

Comparative Analysis of Air Quality in Major Cities vs. Minor Cities in India (2022-2025)

Wesley Wende¹, Vongaiishe Moyo², Dr Shiv Kumar³

¹Sharda University, School of Engineering and Technology

²Sharda University, School of Engineering and Technology

³Sharda University, School of Engineering and Technology

Abstract - Air quality assessment and forecasting is a vital field that connects environmental science, data engineering, and artificial intelligence, particularly due to the growing health risks posed by pollution-related respiratory and cardiovascular diseases. To address this challenge, comprehensive frameworks are required for monitoring and predicting pollutant concentrations, which in turn support regulatory agencies in enforcing standards like the Air Quality Index (AQI) and shaping effective mitigation policies. This study draws on diverse data sources such as ground-based monitoring stations and satellite observations applying preprocessing techniques to handle missing data, fix errors, and spatial-temporal inconsistencies. Advanced feature engineering highlights temporal trends, spatial variations, and pollutant-weather interactions, while methods like time-series decomposition, spatial interpolation, and correlation analysis help explain variability. Predictive modelling combines traditional statistical tools (e.g., ARIMA, Prophet) and machine learning approaches (e.g., Logistic Regression). Additionally, ensemble approaches that integrate data-driven insights with physical dispersion models improve accuracy, and model performance is validated using statistical measures such as mean absolute error, root mean squared error, R^2 , and AQI category hit rates to ensure both precision and practical interpretability.

Key Words: Air Quality Index (AQI), ARIMA, Logic Regression, Prophet

1. INTRODUCTION (Size 11, Times New roman)

India is amongst the developing countries category in the world, making its air quality have vigorous impact on human nature and lifestyle, agriculture, variations in climate and overall changes in the ecosystem [20]. Over the previous 20 years, the number of deaths from the modern forms of air and the toxic pollution have incremented by 66%, mostly affecting the population of the elderly and children under 5 years [3, 8]. The contamination of the atmosphere is caused by dangerous materials like gases (NO_2 , SO_2 , CO), particulate matter ($\text{PM}_{2.5}$, PM_{10}), and biological molecules, and also other factors like deforestation, industrial discharges, fossil fuels, and vehicle emissions [5, 3]. This pollution is truer and more effective in big cities, which not only have large populations but also find it difficult to balance the air quality and the environment, and the same is true for less polluted cities, just with a difference in the impact [4, 20]. Outdoor pollution had been classified as a group 1

carcinogen by the International Agency for Research in Cancer (IARC) [14]. Also, the National Clean Air Program [NCAP] in 2019 was launched for the same concern; it particularly targets the reduction in the PM_{10} and $\text{PM}_{2.5}$ [7, 13]. All these indicate how bad the effects are and why the concerns are relative. The Air Quality Index (AQI) is a tool created to reflect the public point of view of AQI and its health-associated risks. It quantifies the air quality levels from different locations in order to determine the effect of air pollution, whether “good” or “bad,” clearly and points out which areas need the increase of air quality monitoring stations to allow more coverage [8, 10, 20].

The difference in the air quality level from city to city is contributed by different factors like population, productivity, and socioeconomic status. A lot of studies, therefore, have been done on the in-depth analysis of AQI but focus more on the populated cities and the reasons causing the increase in AQI levels. While previous studies have been done on this topic of Air Quality and how it has affected the citizens of India economically, socially and physically, a lot of research focus on the Air Pollution of the major cities using different data sets available and the techniques and Artificial Intelligence. However, limited research has explored that less major cities are not studied often. Therefore the study aim to investigate the air quality in other cities and compare with the major ones, to bring an insight of what could be the cause. Will achieve this by analysing existing dataset from AQI.india and use the preprocessing methods to clean the data, visualise it, and analyse it compared to the major cities. We also going to use models to predict AQI of the future. Aim is to contribute more on the topic of Air Quality from both angles, the major cities and the minor ones and then predict the AQI for the future. The following how our paper is structured.

In section 1, we have the introduction of the topic, history of how air quality in India has progressed from past decades, what pollutants contributes to the air quality, the effects of the rise and how it has been dealt with in terms of policies and laws. Section 2 is about the Literature review, we have studied 20 past papers, related to the subject and summarised only 10 papers to get the insight and the methods they have used and found and concluded on the research gap. Our dataset of Air Quality

Analysis, we got it from Kaggle, an open source for data sets and that is what we used in section 2.1, first we give a background Minor cities and Major cities then we used preprocessing techniques like encoding, standardisation, normalisation to analyse the data and get the results to compare the air quality between the two categories. Section 3 is all the methodologies we have used and how we have analysed the data to reach conclusion. We went on to create a pseudo code for the prediction of the air quality using the Prophet model. In this research we have taken an angle of comparing the

pollutants in Major cities vs Minor cities, the prophet model we used is to predict the air quality of the future, with the aim to make the public aware of the impacts of the air quality in their life and how it can be predicted for the future.

2. Literature Review

Title	Author(s)	Year	Source /Journal	Aim	Methodology	Key findings
1. A framework for city-specific air quality health index: a comparative assessment of Delhi and Varanasi, India	Franciosalge o George, Pallavi Joshi, Sagnik Dey, R K Mall and Santu Ghosh	2025	Environment Research Letters	Propose a framework for developing city-specific AQHI that reflects local air quality	Generalised Additive Models (GAMs) Segmented Regression	Misclassification of air quality days was reduced using localized indices
2. AI-Based Predictive Modeling For Air Quality Assessment And Environmental Risk Forecasting In Urban Ecosystems	Ashok Kumar Panda, Sonali Pradhan, Chinmayee Pati, Naba Kumar Rath, Deepak Kumar Baral	2025	International Journal of Environmental Sciences	Create an AI-based framework for predictive modeling aimed at assessing air quality and anticipating environmental risks in urban ecosystems	Data preprocessing techniques (normalization, outlier detection)	AI-driven models outperform classical statistical approaches in predicting pollutants
3. Assessing AQI of air pollution crisis 2024 in Delhi:	Abhranil Bhuyan, Tapoban Bordoloi, Ra	2024	Discover Atmosphere	Address the difficulties that	Review of CPCB reports, peer-reviewed literature, and	Delhi's AQI frequently exceeded 400 ("Severe"), especially in winter.

its health risks and nationwide impact	bin Debnath, Abu Md Ashif Ikbai, Bikash Debnath, Waikhom Somraj Singh			comes with pollution and ensuring of economic productivity, and improvement of the quality of life in the National Capital Region.	real-time AQI data.	
4.Optimized machine learning model for air quality index prediction in major cities in India	Suresh Kumar Natarajan, Prakash Shanmurthy , Daniel Arockiam , Balamurugan Balusamy, Shitharth Selvarajan	2024	Scientific Reports	Major cities AQI analysis is essential for the government to take proper preventive, proactive measures to reduce air pollution.	Grey Wolf Optimization (GWO) Decision Tree (DT) algorithm	The proposed model attained better prediction performance compared to traditional machine learning algorithms with maximum accuracy
5. Analysis of spatiotemporal distribution of air quality index (AQI) in the state of West Bengal, India from 2016 to 2021	Buddhadev Ghosh, Harish Chandra Barman, Pratap Kumar Padhy	2023	Discover Atmosphere	Evaluate the AQI status from a different point of view and estimate the future	Evaluation of AQI across 22 districts in West Bengal (2016–2021)	AQI ranged from “Satisfactory” to “Moderate,” highest in winter Temperature, RH, and precipitation were negatively correlated with AQI.

				AQI.		
6.An environmental justice analysis of air pollution in India	Priyanka N. deSouza , Ekta Chaudhary , Sagnik Dey, Soohyeon Ko, Jeremy Németh , SarathGuttikunda, Sourangsu Chowdhury , Patrick Kinney, S.V. Subramanian , Michelle L. Bell , Rockli Kim	2023	Scientific Reports	Evaluate associations between total, anthropogenic and source-specific PM2.5 exposures and SES variables .	Socioeconomic status(SES) and demographics Regression models	Socioeconomic factors and access to various household amenities are risk factors of PM2.5 exposures
7.A Study during Lockdown Period Based on AQI over Indian Mega cities during COVID-19	M Chakraborty , S Debnath and S Ghosh	2021	Journal of Physics: Conference Series	Assess the impact of the stay in,lockdown during COVID-19 in relation to air pollution	Comparative analysis of AQI and pollutant levels (PM2.5, SO2, NO2, CO, and O3) Data visualization using Python Data segmentation	Significant reduction in PM2.5, NO2, and CO across all cities during lockdown.
8.Analysis of Air Pollution Data in India between 2015 and 2019	Disha Sharma, Denise Mauzerall	2021	Aerosol and Air Quality Research	Address inconsistencies and data gaps in datasets using a	Data Quality Control Case studies Criteria pollutant data	Particulate pollution dominates pollution, across India

				rigorous procedure to ensure data representa tiveness		
9.Evolution of air pollution management policies and related research in India	Sunil Gulia , Nidhi Shukla , Lavanya Padhi, Parthaa Bosub , S.K. Goyal , Rakesh Kumar	2021	Elsevier B.V	Provide informati on about the evolution of air quality managem ent policies in India .	Geographical Information systems(GIS)	Highly polluted entries are the most studied foresaking the less polluted cities
10.Monitoring particulate matter in India: recent trends and future outlook	Pallavi Pant, Raj M. Lal,Sarath K. Guttikunda, Armistead G. Russell, Ajay S. Nagpure ,Anu Ramaswami, Richard E. Peltier	2018	Springer B.V	Provide the regulatory monitorin g landscape in India, on measurem ent methods on pollution	Comparisongs of multiple sources of Air Quality Data	Lack of centralised Database on AQ makes it hard to utilize the data for long-term trends analysis

2.1 Minor Cities vs Major Cities

Minor Cities

Background

India is amongst the top 10 largest country in the world by area, to be precise it is positioned the seventh. That being said its size also explain its largest populations , it has more than 40 cities with more than a million people and 396 cities with between 100 ,000 and 1 million peoople and 2500 cities with between 10 000 and 100,000 people.(World Population Review,2025). Our researc h is based from the Air Quality of India dataset from Kaggle region by region.

First we selected the least populated cities of India that have approximately statistics from 10 000 and 1 million using the guideline of the World Population Review ,2025.Next step was the analyses of th air quality ,in these regions and what could be the probable pollutants contributing. It was a discovery that there are numerous pollutants which contribute to the Air Quality for it it be measurerd very severe to satisfactory from a scale of 0-5 .The main pollutants being PM2.5,PM10,S02,O3,NO2,CO,NH3 and then the various combinations which come out from theses main on

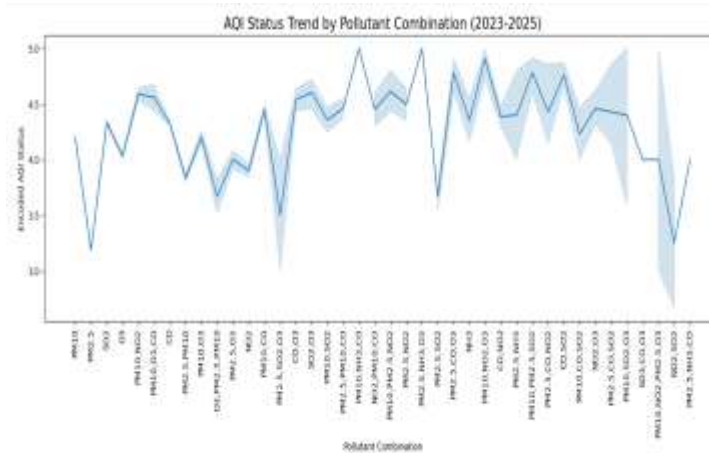


Fig-1: Pollutant combinatins and the corresponding AQI Status result

Major Cities

These are cities with the highest populations in Indi populations that is Mumbai, Delhi, Bengaluru,Kolkata,Chennai with numbers from approximately a .Statistically wise ,they are 5 main big cities with the highest4 million and above in each state(World Population Review-2025) and that are the ones we are focusing on..The same main pollutantPM2.5,PM10,S02,O3,NO2,CO,NH3 and their various combinations which come out from them were analysed for equal analysis between the variations.

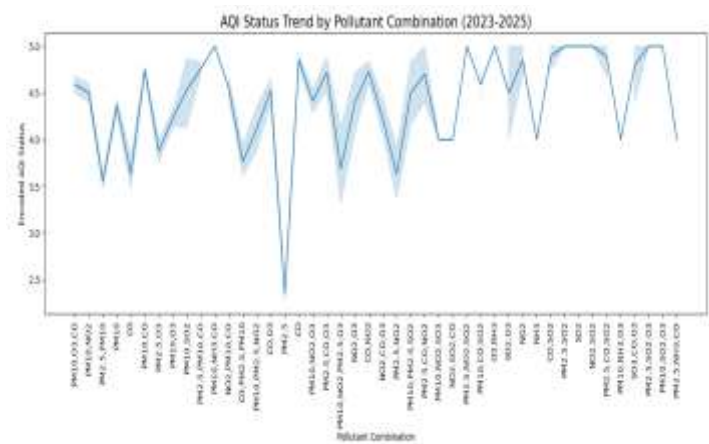


Fig -2: Pollutant combinatins and the corresponding AQI Status result

Table -1: Comparison Table for Major and Minor Cities

Aspect	Major cities	Minor cities	Interpretati on
Frequent Pollutant Combinatio ns	PM2.5, NO2, SO2	PM2.5, S02,O3	Across both cities, PM2.5 , SO2 are common

			though it could be from different soources
Average AQI status	Slightly worse air quality (3.5)	Slightlyl better air quality(3.8)	Minor cities show slightly better air quality though both are in moderate category
Most Hazardous combo	PM2.5→A QI 2.3	NO2, SO2→ AQI 3.2	Lower values result in more toxic combinatio ns
Least Hazardous combo	O3, NH3→AQI 5.0	PM2.5, NH3,CO→A QI 5.0	Higher values mean cleaner combinatio ns
Trend over time	Flattening and Declining	Fluctuating	Major cities remain poor and minor cities have instability

3. Methodology

Data Acquisition and Cleaning

Essential libraries: pandas, numpy, matplotlib.pyplot, and seaborn were imported to allow operations and manipulation of the data. To read and work on the data, it was loaded from CSV file into pandas DataFrame.Columns like (unit and note) were removed from the dataset as they were of lesser importance and so as to reduce the dataset from the noisy data.

Temporal Feature Engineering

To work with the preffered date formats, the year was extracted into new columns for time- based analysis.

Data Reshaping and Scaling

Dataset was pivoted to create a multi-index structure (year, state, area) with pollutants as columns and Standardisation was performed to reduce the huge gap of values in the dataset and to normalise the pollutant values across the regions.

Visualization and Exploratory Analysis

For visualization purposes, Heatmaps were created to analyze pollutant distributions and for filtering data by state and year (e.g., Uttar Pradesh, 2022). Histograms were specifically created to explore AQI distributions by area, pollutant, and air quality status.

Categorical Encoding

To convert categorical variables (state, area) into numerical format, LabelEncoder was used so as to unify the data into a type which was suit for operations and manipulations.

Machine Learning Preparation: To prepare the data for the model, pollutant features were isolated by excluding identifiers (year, state, area, and encoded labels) so as to separate between the data to be trained and the one to be tested. The train test split used at the end to train the model was (80% training, 20% testing).

Model Training and Evaluation

To assess regional air quality patterns, a supervised classification model was implemented. A Logistic Regression classifier was initialized with an increased iteration threshold () to ensure convergence. The model was trained to predict the encoded state labels based on pollutant features. Performance was evaluated using two metrics:

- Root Mean Squared Error (RMSE): Quantifies the average magnitude of prediction errors.
- Coefficient of Determination (R^2 Score): Measures the proportion of variance in the target variable explained by the model.

Time Series Forecasting with Prophet

To analyze temporal trends in air quality, the dataset was aggregated at a monthly level. Columns were renamed to align with Prophet's input format (for date, for AQI value). A Prophet model was trained with yearly seasonality enabled to capture recurring patterns. The model generated 12-month forecasts, which were visualized alongside historical data, including confidence intervals to reflect uncertainty.

Area-Level Forecasting

For granular insights, AQI data was further aggregated monthly by area. Separate Prophet models were trained for each area with sufficient historical data. Forecasts were generated and merged with past observations to produce individualized AQI trend visualizations. This approach enabled localized forecasting and comparative analysis

Pseudo code for Prediction

Supervised Learning for State Classification

```
CREATE pivot table "aqi_pivot" from "aqi_data"
GROUP BY year, state, area AS ROW INDEX
COLUMNS: prominent_pollutants, air_quality_status
VALUES: aqi_value
DISPLAY aqi_pivot
```

```
IF aqi_pivot IS NOT NULL THEN
    RESET the index of aqi_pivot → aqi_pivot_reset
```

```
INITIALIZE label encoders: state_le, area_le
```

```
ENCODE "state" column → state_encoded
    USING state_le.fit_transform()
```

```
ENCODE "area" column → area_encoded
    USING area_le.fit_transform()
```

```
EXTRACT all column names from aqi_pivot_reset →
features_list
```

```
DEFINE columns_to_exclude = [year, state, area,
state_encoded, area_encoded]
```

```
FILTER features_list to EXCLUDE columns_to_exclude
RESULT → feature_columns
```

```
INITIALIZE StandardScaler → scaler
    (Note: scaling will be applied later after train-test split)
```

```
DISPLAY aqi_pivot_reset
```

```
ELSE
    PRINT "Error: aqi_pivot is None, cannot apply encoding."
END IF
```

```
END
```

Pseudo code for Model Training:

```
IMPORT train_test_split from model_selection module
SPLIT the dataset into training and testing sets:
```

```
    INPUT: Features X, Target y
```

```
    PARAMETERS: test_size = 20%, random_state = 42
```

```
    OUTPUT: X_train, X_test, y_train, y_test
```

```
IMPORT LogisticRegression from linear_model module
```

```
INITIALIZE logistic regression model:
    model = LogisticRegression(max_iterations = 10000)
```

```
TRAIN model using training data:
    CALL model.fit(X_train, y_train)
```

```
IMPORT evaluation metrics: mean_squared_error, r2_score
```

```
DISPLAY model (indicating successful initialization)
```

```
END
```

Model Evaluation

```
IMPORT mean_squared_error and r2_score
```

```
CHECK if model exists AND can make predictions:
```

```
    IF model does not exist OR does not have a predict function
    THEN
```

```
        PRINT "Model not found or not properly initialized."
```

```
        INITIALIZE model as LogisticRegression with
max_iterations = 1000
```

```
    END IF
```

CHECK if model is fitted:

```
IF model does not have coefficients (coef_) THEN
  PRINT "Model is not fitted. Fitting the model now."
  FIT model using X_train and y_train
END IF
```

PREDICT target values for test dataset:

```
y_pred = model.predict(X_test)
```

CALCULATE MSE = mean_squared_error(y_test, y_pred)

CALCULATE RMSE = square root of MSE

PRINT "RMSE:", RMSE

PRINT "R² Score:", r₂_score(y_test, y_pred)

END

Future State Prediction

To simulate future air quality scenarios, a hypothetical dataset comprising projected pollutant concentrations was constructed. These synthetic inputs were preprocessed using the previously trained StandardScaler to ensure consistency with the model's training distribution. The scaled data was then passed to the trained Logistic Regression classifier to predict the most probable encoded state label based on pollutant profiles. Finally, the predicted labels were decoded to their corresponding state names using the fitted LabelEncoder, enabling interpretability of the model's output.

4.Results and Discussion

The results from this analysis show the trend of air over time; this takes historical data and predicts the future outcome. It also shows which pollutants are active in a state or an area along with their AQI values. It is clear from the analysis that the major cities have higher AQI status because of the large populations, therefore there is increase in activities which affect the air quality, like industrial gases and wastes, street food stalls, stubble burning and many more. However, it is not an exception with the Minor cities, they have a concerning AQI status but it is lesser than the major cities, its contributions are different, it comes from like stubble burnig, house emmissions, unregulated industries and more other. After the analysis we can conclude that PM_{2.5}, PM₁₀, SO₂, NO₂, O₃, are the most dangerous contributions to this air quality, it makes it worse when they are combined with other pollutants. We then went on to figuring out, how we can predict the air quality status, using the different models, this helps in educating the public on the environment and the effects of it if not handled well. It also can help the decision making process when making mitigating policies, and it can help scholars and students to learn more on this topic and the method used to draw the conclusions.

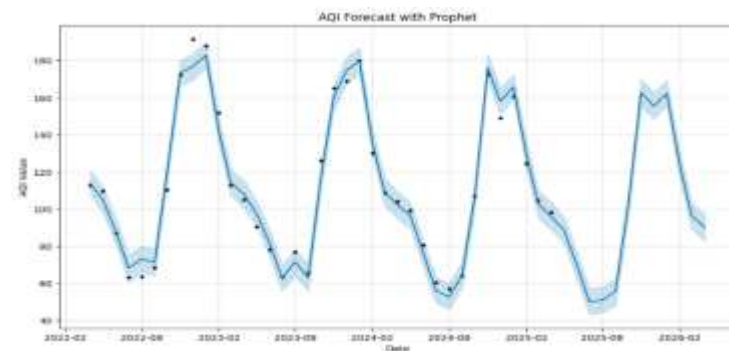


Fig -3: Shows a time series forecast of AQI values from 2022 up to 2026 (future)

Limitations:

Prophet can be quite sensitive to unusual values in the data. Sometimes, these aren't actual errors but simply periods of especially good or bad air quality, and they can throw off how well the model captures the true trends and seasonal patterns. Secondly, it assumes the seasonal patterns stay the same every year, but in reality, things like climate change or new policies can shift those patterns over time. The model might not always pick up on these changes. While Prophet lets us add extra factors (like weather data or events), it may still struggle to capture the full complexity of how different pollutants and outside influences interact to affect air quality. Lastly, in this analysis, we looked at the combined AQI values rather than focusing on individual pollutants. Since Prophet doesn't automatically consider the unique patterns of each pollutant, important differences can get overlooked.

Future directions

To improve our forecasts, we can take a few practical steps. First, it helps to spot and deal with unusual AQI values before modeling; sometimes these are genuine spikes, but sometimes they're just errors. Next, rather than assuming the seasons always follow the same pattern, we could use models that adapt to shifting trends, like those influenced by weather changes. Trying out other forecasting approaches, such as ARIMA or machine learning models like LSTMs, might capture complex patterns better. We can also add more information like weather details, major events, or even individual pollutant levels to help the model make smarter predictions. Instead of only looking at overall AQI, analyzing each key pollutant separately could give us deeper insights. Combining forecasts from different models and regularly checking how well our approach is working ensures we keep our predictions accurate and up-to-date.

5. CONCLUSIONS

This study presents a comprehensive comparative analysis of air quality across major and minor cities in India from 2022 to 2025, which is not a common study among literatures as there is more focus on the major cities. It highlights the spatial and temporal dynamics of pollutant concentrations. By integrating statistical, machine learning, and deep learning models with robust preprocessing and feature engineering, the research reveals that while major cities consistently exhibit poorer air quality due to dense populations and industrial activity, minor cities show fluctuating but generally better AQI levels. These findings underscore the importance of inclusive monitoring and forecasting frameworks that account for both urban and semi-urban environments. This dual-city perspective offers actionable insights for policymakers, enabling targeted interventions and equitable air quality management across diverse regions. While the findings offer valuable insights into spatial and temporal AQI dynamics, the study is subject to limitations related to data availability, model generalizability, and pollutant scope. Future research could expand the temporal range, incorporate additional environmental variables, and refine city classifications to enhance predictive accuracy and policy relevance.

REFERENCES

- Panda, A. K., Pradhan, S., Pati, C., Rath, N. K., & Baral, D. K. (2025). AI-Based Predictive Modeling for Air Quality Assessment and Environmental Risk Forecasting in Urban Ecosystems. *International Journal of Environmental Sciences*, 11(12s), 163-171.
- Jayant, J., Tiwari, A. C., & Jayant, A. B. (2025). Integrating winter's method and white noise for advanced air quality forecasting: a case study of Indore, India. *Environmental Research Communications*, 7(3), 035031.
- George, F., Joshi, P., Dey, S., Mall, R. K., & Ghosh, S. (2025). A framework for city-specific air quality health index: a comparative assessment of Delhi and Varanasi, India *Environmental Research Letters*, 20(8), 084067.
- Natarajan, S. K., Shanmuthy, P., Arockiam, D., Balusamy, B., & Selvarajan, S. (2024). Optimized machine learning model for air quality index prediction in major cities in India. *Scientific reports*, 14(1), 6795.
- Bhuyan, A., Bordoloi, T., Debnath, R., Iqbal, A. M. A., Debnath, B., & Singh, W. S. (2025). Assessing AQI of air pollution crisis 2024 in Delhi: its health risks and nationwide impact. *Discover Atmosphere*, 3(1), 13.
- Ravindiran, G., Hayder, G., Kanagarathinam, K., Alagumalai, A., & Sonne, C. (2023). Air quality prediction by machine learning models: A predictive study on the Indian coastal city of Visakhapatnam. *Chemosphere*, 338, 139518.
- Guttikunda, Sarath K and Dammalapati, Sai Krishna, Air Quality Index (AQI) Analysis for India: 2015-2023 (February 01, 2024).
- Ghosh, B., Barman, H. C., & Padhy, P. K. (2023). Analysis of spatiotemporal distribution of air quality index (AQI) in the state of West Bengal, India from 2016 to 2021. *Discover Atmosphere*, 1(1),
- deSouza, P. N., Chaudhary, E., Dey, S., Ko, S., Németh, J., Guttikunda, S., ... & Kim, R. (2023). An environmental justice analysis of air pollution in India. *Scientific reports*, 13(1), 16690.
- Gupta, N. S., Mohta, Y., Heda, K., Armaan, R., Valarmathi, B., & Arulkumaran, G. (2023). Prediction of air quality index using machine learning techniques: a comparative analysis. *Journal of Environmental and Public Health*, 2023(1), 4916267.
- Sharma, M., Jain, S., Mittal, S., & Sheikh, T. H. (2021). Forecasting and prediction of air pollutant concentrations using machine learning techniques: the case of India. In *IOP conference series: Materials science and engineering* (Vol. 1022, No. 1, p. 012123). IOP Publishing.
- Gangwar, A., Singh, S., Mishra, R., & Prakash, S. (2023). The state-of-the-art in air pollution monitoring and forecasting systems using IoT, big data, and machine learning. *Wireless Personal Communications*, 130(3), 1699-1729.
- Sharma, D., & Mauzerall, D. (2022). Analysis of air pollution data in India between 2015 and 2019. *Aerosol and Air Quality Research*, 22(2), 210204.
- Gulia, S., Shukla, N., Padhi, L., Bosu, P., Goyal, S. K., & Kumar, R. (2022). Evolution of air pollution management policies and related research in India. *Environmental Challenges*, 6, 100431.
- Maji, S., Ahmed, S., Ghosh, S., & Garg, S. K. (2020). Evaluation of air quality index for air quality data interpretation in Delhi, India. *Current Science*, 119(6), 1019-1026.
- Chakraborty, M., Debnath, S., & Ghosh, S. (2021, February). A study during lockdown period based on AQI over Indian mega cities during COVID-19. In *Journal of Physics: Conference Series* (Vol. 1797, No. 1, p. 012056). IOP Publishing.
- Sharma, R., Kumar, R., Sharma, D. K., Son, L. H., Priyadarshini, I., Pham, B. T., ... & Rai, S. (2019). Inferring air pollution from air quality index by different geographical areas: case study in India. *Air Quality, Atmosphere & Health*, 12(11), 1347-1357.
- Gordon, T., Balakrishnan, K., Dey, S., Rajagopalan, S., Thornburg, J., Thurston, G., ... & Nadadur, S. (2018). Air pollution health research priorities for India: Perspectives of the Indo-US Communities of Researchers. *Environment international*, 119, 100.
- Pant, P., Lal, R. M., Guttikunda, S. K., Russell, A. G., Nagpure, A. S., Ramaswami, A., & Peltier, R. E. (2019). Monitoring particulate matter in India: recent trends and future outlook. *Air Quality, Atmosphere & Health*, 12(1), 45-58.
- Nasir, H., Goyal, K., & Prabhakar, D. (2016). Review of air quality monitoring: case study of India. *Indian Journal of science and technology*, 9(44), 1-8.
- World Population Review. (2025). Population of cities in India. World Population Review. <https://worldpopulationreview.com/cities/india>
- India Air Quality Index (AQI) Dataset [2023–2025]