**Climate Change Prediction Based on Human Activities**

Snehil Srivastav

*Department of Engineering And Technology*

*Sharda University*

Greater Noida, India

[snehilsri925@gmail.com](mailto:snehilsri925@gmail.com)

Lalit Kumar

*Department of Engineering And Technology*

*Sharda University*

Greater Noida, India

[Lalitkumar2@gmail.com](mailto:Lalitkumar2@gmail.com)

Aryan Chaudhary

*Department of Engineering And Technology*

*Sharda University*

Greater Noida, India

[aryanchaudhary2003@gmail.com](mailto:aryanchaudhary2003@gmail.com)

Dr. Kalicharan

*Department of Engineering And Technology*

*Sharda University*

Greater Noida, India

[kalicharan.2@sharda.ac.in](mailto:kalicharan.2@sharda.ac.in)

***ABSTRACT***  
**This study presents a predictive framework to evaluate climate change impacts on Delhi's environment using machine learning models, particularly Random Forest (RF) and Extreme Gradient Boosting (XGBoost). By integrating diverse datasets—including atmospheric pollutants, land surface temperature, vegetation indices, and air quality indicators—the model captures dynamic interactions between human activities and climate variables. Leveraging advancements in climate prediction techniques, the study assesses RF and XGBoost for forecasting critical factors such as temperature, precipitation, CO, NO₂, and particulate matter concentrations. Model performance is optimized through a Randomized Search Cross-Validation approach, with XGBoost demonstrating superior accuracy and robustness. This highlights its ability to handle complex, multidimensional data while minimizing the effects of variability and outliers. The research provides a reliable approach for climate impact assessment and supports urban planning, environmental policy-making, and adaptive climate governance, contributing valuable insights into the interplay between anthropogenic activities and climate dynamics in urban regions.**

***Keywords:*** *Climate change prediction, machine learning, XGBoost, Random Forest, environmental modeling, Delhi, air quality, greenhouse gases, anthropogenic impacts*

**I INTRODUCTION**

The growing impact of human activities on global climate has accelerated environmental shifts, prompting extensive research into climate change drivers and adaptation strategies. Industrialization, urbanization, and energy consumption have contributed to heightened greenhouse gas emissions, exacerbating climate conditions and impacting human and ecological health worldwide. Studies indicate that human activities are a significant factor in climate alterations, with urban regions facing specific challenges due to pollution and population density [1]. According to the Intergovernmental Panel on Climate Change (IPCC), climate change has become a critical factor in regional risk assessments and environmental management strategies, emphasizing the need for local adaptation measures [2].

Adaptation strategies, particularly in urban environments, are crucial for mitigating climate-related risks. Researchers advocate for actions tailored to urban needs to counteract the adverse effects of climate change, ranging from infrastructure improvements to ecological resilience measures [3]. Furthermore, NO₂ and other pollutants are key indicators of urban air quality, with recent advances in modeling tools for accurately quantifying emissions and assessing their impacts on health and environmental systems [4]. Studies also highlight the role of climate change in intensifying extreme weather events, such as torrential rains, which require urgent analysis to inform disaster response and planning [5].

Efforts to understand the interplay between climate factors such as soil moisture and vegetation health, as well as their implications for nitrogen dioxide (NO₂) concentration, have led to recent advancements in climate data modeling [6]. Specifically, vegetation indices such as the Normalized Difference Vegetation Index (NDVI) are used to assess ecosystem health and predict NO₂ concentration shifts, which are essential for air quality management in densely populated cities [7]. In areas like New Delhi, the economic and health costs associated with air pollution are largely driven by agricultural practices such as stubble burning, further illustrating the complex economic and environmental repercussions of human activity on climate [8,9].

Cumulative evidence underscores the urgent need for integrated climate prediction models that account for various factors, including pollution, vegetation cover, and temperature fluctuations. Such models can aid in identifying and addressing climate vulnerabilities, facilitating informed policy decisions, and supporting public awareness initiatives [10].

**II LITREATURE REIVIEW**

TABLE I LITERATURE REVIEW

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Author(s)** | **Year** | **Title** | **Source** | **Main Findings** | **Methodology/Approach** | **Limitations** |
| Rahmani, Z. & Ahmadi, J. | 2024 | The impact of human activities on climate change | Sprin Journal of Arts, Humanities and Social Sciences | Human activities, especially fossil fuel combustion, significantly contribute to greenhouse gas emissions, exacerbating climate change. | Literature review and analysis | Limited to secondary data; lacks empirical data on mitigation effectiveness. |
| Intergovernmental Panel on Climate Change (IPCC) | 2023 | Climate Change Information for Regional Impact and for Risk Assessment | IPCC Sixth Assessment Report | Provides a comprehensive assessment of climate science, human impact, and future projections, with a focus on reaching net zero emissions. | Observational analysis, climate modeling | Regional data gaps; reliance on projections that may not capture all local dynamics. |
| He, B.-J. et al. | 2024 | Actions for Urban Climate Change Adaptation | Climate Change and Environmental Sustainability | Stresses the urgent need for urban adaptation strategies to address climate impacts, highlighting 21 best practices from global conferences. | Case studies and conference proceedings compilation | Limited to urban contexts; adaptability in rural areas is unexplored. |
| Meier, S. et al. | 2024 | A lightweight NO₂-to-NOₓ conversion model for quantifying NOₓ emissions | Atmospheric Chemistry and Physics | Developed a model to convert NO₂ to NOₓ, reducing biases in NOₓ emission estimates based on satellite data. | Simulation and satellite data analysis | Applicability may be limited to regions with high NO₂ emissions data availability. |
| Hejazizadeh, Z. et al. | 2022 | Investigating the effects of climate change on torrential rains in Tehran province | Journal Article | Climate change is likely to increase the frequency and severity of torrential rains in Tehran, leading to heightened flood risks. | Statistical analysis of meteorological data | Limited regional scope; results may not generalize to areas with different climatic conditions. |
| Le, M.-H. et al. | 2024 | On the Use of SMAP Soil Moisture for Forecasting NDVI | Geophysical Research Letters | Demonstrated that satellite soil moisture data can improve vegetation health forecasting, especially in water-limited regions. | Satellite data analysis and regression modeling | Primarily effective in water-limited areas; other variables may influence NDVI in different regions. |
| Rahaman, S. N. et al. | 2023 | Effect of vegetation and land surface temperature on NO₂ concentration | Urban Climate | Examines the relationship between NO₂, land surface temperature, and vegetation in Delhi and Dhaka, underscoring urban air quality challenges. | Remote sensing and statistical analysis | Limited to specific urban areas; effects may vary significantly in non-urban regions. |
| Agarwal, A. & Tiwari, N. | 2024 | The Economic Cost of Air Pollution Due to Stubble Burning | Journal of Pollution Effects & Control | Field fires in North India elevate PM2.5 and PM10 levels in New Delhi, impacting economic productivity in the region. | Instrumental variable analysis using NASA data | Limited to stubble burning effects; doesn’t account for other pollution sources. |
| Chen | 2024 | Severe Air Pollution Crisis in New Delhi | Journal Article | Air pollution in New Delhi is responsible for an estimated 10,000 premature deaths annually; PM1 underestimations are highlighted. | In-situ observational data analysis | Underestimation of PM1; results are based on observational data without longitudinal studies on health impacts. |
| Robinson et al. | 2024 | Climate Change, Pollution, and the Aerobiome | Mini-Review | Examines how climate change and pollution affect airborne microbial life, with impacts on health and ecosystems; advocates interdisciplinary collaboration. | Review and synthesis of existing research | Limited empirical data on long-term aero biome shifts; lacks data on specific health outcomes. |

**IV METHODOLOGY**

This study develops a climate prediction model that focuses on multiple environmental factors in Delhi. The model utilizes monthly averaged data on temperature, CO and NO₂ densities, precipitation, vegetation indices, and air pollutants to predict climate trends and air quality in Delhi for the year 2024. The dataset spans 2020 to 2023 and offers historical data essential for training an accurate predictive model.

TABLE II DATASET DESCRIPTION

|  |  |  |  |
| --- | --- | --- | --- |
| **Dataset** | **Source** | **Attribute** | **Purpose** |
| Delhi\_CO\_Data\_2020\_2023 | Sentinel-5P (TROPOMI) | CO\_column\_number\_density | Monitoring carbon monoxide concentration |
| Delhi\_EVI\_Data\_2020\_2023 | MODIS | EVI | Vegetation health and changes |
| Delhi\_NDVI\_Data\_2020\_2023 | MODIS | NDVI | Vegetation index for land cover analysis |
| Delhi\_NO2\_Data\_2020\_2023 | Sentinel-5P (TROPOMI) | NO2\_column\_number\_density | Air quality monitoring (NO2 concentration) |
| Delhi\_Precipitation\_Data\_2020\_2023 | CHIRPS) | Precipitation | Precipitation analysis and trends |
| Delhi\_Temperature\_Data\_2020\_2023 | ERA5 (ECMWF Reanalysis) | temperature\_2m | Temperature monitoring at 2m height |

***Descriptions of Each Flow Chart Step:***

* **Start**: Begin the climate change prediction process.
* **Collect Data**: Gather climate-related data.
  + **Is Data Clean?**
  + If Yes, analyze the data.
  + If No, clean the data.
  + **Clean Data**: Fix errors or inconsistencies.
* **Analyze Data**: Identify patterns or insights.
* **Is Analysis Complete?**
  + If Yes, build the model.
  + If No, continue analysis.
* **Build Prediction Model**: Create the prediction model.
* **Is Model Validated?**
  + If Yes, make predictions.
  + If No, refine the model.
* **Predictions**: Forecast future outcomes.

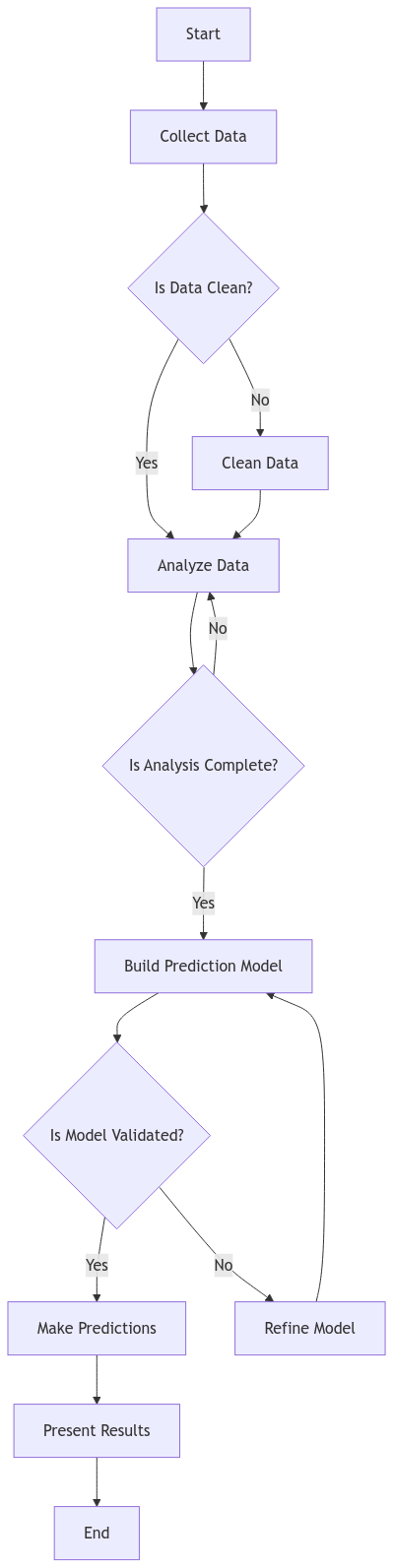


Fig. 1. Flow Chart of the Model

***1. Data Collection and Preprocessing***

The datasets included multiple variables measured monthly or daily, such as temperature (in Kelvin), CO and NO₂ concentrations, precipitation, vegetation indices (EVI and NDVI), and various air pollutants (e.g., SO₂, PM2.5, and PM10). The datasets were obtained from Google Earth Engine and Kaggle for AQI data are loaded and merged based on dates using pandas, creating a consolidated data frame where each row represents aggregated monthly measurements.

TABLE III : DATA PROCESSING MODULES

|  |  |  |  |
| --- | --- | --- | --- |
| **Module** | **Description** | **Inputs** | **Outputs** |
| Data Collection & Integration | Aggregates raw datasets from various sources into a unified format. | External climate datasets | Collected datasets (Temperature, CO, etc.) |
| Data Cleaning | Cleanses data by handling missing values, outliers, and inconsistencies. | Collected datasets | Cleaned datasets |
| Data Transformation | Standardizes data formats, converts units (e.g., Kelvin to Celsius), and prepares data for modeling. | Cleaned datasets | Transformed datasets |
| Feature Engineering | Derives additional features that may enhance model performance, such as month or year indicators. | Transformed datasets | Engineered datasets |
| Model Selection & Training | Trains and fine-tunes models (e.g., XGBoost) using cross-validation to ensure optimal performance. | Engineered datasets | Trained model |
| Prediction Generation | Generates predictions for future climate variables using the trained model. | Trained model, input features | Predicted climate data |
| Visualization & Reporting | Creates visualizations for predicted data and generates analysis reports. | Predicted climate data | Visual reports and analysis |

To address missing values, the mean imputation approach was used for each feature, helping to maintain the continuity of time-series data without biased predictions.

**2. *Data Aggregation and Feature Engineering***

After handling missing values, monthly data was aggregated by calculating mean values for each climate variable. Feature engineering extracted year and month as primary features to capture seasonal temporal trends, creating an input dataset suitable for monthly climate prediction.

***3. Model Selection and Hyperparameter Tuning***

This study adopted XGBoost in a multi-output regression setup to predict multiple dependent variables simultaneously. Multi-output regression is an essential model because of the interdependencies between pollutants and climatic variables.

|  |  |  |  |
| --- | --- | --- | --- |
| **Model** | **Parameter** | **Description** | **Selected Value(s)** |
| XGBoost | n\_estimators | Number of boosting rounds | Selected via cross-validation |
|  | learning\_rate | Step size shrinkage to prevent overfitting | Selected via hyperparameter tuning |
|  | max\_depth | Maximum depth of trees | Selected via hyperparameter tuning |
|  | subsample | Fraction of data to be used in each boosting round | Selected via hyperparameter tuning |
|  | colsample\_bytree | Fraction of features to be used in each tree | Selected via hyperparameter tuning |

TABLE IV: MODEL PARAMETERS AND EVALUATION METRICS

***Hyperparameter Optimization***

To fine-tune the model, RandomizedSearchCV, which tests a subset of parameter combinations to identify the optimal values, was employed. The hyperparameters tuned were as follows:

* ***n\_estimators***: Number of trees
* ***learning\_rate***: Step size for each boosting step
* ***max\_depth***: Maximum depth of each tree, controlling complexity
* ***Subsample*** *and* ***colsample\_bytree***: Sampling parameters for both rows and features, enhancing model generalization and preventing overfitting [15].

The tuning process focused on minimizing the root mean square error (RMSE) using cross-validation, calculated as:

RandomizedSearchCV was employed to iteratively refine the XGBoost model's hyperparameters. By randomly sampling from a predefined grid of parameter combinations, the method efficiently identified the optimal configuration that minimized cross-validation mean squared error, enhancing the model's predictive performance across climate variables.

The selected model was evaluated using the cross-validated root mean square error (RMSE). The Multioutput Regressor framework enables the XGBoost model to jointly predict all target climate variables for each month of 2024.

TABLE V : EVALUATION RESULTS

|  |  |  |
| --- | --- | --- |
| **Model** | **Cross-validated RMSE** | **Description** |
| Random Forest | 224.94 | Baseline model with multiple outputs |
| XGBoost | 133.85 | Selected model with tuned parameters for prediction accuracy |

Predictions for 2024 were generated by iterating through each month, del, to forecast temperature, precipitation, CO and NO₂ densities, EVI, NDVI, and air pollutants.

**5*. Visualization of Predicted Data***

Visualization aids in understanding the temporal changes in climatic variables. A 6 × 2 subplot grid displayed each predicted climate factor, showing month-over-month trends. This method highlights seasonal patterns and extreme values for indicators, such as temperature, precipitation, and air pollutants.

**V RESULT AND DESCUSSION**

The model’s predictions for 2024 provide a detailed monthly projection of several climate indicators for Delhi, including temperature, precipitation, CO and NO₂ levels, and vegetation indices (EVI and NDVI).

**Temperature**:

* **Summer Peaks (May-June)**: Predicted temperatures show an upward trend, exceeding historical averages, relevant for heat resilience planning.
* **Winter Anomalies (Jan-Dec)**: Slightly higher-than-normal averages due to urban heat islands and reduced vegetation cover.

**Precipitation**:

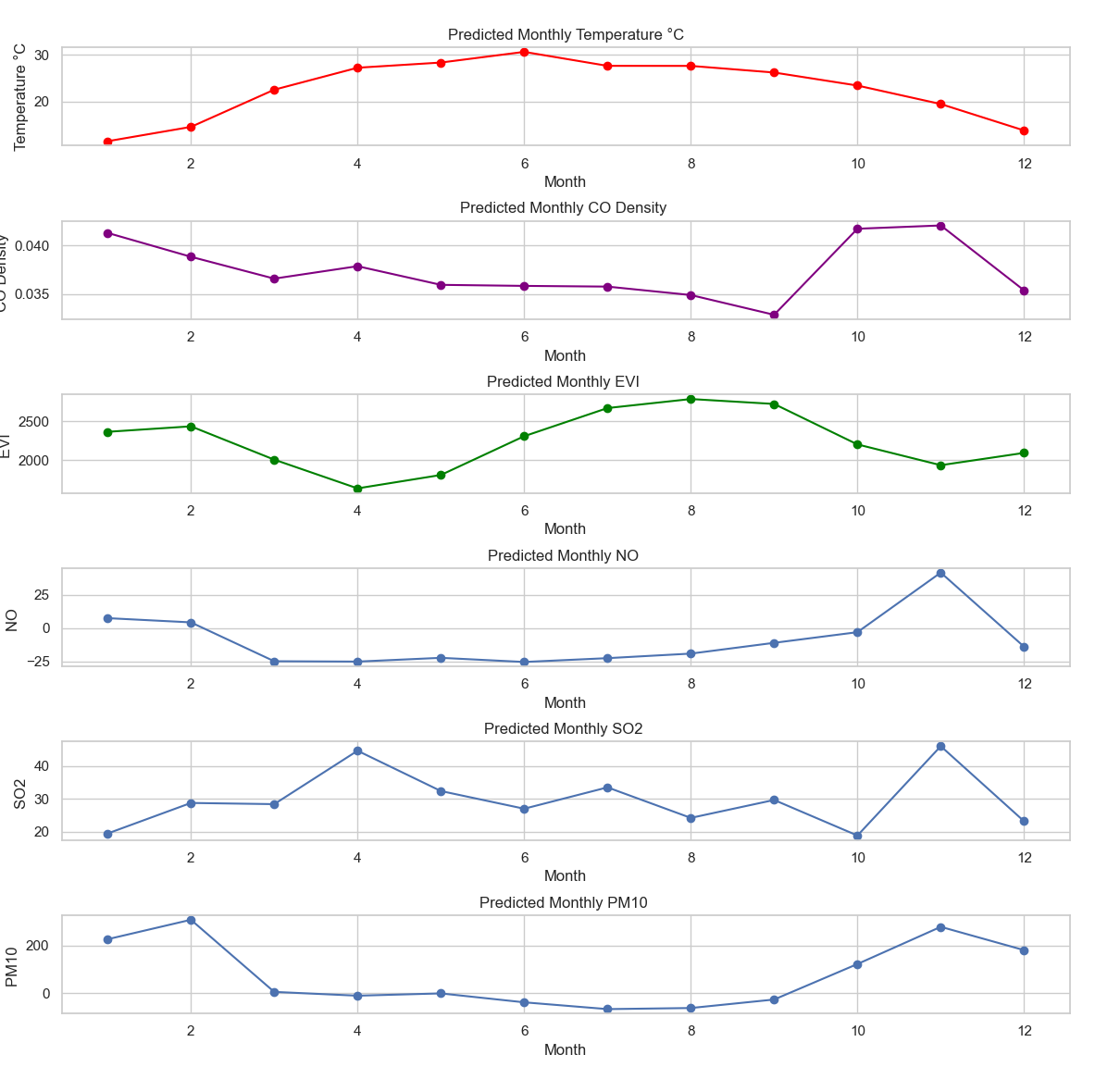
* **Monsoon Intensity (July-August)**: Increased rainfall suggests a potential shift toward intense events, impacting water management and flood risks.
* **Winter Dryness**: Lower precipitation highlights seasonal dryness with rising unpredictability.

**Air Quality (CO & NO₂)**:

* **Winter Peaks**: Elevated pollution due to seasonal cycles and temperature inversions, emphasizing the need for winter air quality controls.
* **Summer & Monsoon Dispersion**: Lower pollutant levels aided by atmospheric dispersion and rainfall cleansing effects.

**Vegetation (EVI & NDVI)**:

* **Monsoon Growth**: Robust vegetation during monsoons aids CO₂ absorption and urban cooling.
* **Winter Declines**: Reduced vegetation due to dryness and pollution stress highlights the need for year-round green cover.

 Fig. 2. Climate Prediction of Delhi(2024)

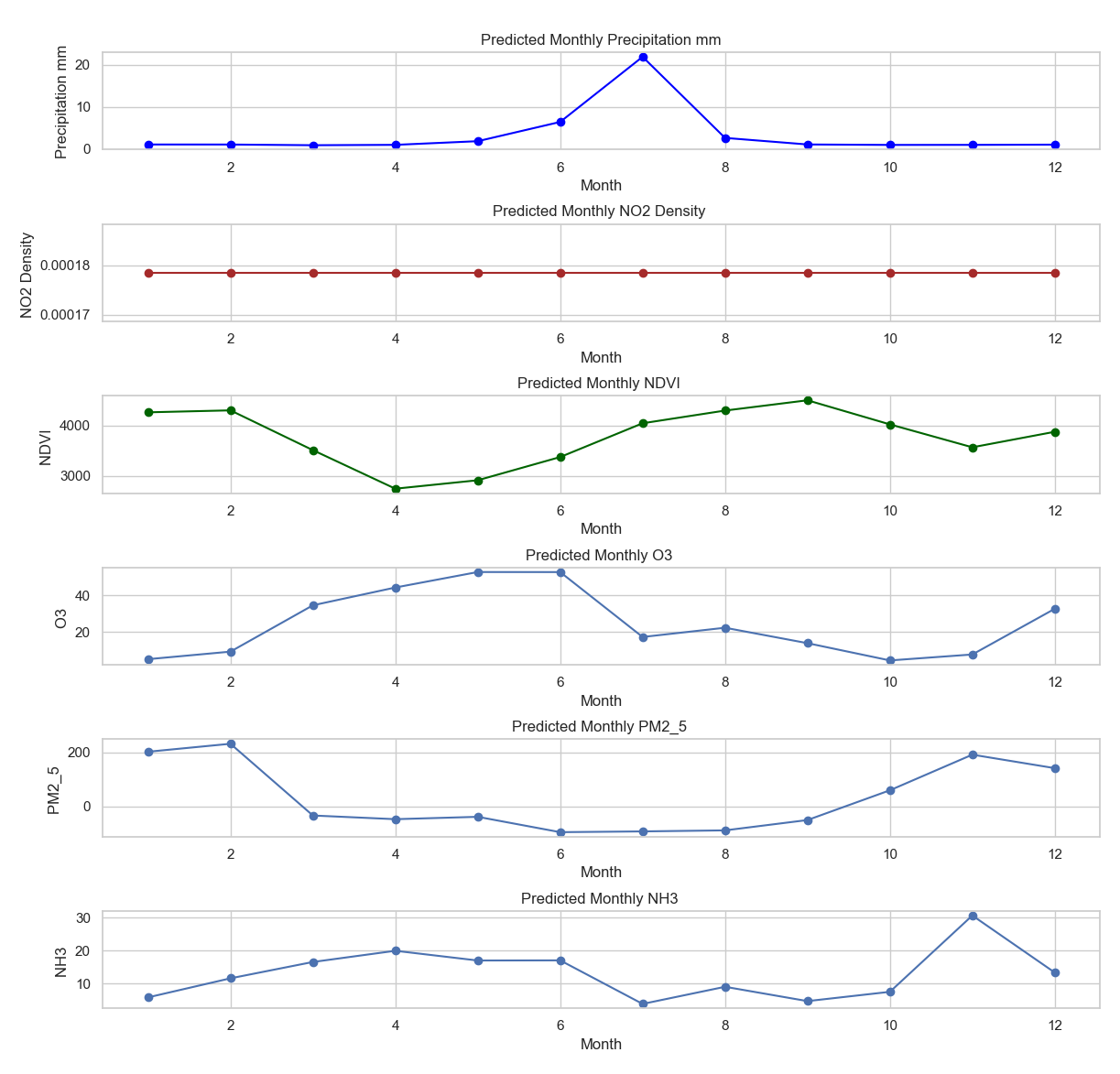


Fig. 3. Climate Prediction of Delhi(2024)

Developed a predictive framework using Random Forest and XGBoost models to assess climate change impacts in Delhi. The XGBoost model outperformed Random Forest, with a cross-validated RMSE of 133.85 compared to 224.94,

Key findings include:

* Temperature Trends: Elevated summer temperatures, slightly higher winter averages, influenced by urban heat island effects and reduced vegetation.
* Precipitation Patterns: Increased rainfall during monsoon months, reduced winter precipitation, suggesting water management challenges.
* ***Air Quality***: Peak pollution in winter, improved pollutant dispersion in summer and monsoon.
* Vegetation Indices: Robust growth in monsoon, decline in winter, indicating ecological stress.

***Limitations***  
While XGBoost performs well, its interpretability is limited compared to traditional models. Feature importance and partial dependence plots can help, but the model's complexity may be challenging for non-experts. Exploring hybrid models or deep learning could further enhance prediction accuracy by capturing spatial and temporal dependencies.

***Future Research Directions***

Studies can build upon this model by:

1. ***Expanding Data Inputs***: Integrating more variables, such as soil moisture or real-time satellite data, can enhance the prediction accuracy.
2. ***Enhancing Temporal Resolution***: Moving from monthly to daily data can improve precision, particularly for pollution metrics that fluctuate over short timescales.
3. ***Long-term Forecasting***: Extending the scope of the model to predict several years into the future could help assess the long-term impact of policy interventions on air quality and climate.

**VI CONCLUSION**

In this study, the XGBoost model demonstrated superior predictive accuracy for Delhi's climate variables, showcasing its effectiveness in handling complex environmental data. By leveraging multiple climate indicators including temperature, precipitation, and pollutant concentrations, the research provides valuable insights for urban climate planning. The model's ability to forecast monthly variations offers critical information for understanding potential climate and pollution trends, supporting proactive environmental management strategies.

**ACKNOWLEDGMENT**

We extend our sincere gratitude to Sharda University for their resources and support, the various data sources that enabled our research, and the climate scientists whose foundational work informed our study. Special thanks to our peers and mentors for their insightful feedback and encouragement throughout this project.

**REFERENCES**

[1] Z. Rahmani and J. The impact of human activities on climate change. Sprin Journal of Arts, Humanities and Social Sciences, 2024.

[2] Intergovernmental Panel on Climate Change (IPCC), Climate Change Information for Regional Impact and Risk Assessment, IPCC Sixth Assessment Report, 2023.

[3] B.-J. He et al., "Actions for Urban Climate Change Adaptation," in Climate Change and Environmental Sustainability, 2024.

[4] S. Meier et al., "A lightweight NO₂-to-NOₓ conversion model for quantifying NOₓ emissions," Atmospheric Chemistry and Physics, 2024.

[5] Z. Hejazizadeh et al., "Investigating the effects of climate change on torrential rain in Tehran Province," Journal Article, 2022.

[6] M.-H. Le et al., "On the Use of SMAP Soil Moisture for Forecasting NDVI," Geophysical Research Letters, 2024.

[7] S. N. Rahaman et al., "Effect of vegetation and land surface temperature on NO₂ concentration," Urban Climate, 2023.

[8] A. Agarwal and N. Tiwari, "The Economic Cost of Air Pollution Due to Stubble Burning," Journal of Pollution Effects & Control, 2024.

[9] Chen, "Severe Air Pollution Crisis in New Delhi," Journal Article, 2024.

[10] R. Robinson et al., "Climate Change, Pollution, and the Aerobiome," Mini-Review, 2024.

[11] Intergovernmental Panel on Climate Change (IPCC). Climate Change: The Physical Science Basis, IPCC Report, 2023.

[12] Svolos, "Public Perception of Climate Change," Journal of Environmental Psychology, 2024.

[13] Yaghoobzadeh, "Performance of Downscaling Methods for Climate Prediction in Iran," Climate Dynamics, 2024.

[14] Ding, "Climate Prediction Using a Random Forest Model," Atmospheric Environment, 2024.

[15] Lv et al., "Global Climate Change Prediction with CNN Models," Journal of Climate Research, 2024.

[16] He and Takemi, "Future Changes in Extreme Precipitation in East Asia," Climate Change, 2024.