Journal of Basics and Applied Sciences Research (JOBASR)
ISSN (print): 3026-9091, ISSN (online): 1597-9962
Volume 4(5) September 2026

DOI: https://dx.doi.org/10.4314/jobasr.v4i5.24

Prediction of Air Pollutants in North-Western Nigeria: A Comparison of LSTM and Facebook Prophet Models

[1]Aminu Bello Zoramawa, [2]Yakubu Musa & [3]Muhammad Zayanu

[1,2,3] Department of Statistics, Usmanu Danfodiyo University, Sokoto
Corresponding author: arimaym@gmail.com

ABSTRACT

Air pollution, driven by urbanization and globalization, is one of the major environmental problems facing society; it affects the physical environment and human health and, as a result, contributes significantly to increased morbidity and mortality. The dynamic, non-stationary nature of air-pollution time series makes air quality forecasting challenging. This has generated increasing interest in the application of advanced time-series and machine-learning models for air-quality forecasting. In this paper, two novel models are proposed: A Long Short-Term Memory (LSTM) and a Facebook Prophet Model (FPM), in which air pollutants were predicted. We used air pollution data from 10 features in the Northwestern states of Nigeria from 2018 through 2023. This research evaluated prediction results using RMSE, MAE, MAPE, and R^2. The study shows that while both models can effectively predict air pollution levels, the LSTM model had the lowest error rates for certain pollutants in designated states. At the same time, the Facebook Prophet Model displayed broader applicability across multiple pollutants and locations. The study also identified common regional factors contributing to air pollution, potentially tied to industrialization, population growth, and deforestation in northern Nigeria. The results suggest that, for comprehensive air quality monitoring, model selection should consider the specific pollutants and geographical location. Overall, although LSTM demonstrated superior performance for some pollutant-location combinations, the prophet model showed greater generalization across multiple pollutants and locations.

Keywords:

Prediction,
Air pollution,
LSTM,
Facebook Prophet Models

INTRODUCTION

Air pollution, exacerbated by urbanization and globalization, is regarded as a critical societal issue due to its severe impacts on both the environment and human health, contributing to significant morbidity and mortality (Manisalidis et al., 2020). It can be categorized as either natural or anthropogenic, with the latter stemming primarily from human activities such as industrial processes and vehicular emissions, which are associated with population growth and increased industrialization, particularly through coal consumption. Notable historical events, like the Great London Smog of 1952, have illustrated the deadly consequences of air pollution, which is estimated to contribute to 8.1 million deaths worldwide and significantly reduced life expectancy by an average of 1.8 years ("Health Impacts of Air Pollution State of Global Air," n.d.).

In Nigeria, where air pollution remains an important environmental and public health concern, Kano was identified as having the worst air pollution in Africa in 2018 (WHO, 2018). Despite severe health impacts, many developing nations, including Nigeria, lack adequate air quality monitoring, with significant air pollution-related deaths attributed to household practices such as burning solid fuels for cooking (Martin, 2018). In Nigeria, urban pollution originates predominantly from vehicular emissions, generator exhaust, and construction activities, among other sources. Furthermore, the traditional practice of burning grassland within cattle-rearing regions as a strategy to stimulate vegetative regrowth at the onset of the wet season following an extended dry period remains widespread. This agricultural method is environmentally

detrimental, and severe air pollution across several regions of the country, particularly in northern Nigeria, has been attributed thereto (Wambebe, N. M., & Duan, X., 2020).

The types of pollutants contributing to air quality issues in Nigeria include sulfur dioxide (SO_2), carbon dioxide (CO_2), carbon monoxide (CO), nitrogen dioxide (NO_2), ozone (O_3), and particulate matter (PM_{10} and $PM_{2.5}$), with the latter being associated with serious health risks (Ojha et al., 2022). Research has also highlighted correlations between air pollution and COVID-19 mortality rates (Niemi et al., 2022). The variability of pollutants is influenced by local meteorological conditions, which significantly affect pollution levels and health outcomes.

Research emphasizes the effectiveness of using combined historical and real-time data on pollutants and meteorological variables to enhance predictions of air pollution levels (Peralta et al., 2022). Various predictive modeling approaches, including time-series models like ARIMA and machine learning techniques such as Artificial Neural Networks, Support Vector Machines, and hybrid strategies, have been employed to forecast air quality effectively (Ha et al., 2015; Huang et al., 2018; Zhou et al., 2018; Balogun & Adebisi, 2021). Future studies aim to refine these models further, considering environmental fluctuations and data integrity issues (Lu et al., 2018).

The body of literature demonstrate application of methods to improve predictive precision, a methodology that can assimilate large datasets and rely less on model-driven or statistical premises is necessary. In this context, scholars have employed many traditional machine learning methodologies, mostly for classification and regression tasks. This body of literature encompasses, among others, Artificial Neural Networks (Ha et al., 2015); Random Forests (Huang et al., 2018); Support Vector Machines (Zhou et al., 2019; Li et al., 2020); K-Nearest Neighbours or Naïve Bayes (Tella et al., 2021); and Extreme Learning Machine (Sharma et al., 2019). These techniques typically surpass traditional statistical approaches.

Nonetheless, these methods require extensive data processing and incur significant computational delays during both training and prediction (Lu et al., 2018), struggle to handle large nonlinear time-series data, and fail to represent the feature distribution in the air quality dataset accurately (Ma et al., 2018). Deep learning overcomes the limitations of traditional machine learning methods by capturing spatial and temporal patterns in a dataset and reducing local minima. Deep learning methodologies can achieve performance levels on par with human specialists (Torres, Hadjout, Sebaa, Martínez-Álvarez, & Troncoso, 2020). Among the cutting-edge deep learning architectures, convolutional neural networks (CNNs) and recurrent neural networks (RNNs) have exhibited exceptional efficacy in the thorough acquisition of spatial and temporal characteristics, respectively, across various applications (Zhu, Dinh, Phung, & Ha, 2021), such as air quality prediction (Du, Li, Yang, & Horng, 2019).

Long short-term memory (LSTM) is widely used as a type of RNN because of its adaptability and strong performance in predicting and forecasting diverse events. Yafouz et al. (2021) forecasted 03 based on diversity data of air pollutants (NO_2, NO_x, SO_2, and CO) and meteorological factors (wind speed and humidity) utilizing a dataset gathered from three locations in Malaysia, where LSTM and CNN surpassed traditional machine learning models such as SVR, MLP, and XGBoost.

Gokul, Mathew, Bhosale, & Nair (2023) evaluated various machine-learning techniques alongside a deep-learning LSTM model using R^2 and minimal error metrics; ultimately, the LSTM emerged as the superior model, achieving R^2 and MAE values of 0.89 and 5.78, respectively, for PM2.5 prediction.

Naz et al. (2023) evaluated various single-step deep-learning forecasting models, such as LSTM and GRU, alongside statistical models to predict pollutants, specifically NO_2, O_3, SO_2, $PM_{2.5}$, and PM10. Overall, deep learning models performed better, with an R2 of 0.856 and an RMSE of 0.59.

Kristiani et al. (2022) used a KNN model to address missing data and improve pollutant estimates, using meteorological factors such as temperature, humidity, wind direction, and wind velocity. LSTM was employed for forecasting and surpassed the comparative models in RMSE. Raheja and Malik (2022) introduced A-LSTM to forecast air quality in New Delhi, India, using various parameter concentrations. In this forecast, A-LSTM surpassed the comparative models, specifically SVR with LSTM and I-LSTM.

Peralta et al. (2022) employed LSTM to tackle space-time forecasting of PM in Santiago de Chile by integrating historical data of PM_{10}, $PM_{2.5}$, and NO_2 with meteorological factors such as temperature, wind speed, wind direction, and relative humidity, collected from stationary monitoring sites. They forecasted PM_{10} levels for sites lacking measurement data. The research highlighted the enhancement of preprocessing. The model produced commendable outcomes in both 1-hour and 24-hour predictions and suggested incorporating more input variables moving forward, which aligns with the aims of this study.

Farhadi, Palacin, and Blythe (2023) utilised LSTM for their forecasting and incorporated sensor data from the Newcastle urban observatory, which encompassed four themes with various variables, such as air quality (CO, PM_1, $PM_{2.5}$, PM_4, PM_{10}, NO_2, NO, O_3), meteorological conditions (wind direction, wind speed, solar radiation, solar diffuse radiation, pressure, rain duration, rain

accumulation, and maximum wind speed), traffic metrics (traffic flow and average speed), and timestamps (year, month, day, hour, minute, and second). The readings were recorded at one-second intervals.

However, several studies have combined LSTM with other attention-driven frameworks or deep learning architectures to improve precision, effectiveness, and forecasting performance. Z. Zhang et al. (2023) used detrended cross-correlation analysis (DCCA) to evaluate the relationship between pollution and meteorological factors across short-, medium-, and long-term predictions, improving LSTM accuracy and efficiency.

Wen et al. (2023) employed the Hunter-prey optimization approach to fine-tune LSTM hyperparameters, resulting in an effective model.

In early 2017, Facebook's primary data team created the Facebook Prophet Model (FPM), a time-series model that uses a data decomposition method for more user-friendly analysis. Since then, scholars worldwide have adopted the method for diverse forecasting applications, such as predicting air pollution. It has demonstrated strong predictive capability for air pollutant levels and can be used to anticipate air pollution in different areas (Hasnain et al. 2022). Khayyat, Laabidi, Almalki, and Al-Zahrani (2021) utilized the Facebook Prophet model to forecast the COVID-19 epidemic in Saudi Arabia. The model aligned with the data and generated forecasts. The model predicted the COVID-19 dataset with greater precision than the alternative models.

Hasnain et al. (2022) used the Prophet model to analyse and forecast pollutants over time, anticipating both short-term and long-term air quality in Jiangsu Province, China. Chérif, Snoun, Bellakhal, and Kanfoudi (2023) employed the Facebook Prophet model to forecast ozone levels in Tunisia. They achieved a 7% improvement over the comparative model.

The long short-term memory (LSTM) (Greff et al., 2016), an adaptation of the recurrent neural network (RNN) that effectively addresses the vanishing gradient problem, has demonstrated its efficacy in modeling long-term dependencies and causal relationships in data, particularly regarding pollutant levels and meteorological factors, while the Facebook Prophet model (FPM) shows significant potential for air quality assessment and prediction (Jin et al., 2021).

Despite decades of studies in air pollution, Northwestern region of Nigeria where farming practice and household activities is the most prominent contributor of carbon emissions, there are still very limited research on prediction of the pollutants of the ambient air. This research aims to evaluate the performance of the Long Short-Term Memory (LSTM) and Facebook Prophet Model (FPM) using available data to identify the most effective model for a specific region and assess their prediction efficacy for air pollution data.

METHODOLOGY

Data Description

The dataset consists of different gaseous pollutants, including SO_2, CO, CO_2, SO_4, and particulate matter ($PM_{2.5}$ and PM_{10}); meteorological variables, including wind speed, minimum temperature, maximum temperature, and rain.

Missing values were checked in this research using the using isnull() method and were filled using ffill() method in pandas. This is particularly useful in time series data where the assumption that the next timestamp will have the same value as the last known is reasonable

LSTM Model Structure

According to Bai et al. (2018), LSTMs are nothing but a stack of neural networks composed of linear layers consisting of weights and biases, just like any other standard neural network. The weights are constantly updated by back-propagation. A typical long short-term memory (LSTM) architecture can be seen in Figure 1 below:

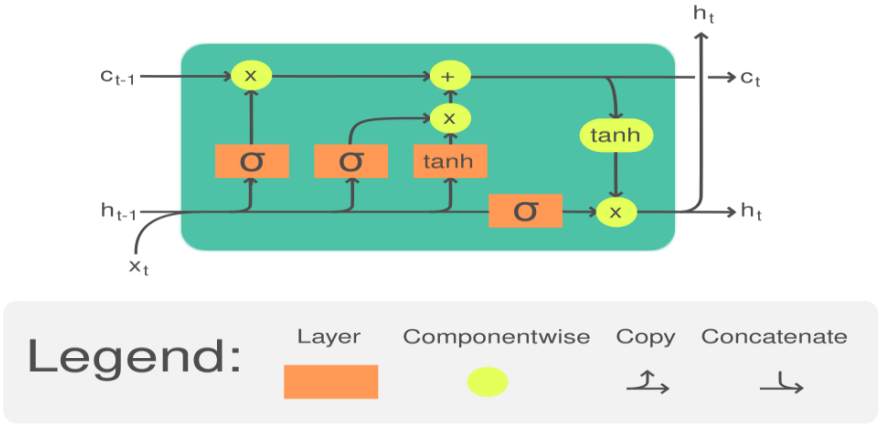


Figure 1: LSTM Architecture
Source: Wikipedia

A typical LSTM has two major components, namely:

(1) Cell: Every unit of the LSTM network is referred to as a cell and consists of three inputs and two outputs. The inputs are
- $x(t)$ = input at time, t
- $h(t-1)$ = previous hidden state at time
- $c(t-1)$ = previous cell state at time

and the outputs are
- $h(t)$ = the updated hidden state to predict the output
- $c(t)$ = the current state

(2) Gates; LSTM uses a special theory of controlling the memorizing process. Popularly referred to as the gating mechanism in LSTM, what the gates in LSTM do is, store the memory components in analog format, and make it a probabilistic score by doing point-wise multiplication using sigmoid activation function, which stores it in the range of 0–1. Gates in LSTM regulate the flow of information in and out of the LSTM cells.

Gates are of 3 types, namely,

(i) Input Gate: The input gate optionally permits information that is relevant to the current cell state. It is the gate that determines which information is necessary for the current input and which isn't by using the sigmoid activation function. It then stores the information in the current cell state. Next, it comes to play the tanh activation mechanism, which computes the vector representations of the input-gate values, which are added to the cell state.

$$i_t = \sigma(W_i \bullet [h_{t-1}, x_t] + b_i) \quad (1)$$

$$\tilde{C}_t = tanh(W_C \bullet [h_{t-1}, x_t] + b_C) \quad (2)$$

(ii) Forget Gate: This gate is responsible for removing unnecessary information. Since the hidden state is responsible for predicting outputs. The output generated from the hidden state at (t-1) timestamp is h (t-1). After the forget gate receives the input x (t) and output from h (t-1), it performs a point-wise multiplication with its weight matrix with an add-on of sigmoid activation, which generates probability scores. These probability scores help it determine what useful information is and what is irrelevant.

$$f_t = \sigma\left(W_f \bullet [h_{t-1}, x_t] + b_f\right) \quad (3)$$

The cell state update mechanism follows after the forget gate. By now, the input gate remembers which tokens are relevant and adds them to the current cell state with tanh activation enabled. Also, the forget gate output, when multiplied with the previous cell state C(t-1), discards the irrelevant information. Hence, by combining these two gates' jobs, our cell state is updated without any loss of relevant information or the addition of irrelevant ones.

$$C_t = f_t \bullet C_{t-1} + i_t \bullet \tilde{C}_t \quad (4)$$

(iii) Output Gate: This gate, which pretty much clarifies from its name that it is about to give us the output, does a quite straightforward job. The output gate decides what to output from our current cell state. The output gate also has a matrix where weights are stored and updated by back-propagation. This weight matrix takes in the input token x(t) and the output from the previously hidden state h(t-1) and does the same old point-wise multiplication task. However, as said earlier, this takes place on top of a sigmoid activation, as we need probability scores to determine what will be the output sequence.

After we get the sigmoid scores, we simply multiply them with the updated cell state, which contains some relevant information required for the final output prediction. A final tanh multiplication is applied at the very last to ensure the values range from [-1,1], and our output sequence is ready!

$$o_t = \sigma(W_0[h_{t-1}, x_t] + b_0) \quad (5)$$

$$h_t = o_t \bullet tanh(C_t) \quad (6)$$

Where $W_i, b_i; W_c, b_c; W_f, b_f; W_0, b_0$ are weights and biases for the input gate, current state, forget gate, and output gate, respectively.

The feature values of the data in this study was normalized using Min-Max Scalar such that the minimum and maximum values of the predictors were mapped to 0 and 1. This can be done using equation (7) below:

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad (7)$$

Where: x_{norm} = normalized value; x = actual value; x_{min}= minimum value; x_{max}= maximum value.

The LSTM model equations from equation (1) to equation (6) were executed using the data from the custody and the parameters used are as follows:

Sequential model: The simplest way of building a model in Keras, where layers are added in sequence.

LSTM layers: Long Short-Term Memory (LSTM) layers are a type of recurrent neural network (RNN) used for time-series data or sequential data. The units' parameter specifies the number of LSTM cells in the layer.

First LSTM layer has 50 units, and return sequences=True ensures that the output is a sequence, allowing the next LSTM layer to receive sequences as input.

Second LSTM layer has 25 units, also with return sequences=True.

Third LSTM layer has 16 units, with return sequences=False as it's the final LSTM layer and is

expected to return a single vector (no need to return sequences).

Dropout (20%): The Dropout layer is a regularization technique to prevent overfitting by randomly setting a fraction of input units to 0 during training.

Output layer with 1 unit for regression task: model.add (Dense (units=1))

Compile the model: model.compile (optimizer='adam', loss='mean_squared_error') model.summary ()

Early stopping to avoid overfitting: early_stopping = EarlyStopping (monitor='val_loss', patience=20, restore_best_weights=True)

Train the model: model.fit (X_train, Y_train, epochs=50,batch_size=8, validation_split=0.2, callbacks=[early_stopping], verbose=1)

Output Layer (Dense):
model.add (Dense (units=1)): This adds a fully connected layer with a single neuron, which is typical for regression tasks where you're predicting a single continuous value.

Since this is a regression task, there's no activation function in the output layer (like sigmoid or softmax for classification), because we want the network to output raw values.

Compile the Model:
optimizer='adam': The Adam optimizer is an adaptive learning rate method widely used in deep learning.

Model Summary: model.summary (): This prints a summary of the model's architecture, including the layer types, output shapes, and the number of trainable parameters.

Early Stopping: This callback monitors the validation loss during training and stops the process if it doesn't improve for 20 consecutive epochs (patience). By setting restore_best_weights=True, the model will revert to the best weights encountered during training, helping to avoid overfitting.

Model Training: This starts training the model on the training 70% (3024) data.

Epochs=50: Train the model for up to 50 epochs.

batch_size=8: The batch size, meaning that updates to the model are made every 8 samples.

validation_split=0.2: 20% of the training data is used as validation data.

Callbacks=[early_stopping]: Early stopping is applied to prevent overfitting and reduce unnecessary epochs.

Verbose=1: Displays detailed output for each epoch.

The test data used was 30% (1296) observations

The model is well-structured for a regression task with built-in mechanisms for preventing overfitting via dropout and early stopping.

Facebook Prophet Model (FPM)
The Facebook Prophet Model is an open-source algorithm developed by Facebook's core data team in 2017 that generates time series models. The model takes the following framework:

$$y(t) = c(t) + s(t) + x(t) + e_t (7)$$

Where y(t): the observed value at time, t.
c(t): the trend component (current growth).
s(t): the seasonal component
x(t): the exogenous impacts.
e_t: the error term (normally distributed noise)
Note: exogenous impacts are split into two: (a) holiday effects, say h(t), and (b) external regressors, say xi(t), which are other external drivers.
The Facebook Prophet Model decomposes the time series into three main components:

1. **Trend:** This captures the long-term increase or decrease in the time series. The trend can be modeled using either a linear or logistic growth function as
   Linear trend $c(t) = c_0 + k \bullet t$ (8)
   Where c_0: initial value at t=0
   k: growth rate
   Logistic trend
   $c(t) \frac{c}{1+exp(-k(t-t_0))}$ (9)
   where c: carrying capacity (maximum value)
t_0: inflection point (where the growth rate changes)

2. **Seasonality**: Captures repeating patterns (e.g., daily, weekly, yearly). Seasonality is typically modelled using Fourier series:

$$s(t) = \sum_{n=1}^{N} \left(a_n \cos(\frac{2\pi nt}{p}) + b_n \sin(\frac{2\pi nt}{p}) \right)$$

Where N: number of Fourier terms
P: period of seasonality (e.g 365 for yearly seasonality)
a_n; b_n: Fourier coefficients

3. **Exogenous impacts**: This allows for the inclusion of specific events that might affect the time series. For instance, the holiday effect is modelled as a set of binary variables indicating the presence of a holiday, each with an associated effect:

$$h(t) = \sum_{k=1}^{K} \Upsilon_k \bullet I_k(t) \quad (11)$$

Where Υ_k: effect of holiday k

$I_k(t)$: Indicator function that is 1 if holiday k occurs and 0 otherwise.

RESULTS AND DISCUSSION

This study investigates air pollution prediction in six (6) states in the Northwestern states of Nigeria including Jigawa, Kano, Kaduna, Katsina, Sokoto, and Zamfara. The observations were taken through six (6) from 1/1/2018 to 12/1/2023 inclusive amounting to 4320 observations. The dataset includes gaseous pollutants (SO_2, CO, CO_2, and SO_4), particulate matter ($PM_{2.5}$ and PM_{10}), which are the most frequently reported, and atmospheric variables, including maximum and minimum temperature and rainfall, with geographical coverage the six aforementioned northern states. The literature review encompasses statistical models for air pollution prediction, machine learning and deep learning algorithms, and Facebook Prophet models. The methodology employs a variant of recurrent neural networks, specifically the long short-term memory (LSTM) model, alongside the Facebook Prophet Model for state-level air pollution forecasting. Predictive performance is evaluated using four metrics: root mean squared error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and R-squared (R^2). RMSE is highlighted for its effectiveness in quantifying regression model predictive performance. The analysis identifies a model suitable for location-specific air pollution prediction and compares the models' generalisation capabilities.

(a) The Time Series Plot of Temporal Distribution of Different Gaseous Pollutants

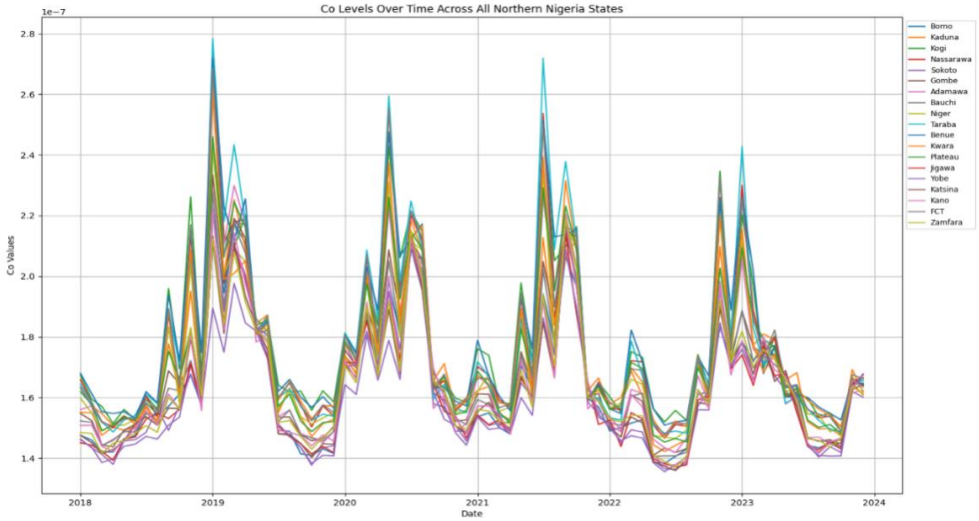


Figure 2: CO_2 levels for all northern Nigeria states from 2018 to 2024 (source: Author)

Figure 2 indicates the temporal distribution of CO emissions across Northern Nigeria and showcases potential cyclical patterns in pollution levels. Dry or wet seasons, agricultural practices (e.g., burning biomass), and variations in industrial activities could influence the cyclical nature of CO levels. Large spikes might be associated with specific events like fires, industrial surges, or increased vehicular emissions during certain periods.



Figure 3: CO_2 levels for all northern Nigeria states from 2018 to 2024 (source: Author)

Figure 3 shows a time plot of observed CO_2 levels across the states. A noticeable upward trend in CO_2 levels is observed during the period, with fluctuations, and the overall pollutant levels have been steadily increasing, with subtle differences between the states suggesting that common regional factors across the states influence CO_2 emissions, potentially attributed to increased industrialization, population growth, or deforestation, which is very common in northern Nigeria.

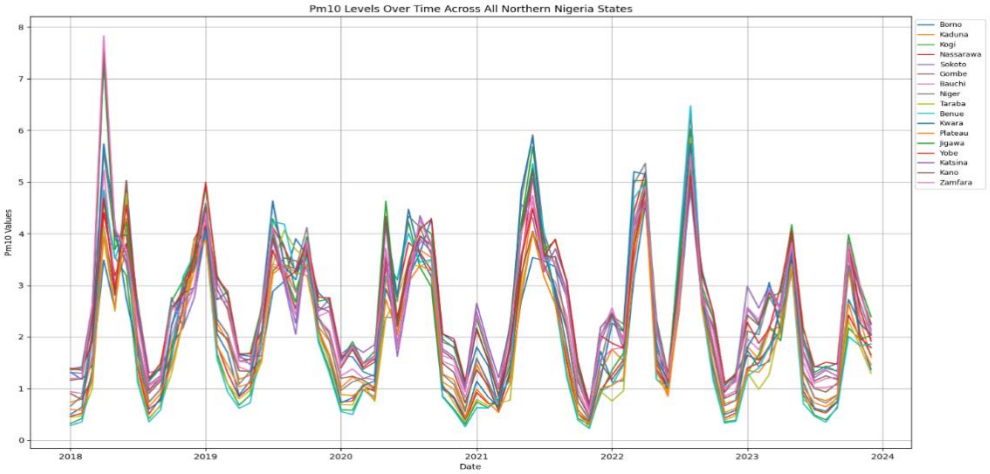


Figure 4: PM_{10} levels for all northern Nigeria states from 2018 to 2024 (source: Author)

The PM10 time plot depicted in Figure 4 shows **seasonal or cyclical fluctuations** in PM10 levels across northern Nigerian states, with **recurrent peaks** likely driven by regional environmental factors such as dust storms during the dusty trade wind or Harmattan season, which blows across the region from the Sahara and may explain the regular spikes in PM10 levels at certain times of the year.

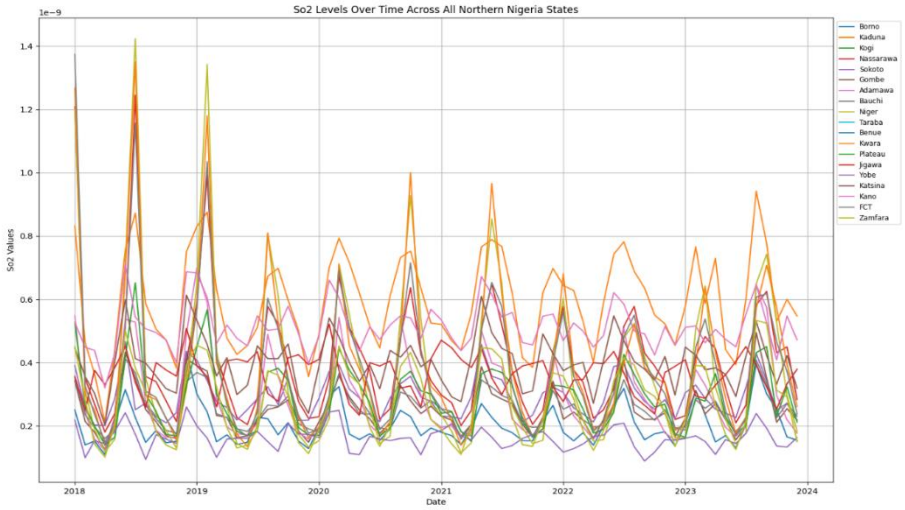


Figure 5: S02 levels for all northern Nigeria states from 2018 to 2024 (source: Author)

Different states have different $S0_2$ concentrations, and Figure 5 below shows higher concentrations between 2018 and early 2019. Seasonal and location-based environmental factors may influence variability in concentration levels.

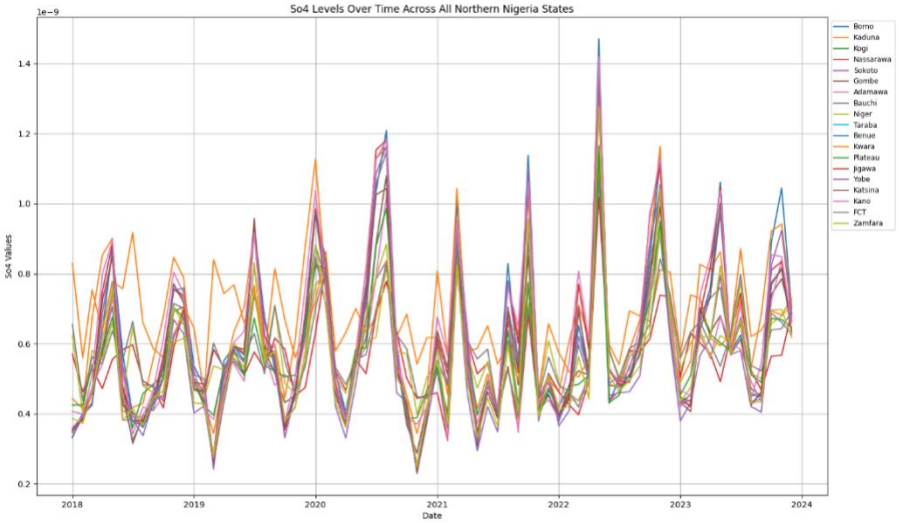


Figure 6: S04 levels for all northern Nigeria states from 2018 to 2024 (source: Author)

In Figure 6, dramatic fluctuations in $S0_4$ across most of the states and fewer deviations in pattern are apparent. Common factors contributing to the pollutant concentration can be reasonably found in the northern states.

(b) Model Predictions

Particulate Matter ($PM_{2.5}$) Model Prediction

Figures 7 and 8, respectively, visualize the predictive performance of LSTM and FPM models for particulate matter ($PM_{2.5}$). The historical data in the LSTM model in Figure 7 indicated strong fluctuations with large spikes (more pronounced around 2019) and drops, showcasing volatile air quality in $PM_{2.5}$ levels across the state; these fluctuations suggest the presence of seasonal or local factors, while the predicted/forecasted $PM_{2.5}$ levels from mid-2023 to 2024 show somewhat stable upward and downward (cyclic) $PM_{2.5}$ levels, aligning with the

historical data's apparent seasonality and suggesting that the model expects fewer extreme pollution events. This prediction means that while extreme spikes are less likely, air quality may remain moderately poor overall, around 50-75 µg/m³ (slightly above WHO guidelines), without dramatic spikes or declines. As the actual data is more volatile, the plot indicates a need for additional model refinement or consideration of external factors.

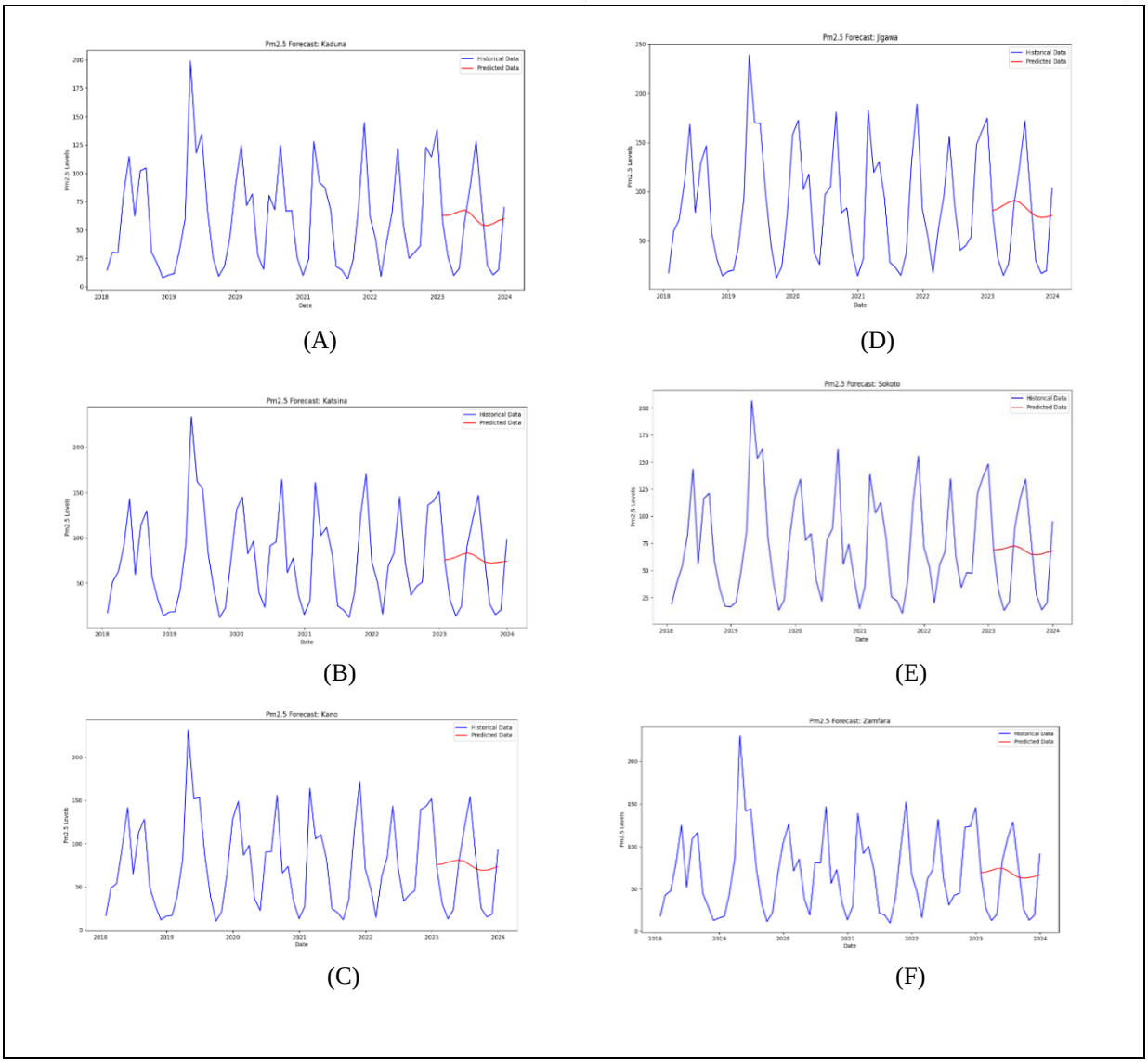


Figure 7: LSTM Prediction and Forecast for $PM_{2.5}$

The black dots represent the observed PM2.5 levels in Figure 8; significant variability and a repeating pattern are also evident, with peaks and troughs occurring almost every year The predicted or smoothed trend also follows a cyclical pattern like the actual data; the line maintained relatively consistent fluctuations around an average of 50 µg/m³ through the forecast (2025). The Prophet model fits the historical data better than the LSTM model, capturing major peaks and troughs. The confidence band tends to be narrow, suggesting low $PM_{2.5}$ levels in the future; this trend may be attributed to sanitation and infrastructure development in the states.

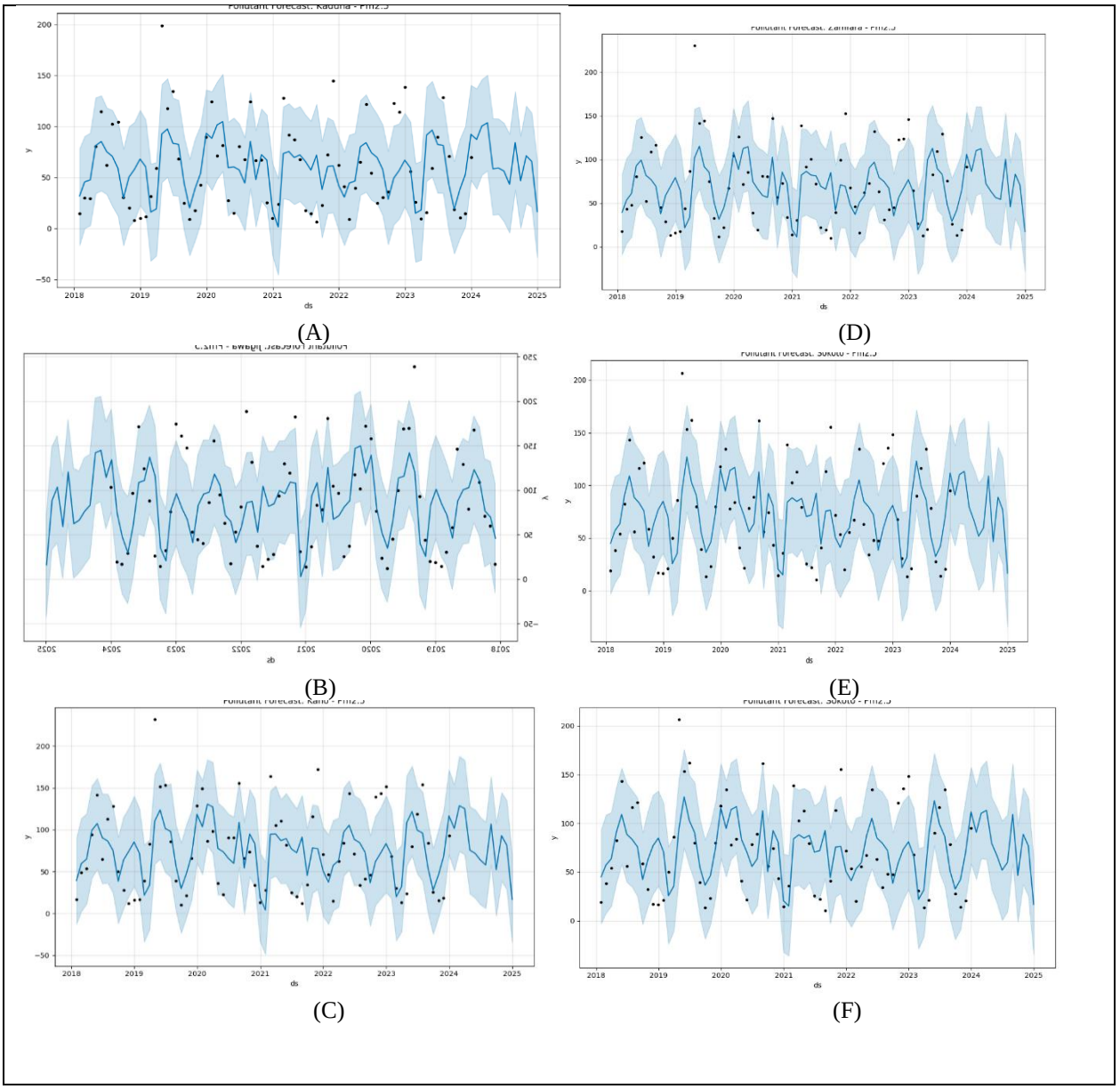


Figure 8: FPM Prediction and Forecast for $PM_{2.5}$

RMSE Performance Comparison of the Models

In general, RMSE represents the difference between the predicted and actual values. The smaller the distance between the two, the better the model's predictive performance.

The graph in Figure 9 illustrates the RMSE (Root Mean Square Error) comparison between LSTM (blue line) and Prophet (green line) across the states in the case study. A notable difference in prediction performance is observed between the contrastive models in Kaduna and Sokoto, where the Prophet model has lower RMSEs of 36, 39, and 51 in Kaduna, Sokoto, and Kano, respectively. These results suggest that the Prophet model performs better at predicting pollutants in these states; the contrastive models performed similarly well in Katsina and Jigawa. However, LSTM obtained notably better accuracy in Zamfara. Overall, the Prophet Model is the best predictor for PM2.5.

Table 1: PM2.5 RMSE VALUES

RMSE FOR $PM_{2.5}$ POLLUTANT		
STATE	LSTM	FPM
KADUNA	46.17125	36.42122
SOKOTO	48.43757	38.82825
JIGAWA	58.06085	58.06085
KATSINA	51.4292	51.4292
KANO	52.41089	51.4292
ZAMFARA	47.74106	52.41089

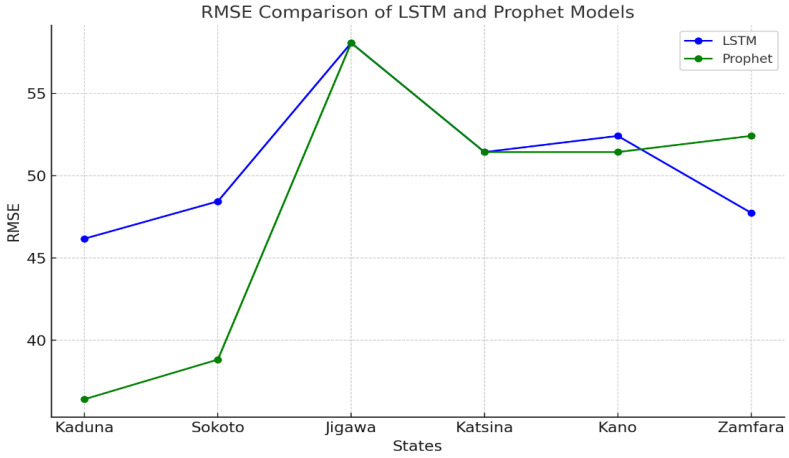
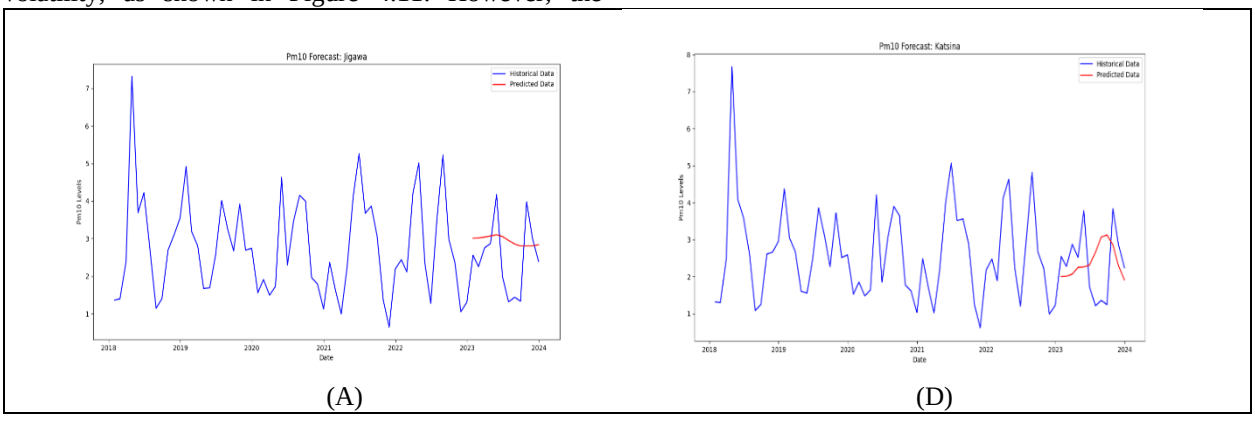


Figure 9: $PM_{2.5}$ RMSE comparison of LSTM and Prophet models

Particulate Matter (PM_{10}) Model Prediction

Similarly, past observations of particulate matter (PM10) showed high initial peaks across all states, indicating extreme PM10 pollution. A moderately high volatility with significant peaks and troughs continued, suggesting varying ambient air pollution conditions across the states, compared with PM2.5, which shows lower volatility, as shown in Figure 4.11. However, the predicted trend appears smoother than the historical trend (and compared with PM2.5), suggesting lower volatility relative to PM2.5 and indicating potentially lower or insignificant air pollution levels in all states except Kano and Katsina, where an upward trend is anticipated.



(A) (D)

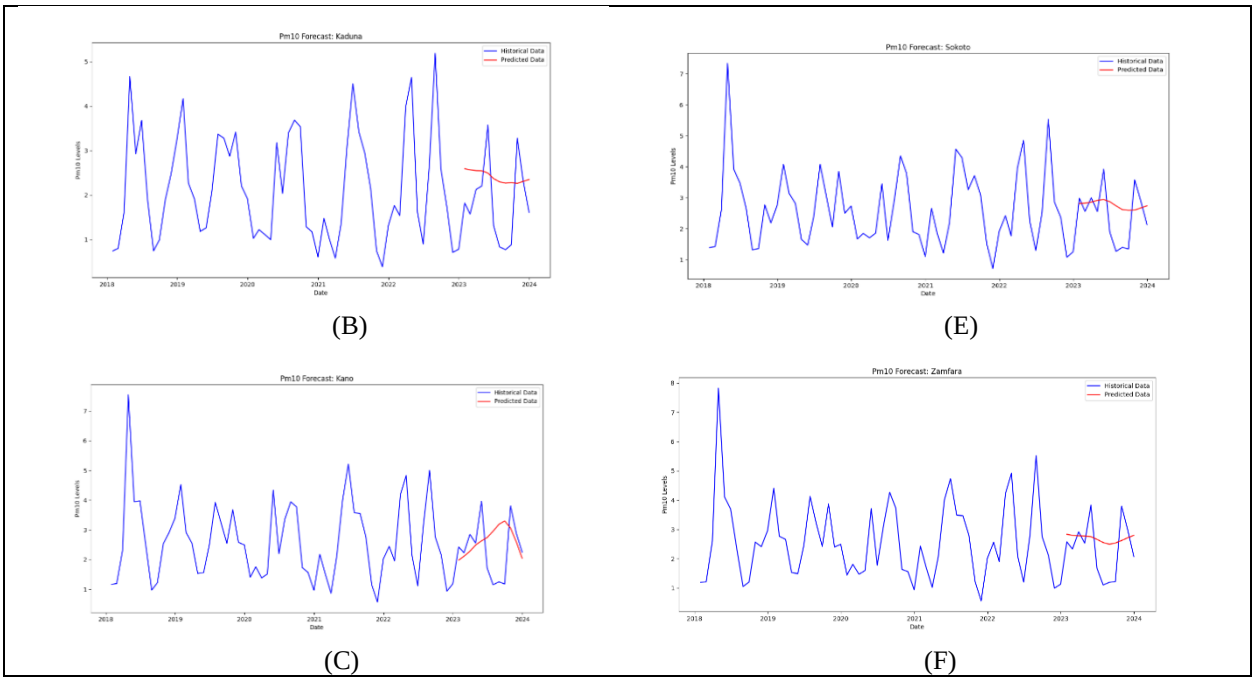
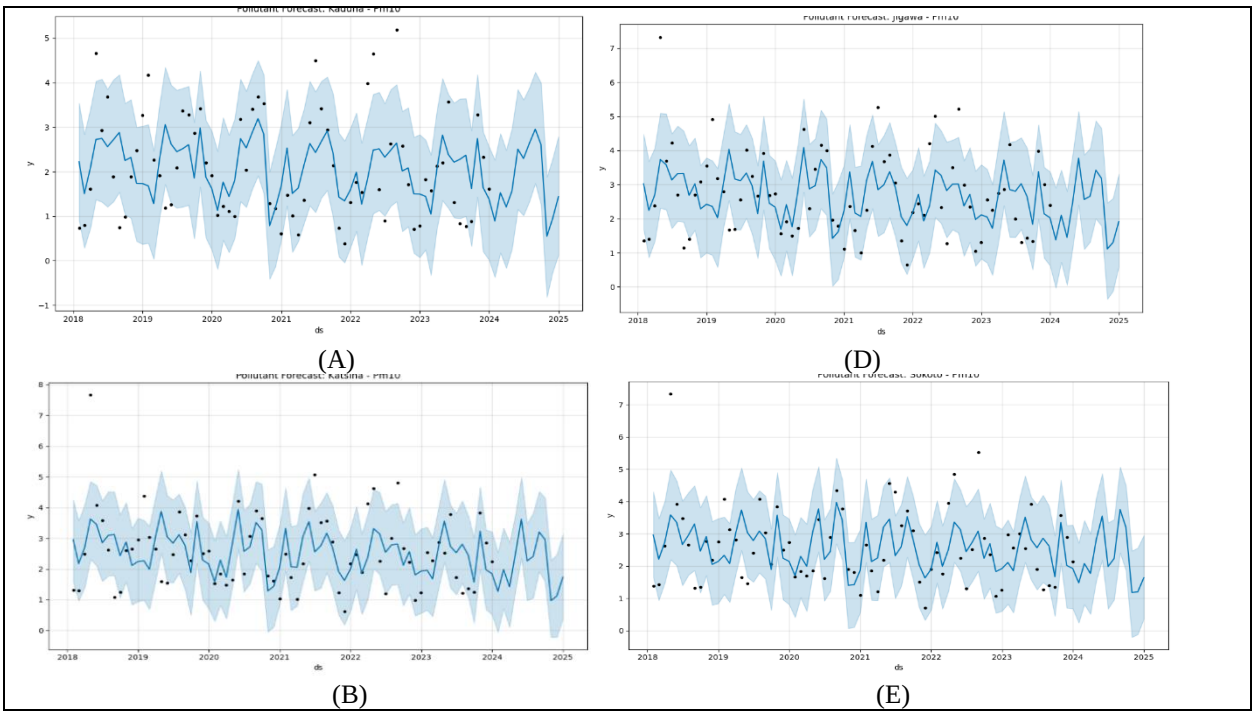
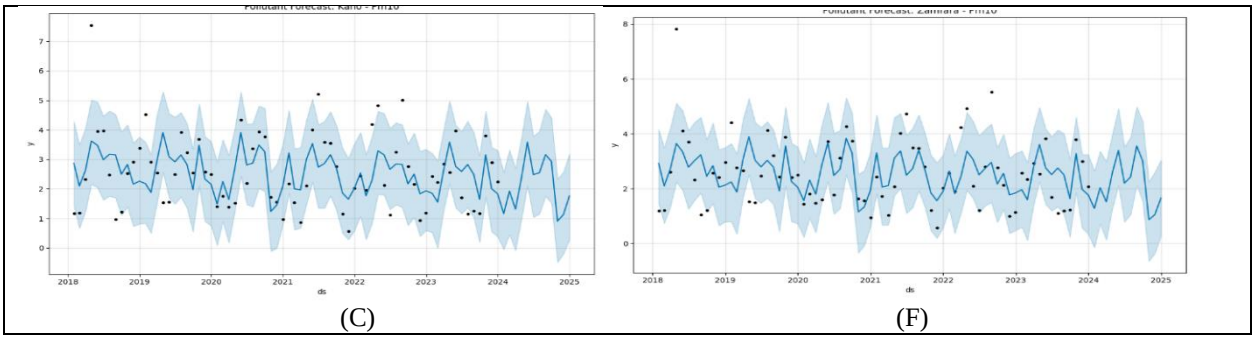


Figure 10: LSTM prediction for PM_{10}.

The graphs in Figure 11 show persistent downward volatility across all states; the decline indicates low particulate matter levels. It is worth noting that the shaded area of the graphs indicates the confidence interval of the forecast; it also reveals that PM_{10} levels range between (-1 and +5), meaning the particulate pollution levels could be zero or as high as 5 µg/m³ owing to future seasonal or weather conditions.





(C) (F)

Figure 11: Prophet Forecast for PM_{10}

RMSE Performance Comparison Of The Models

Figure 12 compares the performance of LSTM (blue line) and Prophet (dashed green line). For most states, including Jigawa, Katsina, Kano, and Zamfara, both models produced nearly identical results, with the highest error in Katsina. However, the Prophet model produced more consistent predictions, with less fluctuation across states.

Table 2: PM_{10} RMSE VALUES

RMSE FOR PM_{10} POLLUTANT		
STATE	LSTM	FPM
KADUNA	1.106547	1.029591
SOKOTO	0.921692	1.004266
JIGAWA	1.075562	1.075562
KATSINA	1.186164	1.186164
KANO	1.224998	1.224998
ZAMFARA	1.010746	1.010746

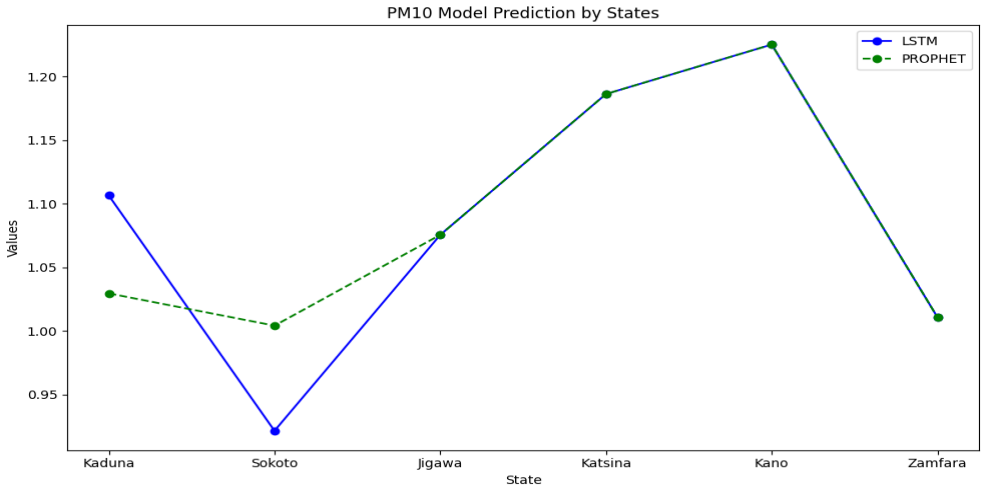


Figure 12: PM_{10} RMSE comparison of LSTM and Prophet models

$C0_2$ Model Prediction

Historical data across all states show an upward trend, indicating a significant increase in CO_2 pollutant levels, likely driven by rising automobile emissions and factory activity. The predicted trends suggest a slight stabilization but a noticeable increase in pollution levels in the future.

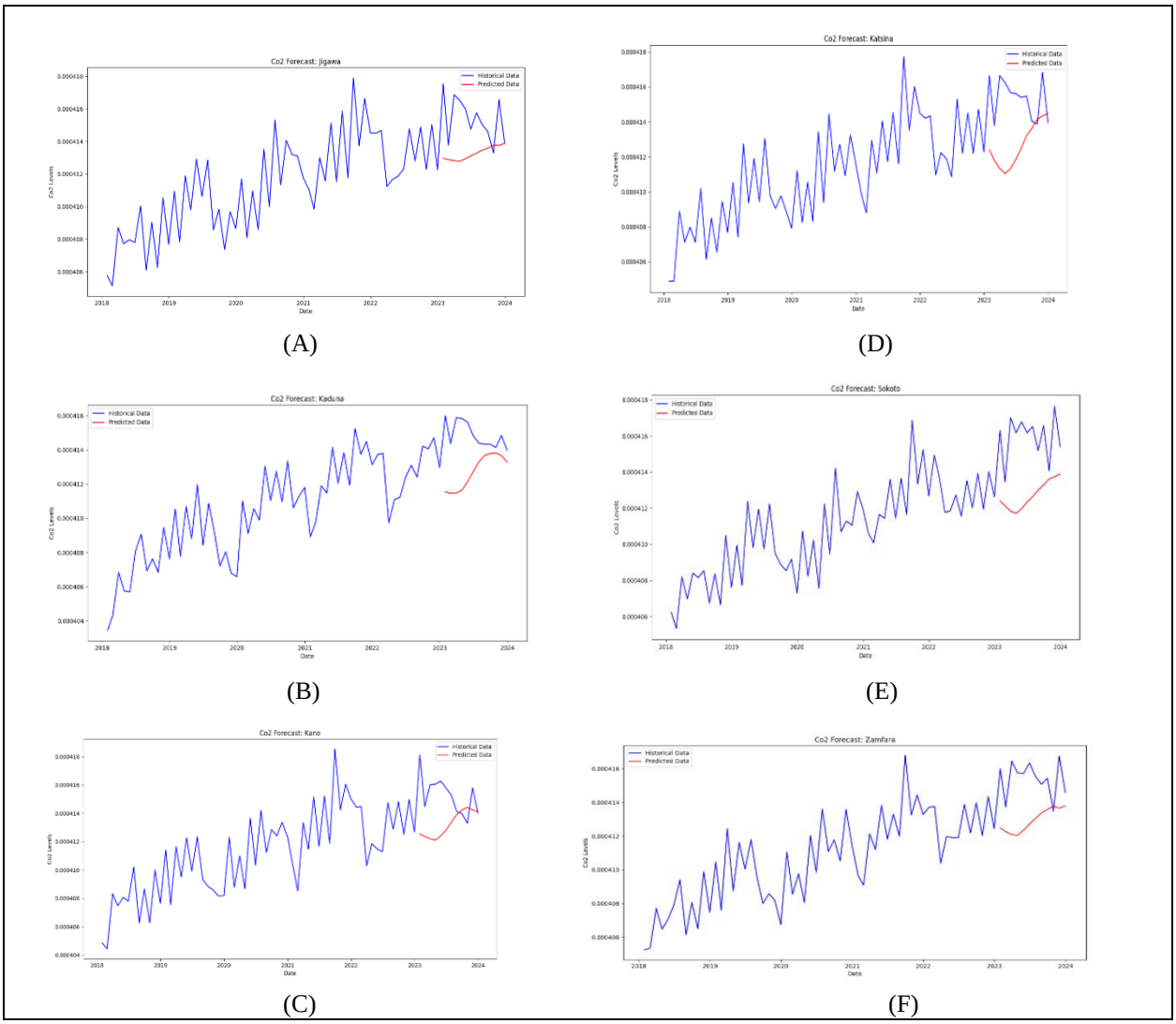
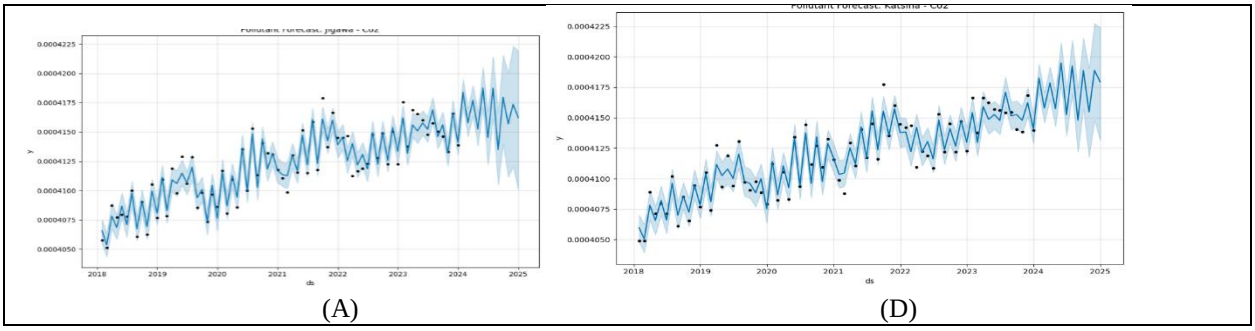


Figure 13: LSTM prediction for $C0_2$.



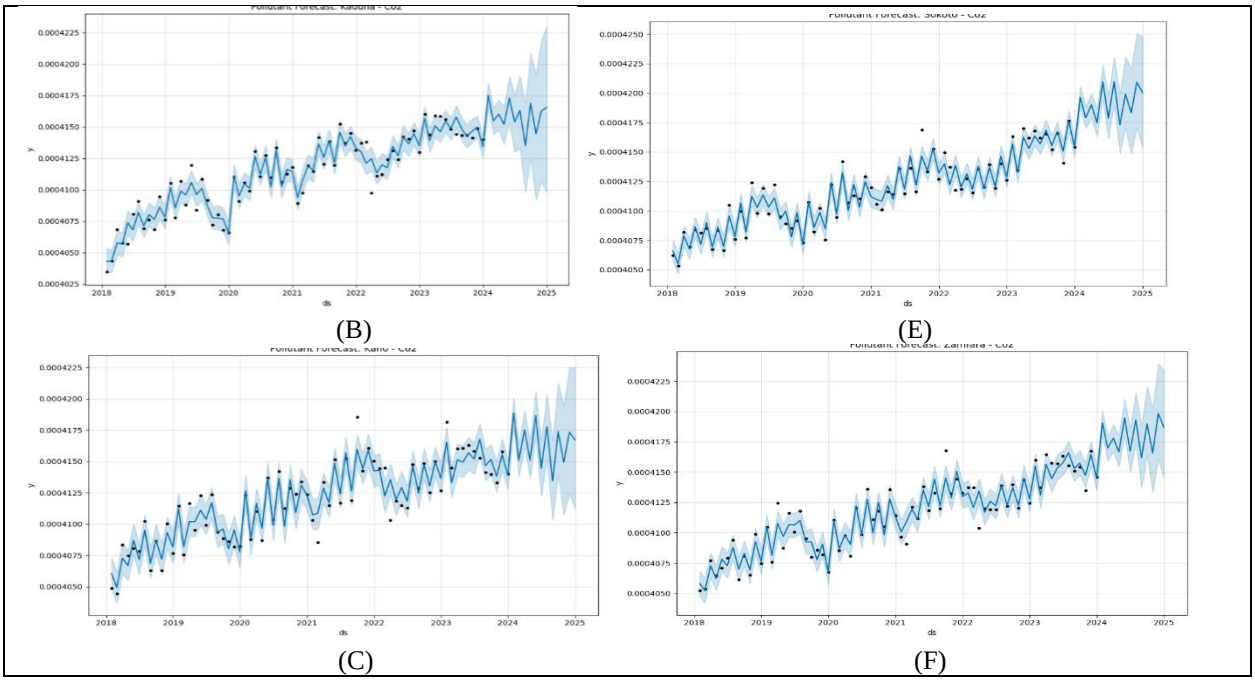


Figure 14: Prophet Prediction for CO_2

C02 RMSE Performance Comparison of the Models

The contrastive models shown in Figure 15 below show similar performance in predicting CO_2 for Jigawa, Katsina, and Zamfara, indicating similar emission patterns across the states. However, the Prophet model outperformed the LSTM prediction for Sokoto and Kaduna with a significant difference. This suggests using the Prophet model, especially in Sokoto and Kaduna, for CO2 prediction and forecasting.

Table 3: $C0_2$ RMSE Values

RMSE FOR $C0_2$ POLLUTANT		
STATE	LSTM	FPM
KADUNA	2.53E-06	7.23E-07
SOKOTO	3.38E-06	6.83E-07
JIGAWA	2.57E-06	2.57E-06
KATSINA	3.21E-06	3.21E-06
KANO	2.75E-06	2.75E-06
ZAMFARA	2.67E-06	2.67E-06

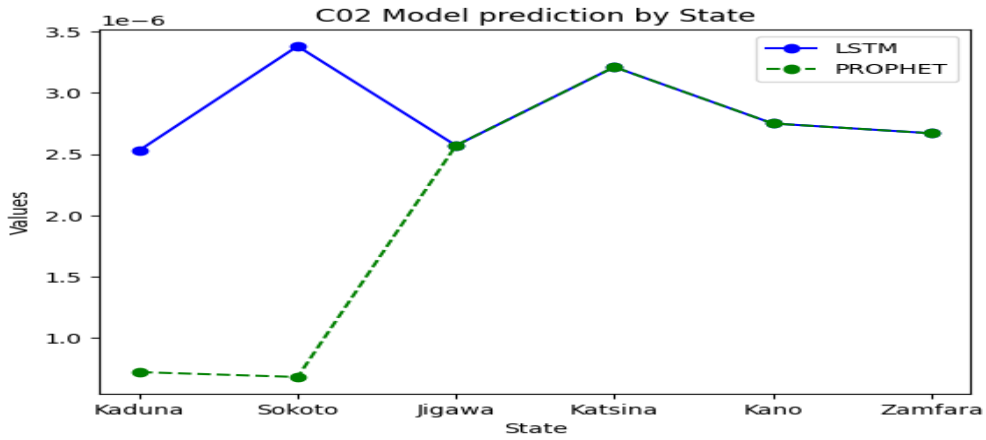


Figure 15: $C0_2$ RMSE comparison of LSTM and FPM Model

CO Model Prediction

Significant peaks and drops in CO pollutant levels are observed throughout the periods. CO levels are high, likely due to increased emissions from more motor vehicles and factories. The predicted data followed the past pattern, with significant fluctuations in both figures.

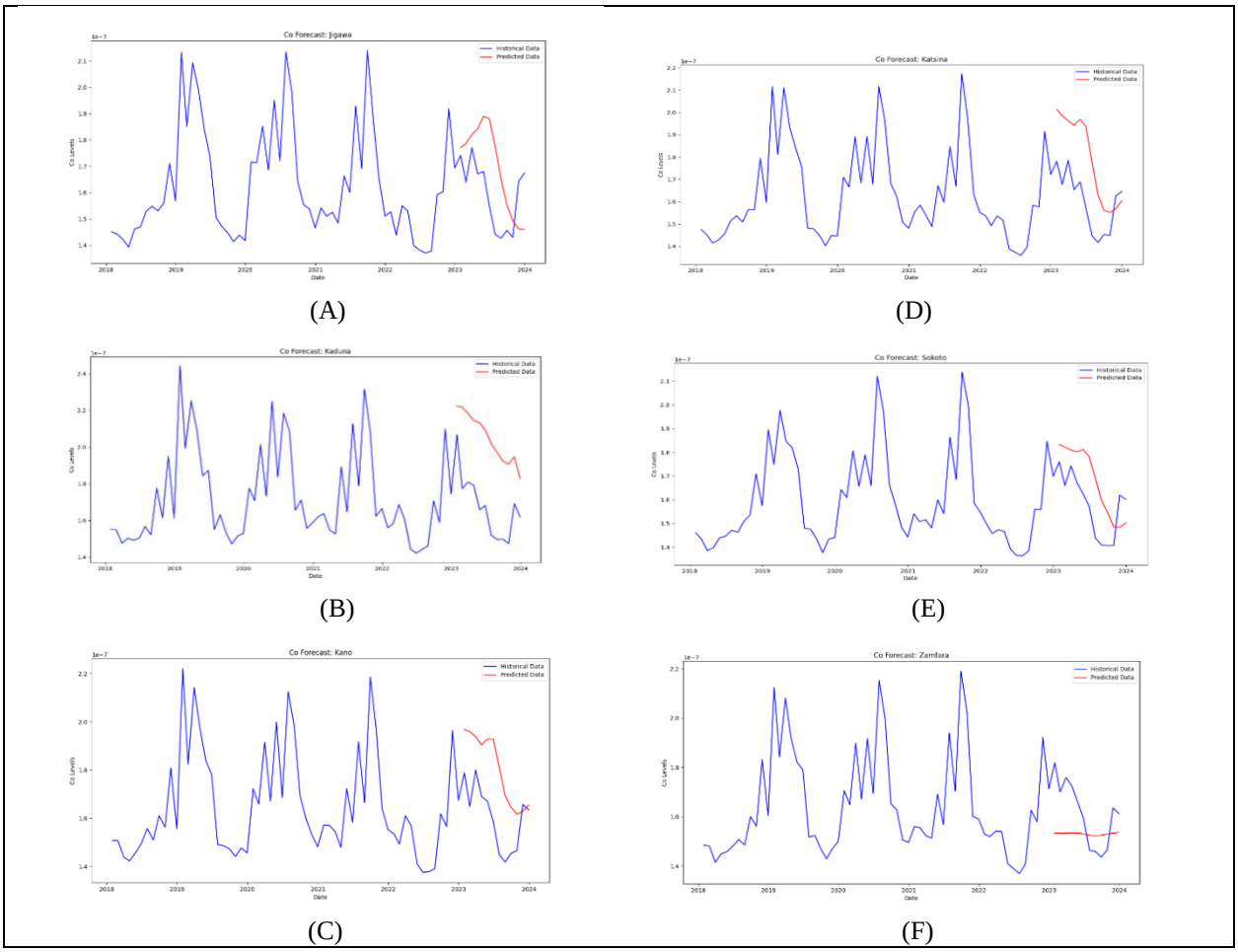


Figure 16: LSTM prediction for CO

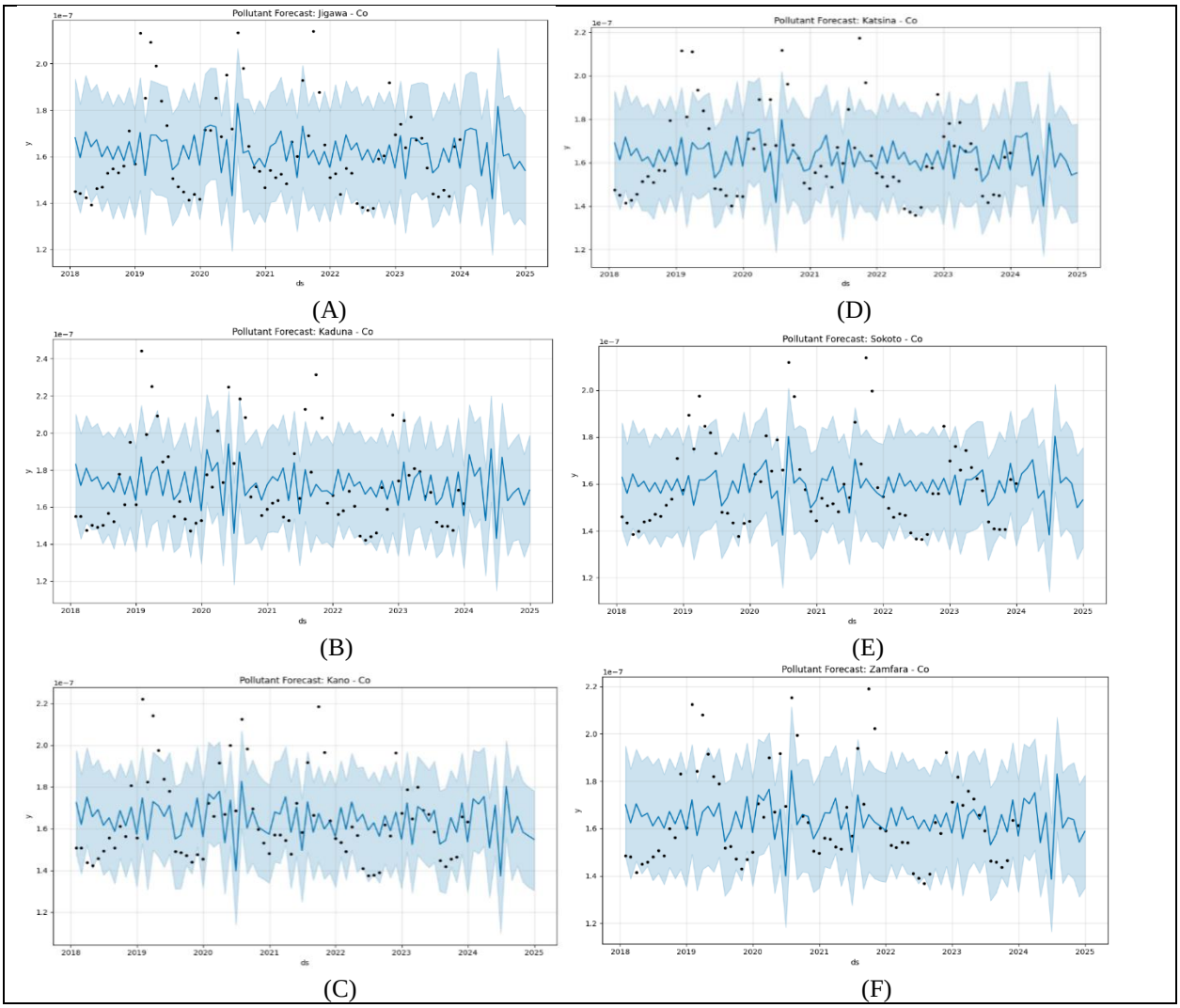


Figure 17: FPM Prediction for CO

CO RMSE Performance Comparison of the Models

From the graph below, a noticeable performance difference between the two models is evident between Kaduna and Sokoto. The LSTM model predicts slightly higher error values, especially in Kaduna State, and the lowest error in Sokoto State, potentially because it captures more complex temporal dependencies or fluctuations in pollutant levels. However, Prophet has lower prediction error in Kaduna while matching LSTM's performance in the remaining states.

Table 4: CO RMSE Values

RMSE FOR CO POLLUTANT		
STATE	LSTM	FPM
KADUNA	3.82E-08	2.21E-08
SOKOTO	1.21E-08	1.73E-08
JIGAWA	1.53E-08	1.53E-08
KATSINA	2.07E-08	2.07E-08
KANO	2.14E-08	2.14E-08
ZAMFARA	1.53E-08	1.53E-08

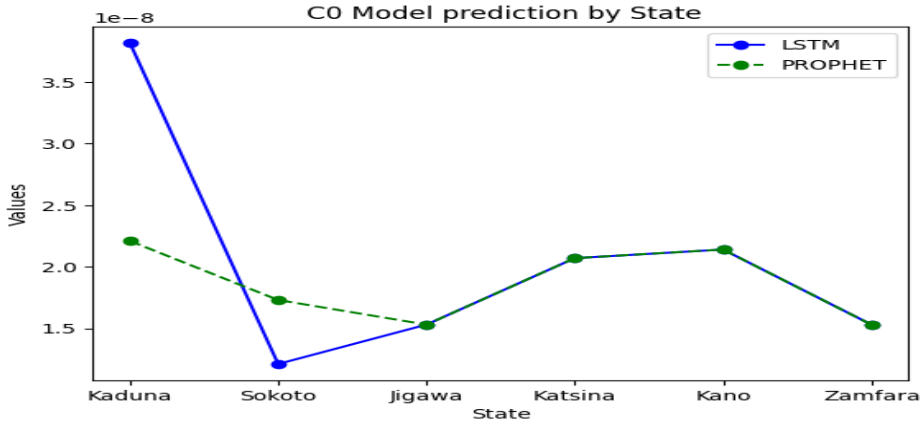


Figure 18: CO RMSE comparison of LSTM and Prophet Model

SO_2 Model Prediction

Figure 19 shows historical SO_2 in Jigawa (A), Kaduna (B), Kano (C), and Katsina (D), indicating strong cyclical behavior, potentially linked to seasonal factors or recurring industrial/urban activities. At the same time, the predictions are smoother, potentially indicating the model's confidence in a gradual shift or in SO_2 levels. The forecast predicts a small increase in SO_2 levels in the same states in 2024. In Sokoto and Zamfara, the predicted data suggest the model anticipates stabilization in SO_2 levels but still reflects the periodic increases seen in historical data; forecast precision decreases further into the future, especially in Sokoto (E) and Zamfara (F).

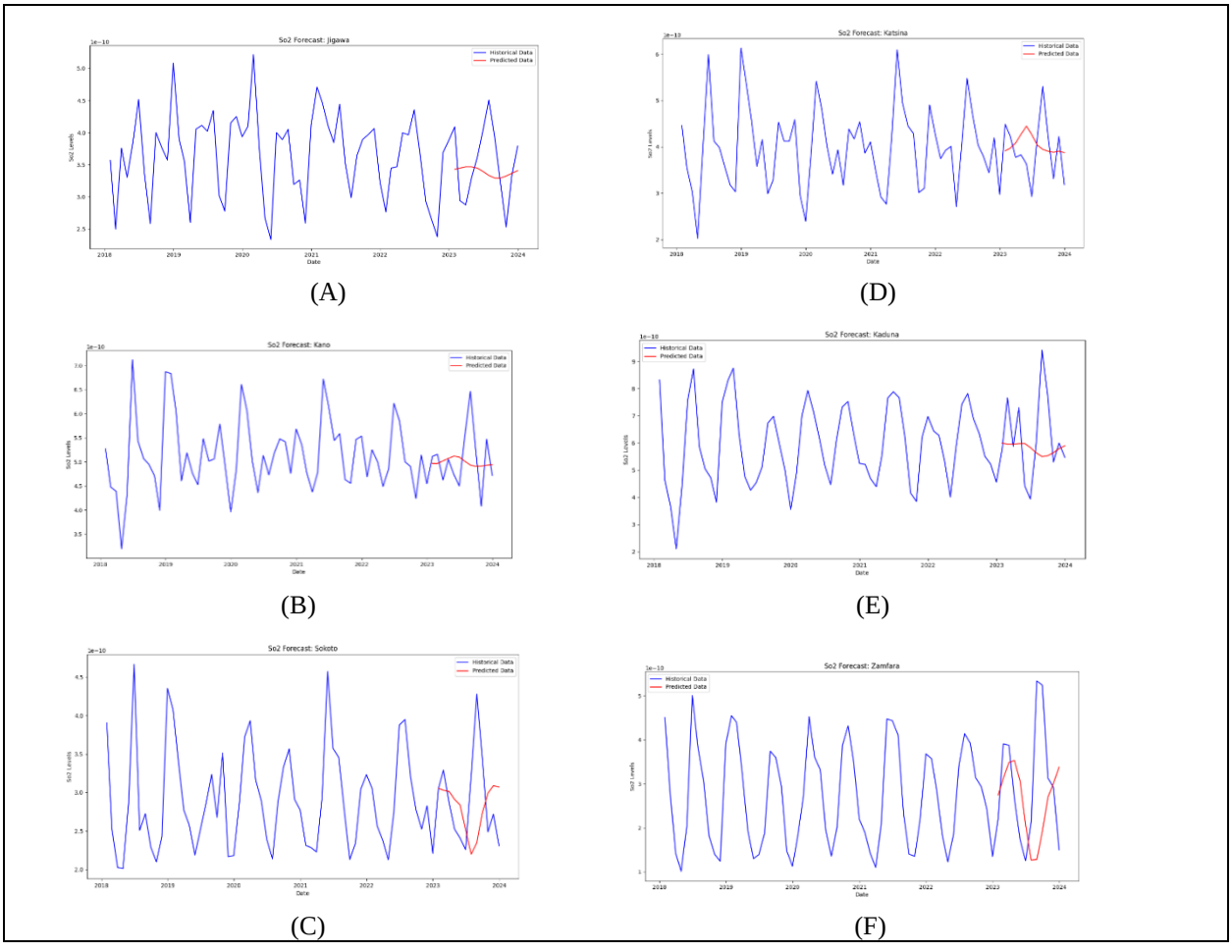


Figure 19: LSTM prediction for $S0_2$.

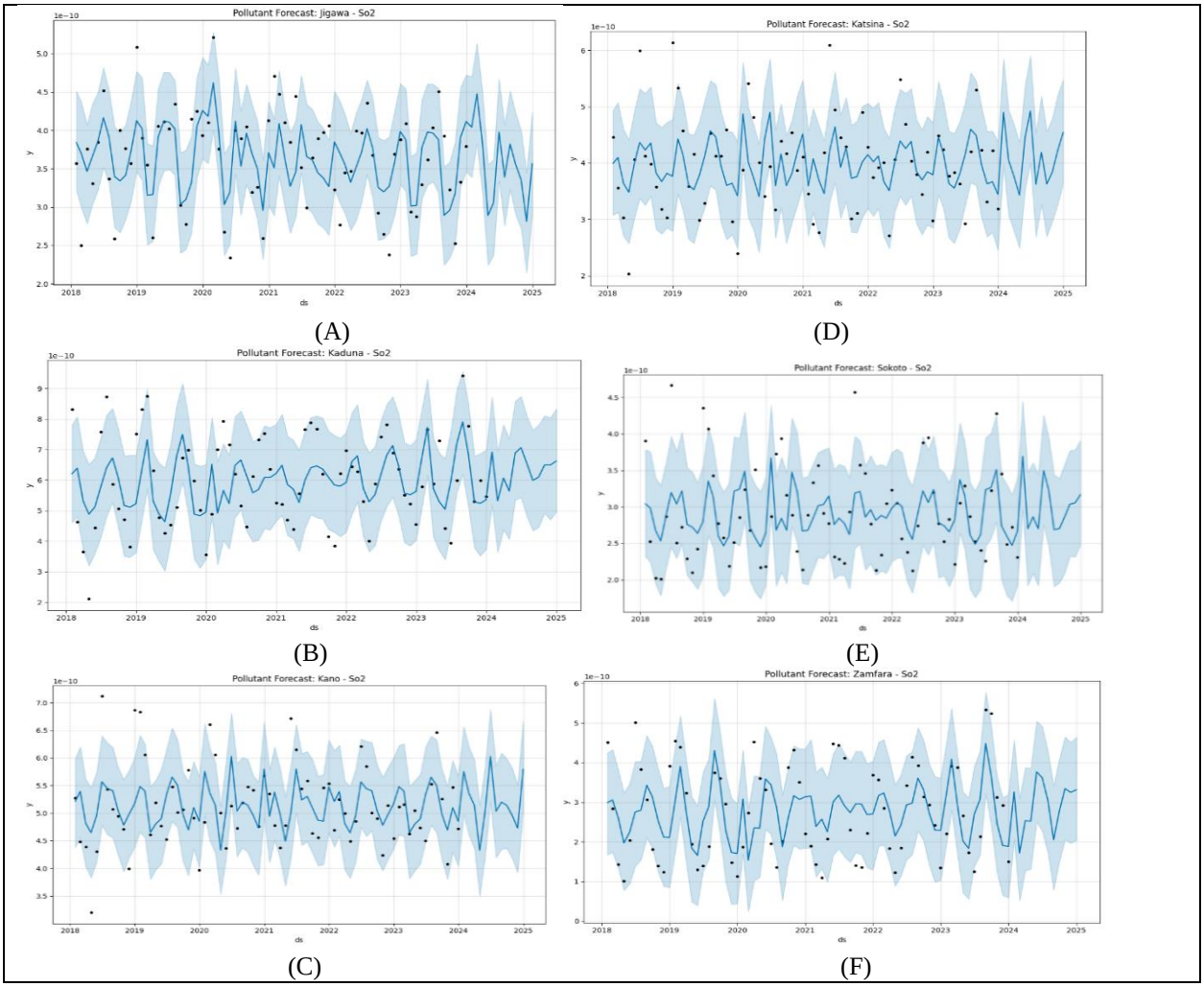


Figure 20: Prophet Prediction for SO_2.

SO_2 RMSE Performance Comparison of the Models

Both models perform relatively poorly in SO_2 prediction in Kaduna and Zamfara. This is evident from the dramatic rise in error in the states. The contrastive models achieve relatively lower prediction errors in Sokoto, Jigawa, Katsina, and Kano, indicating strong performance. Overall, LSTM was also slightly outperformed.

Table 5: SO_2 RMSE Values

RMSE FOR SO2 POLLUTANT		
STATE	LSTM	FPM
KADUNA	1.63E-10	1.28E-10
SOKOTO	6.30E-11	5.80E-11
JIGAWA	6.06E-11	6.06E-11
KATSINA	7.10E-11	7.10E-11
KANO	6.28E-11	6.28E-11
ZAMFARA	1.37E-10	1.37E-10

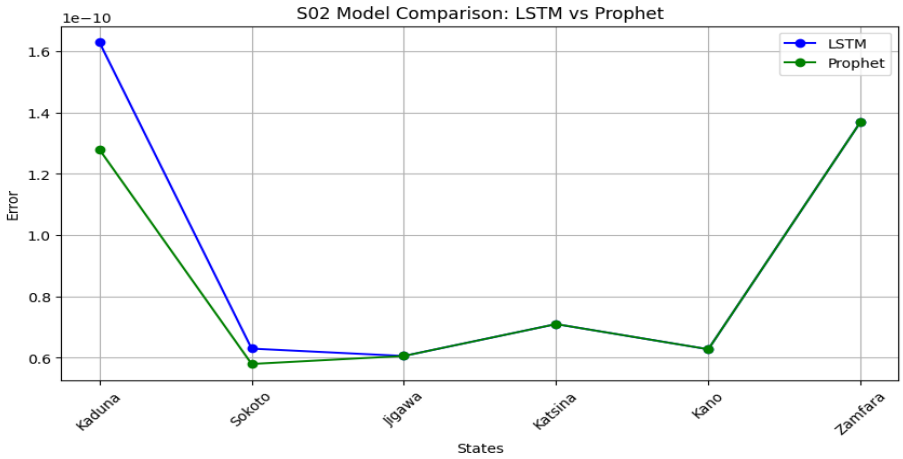


Figure 21: SO_2 RMSE comparison of LSTM and Prophet Model

SO_4 Model Prediction

A significant drop in SO_4 pollutant levels can be seen between 2020 and 2021, and peaks in 2022 to 2023 show possible effects of the lockdown, as a pronounced decrease in burning coal and oil in power plants and Fossil fuels which are major sources of SO_2 emissions that combine with oxygen to form SO_4 and cause ozone layer degradation which in turn leads to global warming. The predicted data were also smooth, suggesting a possible decrease in pollutant levels.

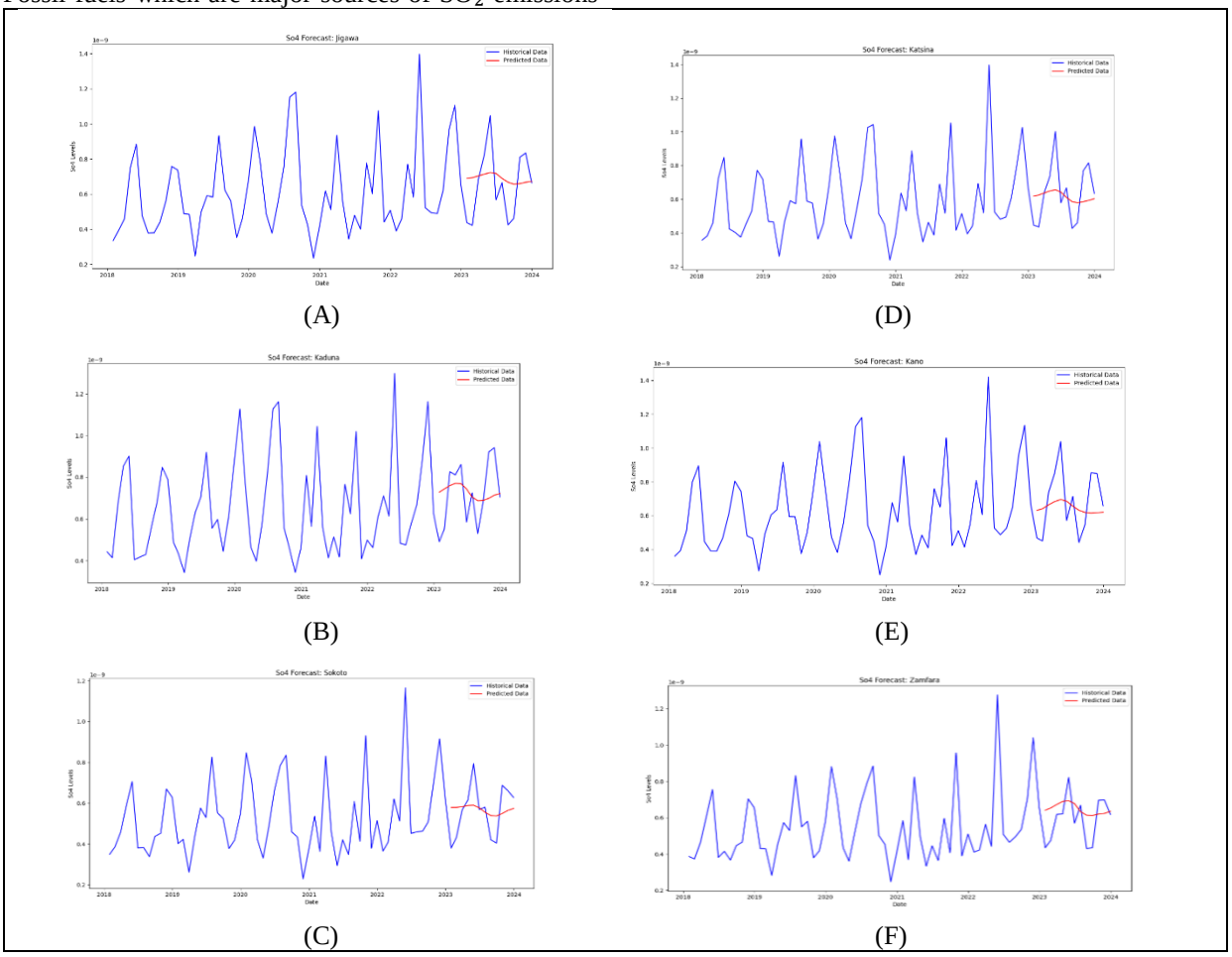


Figure 22: LSTM prediction for $S0_4$.

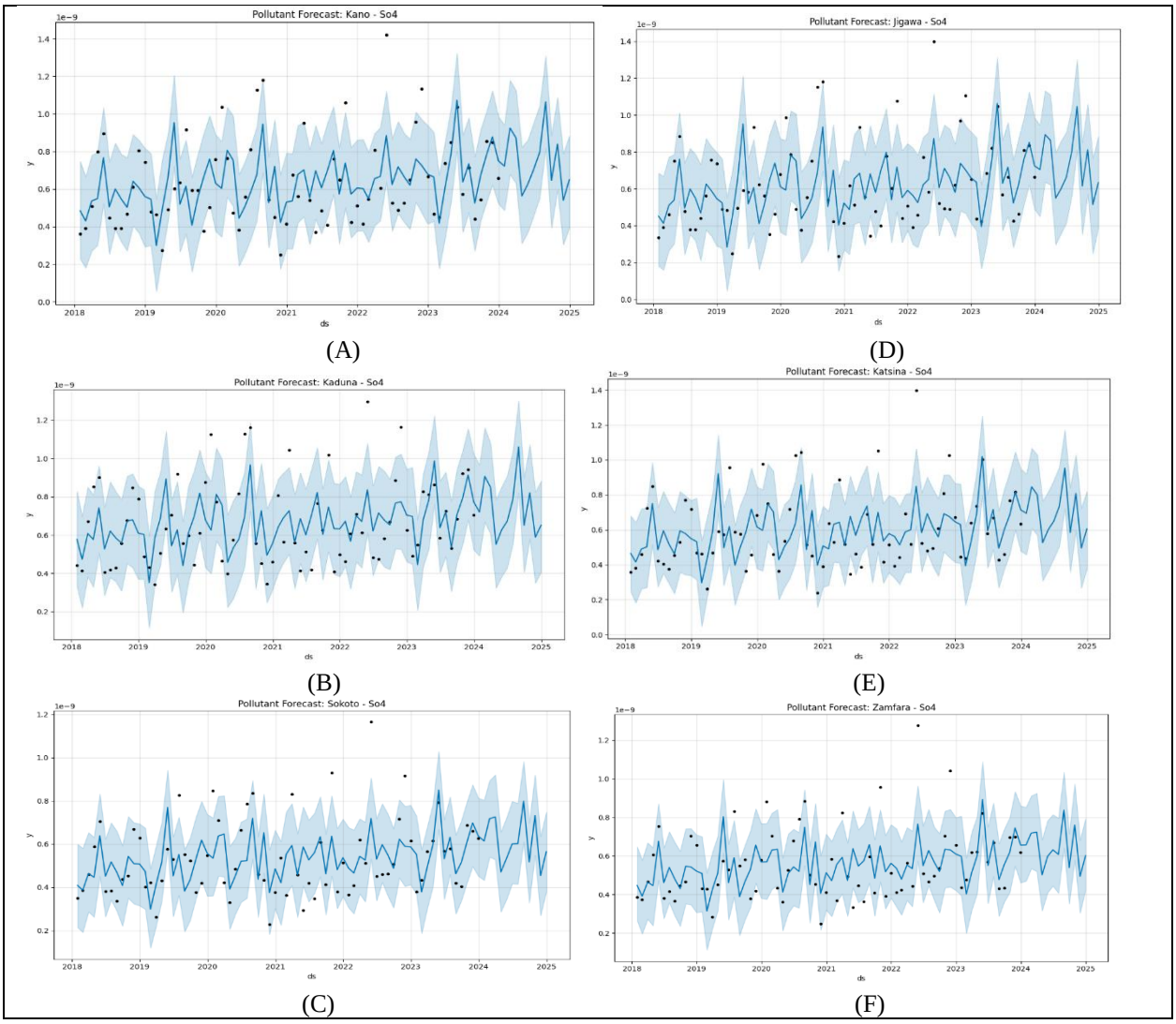


Figure 23: FPM Prediction for SO_4.

SO_4 RMSE Performance Comparison of the Models

LSTM produces lower prediction errors for the SO_4 pollutant in Kaduna and Sokoto; however, the two models perform equally well in the remaining states. This shows that LSTM proved more effective in SO_4 prediction in the case study.

Table 6: SO_4 RMSE VALUES

RMSE FOR SO_4 POLLUTANT		
STATE	LSTM	FPM
KADUNA	1.50E-10	1.88E-10
SOKOTO	1.19E-10	1.45E-10
JIGAWA	1.87E-10	1.87E-10
KATSINA	1.67E-10	1.67E-10
KANO	1.79E-10	1.79E-10
ZAMFARA	1.30E-10	1.30E-10

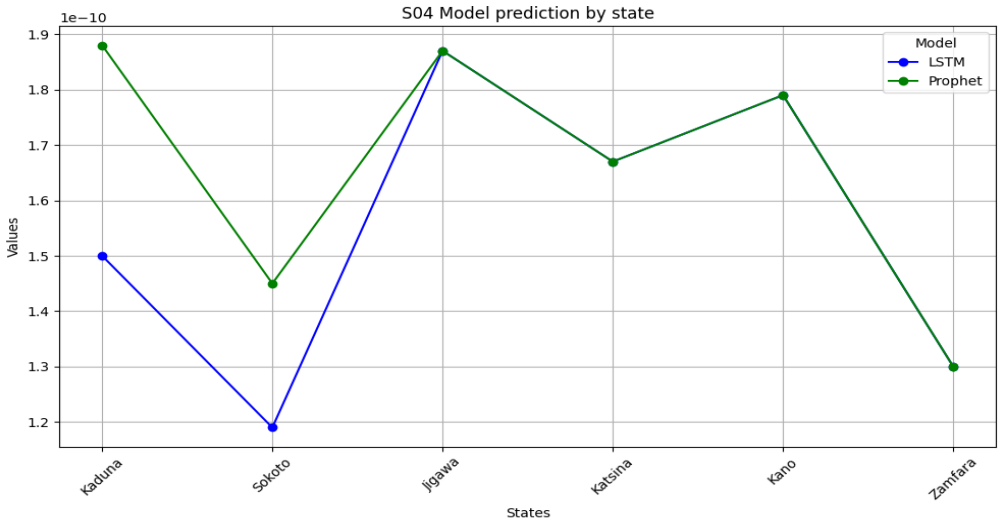


Figure 24: SO_4 RMSE comparison of LSTM and FPM Model

CONCLUSION

This study investigated air pollution forecasting for various pollutants including $S0_2$, $C0_2$, C0, $S0_4$, $PM_{2.5}$, and PM_{10} in selected northern states of Nigeria. The dataset encompassed multiple gaseous pollutants such as SO_2, CO_2, and SO_4, along with particulate matter ($PM_{2.5}$ and PM_{10}) and relevant atmospheric variables like temperature and rainfall, reflecting geographical coverage in these states.

The research reviewed statistical and machine learning models for air pollution forecasting, including deep learning algorithms and the Facebook Prophet model. Specifically, the Long Short-Term Memory (LSTM) model, a variant of recurrent neural network (RNN), was employed alongside the Facebook Prophet model for state-wise predictions. The predictive performance of both models was evaluated using four metrics: root mean squared error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and R-squared (R^2), with a focus on RMSE as a novel measure for regression accuracy.

LSTM was found to deliver the lowest prediction errors for SO_4, PM10, and PM2.5 in Kaduna and Sokoto, making it the preferred choice for these pollutants. Conversely, the Facebook Prophet Model demonstrated superior accuracy for S02, CO, CO_2, and PM10 in the sam e states, indicating its utility for those specific pollutants. The two models exhibited comparable performance in several states, including Jigawa, Katsina, Kano, and Zamfara, where their predictive capabilities aligned closely for SO_4, S02, CO, and PM_{10}.

The study concluded that while both models can effectively predict air pollution levels, the LSTM model had the least error rates for certain pollutants in designated states, while the Facebook Prophet Model displayed broader applicability across multiple pollutants and locations. Common regional factors contributing to air pollution were noted, potentially tied to industrialization, population growth, and deforestation in northern Nigeria. The results suggest that, for comprehensive air quality monitoring, each model's strengths should be considered for specific pollutants and geographical areas.

REFERENCE

Bai, L., Wang, J., Ma, X., & Lu, H. (2018). Air Pollution Forecasts: An Overview. *International Journal of Environmental Research and Public Health*, *15*(4), 780. https://doi.org/10.3390/ijerph15040780

Balogun, A., & Adebisi, N. (2021). Sea level prediction using ARIMA, SVR and LSTM neural network: assessing the impact of ensemble Ocean-Atmospheric processes on models' accuracy. *Geomatics Natural Hazards and Risk*, *12*(1), 653–674. https://doi.org/10.1080/19475705.2021.1887372

Chérif, H., Snoun, H., Bellakhal, G., & Kanfoudi, H. (2023). Forecasting of ozone concentrations using the Neural Prophet model: application to the Tunisian case. *Euro-Mediterranean Journal for Environmental Integration*, *8*(4), 987–998. https://doi.org/10.1007/s41207-023-00414-x

Du, S., Li, T., Yang, Y., & Horng, S. (2019). Deep air quality forecasting using hybrid Deep Learning

Framework. *IEEE Transactions on Knowledge and Data Engineering*, *33*(6), 2412–2424. https://doi.org/10.1109/tkde.2019.2954510

Farhadi, F., Palacin, R., & Blythe, P. (2023). Machine learning for transport Policy interventions on air quality. *IEEE Access*, *11*, 43759–43777. https://doi.org/10.1109/access.2023.3272662

Gokul, P. R., Mathew, A., Bhosale, A., & Nair, A. T. (2023). Spatio-temporal air quality analysis and PM2.5 prediction over Hyderabad City, India using artificial intelligence techniques. *Ecological Informatics*, *76*, 102067. https://doi.org/10.1016/j.ecoinf.2023.102067

Greff, K., Srivastava, R. K., Koutnik, J., Steunebrink, B. R., & Schmidhuber, J. (2016). LSTM: a Search space Odyssey. *IEEE Transactions on Neural Networks and Learning Systems*, *28*(10), 2222–2232. https://doi.org/10.1109/tnnls.2016.2582924

Ha, Q., Wahid, H., Duc, H., & Azzi, M. (2015). Enhanced radial basis function neural networks for ozone level estimation. *Neurocomputing*, *155*, 62–70. https://doi.org/10.1016/j.neucom.2014.12.048

Hasnain, A., Sheng, Y., Hashmi, M. Z., Bhatti, U. A., Hussain, A., Hameed, M. Zha, Y. (2022). Time Series analysis and forecasting of air pollutants based on Prophet Forecasting model in Jiangsu Province, China. *Frontiers in Environmental Science*, *10*. https://doi.org/10.3389/fenvs.2022.945628

Health Impacts of Air Pollution | State of Global air. (n.d.). Retrieved from https://www.stateofglobalair.org/hap

Huang, K., Xiao, Q., Meng, X., Geng, G., Wang, Y., Lyapustin, A., Liu, Y. (2018). Predicting monthly high-resolution PM2.5 concentrations with random forest model in the North China Plain. *Environmental Pollution*, *242*, 675–683. https://doi.org/10.1016/j.envpol.2018.07.016

Jin, N., Zeng, Y., Yan, K., & Ji, Z. (2021). Multivariate air quality forecasting with nested long short term memory neural network. *IEEE Transactions on Industrial Informatics*, *17*(12), 8514–8522. https://doi.org/10.1109/tii.2021.3065425

Khayyat, M., Laabidi, K., Almalki, N., & Al-Zahrani, M. (2021). Time series Facebook Prophet Model and Python for COVID-19 outbreak prediction. *Computers, Materials & Continua/Computers, Materials & Continua (Print)*, *67*(3), 3781–3793. https://doi.org/10.32604/cmc.2021.014918

Kristiani, E., Lin, H., Lin, J., Chuang, Y., Huang, C., & Yang, C. (2022). Short-Term prediction of PM2.5 using LSTM deep learning methods. *Sustainability*, *14*(4), 2068. https://doi.org/10.3390/su14042068

Li, C., Ingersoll, A., Bolton, S., Levin, S., Janssen, M., Atreya, S. Zhang, Z. (2020). The water abundance in Jupiter's equatorial zone. *Nature Astronomy*, *4*(6), 609–616. https://doi.org/10.1038/s41550-020-1009-3

Lu, J., Liu, A., Dong, F., Gu, F., Gama, J., & Zhang, G. (2018). Learning under Concept Drift: A Review. *IEEE Transactions on Knowledge and Data Engineering*, 1. https://doi.org/10.1109/tkde.2018.2876857

Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., & Johnson, B. A. (2018). Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, *152*, 166–177. https://doi.org/10.1016/j.isprsjprs.2019.04.015

Manisalidis, I., Stavropoulou, E., Stavropoulos, A., & Bezirtzoglou, E. (2020). Environmental and Health Impacts of Air Pollution: A review. *Frontiers in Public Health*, *8*. https://doi.org/10.3389/fpubh.2020.00014

Martin. (2018, July 18). 90 per cent of the planet is breathing in polluted air – World Health Organization. Retrieved from https://www.un.org/sustainabledevelopment/blog/2018/05/90-per-cent-of-the-planet-is-breathing-in-polluted-air-world-health-organization/

Naz, F., Mccann, C., Fahim, M., Cao, T., Hunter, R., Viet, N. T. Duong, T. Q. (2023). Comparative analysis of deep learning and statistical models for air pollutants prediction in urban areas. *IEEE Access*, *11*, 64016–64025. https://doi.org/10.1109/access.2023.3289153

Niemi, M. E. K., Daly, M. J., & Ganna, A. (2022). The human genetic epidemiology of COVID-19. *Nature Reviews Genetics*, *23*(9), 533–546. https://doi.org/10.1038/s41576-022-00478-5

Ojha, N., Soni, M., Kumar, M., Gunthe, S. S., Chen, Y., & Ansari, T. U. (2022). Mechanisms and pathways for Coordinated control of fine particulate matter and ozone. *Current Pollution Reports*, *8*(4), 594–604. https://doi.org/10.1007/s40726-022-00229-4

Peralta, B., Sepúlveda, T., Nicolis, O., & Caro, L. (2022). Space-Time Prediction of PM2.5 Concentrations in Santiago de Chile Using LSTM Networks. *Applied Sciences*, *12*(22), 11317. https://doi.org/10.3390/app122211317

Raheja, S., & Malik, S. (2022). Prediction of air quality using LSTM Recurrent neural network. *International Journal of Software Innovation*, *10*(1), 1–16. https://doi.org/10.4018/ijsi.297982

Sharma, E., Deo, R. C., Prasad, R., & Parisi, A. V. (2019). A hybrid air quality early-warning framework: An hourly forecasting model with online sequential extreme learning machines and empirical mode decomposition algorithms. *The Science of the Total Environment*, *709*, 135934. https://doi.org/10.1016/j.scitotenv.2019.135934

Tella, A., Balogun, A., Adebisi, N., & Abdullah, S. (2021). Spatial assessment of PM10 hotspots using Random Forest, K-Nearest Neighbour and Naïve Bayes. *Atmospheric Pollution Research*, *12*(10), 101202. https://doi.org/10.1016/j.apr.2021.101202

Torres, J. F., Hadjout, D., Sebaa, A., Martínez-Álvarez, F., & Troncoso, A. (2020). Deep learning for Time Series Forecasting: A survey. *Big Data*, *9*(1), 3–21. https://doi.org/10.1089/big.2020.0159

Wambebe, N. M., & Duan, X. (2020). Air Quality Levels and Health Risk Assessment of Particulate Matters in Abuja Municipal Area, Nigeria. *Atmosphere*, *11*(8), 817. https://doi.org/10.3390/atmos11080817

Wen, D., Zheng, S., Chen, J., Zheng, Z., Ding, C., & Zhang, L. (2023). Hyperparameter-Optimization-Inspired Long Short-Term memory network for air quality grade prediction. *Information*, *14*(4), 243. https://doi.org/10.3390/info14040243

World Health Organization. (2018). Health and the environment: draft road map for an enhanced global response to the adverse health effects of air pollution: report by the Secretariat. Retrieved from https://apps.who.int/iris/handle/10665/252673

Yafouz, A., Ahmed, A. N., Zaini, N., Sherif, M., Sefelnasr, A., & El-Shafie, A. (2021). Hybrid deep learning model for ozone concentration prediction: comprehensive evaluation and comparison with various machine and deep learning algorithms. *Engineering Applications of Computational Fluid Mechanics*, *15*(1), 902–933. https://doi.org/10.1080/19942060.2021.1926328

Zhang, Z., Chen, H., & Huang, X. (2023). Prediction of air quality combining Wavelet Transform, DCCA correlation analysis and LSTM model. *Applied Sciences*, *13*(5), 2796. https://doi.org/10.3390/app13052796

Zhou, Y., Chang, F., Chang, L., Kao, I., Wang, Y., & Kang, C. (2018). Multi-output support vector machine for regional multi-step-ahead PM2.5 forecasting. *The Science of the Total Environment*, *651*, 230–240. https://doi.org/10.1016/j.scitotenv.2018.09.111

Zhou, Y., Chang, F., Chang, L., Kao, I., Wang, Y., & Kang, C. (2019). Multi-output support vector machine for regional multi-step-ahead PM2.5 forecasting. *Science of the Total Environment*, *651*, 230–240. https://doi.org/10.1016/j.scitotenv.2018.09.111

Zhu, Q., Dinh, T. H., Phung, M. D., & Ha, Q. P. (2021). Hierarchical convolutional neural network with feature preservation and autotuned thresholding for crack detection. *IEEE Access*, *9*, 60201–60214. https://doi.org/10.1109/access.2021.3073921

APPENDIX

TABLE 4.7: COMPARISON METRICS FOR LSTM and FPM MODEL

		LSTM METRICS				FPM METRICS			
State	Pollutant	MAE	RMSE	MAPE	MAE	RMSE	MAPE	R-squared	
Kaduna	Pm2.5	40.5673	46.17125	184.1296	-0.10383	29.36896	36.42122	108.2474	0.30256
Kaduna	Pm10	0.980172	1.106547	86.3251	-0.45899	0.832843	1.029591	57.57801	0.234095
Kaduna	So2	1.25E-10	1.63E-10	19.97189	-0.12347	1.08E-10	1.28E-10	20.01259	0.247656
Kaduna	Co2	2.09E-06	2.53E-06	0.502536	-7.63343	5.42E-07	7.23E-07	0.131986	0.943617
Kaduna	Co	3.65E-08	3.82E-08	22.29141	-4.34223	1.78E-08	2.21E-08	10.14816	0.145923
Kaduna	So4	1.28E-10	1.50E-10	19.59965	0.017974	1.49E-10	1.88E-10	23.94146	0.286156
Sokoto	Pm2.5	43.26678	48.43757	146.3255	-0.0441	32.07154	38.82825	81.88261	0.316407
Sokoto	Pm10	0.770978	0.921692	45.71467	-0.05357	0.768314	1.004266	35.86786	0.274595
Sokoto	So2	4.59E-11	6.30E-11	15.18419	-0.20041	4.71E-11	5.80E-11	16.17857	0.197375
Sokoto	Co2	3.00E-06	3.38E-06	0.72033	-4.18547	5.34E-07	6.83E-07	0.129645	0.948895
Sokoto	Co	1.16E-08	1.21E-08	7.365453	0.120694	1.40E-08	1.73E-08	8.620566	0.109943
Sokoto	So4	9.81E-11	1.19E-10	19.39914	0.050093	1.14E-10	1.45E-10	22.82725	0.317987
Jigawa	Pm2.5	50.49055	58.06085	154.9275	-0.04705	50.49055	58.06085	154.9275	-0.04705
Jigawa	Pm10	0.941809	1.075562	53.03398	-0.26043	0.941809	1.075562	53.03398	-0.26043
Jigawa	So2	5.13E-11	6.06E-11	14.64421	-0.16694	5.13E-11	6.06E-11	14.64421	-0.16694
Jigawa	Co2	2.18E-06	2.57E-06	0.52361	-1.8443	2.18E-06	2.57E-06	0.52361	-1.8443
Jigawa	Co	1.34E-08	1.53E-08	8.409705	-0.55393	1.34E-08	1.53E-08	8.409705	-0.55393
Jigawa	So4	1.57E-10	1.87E-10	27.88041	0.042593	1.57E-10	1.87E-10	27.88041	0.042593
Katsina	Pm2.5	45.34666	51.4292	157.2147	-0.07572	45.34666	51.4292	157.2147	-0.07572
Katsina	Pm10	1.044535	1.186164	58.82069	-0.69387	1.044535	1.186164	58.82069	-0.69387
Katsina	So2	6.17E-11	7.10E-11	16.51051	-0.21955	6.17E-11	7.10E-11	16.51051	-0.21955
Katsina	Co2	2.66E-06	3.21E-06	0.638388	-4.59652	2.66E-06	3.21E-06	0.638388	-4.59652
Katsina	Co	1.90E-08	2.07E-08	11.70248	-1.56761	1.90E-08	2.07E-08	11.70248	-1.56761
Katsina	So4	1.39E-10	1.67E-10	22.68595	0.046791	1.39E-10	1.67E-10	22.68595	0.046791
Kano	Pm2.5	46.04929	52.41089	159.6657	-0.0725	46.04929	52.41089	159.6657	-0.0725
Kano	Pm10	1.00829	1.224998	63.79993	-0.62829	1.00829	1.224998	63.79993	-0.62829
Kano	So2	4.99E-11	6.28E-11	9.751531	-0.11365	4.99E-11	6.28E-11	9.751531	-0.11365
Kano	Co2	2.18E-06	2.75E-06	0.523232	-2.62342	2.18E-06	2.75E-06	0.523232	-2.62342
Kano	Co	1.98E-08	2.14E-08	12.45924	-1.83225	1.98E-08	2.14E-08	12.45924	-1.83225
Kano	So4	1.55E-10	1.79E-10	23.57175	0.046586	1.55E-10	1.79E-10	23.57175	0.046586
Zamfara	Pm2.5	42.57575	47.74106	157.8427	-0.08001	42.57575	47.74106	157.8427	-0.08001
Zamfara	Pm10	0.850157	1.010746	55.24954	-0.09983	0.850157	1.010746	55.24954	-0.09983
Zamfara	So2	9.35E-11	1.37E-10	31.00141	-0.07046	9.35E-11	1.37E-10	31.00141	-0.07046
Zamfara	Co2	2.33E-06	2.67E-06	0.56008	-3.4252	2.33E-06	2.67E-06	0.56008	-3.4252
Zamfara	Co	1.36E-08	1.53E-08	8.123359	-0.43107	1.36E-08	1.53E-08	8.123359	-0.43107
Zamfara	So4	1.12E-10	1.30E-10	21.56961	-0.14786	1.12E-10	1.30E-10	21.56961	-0.14786