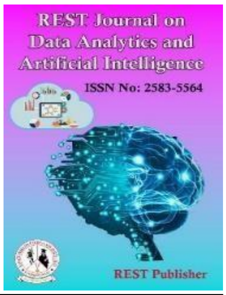
REST Journal on Data Analytics and Artificial Intelligence
Vol: 3(3), September 2024
REST Publisher; ISSN: 2583-5564
Website: http: //restpublisher.com/journals/jdaai/
DOI: https://doi.org/10.46632/jdaai/3/3/2



Assessing Statistical Models for Predictive Accuracy of $PM_{2.5}$ Pollution in Delhi, India

[*1] Divyansh Sharma, [2] Sapan Thapar, [1]Kamna Sachdeva
[1] TERI School of Advanced Studies, New Delhi, India.
[2] Delhi Skill and Entrepreneurship University, Dwarka, New Delhi, India.
**Corresponding author: divyansh.sharma@terisas.ac.in*

***Abstract**: Particulate matter is a significant atmospheric pollutant that poses substantial health risks. Reliable and precise air quality forecasts are essential for the timely implementation of preventive measures to minimize these health risks. This study examines the effectiveness of various statistical methods in forecasting long-term trends of particulate matter ($PM_{2.5}$) pollution. Using historical data from government-operated monitoring stations in Delhi, the research applies a range of time-series analysis techniques to identify patterns and predict future pollution levels. The analysis reveals that the Seasonal Autoregressive Integrated Moving Average model with exogenous variables (SARIMAX) significantly outperforms other models, such as ARIMA, SARIMA, and ARIMA with exogenous variables (ARIMAX). The exceptional performance of SARIMAX demonstrates its potential as a robust early warning system, which can facilitate the implementation of preventive measures to mitigate the impact of pollution on public health. This emphasizes the model's significance in supporting proactive environmental and health policy strategies.*

***Keywords:** Particulate matter; $PM_{2.5}$; Forecasting; SARIMA; ARIMA; SARIMAX; ARIMAX*

1. INTRODUCTION

The rapid urbanization and development of cities pose significant challenges, leading to population surges and inadequate public services. This uncontrolled urban expansion in developing countries pressures natural resources, resulting in environmental degradation at multiple levels [1]. Among global environmental concerns, ambient (outdoor) air pollution is notably exacerbated by these unsustainable practices. Poor ambient air quality, due to increased levels of pollutants, is primarily attributed to anthropogenic emissions from vehicles, industry, construction, and domestic burning, though natural emissions also contribute [2], [3]. These sources release particulate matter (PM10, PM2.5) and gaseous pollutants like nitrogen and sulfur oxides, carbon monoxide, and ozone [4]. The concentrations of these pollutants are influenced by synoptic patterns and meteorological parameters [5].

Concerns about poor air quality in India have surged recently due to its adverse effects on human health, agricultural productivity, and the economy [6], [7]. India's rapid urban and industrial development has led to some of the world's most polluted air. With 34% of its 1.3 billion population living in urban areas, air quality is deteriorating rapidly [8], [9]. A study by [10] highlighted that over 70% of India's population is exposed to particulate concentrations exceeding National Ambient Air Quality Standards (NAAQS), contributing to increased mortality and morbidity, with approximately 0.7 million deaths annually linked to ambient air pollution. The economic impact of poor air quality accounts for about 1.4% of the GDP, including expenditures on health issues related to pollution [11].

Many metropolitan and Tier-I cities in India, such as Delhi, Kolkata, and Mumbai, are experiencing worsening air quality [12]. This deterioration is due to various factors, including vehicle exhaust emissions (VEEs), resuspended dust, biomass burning, and industrial pollution [13], [14]. Consequently, these cities are struggling to meet health-based air quality standards. Regularly examining changes in air pollutant concentrations is essential to develop reliable solutions to mitigate and assess the health and environmental risks associated with poor air quality in urban areas [15]. Predicting and forecasting air pollution is crucial not only for enabling residents to plan their daily activities and avoid high-pollution areas but also for supporting urban planning and public health strategies [16]. Time series modeling is pivotal in this process, offering benefits such as data cleaning, understanding, and forecasting. As a vital quantitative technique, time series forecasting involves collecting and analyzing historical data to develop models that predict future scenarios [17]. Given its significance, research on pollution forecasting has become a critical area in environmental protection, aiming to evaluate and implement necessary measures to mitigate the long-term effects of pollution [18].

This research addresses significant gaps in predicting and forecasting $PM_{2.5}$ levels in Delhi by conducting a comparative analysis of traditional (ARIMA & SARIMA) and advanced (ARIMAX & SARIMAX) time series models. The study explores spatio-temporal trends in $PM_{2.5}$ data from 39 monitoring stations (2019-2023). It evaluates model accuracy with metrics like RMSE and R-squared, integrating a broad range of exogenous variables such as temperature, wind speed, rainfall, solar radiation, barometric pressure, and humidity. This comprehensive approach aims to improve predictive accuracy and provide actionable insights for policymakers and urban planners, enhancing air quality forecasting in megacities prone to pollution.

2. METHODS

A. Location Description: Delhi, situated at 28°34'N and 77°12'E, is India's capital city. As of the 2011 census, it had a population of 16.8 million and a growth rate of 1.92%, making it India's second-largest city by population. The city's climate is semi-arid and subtropical, with prevailing wind patterns mainly from the west and northwest. The average wind speed annually is between 0.9 to 2 m/s, according to [19]. Temperature-wise, Delhi has an average yearly temperature of 31.5°C. During the hot months from March to June, temperatures can soar up to 45°C. In contrast, temperatures significantly drop during the winter months of December and January, which helps in trapping emissions and raising pollution levels [20]. The monsoon season, typically from July to September, brings most of Delhi's rainfall, significantly influencing the city's annual weather patterns.

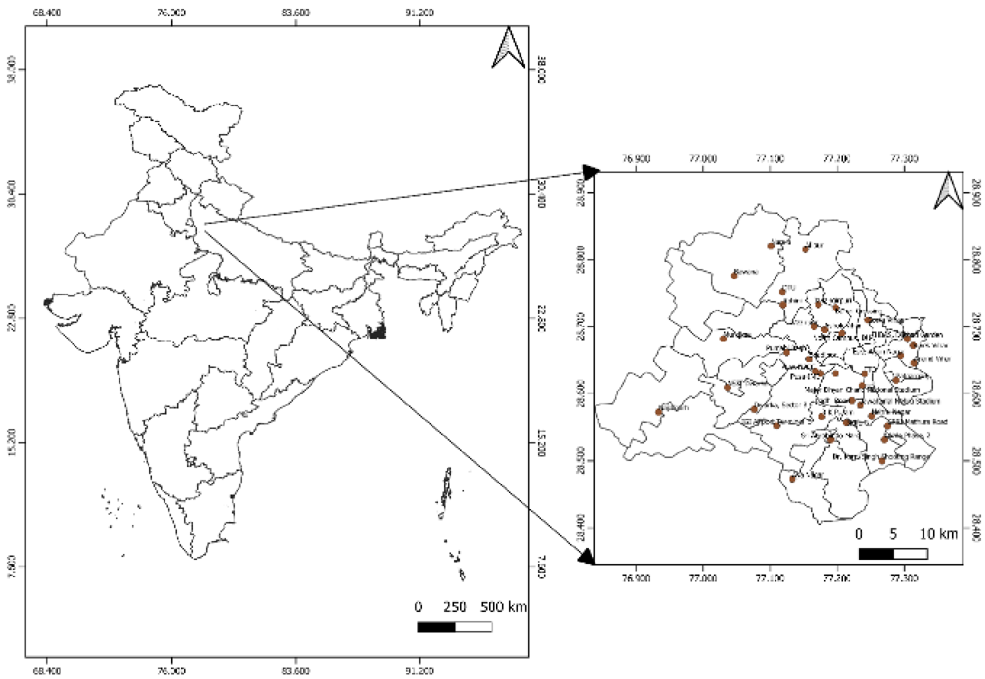


FIGURE 1. The location of Delhi city

B. Data Collection and Pre-Processing: Daily $PM_{2.5}$ concentration data were collected from the Central Pollution Control Board's digital repository (http://app.cpcbccr.com/ccr/#/login) for 39 monitoring stations across Delhi. This dataset covers five years, starting January 1, 2019, and ending December 31, 2023. $PM_{2.5}$ levels are measured using gravimetric, TEOM, or beta attenuation methods at stations managed by the CPCB, IMD, and DPCC. These measurements adhere to the [21] for ambient air pollutant measurement. Regular calibration of the monitors by the managing authorities ensures data reliability [3], [22].

Meteorological data, including daily measurements of temperature, humidity, wind speed, rainfall, solar radiation, and barometric pressure, were also sourced from the same CPCB repository for the same stations. These instruments are regularly calibrated according to guidelines to ensure data accuracy. Data pre-processing has been conducted to improve the quality and usability of this information.

To ensure robust analysis, we pre-processed air quality and meteorological data from 39 Delhi stations, spanning 2019 to 2023. We eliminated stations with less than 75% data completeness, interpolated isolated missing values, and applied mean imputation for continuous gaps using historical seasonal patterns. Meteorological data from nearby stations filled gaps, and monthly data aggregation allowed for effective trend analysis and model development

C. Data Analysis: We analysed $PM_{2.5}$ concentrations at various Delhi stations from 2019 to 2023 using both traditional and advanced time series models to adapt to the dynamic air quality data. Traditional models like ARIMA and SARIMA were used to identify data patterns, while ARIMAX and SARIMAX incorporated meteorological variables like temperature, wind speed, and humidity to assess external effects on $PM_{2.5}$ levels [23], [24].

Before model application, we confirmed data stationarity using the Augmented Dickey-Fuller (ADF) test, as recommended by [25]. This step ensured the data's mean, variance, and covariance were constant over time, a crucial factor since non-stationary data can lead to inaccurate time series analysis results.

Our study employed ARIMA and SARIMA models to predict $PM_{2.5}$ levels, adapting them to capture both the typical and seasonal variations in air quality data. These models were fine-tuned using the auto_arima function from the pmdarima library, optimizing parameter selection based on the Akaike Information Criterion (AIC) to ensure the best model fit [26].

ARIMA, widely used for forecasting stationary time series, combines three elements: autoregressive (AR), differencing (I), and moving average (MA). The AR component uses previous values to predict future ones with p as the number of lags. Differencing, denoted by d, stabilizes the series, while MA, represented by q, models the error using past forecast errors.

$$\left(1 - \sum_{i=1}^{p} \emptyset_i L^i\right)(1 - L)^d Y_t = \left(1 + \sum_{j=1}^{q} \theta_j L^j\right)\varepsilon_t$$

where ϕ_i and θ_j are the coefficients for the AR and MA parts, L is the lag operator, Yt is time series data at time t, and ϵ_t is the noise [27]. Other models incorporated in the study are extension of the ARIMA model which induces seasonality (SARIMA) as well as exogenous features (ARIMAX and SARIMAX)

To broaden our analysis, we integrated ARIMAX and SARIMAX models, which included meteorological factors as exogenous variables, enhancing the predictive accuracy by accounting for external environmental impacts. Parameter optimization was achieved using the auto_arima function, with additional tests on variable combinations to pinpoint those most impactful on model performance. Our comprehensive and methodical approach in fine-tuning the models allowed us to leverage time series analysis effectively, providing reliable forecasts of $PM_{2.5}$ concentrations. This helped us not only understand but also predict air quality trends, supporting urban air quality management strategies.

In the final analysis phase, we assessed the models' performance using R-squared and Root Mean Square Error (RMSE) across each station to identify the most accurate models. The best models were then applied to test datasets to validate their effectiveness in real-world scenarios, proving essential for enhancing air quality management [28], [29]

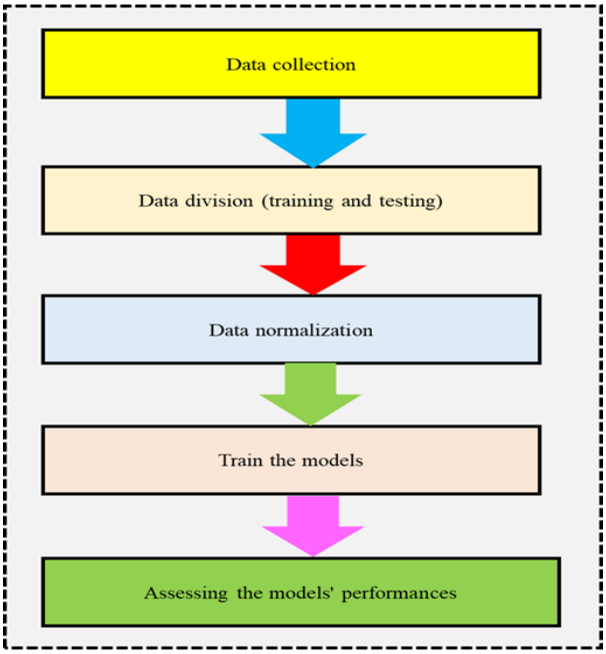


FIGURE 2. Model development steps.

3. RESULT AND DISCUSSION

We conducted an extensive comparative study of four time series models—ARIMA, ARIMAX, SARIMA, and SARIMAX—to predict $PM_{2.5}$ levels at the air quality monitoring network in Delhi from 2019 to 2023. Utilizing detailed datasets, we evaluate each model's effectiveness using the Root Mean Square Error (RMSE) and R-squared (R^2) metrics, which measure the accuracy of the predictions and the variance explained by the models, respectively.

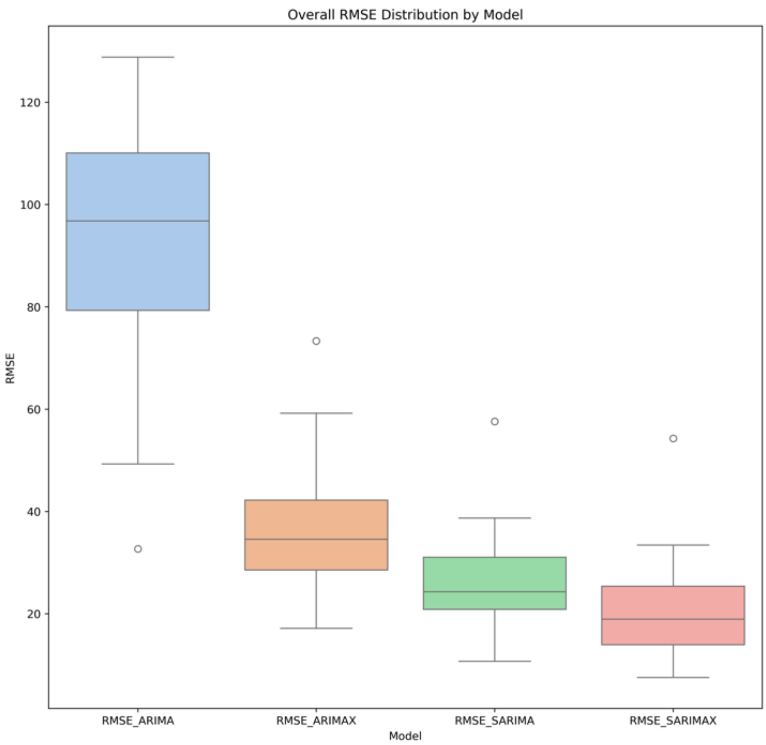


FIGURE 3. Overall RMSE of Time Series Models for PM2.5 at Delhi Monitoring Stations.

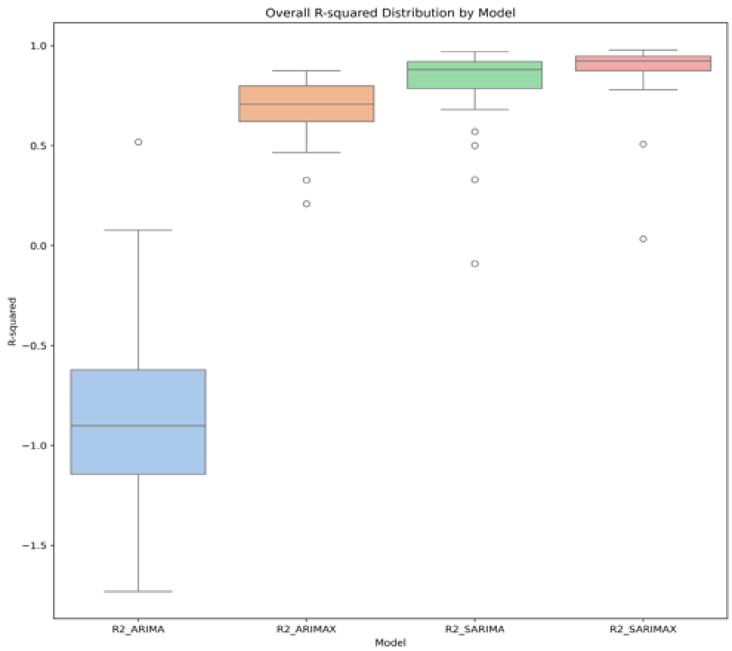


FIGURE 4. Overall R2_Score of Time Series Models for PM2.5 at Delhi Monitoring Stations

The distribution of RMSE and R² values for each model type, as visualized in the box plots (Figure 3 & 4), reveals significant insights into their predictive capabilities. The SARIMAX model generally exhibited the most favorable performance, achieving the lowest RMSE values across most stations, suggesting a robust fit to the data. Conversely, the ARIMA model frequently showed higher RMSE values, indicating a less accurate fit. This pattern was consistent with the R-squared values, where SARIMAX models also tended to explain a higher variance, demonstrating superior predictive power compared to the ARIMA models.

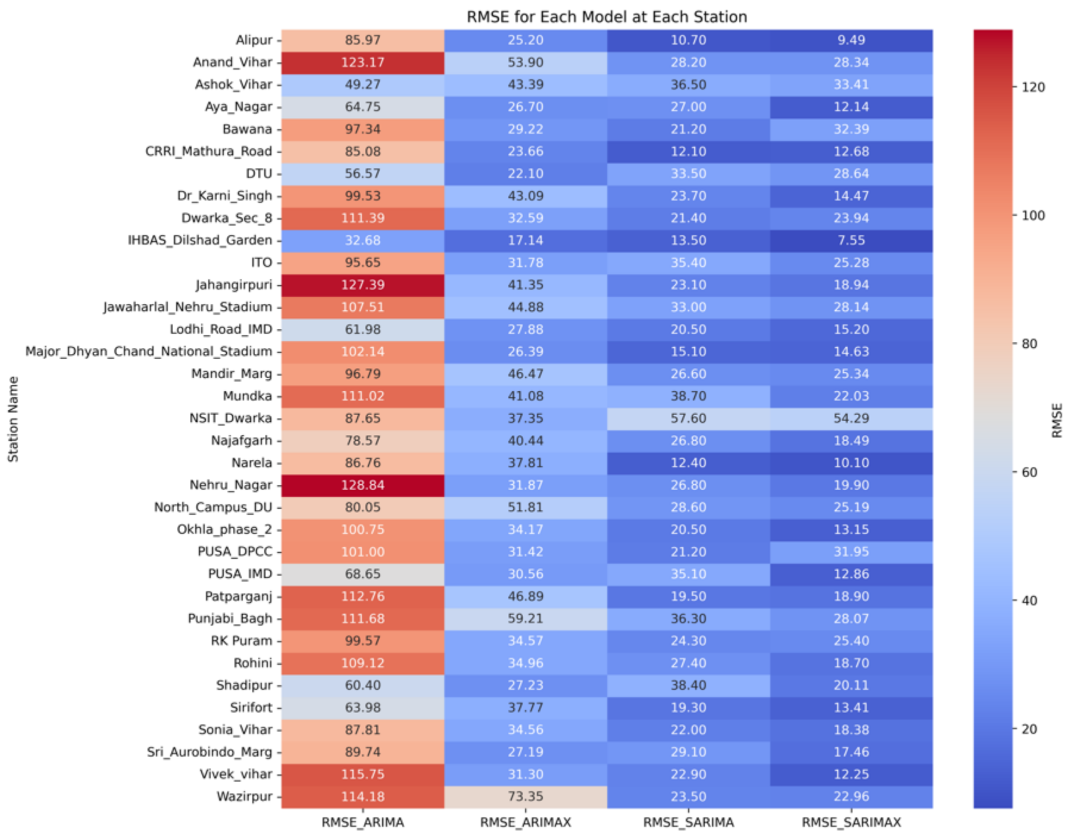


FIGURE 5. RMSE Scores for Time Series Models at Each Station

Detailed examination of model performance at each station further highlighted the consistency of SARIMAX models in achieving lower RMSE and higher R^2 values. For instance, at the Alipur station, the SARIMAX model produced an RMSE of 9.49 and an R^2 of 0.98, markedly outperforming other models. Similarly, at the Vivek Vihar, which is critical given its traffic density, the SARIMAX model not only achieved the RMSE of 12.25 but also managed an R^2 of 0.97, underscoring its effectiveness in high pollution areas.

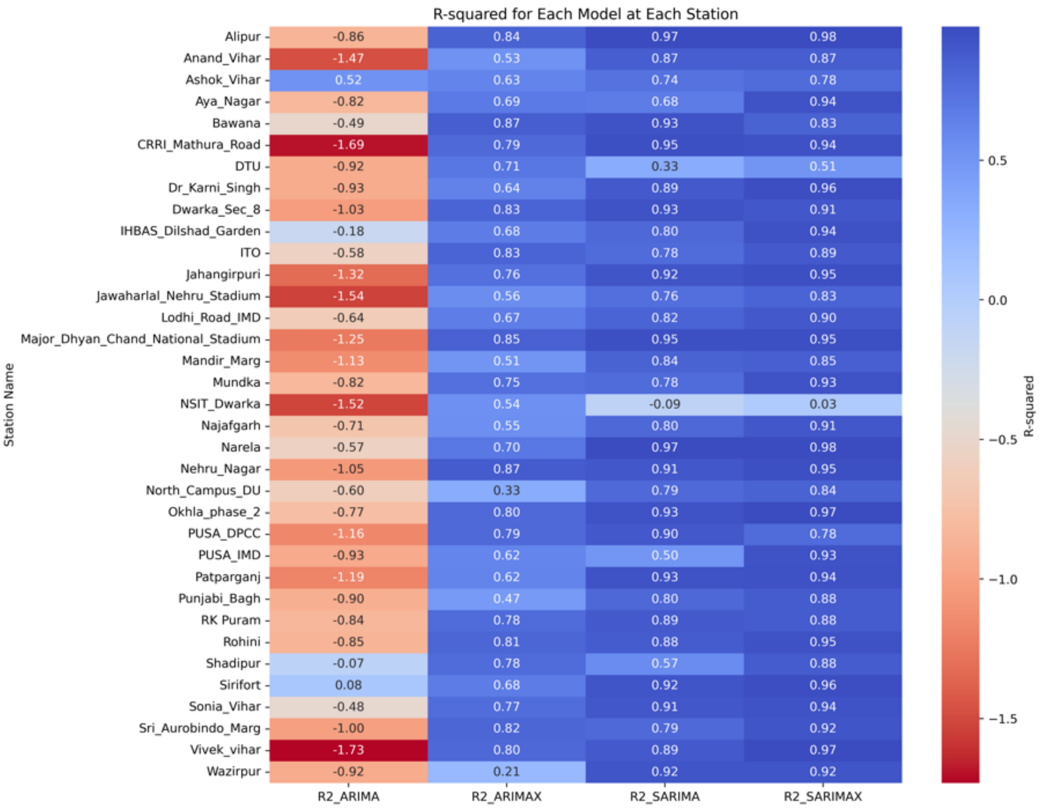


FIGURE 6. R2 Scores for Time Series Models at Each Station

The integration of exogenous variables in the ARIMAX and SARIMAX models significantly impacted their performance. Variables such as ambient temperature ('amb_temp'), wind speed ('wind_spd'), and rainfall were frequently used across stations, enhancing model predictions. The station-specific results illustrate that the inclusion of variables like solar radiation ('sol_rad') and barometric pressure ('bar_press') in the SARIMAX model notably improved its predictive accuracy at stations like Jawaharlal Nehru Stadium and Wazirpur, respectively.

Our evaluation revealed clear differences in performance among traditional (ARIMA and SARIMA) versus advanced (ARIMAX and SARIMAX) time series models for predicting $PM_{2.5}$ levels. Traditional models effectively identified seasonal trends but fell short in accounting for external factors that critically affect pollution levels. This was evident from their generally higher RMSE and lower R^2 values. On the other hand, ARIMAX and SARIMAX models incorporated crucial meteorological data as exogenous variables, markedly improving their predictive capabilities. The SARIMAX model, in particular, displayed enhanced accuracy, proving highly effective in managing complex urban air quality scenarios and deepening our comprehension of air pollution dynamics in large cities like Delhi.

This study not only addresses a significant gap in environmental modeling but also lays a methodological foundation for future research, aiming to enhance the predictive accuracy and real-world utility of air quality forecasts in urban areas. It recommends expanding the integration of broader datasets, including real-time traffic, biomass burning, construction activities, and industrial outputs, to refine predictions. Additionally, it suggests enhancing collaborations with local communities and stakeholders, aligning scientific efforts with urban governance to create more sustainable, health-conscious urban environments.

TABLE 1. Station Specific Optimal Time Series Model

Station Name	Best Model	Model Paramete	Exogenous Features
Alipur	SARIMAX	((0, 1, 1),	('wind_spd', 'rainfall')
Anand_Vihar	SARIMA	(0, 1, 0)x(1, 0, 1,	None
Ashok_Vihar	SARIMAX	((0, 1, 0), (1, 0, 0,	('amb_temp', 'rainfall')
Aya_Nagar	SARIMAX	((1, 1, 2), (1, 0, 2,	('amb_temp', 'wind_spd', 'sol_rad')
Bawana	SARIMA	(0, 1,	None
CRRI_Mathura_Road	SARIMA	(0, 1, 0)x(2, 0, 0, 12)	None
DTU	ARIMAX	(0, 1, 1)	('amb_temp',)
Dr_Karni_Singh	SARIMAX	((2, 1, 2), (1, 0, 1,	('amb_temp', 'rel_hum', 'wind_spd')
Dwarka_Sec_8	SARIMA	(1, 1, 0)x(2, 0, 1,	None
IHBAS_Dilshad_Garden	SARIMAX	((1, 0, 0), (2, 0, 0, 12))	('amb_temp', 'rainfall', 'bar_press')
ITO	SARIMAX	((0, 1, 1), (1, 0, 1,	('amb_temp', 'wind_spd', 'rainfall', 'sol_rad')
Jahangirpuri	SARIMAX	((3, 1, 0), (2, 1, 1,	('wind_spd', 'rainfall')
Jawaharlal_Nehru_Stadium	SARIMAX	((0, 1, 0), (1, 0, 2, 12))	('rel_hum', 'sol_rad')
Lodhi_Road_IMD	SARIMAX	((0, 1, 0), (1, 0, 2,	('rainfall', 'sol_rad', 'bar_press')
Major_Dhyan_Chand_National_Stadium	SARIMAX	((0, 1, 1), (2, 1, 0, 12))	('amb_temp', 'rel_hum', 'wind_spd', 'rainfall', 'sol_rad')
Mandir_Marg	SARIMAX	((0, 1, 0), (2, 0, 0,	('rel_hum', 'sol_rad')
Mundka	SARIMAX	((2, 1, 2),	('rel_hum', 'sol_rad')
NSIT_Dwarka	ARIMAX	(0, 1, 1)	('amb_temp', 'rel_hum', 'wind_spd', 'rainfall', 'sol_rad')
Najafgarh	SARIMAX	((3, 1, 1), (1, 0, 1,	('sol_rad',)
Narela	SARIMAX	((0, 1, 2), (0, 1, 1,	('amb_temp', 'wind_spd', 'rainfall')
Nehru_Nagar	SARIMAX	((2, 1, 2), (2, 0, 0,	('rainfall', 'sol_rad')
North_Campus_DU	SARIMAX	((2, 1, 0), (2, 1, 0,	('rel_hum', 'rainfall')
Okhla_phase_2	SARIMAX	((0, 1, 1), (2, 0, 1,	('amb_temp', 'sol_rad', 'bar_press')
PUSA_DPCC	SARIMA	(0, 1, 0)x(1, 0, 2,	None
PUSA_IMD	SARIMAX	((0, 1, 0), (1, 0, 2,	('amb_temp', 'rainfall')
Patparganj	SARIMAX	((2, 1, 0), (2, 1, 1,	('bar_press',)
Punjabi_Bagh	SARIMAX	((0, 1, 0), (1, 0, 1,	('amb_temp', 'rel_hum', 'rainfall', 'bar_press')
RK Puram	SARIMA	(0, 1,	None
Rohini	SARIMAX	((0, 1, 1),	('rainfall',)
Shadipur	SARIMAX	((1, 0, 1), (1, 0, 1,	('wind_spd', 'rainfall', 'sol_rad', 'bar_press')
Sirifort	SARIMAX	((0, 0, 2),	('wind_spd',)
Sonia_Vihar	SARIMAX	((0, 1, 0), (2, 1, 0,	('wind_spd', 'rainfall', 'sol_rad')
Sri_Aurobindo_Marg	SARIMAX	((0, 1, 0), (1, 0, 2, 12))	('amb_temp', 'wind_spd', 'bar_press')
Vivek_vihar	SARIMAX	((0, 1, 1), (1, 0, 2,	('amb_temp', 'wind_spd', 'sol_rad')
Wazirpur	SARIMAX	((0, 1, 1), (2, 0, 0,	('rel_hum', 'wind_spd', 'rainfall')

4. CONCLUSION

This detailed analysis across multiple time series models and a variety of exogenous factors underlines the nuanced understanding required to effectively predict PM2.5 levels in urban settings like Delhi. The SARIMAX model, with its ability to incorporate external influences, stands out as particularly effective, offering significant potential for policymakers and environmental scientists in crafting more accurate air quality forecasts and better-informed pollution control strategies.

Acknowledgment: The authors would like to thank the CPCB for providing data this research.

REFERENCES

[1]. Y. Shi, M. Bilal, H. C. Ho, and A. Omar, "Urbanization and regional air pollution across South Asian developing countries – A nationwide land use regression for ambient PM2.5 assessment in Pakistan," Environmental Pollution, vol. 266, p. 115145, 2020, doi: 10.1016/j.envpol.2020.115145.

[2]. R. Ganguly, D. Sharma, and P. Kumar, "Trend analysis of observational PM<inf>10</inf> concentrations in Shimla city, India," Sustain Cities Soc, vol. 51, 2019, doi: 10.1016/j.scs.2019.101719.

[3]. S. M. L. Hama et al., "Four-year assessment of ambient particulate matter and trace gases in the Delhi-NCR region of India," Sustain Cities Soc, vol. 54, no. December 2019, p. 102003, 2020, doi: 10.1016/j.scs.2019.102003.

[4]. S. K. Guttikunda, R. Goel, and P. Pant, "Nature of air pollution, emission sources, and management in the Indian cities," Atmos Environ, vol. 95, pp. 501–510, 2014, doi: 10.1016/j.atmosenv.2014.07.006.

[5]. H. Zhang, Y. Wang, J. Hu, Q. Ying, and X. M. Hu, "Relationships between meteorological parameters and criteria air pollutants in three megacities in China," Environ Res, vol. 140, pp. 242–254, Jul. 2015, doi: 10.1016/j.envres.2015.04.004.

[6]. A. Pandey et al., "Health and economic impact of air pollution in the states of India: The Global Burden of Disease Study 2019," Lancet Planet Health, vol. 5, no. 1, pp. e25–e38, Jan. 2021, doi: 10.1016/S2542-5196(20)30298-9.

[7]. S. D. Ghude et al., "Reductions in India's crop yield due to ozone," Geophys Res Lett, vol. 41, no. 15, pp. 5685–5691, Aug. 2014, doi: 10.1002/2014GL060930.

[8]. A. Kumar, H. Hakkim, S. D. Ghude, and V. Sinha, "Probing wintertime air pollution sources in the Indo-Gangetic Plain through 52 hydrocarbons measured rarely at Delhi & Mohali," Science of the Total Environment, vol. 801, Dec. 2021, doi: 10.1016/j.scitotenv.2021.149711.

[9]. U. Nations, D. of Economic, S. Affairs, and P. Division, "World Urbanization Prospects The 2018 Revision," 2018.

[10]. K. Balakrishnan et al., "The impact of air pollution on deaths, disease burden, and life expectancy across the states of India: The Global Burden of Disease Study 2017," Lancet Planet Health, vol. 3, no. 1, pp. e26–e39, Jan. 2019, doi: 10.1016/S2542-5196(18)30261-4.

[11]. A. Pandey et al., "Health and economic impact of air pollution in the states of India: The Global Burden of Disease Study 2019," Lancet Planet Health, vol. 5, no. 1, pp. e25–e38, 2021, doi: 10.1016/S2542-5196(20)30298-9.

[12]. D. Sharma and D. Mauzerall, "Analysis of Air Pollution Data in India between 2015 and 2019," Aerosol Air Qual Res, vol. 22, no. 2, Mar. 2022, doi: 10.4209/aaqr.210204.

[13]. B. R. Gurjar, K. Ravindra, and A. S. Nagpure, "Air pollution trends over Indian megacities and their local-to-global implications," Oct. 01, 2016, Elsevier Ltd. doi: 10.1016/j.atmosenv.2016.06.030.

[14]. P. Pant et al., "Monitoring particulate matter in India: recent trends and future outlook," Air Qual Atmos Health, vol. 12, no. 1, pp. 45–58, Jan. 2019, doi: 10.1007/s11869-018-0629-6.

[15]. P. E. Lang, D. C. Carslaw, and S. J. Moller, "A trend analysis approach for air quality network data," Atmos Environ X, vol. 2, Apr. 2019, doi: 10.1016/j.aeaoa.2019.100030.

[16]. R. Rakholia, Q. Le, B. Quoc Ho, K. Vu, and R. Simon Carbajo, "Multi-output machine learning model for regional air pollution forecasting in Ho Chi Minh City, Vietnam," Environ Int, vol. 173, Mar. 2023, doi: 10.1016/j.envint.2023.107848.

[17]. H. Hajmohammadi and B. Heydecker, "Multivariate time series modelling for urban air quality," Urban Clim, vol. 37, May 2021, doi: 10.1016/j.uclim.2021.100834.

[18]. P. Nath, P. Saha, A. I. Middya, and S. Roy, "Long-term time-series pollution forecast using statistical and deep learning methods," Neural Comput Appl, vol. 33, no. 19, pp. 12551–12570, Oct. 2021, doi: 10.1007/s00521-021-05901-2.

[19]. A. Masood and K. Ahmad, "A model for particulate matter (PM2.5) prediction for Delhi based on machine learning approaches," in Procedia Computer Science, Elsevier B.V., 2020, pp. 2101–2110. doi: 10.1016/j.procs.2020.03.258.

[20]. P. Rai et al., "Real-time measurement and source apportionment of elements in Delhi's atmosphere," Science of the Total Environment, vol. 742, Nov. 2020, doi: 10.1016/j.scitotenv.2020.140332.

[21]. "Guidelines for the Measurement of Ambient Air Pollutants," Delhi, 2013.

[22]. S. Tiwari, A. Dahiya, and N. Kumar, "Investigation into relationships among NO, NO2, NOx, O3, and CO at an urban background site in Delhi, India," Atmos Res, vol. 157, pp. 119–126, Apr. 2015, doi: 10.1016/j.atmosres.2015.01.008.

[23]. I. Kalate Ahani, M. Salari, and A. Shadman, "Statistical models for multi-step-ahead forecasting of fine particulate matter in urban areas," Atmos Pollut Res, vol. 10, no. 3, pp. 689–700, May 2019, doi: 10.1016/j.apr.2018.11.006.

[24]. F. Sánchez Lasheras, P. J. García Nieto, E. García Gonzalo, L. Bonavera, and F. J. de Cos Juez, "Evolution and forecasting of PM10 concentration at the Port of Gijon (Spain)," Sci Rep, vol. 10, no. 1, Dec. 2020, doi: 10.1038/s41598-020-68636-5.

[25]. R. J. Hyndman and G. Athanasopoulos, Forecasting: Principles and Practice, 3rd ed. Melbourne, Australia: OTexts, 2021. Accessed: May 17, 2024. [Online]. Available: https://otexts.com/fpp3/

[26]. T. G. Smith and A. Smith, "pmdarima: ARIMA estimators for Python," http://www.alkaline-ml.com/pmdarima.

[27]. G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, Time series analysis: forecasting and control, 5th ed. John Wiley & Sons, 2015.

[28]. K. L. S. da Silva, J. L. López-Gonzales, J. E. Turpo-Chaparro, E. Tocto-Cano, and P. C. Rodrigues, "Spatio-temporal visualization and forecasting of PM 10 in the Brazilian state of Minas Gerais," Sci Rep, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-30365-w.

[29]. J. Kaur, K. S. Parmar, and S. Singh, "Autoregressive models in environmental forecasting time series: a theoretical and application review," Feb. 01, 2023, Springer Science and Business Media Deutschland GmbH. doi: 10.1007/s11356-023-25148-9.

