D. and Laskar, Y. 2021. Analysis and Prediction of Air Pollution in Assam Using ARIMA/SARIMA and Machine Learning. In Muthukumar, P., Sarkar, D.K., De, D.and De, C.K. (eds), Innovations in Sustainable Energy and Technology, Springer, Cham, pp. 317-330.

Li, Y. and He, J. 2017. Design of an intelligent indoor air quality monitoring and purification device. IEEE Information Technology and Mechatronics Engineering Conference (ITOEC), 3-5 October 2017, Chongqing, China, IEEE, NY, pp. 1147 -1150.

Madan, T., Sagar, S. and Virmani, D. 2020. Air quality prediction using machine learning algorithms: A review. International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), 18-19 December 2020, Greater Noida, India, IEEE, NY, pp. 140-145.

Mahendra, H.N. and Mallikarjunaswamy, S. 2022. An efficient classification of hyperspectral remotely sensed data using support vector machine, International Journal of Electronics and Telecommunications, 68(3): 609-617.

Mahendra, H. N., Mallikarjunaswamy, S., Rekha, V., Puspalatha, V. and Sharmila, N. 2019. Performance analysis of different classifier for remote sensing application. International Journal of Engineering and Advanced Technology, 9: 2249-8958. DOI: 10.35940/ijeat.A1879.109119.

Marjovi, A., Arfire, M. and Martinoli, A. 2015. High-Resolution Air Pollution Maps in Urban Environments Using Mobile Sensor Networks, International Conference on Distributed Computing in Sensor Systems, 10-12 June 2015, Fortaleza, Brazil, IEEE, NY, pp. 11-20.

Mudholkar, A., Akash, S., Ajay, C. and Gowramma, G.S. 2019. Air pollution data analysis using the ARIMA model. In Wang, J., Reddy, G.R., Prasad, K.V. and Reddy, V.S. (eds), Soft Computing and Signal Processing Proceedings, Springer, Cham, pp. 74-78.

Nandini, K. and Fathima, G. 2019. Urban air quality analysis and prediction using machine learning. International Conference on Advanced Technologies in Intelligent Control, Environment, Computing & Communication Engineering (ICATIECE), 19-20 March 2019, Bengaluru, IEEE, NY, pp. 98-102.

Pandey, G., Zhang, B. and Jian, L. 2018. Predicting submicron air pollution indicators: a machine learning approach. Environ. Sci. Process. Impacts, 15(5): 996-1005.

Pasupuleti, V.R., Uhasri, P., Kalyan, S. and Reddy, H.K. 2020. Air Quality Prediction of Data Log by Machine Learning. International Conference on Advanced Computing and Communication Systems (ICACCS), 6-7 March 2020, IEEE, NY, pp. 1395-1399.

Qingping, Z., Haiyan, J., Jianzhou, W. and Jianling, Z. 2014. A hybrid model for PM2.5 forecasting based on ensemble empirical mode decomposition and a general regression neural network. Sci. Total Environ., 496(1): 264-274.

Samal, K.K., Babu, K.S., Das, S.K., and Acharaya, A. 2019. Time Series-Based Air Pollution Forecasting using SARIMA and Prophet Model. Proceedings of the International Conference on Information Technology and Computer Communications, 16-18, 2019, Singapore, IEEE, NY, pp. 124-129.

San José, R., Pérez, J.L., Pérez, L. and Barras, R.M.G. 2019. Global climate-driven effects on urban air pollution simulations using very high spatial resolution. Int. J. Environ. Pollut., 66(1): 143-161.

Santos, J.A., Jiménez, M. and Espinosa, F. 2020. Effect of event-based sensing on IoT node power efficiency: Case study: Air quality monitoring in smart cities. IEEE Access, 7: 132577-132586.

Savita, V.M., Mohurle, R.P. and Manisha, P. 2018. A study of fuzzy clustering concept for measuring air pollution index. Int. J. Adv. Sci. Res., 3(1): 43-45.

Singh, T., Narasimhan, T.L. and Lakshminarayanan, C.S. 2020. Deep Air: Air quality prediction using deep neural network. TENCON IEEE Region Conference (TENCON), 16-19 Nov 2020, 124-129.

Stellwagen, E. and Tashman, L. 2013. ARIMA: The models of box and Jenkins. Int. J. Appl. Forecast, 16: 28-33.

Xiang, L., Peng, L., Hu, Y., Shao, J. and Chi T. 2016. Deep learning architecture for air quality predictions. Environ. Sci. Pollut. Res. Int., 23(22): 22408-22417.

Zhu, C., Cai, T., Yang, C. and Zhou, X. 2018. A machine learning approach for air quality prediction: model regularization and optimization. Big Data Cognit. Comp., 2(1): 5-10.

Zhu, D., Changjie, C., Tianbao, Y. and Xun, Z. 2018. A machine learning approach for air quality prediction: Model regularization and optimization. Big Data Cognit. Comput, 11: 1-5