Journal of Computer Science and Technology Studies
ISSN: 2709-104X
DOI: 10.32996/jcsts
Journal Homepage: www.al-kindipublisher.com/index.php/jcsts

JCSTS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Modeling PM2.5 Concentrations in the Export Processing Area of Dhaka Over a 12-Month Period Using Time Series

Gordon Hung
Independent Researcher, Computer Science, Hsinchu, Taiwan
Corresponding Author: Gordon Hung, **E-mail**: gordonh741@gmail.com

| ABSTRACT

Air pollution, particularly PM2.5, poses significant health risks, with Dhaka, Bangladesh, experiencing some of the highest concentrations of fine particulate matter. This study employs 11 time-series models to analyze PM2.5 concentrations in the export processing area of Dhaka. Using a recently published dataset from January 2019 to December 2023, the models were trained to forecast PM2.5 concentrations for a given day of each month over a 12-month period. The results indicated that the Holt-Winters model, followed by a Neural Network and SARIMA, achieved the best performance. An ensemble model incorporating these top three models was developed, demonstrating high forecasting accuracy compared to the other models. This study provides valuable insights for policymakers and businesses, offering an advanced framework for forecasting PM2.5 concentrations to address the ongoing air pollution issue in Dhaka.

| KEYWORDS

PM2.5; environmental modelling; time series forecasting; model optimization; ensemble model;

| ARTICLE INFORMATION

ACCEPTED: 01 February 2025 **PUBLISHED:** 11 February 2025 **DOI:** 10.32996/jcsts.2025.7.1.13

1. Introduction

Air pollution, especially fine particulate matter (PM2.5), is one of the most pressing environmental and public health challenges in many urban areas [1, 2, 3]. PM2.5 refers to airborne particles with a diameter of 2.5 micrometers or less, small enough to penetrate deep into the lungs and even enter the bloodstream [4, 5]. Its small size makes it highly dangerous as it is more likely to travel into deeper parts of the lungs and cause severe tissue damage and lung inflammation [6].

Exposure to high levels of PM2.5 has been linked to a variety of health issues. For example, respiratory diseases such as asthma and chronic obstructive pulmonary disease (COPD) are common in urban areas with high concentrations of PM2.5 [7, 8]. Cardiovascular conditions like heart attacks and strokes, and increased risks of lung cancer are also common conditions associated with prolonged exposure to high levels of PM2.5 [9, 10]. Globally, the World Health Organization (WHO) estimates that air pollution contributes to approximately 6.7 million premature deaths annually, with PM2.5 being a major contributor [11].

Among the cities grappling with severe PM2.5 pollution, Dhaka, the capital of Bangladesh, stands out as one of the most critically affected [12, 13]. Rapid urbanization, unchecked industrialization, and burgeoning vehicular emissions have transformed Dhaka into a hotspot for dangerously high concentrations of PM2.5. The city's air pollution levels frequently surpass the WHO's recommended limits, posing significant risks to its residents. Brick kilns, a staple of the region's industrial output, are one of the primary sources of fine particulate matter, alongside emissions from unregulated vehicles and construction activities [14, 15]. Dhaka's severe traffic problems further exacerbate the problem, with countless vehicles releasing harmful emissions on a daily basis. Compounding these issues is the city's lack of robust air quality management systems, leading to its deteriorating air quality [16, 17, 18].

Efforts to combat air pollution in Dhaka have been initiated over the years but with limited success. The Bangladesh National Environment Policy, introduced in 1992, laid the groundwork for environmental governance, aiming to control pollution and promote sustainable practices [19]. More recently, projects like the Clean Air and Sustainable Energy (CASE) initiative, launched in 2019, sought to monitor air quality and develop cleaner energy solutions [20]. However, these measures have fallen short in addressing the city's rapidly worsening air quality. The enforcement of regulations remains inconsistent, and the absence of coordinated urban planning further undermines progress. Consequently, PM2.5 concentrations in Dhaka continue to rise, highlighting the need for more comprehensive approaches to effectively target the issue.

The continuous increase in PM2.5 pollution in Dhaka is an urgent issue that must be addressed as millions of lives are at risk [21, 22, 23]. Therefore, this study utilizes Artificial Intelligence (AI) to offer a state-of-the-art approach for effectively monitoring PM2.5 concentrations in Dhaka. This ensures that policymakers and business owners can make the most informed decisions in targeting the issue of rising PM2.5 concentrations.

The remainder of this paper is organized as follows: Section 2 provides a comprehensive literature review of previous studies. Section 3 discusses the detailed descriptions of the methodology, including the dataset and the models. Section 4 offers the results of the study, presenting the performances of the models and proposing a highly accurate ensemble model. Finally, this paper concludes in Section 5.

2. Literature Review

There have been several notable studies done on a similar topic. This section offers a comprehensive analysis of the existing literature.

Huang et al. proposed a deep CNN-LSTM model for forecasting PM2.5 concentrations, demonstrating its superior accuracy compared to other machine learning methods [24]. Gao and Li introduced a Graph-based LSTM (GLSTM) model that incorporates a parameterized adjacency matrix to predict PM2.5 concentrations, demonstrating improved accuracy and interpretability by using spatiotemporal data across multiple monitoring stations [25]. Mahajan et al. explored the use of Neural Network Autoregression (NNAR) for forecasting PM2.5 levels, comparing its performance with Holt-Winters and ARIMA models [26]. Mahajan et al. also proposed a cluster-based hybrid neural network model to improve PM2.5 forecasting accuracy and efficiency, demonstrating that clustering monitoring stations by geographical distance reduces both forecasting error and computation time using data from 557 Airbox devices across Taiwan [27]. Zaini et al. proposed a hybrid EEMD-LSTM model for hourly PM2.5 forecasting in Malaysia, demonstrating that the model, which combines data decomposition with deep learning, outperforms other models in forecasting accuracy [28]. Mahajan et al. further proposed an exponential smoothing method with drift for short-term PM2.5 forecasting, demonstrating its effectiveness in forecasting PM2.5 concentrations with low error and acceptable computation time using data from 132 IoT monitoring stations in Taichung, Taiwan [29].

The literature suggests that complex models, such as hybrid and ensemble models, tend to exhibit higher predictive capacities. While multiple studies have analyzed PM2.5 trends across various regions, many rely on outdated datasets that fail to capture current PM2.5 concentration patterns. This study addresses this gap by utilizing a recently published dataset to analyze PM2.5 concentration fluctuations in Dhaka, Bangladesh—a region often overlooked despite experiencing severely high PM2.5 levels.

3. Materials and Methods

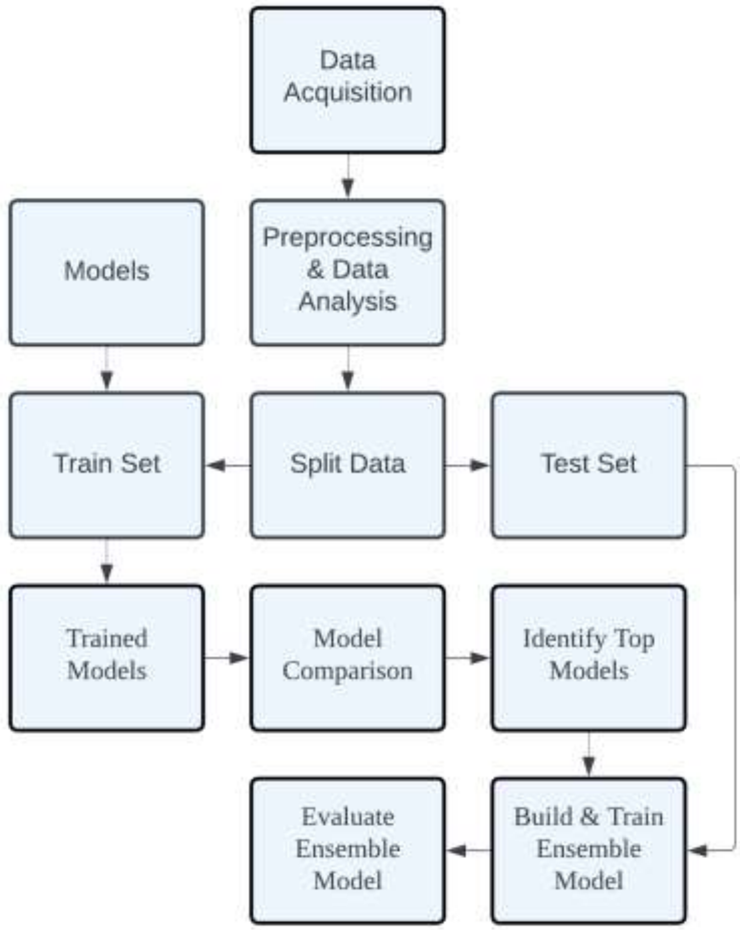
3.1 Objectives

Building upon previous research, this paper employs 11 time-series models to analyze and forecast PM2.5 concentrations in Dhaka, Bangladesh. For comprehensive evaluation, five error metrics were applied to assess the models. Subsequently, an ensemble model is proposed, achieving 90% accuracy. The study utilizes a recently published dataset spanning January 2019 to December 2023, offering valuable insights to policymakers and business organizations [30]. The contributions of this study are as follows:

- Provide a comprehensive analysis of a recently published dataset.
- Employ 11 time-series models to analyze PM2.5 trends in Dhaka and offer a robust framework for PM2.5 concentration prediction.
- Present an ensemble model capable of achieving high accuracy.
- Offer insights into PM2.5 concentrations in Dhaka from 2019 to 2023.

The outline of the steps taken is shown in Figure 1.

Figure 1. Workflow overview.



3.2 Data and Preprocessing

This study utilizes a recently published dataset containing measurements of Particulate Matter (PM2.5, PM10), CO concentrations, and Air Quality Index (AQI) values in the export processing area of Dhaka, Bangladesh. Our research focused on analyzing PM2.5 levels because they are a critical indicator of air pollution and have significant health impacts, particularly on respiratory and cardiovascular systems. The data for PM2.5 was acquired using a respirable dust sampler (APS-113NL), which collected 24-hour samples once a month. The dataset spans from January 2019 to December 2023, with no missing values. The data was collected at a precise location, with coordinates latitude 23.94821 and longitude 90.27727. Table 1 presents a descriptive analysis of the dataset.

Table 1. Summary of dataset.

Statistics	PM2.5
Standard Deviation	36.03
Minimum	9.10
25%	16.75
50%	34.95
75%	75.75
Maximum	140.70

After acquiring the dataset, we standardized the dates and PM2.5 concentration values. We then checked the stationarity of the dataset to meet the requirements of the time series models. The Augmented Dickey-Fuller (ADF) Test and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) Test were applied. The ADF test determines whether a time series is stationary by testing for the presence of a unit root, while the KPSS test evaluates stationarity by assessing whether the series has a constant mean or exhibits a trend. The ADF test can be described mathematically as follows:

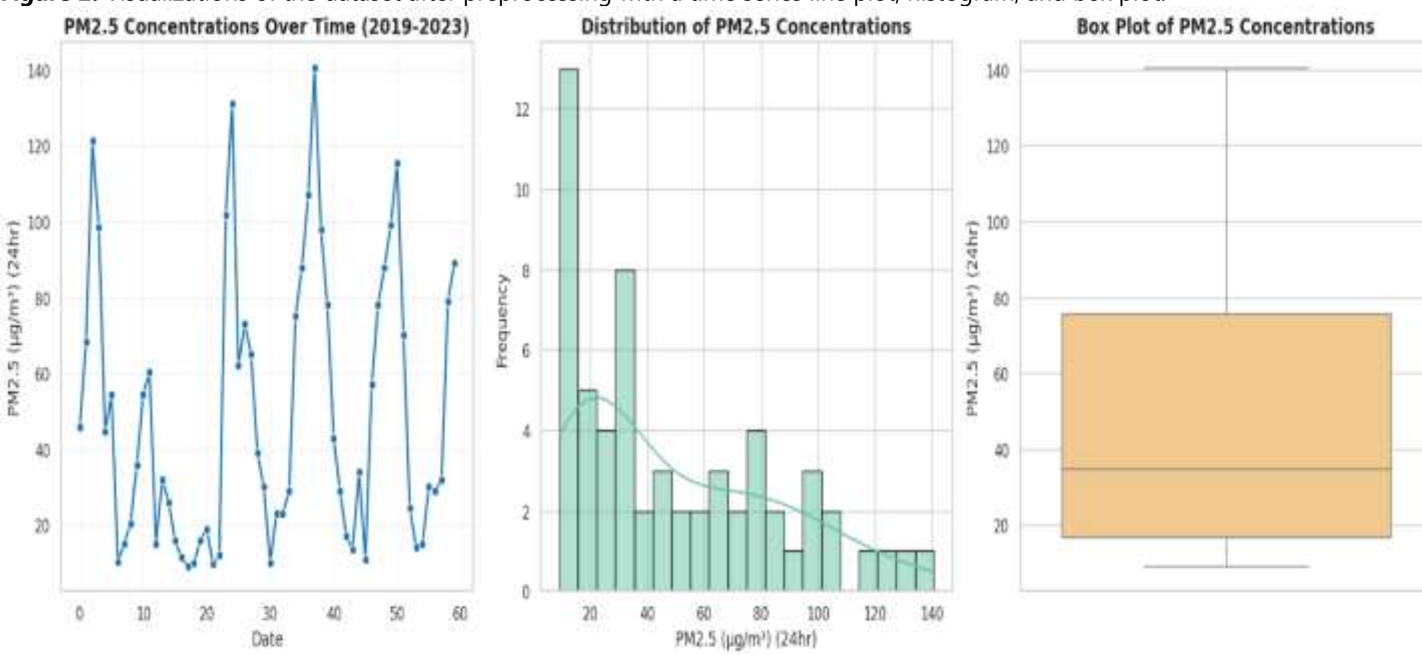
$$\Delta Y_t = \alpha + \beta t + \gamma Y_{t-1} + \sum_{i=1}^{p} \delta_i \Delta Y_{t-1} + \epsilon_t \quad (1)$$

Where ΔY_t represents the change in the series at time t, γY_{t-1} tests for a unit root, $\alpha + \beta t$ accounts for trends and drifts, $\sum_{i=1}^{p} \delta_i \Delta Y_{t-1}$ is the lagged difference to correct for autocorrelation, and ϵ_t is the error term. On the other hand, the KPSS test can be described as follows:

$$\text{KPSS} = \frac{1}{r^2} \sum_{t=1}^{T} \frac{S_t^2}{\widehat{\sigma^2}} \quad (2)$$

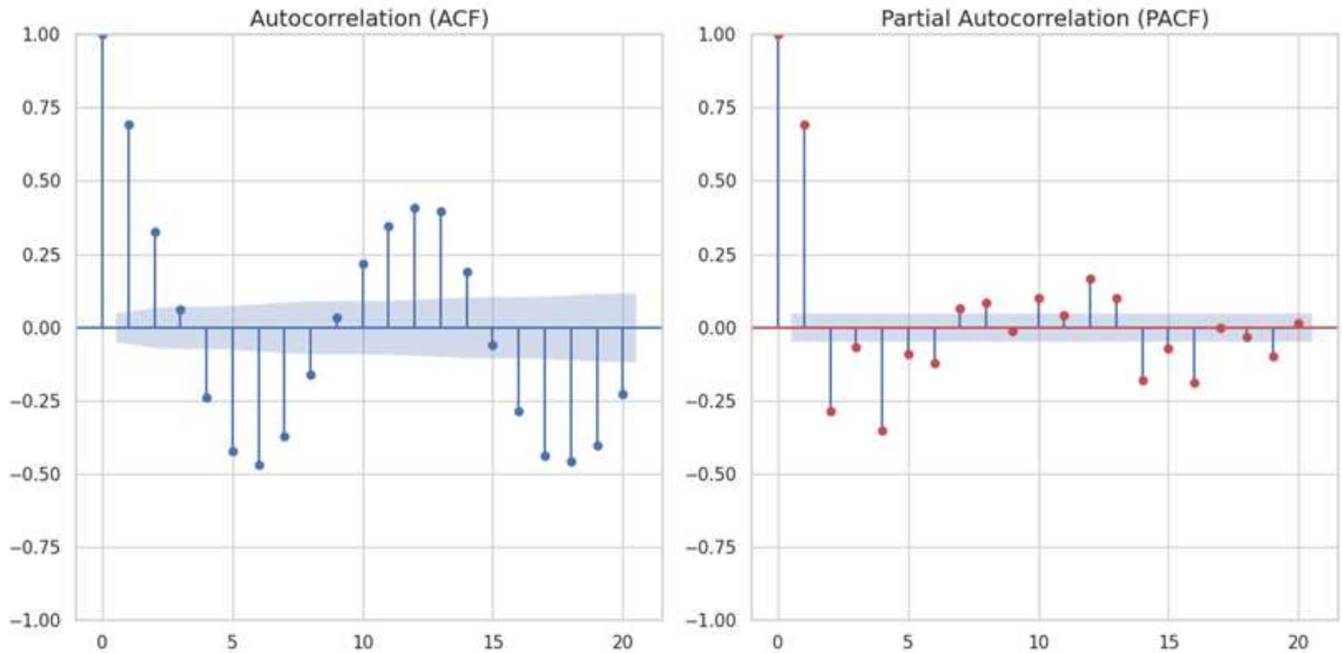
Where S_t represents the partial sum of the series at time t, $\widehat{\sigma^2}$ is the estimated variance of the residuals, r is the number of observations. Together, the ADF test checks for unit roots (non-stationarity), while the KPSS test checks for trend stationarity, providing complementary information. Both tests confirmed that our dataset is stationary. Figure 2 shows the dataset after preprocessing.

Figure 2. Visualizations of the dataset after preprocessing with a time series line plot, histogram, and box plot.



To more effectively understand the underlying dependencies of the dataset, we used the Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) graphs to visualize our dataset. The ACF helps identify the overall correlation between the current and past values, providing insights into the persistence of trends. The PACF, on the other hand, isolates the correlation at specific lags, accounting for the influence of intermediate lags, and helps determine the order of autoregressive models. The graphs are shown below:

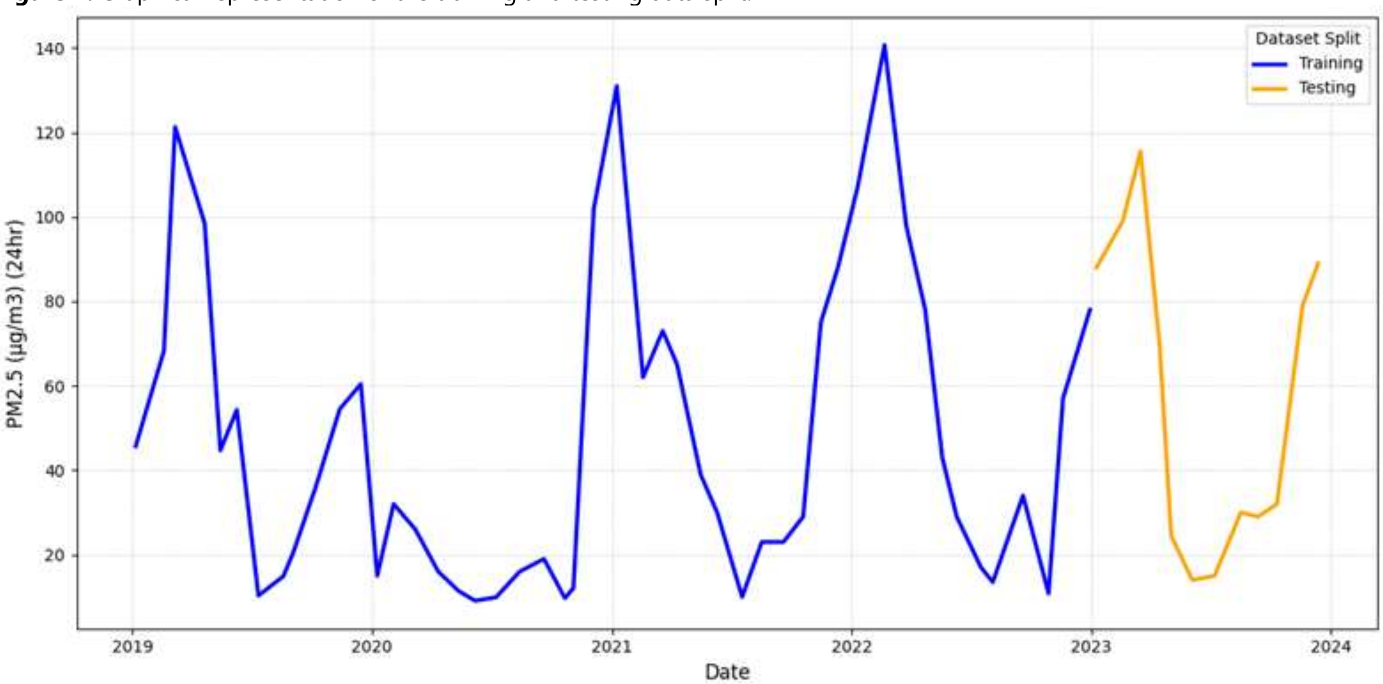
Figure 3. ACF and PACF graphs with a lag value of 20.



The ACF plot shows significant positive correlations for the first few lags, followed by a decaying pattern, indicating a potential dependency on past values that diminishes over time. The PACF plot reveals a sharp drop after lag 1, suggesting that the series primarily relies on its most recent value for forecasting. Together, these plots suggest that the time series may exhibit both short-term and longer-term dependencies that can inform model selection and structure.

After preprocessing, we split our dataset into training and testing sets based on an 80:20 split. The training set ranges from January 2019 to December 2022, while the testing set ranges from January 2023 to December 2023. Figure 4 shows the data split.

Figure 4. Graphical representation of the training and testing data split.



3.3 Evaluation Metrics

To comprehensively analyze our models, we applied four standard evaluation metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Symmetric Mean Absolute Percent Error (SMAPE), and Max Error. MAE measures the average magnitude of errors between the predicted and the actual. The mathematical representation is given as follows.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \widehat{y_i}| \tag{3}$$

Where n is the number of observations. RMSE measures the square root of the average squared differences between predicted and actual values, emphasizing larger errors. The mathematical representation is given as follows.

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y_i})^2} \tag{4}$$

SMAPE is a percentage-based error metric that normalizes the absolute error by the sum of actual and predicted values. The mathematical representation is given as follows.

$$\text{SMAPE} = \frac{100}{n}\sum_{i=1}^{n}\frac{|y_i - \widehat{y_i}|}{(y_i + \widehat{y_i})/2} \tag{5}$$

Max Error identifies the largest absolute difference between predicted and actual values, focusing on the single worst prediction. The mathematical representation is given as follows.

$$\text{Max Error} = \max_{i=1,\ldots,n} |y_i - \widehat{y_i}| \tag{6}$$

3.4 Models

We employed 11 time-series models—Naive, Drift, Simple Exponential Smoothing (SES), Holt's Linear, ARIMA, Auto ARIMA, Holt-Winters, SARIMA, FFT, Theta, and a Neural Network (NN)—to forecast PM2.5 concentrations in Dhaka for the year 2023.

The Naive model assumes that the forecast for the next period is simply the value from the most recent period. The Drift model assumes that the time series data follows a random walk with a constant drift. The forecast is made by adding a fixed trend to the most recent value, usually calculated as the average change over time. SES is a forecasting method that uses weighted averages of past observations to make predictions, giving more weight to recent observations. It is best suited for time series data without a trend or seasonality. Holt's Linear method is an extension of SES that accounts for linear trends in the data. It adds a trend component to the model, making it suitable for time series with a trend but no seasonality. ARIMA is a time series forecasting method that combines autoregression (AR), differencing (I), and moving average (MA) components. It is used for modeling time series data with a trend or seasonality. Auto ARIMA is an automated method for selecting the best ARIMA model by performing a grid search over possible values of p, d, and q (the AR, differencing, and MA terms). Holt-Winters is an extension of exponential smoothing that handles seasonality and trends in time series data. It has three components: level, trend, and seasonality, and can be applied in both additive and multiplicative forms depending on the nature of the series. SARIMA is an extension of ARIMA that explicitly handles seasonality in time series data. It includes seasonal autoregressive (SAR), seasonal differencing (SI), seasonal moving average (SMA), and seasonal period components in addition to the regular ARIMA components. FFT is a signal processing method used to transform time series data into frequency components. By identifying dominant cycles in the data, FFT can help model and forecast periodic behavior in time series. The Theta model is a forecasting technique that combines exponential smoothing and a decompositional approach. Lastly, Neural Networks (NN) are machine learning models that are adept at recognizing patterns over time.

4. Results and Dicussion

4.1 Top Models

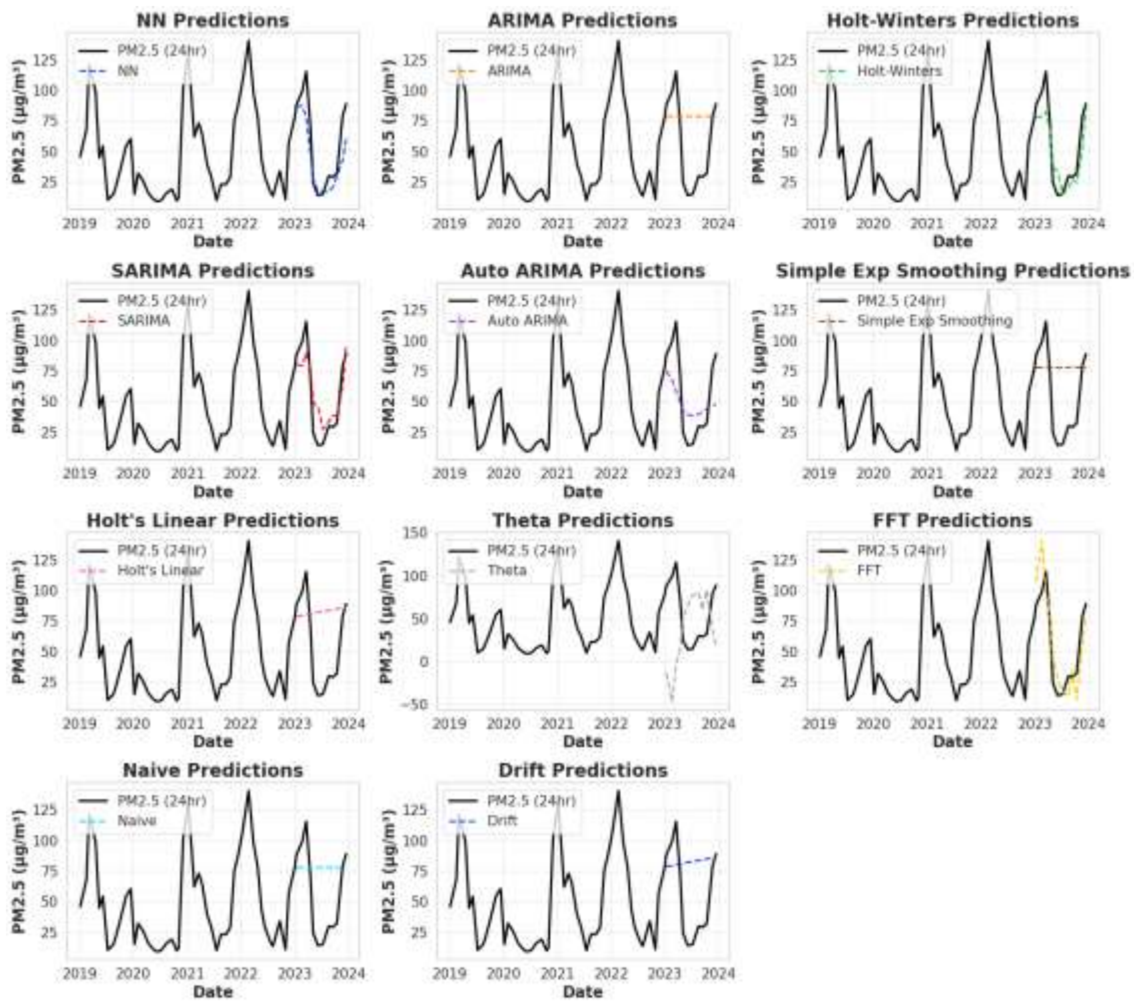
After training our 11 models, we employed them to forecast PM2.5 concentrations for the year 2023. Their results were compared to the test set and the four error metrics were used to evaluate their performances. Table 2 presents the performances of the models.

Table 2. Performances of the models.

Models	MAE	RMSE	SMAPE	Max Error	Rank
Holt-Winters	12.456	15.94	13.287	32.952	1
NN	14.138	19.114	13.891	37.284	2
SARIMA	14.242	16.78	15.997	32.213	3
FFT	16.458	19.142	18.97	41.7	4
Auto ARIMA	24.82	28.254	24.975	57.533	5
Naive	34.342	40.715	30.967	64	6
ARIMA	34.563	41.325	30.999	65.164	7
SES	33.008	34.943	31.692	57.333	8
Drift	36.58	43.864	31.885	68.123	9
Holt's Linear	36.585	43.87	31.887	68.131	10
Theta	67.292	75.767	63.37	145.479	11

While most models were able to capture the seasonality of the dataset, some models merely assumed linearity. Figure 5 shows the predictions of the models.

Figure 5. PM2.5 concentration predictions of the models.



We identified the top three models through a comprehensive comparison of all the evaluation metrics, with Holt-Winters ranking first, followed by NN and SARIMA. The sections below offer a detailed explanation of each model.

4.1.1 Holt-Winters

The Holt-Winters method is also known as Triple Exponential Smoothing; it is an extension of the simple exponential smoothing model to capture both trend and seasonality in time series data. The model has three primary components: level (L_t), the smoothed value of the series; trend (T_t), the rate of change in the level; and seasonality (S_t), the repeating cycle of fluctuations in the data. The model can be described by the following equations:
Level Equation:

$$L_t = \alpha(y_t - S_{t-m}) + (1-\alpha)(L_{t-1} + T_{t-1}) \tag{7}$$

Where α is the smoothing constant for the level.
Trend Equation:

$$T_t = \beta(L_t - L_{t-1}) + (1-\beta)T_{t-1} \tag{8}$$

Where β is the smoothing constant for the trend.
Seasonal Equation:

$$S_t = \gamma(y_t - L_t) + (1-\gamma)S_{t-m} \tag{9}$$

Where γ is the smoothing constant for the seasonal component.
Forecast Equation:

$$\hat{y}_{t+h} = L_t + hT_t + S_{t+m-h} \tag{10}$$

Where h is the forecast horizon.

4.1.2 Neural Network (NN)

NNs are inspired by the structure of the brain, capable of learning complex pat-terns in data. They are effective at capturing both linear and nonlinear relationships between past observations and future values. NNs have a single input and output layer, with multiple hidden layers in between, used for processing data and learning patterns. The equations below model a forward pass in a NN with only one hidden layer.
Hidden Layer:

$$h_t = f(W^{(1)}x_t + b^{(1)}) \tag{11}$$

Where x_t is the input vector at time t, $W^{(1)}$ represents the weight matrix for the input layer to the first hidden layer, and $b^{(1)}$ is the bias term for the first hidden layer. The function f is an activation function.
Output Layer:

$$\hat{y_t} = g(W^{(2)}h_t + b^{(2)}) \tag{12}$$

Where h_t is the output from the hidden layer, $W^{(2)}$ represents the weight matrix for the hidden layer to the output layer, and $b^{(2)}$ is the bias term for the output layer. The function g is the identity function.

4.1.3 SARIMA

The SARIMA model is a combination of the non-seasonal ARIMA model and seasonal components. It can be represented as follows:

$$\emptyset_p = (B)(1-B)^d(1-B^m)^D y_t = \theta_q(B)\epsilon_t + \varphi_p(B^m)(1-B)^D(1-B^m)^D\epsilon_t \tag{13}$$

Where $\emptyset_p(B)$ is the autoregressive polynomial of order p, $\theta_q(B)$ is the moving average polynomial of order q. $\varphi_p(B^m)$ is the seasonal autoregressive polynomial of order P, ϵ_t is the white noise error term. To account for seasonality, seasonal differencing is applied:

$$\Delta_s y_t = y_t - y_{t-m} \tag{14}$$

Where Δ_s denotes the seasonal difference, and m is the length of the seasonal period. The forecast for h-steps ahead using SARIMA is:

$$\hat{y}_{t+h} = \mu + \sum_{i=1}^{p} \emptyset_i y_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \sum_{i=1}^{P} \varphi_i y_{t-i-m} + \sum_{j=1}^{Q} \Theta_i \epsilon_{t-j-m} \tag{15}$$

Where μ is the mean of the series, and ϵ_t is the white noise error term.

4.2 Ensemble Model

Our top models—Holt-Winters, Neural Network, and SARIMA—were then used to construct an ensemble model that utilizes stacked generalization to strategically combine the predictions from all three models. Stacked generalization is an ensemble learning technique that uses a meta-model to integrate the predictions of multiple base models. Our ensemble model builds upon this structure by using the top models as the base models and a simple Linear Regression as the meta-model. Figure 6 shows the general architecture of our ensemble model. Table 3 presents the performance of the ensemble model compared to the top models.

Figure 6. General architecture of ensemble model.

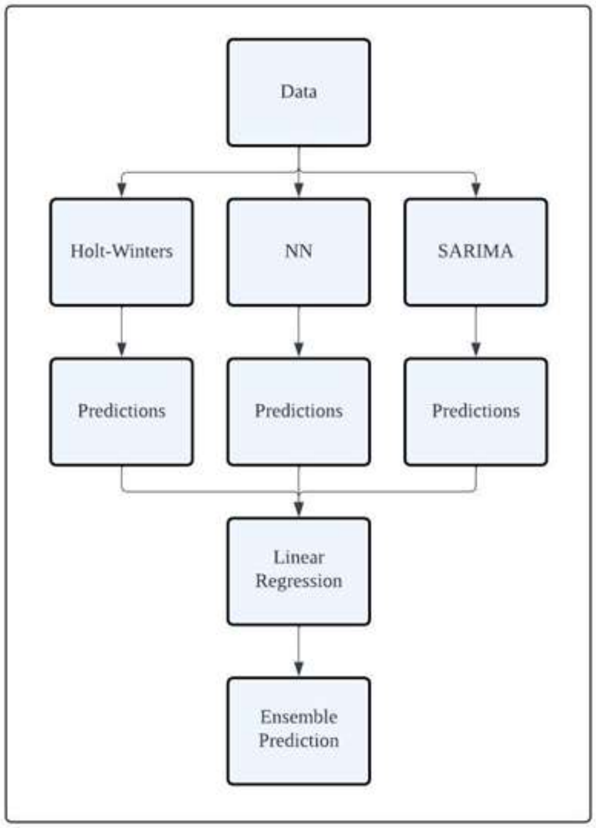
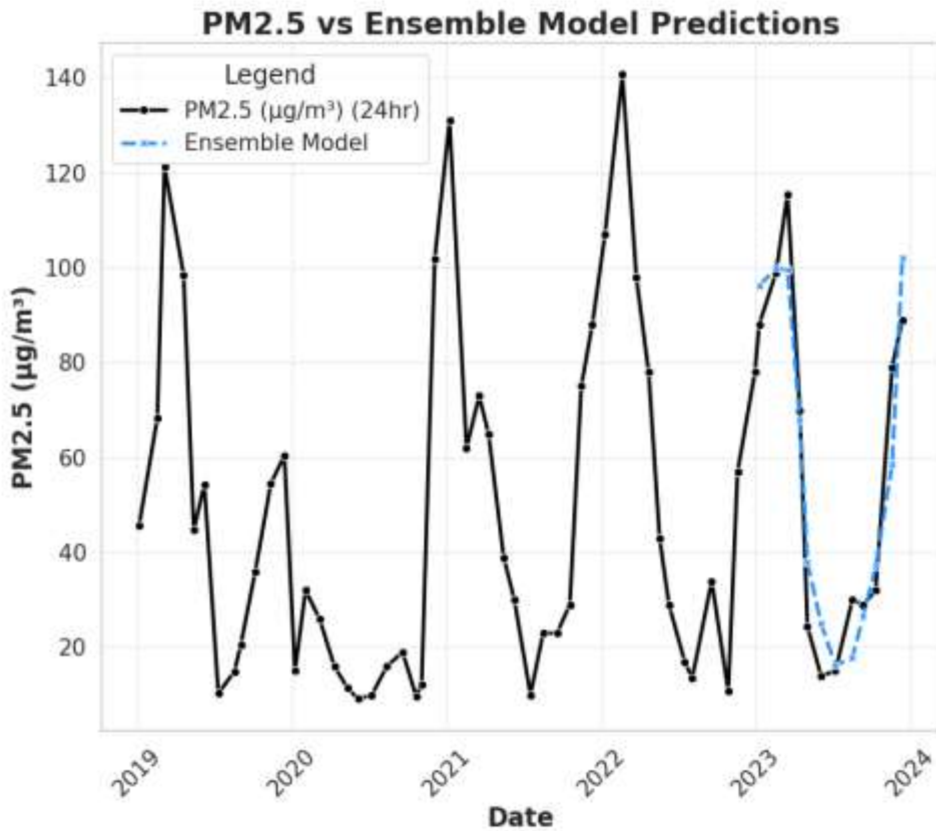


Table 3. Performances of the models.

Models	MAE	RMSE	SMAPE	Max Error	Rank
Ensemble Model	7.801	10.740	9.892	24.867	1
Holt-Winters	12.456	15.94	13.287	32.952	2
NN	14.138	19.114	13.891	37.284	3
SARIMA	14.242	16.78	15.997	32.213	4

The ensemble model achieved the highest accuracy due to its ability to process and learn from the predictions of the base models. Figure 7 shows the predictions of the model.

Figure 7. PM2.5 concentration predictions of the proposed ensemble model.



5. Conclusion

Accurate forecasting of PM2.5 concentrations is crucial for addressing health issues and implementing effective measures. This study analyzes 11 time-series models to forecast PM2.5 concentrations in Dhaka, Bangladesh, based on a recently published dataset, and presents an ensemble model that achieves a SMAPE of 10.489%, which is highly accurate considering the sparsity of the dataset and the fluctuating nature of PM2.5 concentrations. A limitation of this study is the use of univariate models in forecasting PM2.5 concentration. Future studies can investigate more complex multivariate models to better understand and utilize feature-target relationships. Additionally, future work could focus on conducting more extensive hyperparameter searches and developing more advanced hybrid models. Future studies could also utilize the findings of this paper to model PM2.5 concentration on a much larger scale, integrating with hardware devices for system simulations. In summary, this paper presents a robust framework for PM2.5 concentration forecasting and provides invaluable insights into the rapidly changing PM2.5 trends for policymakers and other researchers.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.
Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

[1] Epa, O. U. (2016). Health and environmental effects of particulate matter (PM). https://www.epa.gov/pm-pollution/health-and-environmental-effects-particulate-matter-pm
[2] Inhalable Particulate Matter and health (PM2.5 and PM10). (n.d.). Arb.ca.gov. Retrieved November 24, 2024, from https://ww2.arb.ca.gov/resources/inhalable-particulate-matter-and-health
[3] (N.d.). Nih.gov. Retrieved November 24, 2024, from https://pmc.ncbi.nlm.nih.gov/articles/PMC4740125/
[4] Us Epa, O. (2016). Particulate matter (PM) basics. https://www.epa.gov/pm-pollution/particulate-matter-pm-basics
[5] (N.d.). Nih.gov. Retrieved November 24, 2024, from https://pmc.ncbi.nlm.nih.gov/articles/PMC4740125/#:~:text=raising%20worldwide%20concerns.-,PM2.,investigate%20the%20impact%20of%20PM2
[6] The dangers of PM2.5. (n.d.). Vibhavadi.com. Retrieved November 24, 2024, from https://www.vibhavadi.com/en/Health-expert/detail/101
[7] (N.d.). Nih.gov. Retrieved November 24, 2024, from https://pmc.ncbi.nlm.nih.gov/articles/PMC10086390/

[8] Zafirah, Y., Lin, Y.-K., Andhikaputra, G., Deng, L.-W., Sung, F.-C., & Wang, Y.-C. (2021). Mortality and morbidity of asthma and chronic obstructive pulmonary disease associated with ambient environment in metropolitans in Taiwan. PloS One, 16(7), e0253814. https://doi.org/10.1371/journal.pone.0253814

[9] Krittanawong, C., Qadeer, Y. K., Hayes, R. B., Wang, Z., Virani, S., Thurston, G. D., & Lavie, C. J. (2023). PM2.5 and cardiovascular health risks. Current Problems in Cardiology, 48(6), 101670. https://doi.org/10.1016/j.cpcardiol.2023.101670

[10] (N.d.). Nih.gov. Retrieved November 24, 2024, from https://pmc.ncbi.nlm.nih.gov/articles/PMC9370264/

[11] Types of pollutants. (n.d.). Who.int. Retrieved November 24, 2024, from https://www.who.int/teams/environment-climate-change-and-health/air-quality-and-health/health-impacts/types-of-pollutants

[12] Dhaka Air Quality Index (AQI) and Bangladesh air pollution. (n.d.). Iqair.com. Retrieved November 24, 2024, from https://www.iqair.com/bangladesh/dhaka?srsltid=AfmBOoo5kdqFqlzqnf2L6MP8uAbbWeyEOJsJ8ggEyOfSE7iAKEN8Kw_q

[13] Find an expert. (2024, June 25). Bangladesh's air quality is among the world's worst. What can be done? Atlantic Council. https://www.atlanticcouncil.org/blogs/energysource/bangladeshs-air-quality-is-among-the-worlds-worst-what-can-be-done/

[14] Guttikunda, S. K., Begum, B. A., & Wadud, Z. (2013). Particulate pollution from brick kiln clusters in the Greater Dhaka region, Bangladesh. Air Quality, Atmosphere, & Health, 6(2), 357–365. https://doi.org/10.1007/s11869-012-0187-2

[15] (N.d.). Pulitzercenter.org. Retrieved November 24, 2024, from https://pulitzercenter.org/stories/bangladeshs-air-pollution-problem-grows-brick-brick

[16] (N.d.). Nih.gov. Retrieved November 24, 2024, from https://pmc.ncbi.nlm.nih.gov/articles/PMC8193182/#:~:text=Literature%20also%20shows%20that%20PM,gradually%20increasing%20in%20Dhaka%20City.&text=Rapid%20development%2C%20industrialization%2C%20and%20pollution,in%20the%20densely%20populated%20city

[17] Bangladesh breathed worst air last year according to World Air Quality Report. (n.d.). Vietnamnews.Vn. Retrieved November 24, 2024, from https://vietnamnews.vn/world/1652340/bangladesh-breathed-worst-air-last-year-according-to-world-air-quality-report.html

[18] World Bank Group. (2024, September 30). High air pollution level is creating physical and mental health hazards in Bangladesh: World Bank. World Bank Group. https://www.worldbank.org/en/news/press-release/2022/12/03/high-air-pollution-level-is-creating-physical-and-mental-health-hazards-in-bangladesh-world-bank

[19] (N.d.). Academia.edu. Retrieved November 24, 2024, from https://www.academia.edu/88517258/National_Environmental_Policy_of_Bangladesh_1992

[20] Development projects : Clean Air and Sustainable Environment Project - P098151. (n.d.). World Bank. Retrieved November 24, 2024, from https://projects.worldbank.org/en/projects-operations/project-detail/P098151

[21] IQAir. (n.d.). Iqair.com. Retrieved November 24, 2024, from https://www.iqair.com/newsroom/dhaka-suffers-record-air-pollution-2021?srsltid=AfmBOorzALg7jhZ1-0NmmFmpwZgj3LDcX0REz8sNG0vbKz3al72GoTvw

[22] (N.d.). Iop.org. Retrieved November 24, 2024, from https://iopscience.iop.org/article/10.1088/2515-7620/acb90d

[23] (N.d.). Edu.Bd. Retrieved November 24, 2024, from https://bsmrau.edu.bd/seminar/wp-content/uploads/sites/318/2018/08/Seminar-Paper-Eva-Armin.pdf

[24] Huang, C. -J., & Kuo, P. -H. (2018). A Deep CNN-LSTM Model for Particulate Matter (PM2.5) Forecasting in Smart Cities. Sensors, 18(7), 2220. https://doi.org/10.3390/s18072220

[25] Gao, X., & Li, W. (2021). A graph-based LSTM model for PM2.5 forecasting. Atmospheric Pollution Research, 12(9), 101150. https://doi.org/10.1016/j.apr.2021.101150

[26] Mahajan, S., Chen, L.-J., & Tsai, T.-C. (2017). An empirical study of PM2.5 forecasting using neural network. 2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), 1–7.

[27] S. Mahajan, H. -M. Liu, T. -C. Tsai and L. -J. Chen, "Improving the Accuracy and Efficiency of PM2.5 Forecast Service Using Cluster-Based Hybrid Neural Network Model," in IEEE Access, vol. 6, pp. 19193-19204, 2018, doi: 10.1109/ACCESS.2018.2820164.

[28] Zaini, N., Ean, L.W., Ahmed, A.N. et al. PM2.5 forecasting for an urban area based on deep learning and decomposition method. Sci Rep 12, 17565 (2022). https://doi.org/10.1038/s41598-022-21769-1

[29] Mahajan, S., Chen, L. -J., & Tsai, T. -C. (2018). Short-Term PM2.5 Forecasting Using Exponential Smoothing Method: A Comparative Analysis. Sensors, 18(10), 3223. https://doi.org/10.3390/s18103223

[30] Rahman, M., Rashid, F., Kumar, D., Habib, M. A., & Ullah, A. (2024). Dataset of air pollutants (PM2.5, PM10, CO) concentrations in the export processing area of Dhaka, Bangladesh. Data in Brief, 55(110594), 110594. https://doi.org/10.1016/j.dib.2024.110594