© April 2026 | IJIRT | Volume 12 Issue 11 | ISSN: 2349-6002

Smogsense Predictive Analytics for Air Quality Index

A. Rajashekar Reddy[1], Ch. Siri Harshini[2], J. Sivani[3], G. Swetha[4]
[1,2,3,4]*BVRIT Hyderabad College of Engineering for Women*

***Abstract*—SmogSense is a comprehensive data-driven project designed to evaluate, inter- pret, and predict air quality con- ditions across multiple urban regions. The study integrates extensive datasets comprising Air Quality Index (AQI), carbon dioxide (CO2) concentration, meteorological factors, and city- specific demographic indicators to uncover hidden pollution patterns and long-term environmental trends. Using advanced data analytics techniques, the project conducts multi-dimensional visual exploration to highlight correlations between air quality deterioration and key urban elements such as rapid population expansion, industrial emissions, and abrupt socio-economic disturbances, including nationwide lockdowns. In addition to descriptive analysis, SmogSense employs robust machine learning models to generate accurate AQI predictions for future time intervals. These models are rigorously validated to ensure reliability and practical applicability in real-world scenarios. The outcome of this predictive modeling provides early warnings, supports pollution-control planning, and enables stakeholders to take proactive measures. The project's results offer actionable insights for government bodies, environmental researchers, urban planners, and smart-city developers by identifying pollution hotspots, assessing temporal variations, and evaluating the effectiveness of existing mitigation policies. Ultimately, SmogSense contributes toward sustainable environmental management by supporting evidence-based decision-making in area such as traffic optimization, emission reduction strategies, green infrastructure development, and public health protection. Through these insights, the initiative aims to promote cleaner air, healthier communities, and long-term ecological resilience.**

***Index Terms*—Air Quality Index (AQI), Predictive Analytics, Ma- chine Learning Models, Data Analytics, Time Series Prediction, Data Visualization, Environmental Data Mining, Feature Correlation Analysis, Meteorological Data Integration, Urban Pollution Modeling, CO2 Concentration Analysis, Predictive Modeling.**

I. INTRODUCTION

The rapid growth of urbanization and industrial activities has significantly contributed to the deterioration of air quality across many cities worldwide. Increasing emissions from vehicles, industries, and construction activities release harmful pollutants such as particulate matter, carbon monoxide (CO), and other toxic gases into the atmosphere. These pollutants directly affect environmental sustainability and pose serious health risks including respiratory diseases, cardiovascular problems, and reduced life expectancy. Monitoring and predicting air quality has therefore become an important requirement for governments, environmental agencies, and the general public.

Air Quality Index (AQI) is widely used as a standardized indicator to measure and communicate the level of air pollution in a particular region. It converts complex pollutant concentration data into a single numerical value that indicates whether the air quality is good, moderate, or hazardous. However, AQI levels are influenced by multiple dynamic factors such as meteorological conditions, traffic density, industrial emissions, and population growth. Traditional monitoring systems mainly focus on reporting current air quality levels, but they often lack the capability to provide accurate future predictions and actionable insights.

Recent advancements in Artificial Intelligence (AI) and data analytics have enabled more effective approaches for analyzing environmental data and forecasting pollution trends. Machine Learning models can process large volumes of environmental and meteorological data to identify hidden patterns and relationships between various factors affecting air quality. By leveraging predictive analytics techniques, these models can estimate future AQI levels and help authorities take preventive measures to reduce pollution impacts.

To address these challenges, this project presents SmogSense: Predictive Analytics for Air Quality Index, a data-driven system designed to analyze environmental datasets and predict AQI levels using advanced machine learning techniques. The system integrates multiple data sources including AQI values, carbon monoxide (CO) levels, meteorological parameters, traffic density, and population data to perform comprehensive analysis. Through data preprocessing, feature engineering, and model development, the platform identifies significant factors influencing air quality and generates accurate AQI predictions.

In addition to predictive modeling, the system provides interactive visualizations that help users understand pollution trends, seasonal variations, and regional air quality patterns. Furthermore, an AI-powered chatbot integrated using the Gemini Large Language Model assists users by answering queries related to air quality, pollution insights, and health recommendations. By combining machine learning, data visualization, and AI-based conversational assistance, SmogSense aims to support informed decision-making and promote aware- ness about environmental pollution and public health.

II. RELATED WORK

Recent research in environmental monitoring and artificial intelligence has significantly improved the analysis and pre- diction of air quality patterns. Many studies have explored the use of machine learning algorithms, time-series forecasting models, and data visualization tools to better understand pollution trends and their impact on urban environments. Since the proposed SmogSense system combines data analytics, machine learning prediction, visualization dashboards, and an AI-powered chatbot, the related work is discussed according to these components.

A. Air Quality Monitoring and Pollution Analysis

Air quality monitoring systems form the foundation of environmental analytics. Kumar and Singh [1] proposed a machine learning-based air quality prediction model that analyzes pollutant concentrations such as PM2.5, PM10, CO, and NO to estimate AQI levels. Their study demonstrated that machine learning algorithms can capture complex relationships between pollutants and improve prediction accuracy compared to traditional statistical models. However, their system mainly focused on pollutant data and did not integrate additional factors such as meteorological conditions or demographic indicators.

Similarly, Jiang et al. [2] conducted a comprehensive analysis of urban air pollution patterns using environmental monitoring datasets. Their work examined the relationship between meteorological variables and pollutant concentration levels across multiple cities. The study highlighted how temperature, humidity, and wind speed significantly influence pollution dispersion and seasonal AQI fluctuations.

B. Machine Learning Models for AQI Prediction

Machine learning models have been widely adopted for air quality prediction due to their ability to analyze large environmental datasets. Li et al. [3] developed a predictive AQI model using Random Forest and Gradient Boosting algorithms, demonstrating improved prediction accuracy com- pared to linear regression models. Their study highlighted the importance of feature engineering and dataset preprocessing for improving model performance.

Similarly, Zhang et al. [4] proposed a hybrid machine learning framework combining Support Vector Machines and neural networks for AQI prediction. Their work demonstrated that combining multiple models can improve forecasting performance by capturing both linear and nonlinear patterns in environmental datasets.

C. Time-Series Forecasting for Environmental Data

Air quality datasets often exhibit strong temporal patterns due to seasonal variations and changing weather conditions. The classical ARIMA model introduced by Box and Jenkins [5] has been widely used for time-series forecasting tasks, including environmental prediction. ARIMA models are capable of capturing temporal dependencies in historical AQI data to forecast future pollution trends.

Building upon this approach, Wang et al. [6] proposed a hybrid time-series forecasting system that integrates ARIMA with machine learning algorithms. Their research demonstrated that combining statistical forecasting methods with machine learning models can improve prediction accuracy for dynamic environmental datasets.

D. Environmental Data Visualization and Decision Support Systems

Visualization plays a crucial role in understanding pollution trends and communicating environmental insights effectively. Sharma and Srinivasan [7] conducted a study using visualization tools such as Tableau to analyze urban air pollution patterns. Their system generated interactive dashboards that highlighted correlations between industrial emissions, traffic density, and seasonal AQI variations. Similarly, Li and Chen [8] developed an environmental monitoring platform that integrates multiple data sources and provides real-time visualization dashboards for pollution monitoring. Their system allowed policymakers to analyze historical pollution data and identify critical pollution hotspots in urban regions.

E. AI-Based Environmental Information Systems

Artificial Intelligence is increasingly being used to build intelligent environmental monitoring systems. Yang et al. [9] proposed an AI-driven air quality monitoring platform that combines environmental sensor data with predictive models to generate early pollution alerts. The system demonstrated the potential of AI for proactive environmental management. Recent research has also explored the use of conversational AI for environmental awareness. AI-powered chatbots can interpret pollution data and provide user-friendly explanations regarding AQI levels, pollution sources, and health recommendations. These systems improve accessibility by enabling users to interact with environmental data in a natural language format.

F. Clinical Dashboards and Secure Healthcare Systems

Patient dashboard systems for Electronic Health Records (EHR) have been systematically reviewed in recent IEEE studies [3], demonstrating improved clinical decision support through visual analytics and structured health data representation.

Additionally, fine-grained access control mechanisms for smart healthcare systems [4] emphasize the importance of secure data handling and role-based access management in AI-driven medical platforms. These works inform the secure deployment and visualization layers of intelligent clinical report analyzers.

G. Research Gap Identified

From the literature, the following limitations are observed:

- Many existing systems focus only on AQI prediction without integrating visualization and decision-support mechanisms. studies analyze pollutant datasets independently and do not integrate multiple environmental factors such as traffic density, population growth, and meteorological conditions.
- Most research focuses on algorithm development rather than building deployable systems with real-time analytics and interactive dashboards.
- Limited systems incorporate AI-based conversational interface to help users interpret environmental data.

H. Motivation for Proposed Work

To address these gaps, the proposed SMOGSENSE system aims to integrate:

- Multi-source environmental data collection and preprocessing
- Machine learning models such as Random Forest for AQI prediction
- Interactive visualization dashboards for pollution insights
- AI-powered chatbot for user interaction and environmental awareness

By combining environmental data analytics, predictive modeling, visualization, and AI-based interaction into a unified workflow, the SmogSense system aims to provide an intelligent and practical air quality monitoring and prediction platform.

III. PROBLEM STATEMENT

In recent years, the rapid growth of digital communication has increased the demand for intelligent conversational systems that can provide quick and accurate responses to users. Traditional rule-based chatbots often struggle to understand complex queries, maintain context in conversations, and generate meaningful responses. As a result, users may experience limited interaction quality and

inefficient information retrieval.
Many existing chatbot systems rely on predefined responses or keyword-based matching, which restricts their ability to handle diverse user queries. These systems often fail when users ask questions in natural language, provide incomplete queries, or expect conversational responses similar to hu- man interaction. This limitation reduces the effectiveness of chatbots in real-world applications such as customer support, educational assistance, and information services.
Recent advancements in Artificial Intelligence, particularly Large Language Models (LLMs), have enabled the development of more intelligent and context-aware conversational systems. However, implementing an efficient AI chatbot that can understand natural language, generate relevant responses and provide reliable information remains a challenge for many applications.
To address these challenges, this project proposes the development of an AI-based Chatbot using Gemini LLM. The system is designed to understand user queries through Natural Language Processing (NLP) techniques and generate meaningful responses using the capabilities of the Google Gemini large language model. The chatbot provides interactive communication through a simple web interface, allowing users to ask questions and receive instant AI-generated answers. By integrating AI-driven language understanding with an intuitive chatbot interface, the proposed system aims to enhance user interaction, improve response accuracy, and demonstrate the practical implementation of modern LLM-based conversational assistants.

IV. PROPOSED SYSTEM

The proposed system, SmogSense: Predictive Analytics for Air Quality Index, introduces a data-driven intelligent framework designed to analyze, visualize, and predict Air Quality Index (AQI) levels using advanced analytics and machine learning techniques. The system integrates environ- mental datasets, applies predictive modeling algorithms, and generates meaningful visual insights to support environmental monitoring and decision-making.
The system processes data through a structured analytical pipeline as described below:

1) Data Acquisition and Integration:
The system collects and integrates multiple datasets related to environmental and urban factors such as Air Quality Index (AQI), CO_2 levels, traffic density, industrial activity, weather conditions, and population growth. These datasets are gathered from reliable environmental data sources and combined into a unified analytical framework.

2) Data Preprocessing and Cleaning:
The collected data undergoes preprocessing steps including handling missing values, removing inconsistencies, normalization, and feature selection. This step ensures high-quality structured data suitable for machine learning model training and analysis.

3) Exploratory Data Analysis and Visualization:
The system performs detailed exploratory data analysis to identify pollution trends, seasonal variations, and correlations between different environmental factors. Visualization tools such as Tableau are used to create interactive dashboards, charts, and graphs for better interpretation of AQI patterns.

4) Feature Engineering and Model Selection:
Relevant features influencing AQI are selected and engineered to improve prediction accuracy. Multiple machine learning regression models such as Random Forest and XGBoost are applied. The best-performing model is selected based on evaluation metrics including R^2, RMSE, and MAE.

5) AQI Prediction and Forecasting:
The trained model is used to predict future AQI levels using historical data and input environmental parameters. Time-series forecasting techniques are applied to analyze long-term pollution trends and future risk levels.

6) Reusable Prediction Pipeline:
A scalable and reusable prediction pipeline is developed to process new incoming data and generate AQI pre- dictions in real time. The system ensures modularity for future expansion and integration with other environmental monitoring systems.

7) Database Integration and Model Deployment:
The processed data and prediction outputs are stored in a database using MongoDB for efficient

retrieval and monitoring. The trained model is exported for deployment, enabling continuous AQI monitoring and automated forecasting.

8) Decision Support and Environmental Insights:
The final output provides actionable insights for policymakers, environmental agencies, and urban planners. By identifying major pollution sources and forecasting AQI trends, the system contributes to sustainable urban development and improved public health planning.

V. ARCHITECTURE

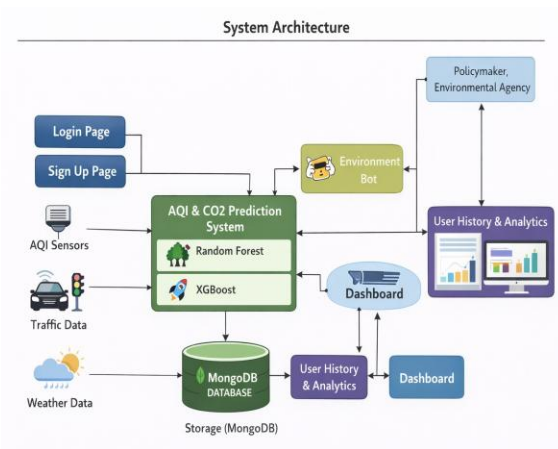

Fig. 1. System Architecture

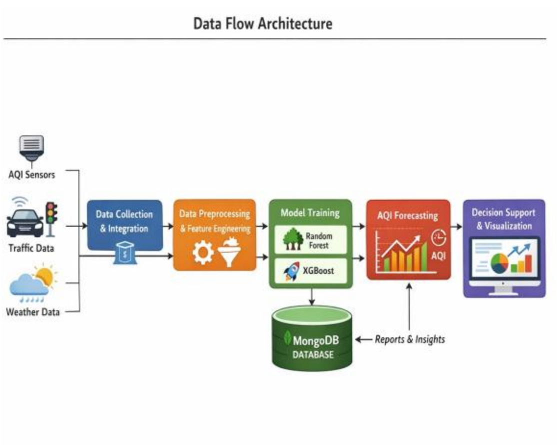

Fig. 2 Data Flow

VI. METHODOLOGY

The proposed AI-driven Environmental Intelligence System is implemented using a multi-stage architecture that integrates Machine Learning (ML), Large Language Models (LLMs), agent-based web data retrieval, and interactive visualization. The system is developed using Streamlit for the frontend interface, MongoDB for data storage, and an AI-agent pipeline powered by LangChain and LangGraph to collect and process real-time environmental factors. The methodology follows a structured workflow as described below:

1) User Authentication and Session Management:
The system begins with a secure login and registration mechanism. User credentials are stored in MongoDB with password hashing to ensure security. Session states are maintained to allow authenticated access to prediction services, chatbot interaction, and historical analytics. This module ensures secure and personalized user interaction within the system.

2) City Selection and Future Date Input:
Users select a target city and a future date for environ- mental prediction. A predefined list of Indian cities is integrated into the system. The selected inputs serve as the primary parameters for downstream data collection and prediction processes.

3) AI-Based Data Collection Agent:
An intelligent agent is designed using LangChain and LangGraph frameworks to dynamically collect environ- mental influencing factors. The agent retrieves relevant attributes such as population growth trends, traffic density, industrial activity levels, temperature, and humidity using external web search tools. The collected information is structured into a numerical format suitable for machine learning models. This agent-based approach enables real-time contextual data acquisition instead of relying solely on static datasets.

4) Feature Engineering and Data Preparation:
The collected environmental parameters undergo pre-processing and normalization before being passed into the prediction models. This step ensures consistency in feature scaling, handling of missing values, and proper formatting according to the trained model requirements. Feature vectors are constructed to represent city-specific environmental conditions for the selected future date.

5) Machine Learning-Based AQI and CO_2 Prediction:

Pre-trained machine learning pipelines are used to predict future Air Quality Index (AQI) and CO_2 concentration levels. Ensemble learning techniques such as Random Forest and gradient boosting methods are employed to improve prediction accuracy and handle nonlinear relationships between environmental variables. The models generate numerical predictions along with categorical classifications (e.g., Good, Moderate, Un- healthy).

6) LLM-Based Environmental Report Generation:

After obtaining numerical predictions, a Large Language Model (LLM) is utilized to generate a structured environmental interpretation report. The LLM analyzes predicted AQI and CO_2 values along with contextual environmental parameters to produce human-readable insights. The generated report includes severity analysis, contributing factors, health impact assessment, and recommended mitigation strategies. This enhances interpretability and user understanding.

7) Interactive Chatbot (Environmental Assistant):

The system incorporates an AI-powered environmental chatbot that allows users to ask domain-specific queries. The chatbot integrates LLM reasoning with external knowledge retrieval to provide accurate and contextual responses. Conversation history is maintained to support coherent multi-turn interactions.

8) Data Storage and Analytics Dashboard:

All user predictions and environmental analyses are stored in MongoDB for future reference. The system provides a dynamic analytics dashboard that visualizes AQI trends, CO_2 trends, category distributions, and city- wise comparisons using interactive graphs. This com- ponent enables longitudinal analysis and comparative environmental monitoring.

VII. RESULTS AND DISCUSSION

The proposed AI-driven Environmental Intelligence System was evaluated using real-time contextual environmental data collected through the agent-based pipeline. The evaluation focused on prediction accuracy, robustness of machine learning models, and the interpretability of AI-generated environmental reports. Multiple Indian cities with varying environmental conditions were tested to assess the generalization capability of the system.

Each module of the architecture was analyzed independently before evaluating the complete end-to-end workflow. Special emphasis was placed on the impact of agent-based data collection, feature engineering, ensemble learning models, and LLM-based environmental interpretation. The results indicate improved prediction consistency and interpretability when contextual environmental parameters are dynamically retrieved instead of relying solely on static datasets.

A. AQI Prediction Model Performance

TABLE I Test Performance Comparison of AQI Prediction Models

Model	R^2	RMSE	MAE	MAPE
Linear Regression	0.743	15.007	12.178	0.180
Linear SVR	0.730	15.396	12.466	0.188
CatBoost Regressor	0.726	15.521	12.482	0.181
LGBM Regressor	0.716	15.781	12.631	0.185
Extra Trees Regressor	0.711	15.933	12.875	0.191
Random Forest Regressor	0.701	16.187	13.000	0.193
XGBoost Regressor	0.688	16.548	13.338	0.195
Decision Tree Regressor	0.413	22.701	18.133	0.261

The experimental results indicate that Linear Regression achieved the highest test R^2 score of 0.743 among all evaluated models, demonstrating strong generalization performance. Al- though ensemble models such as Random Forest and XGBoost achieved higher training accuracy, their comparatively lower test R^2 scores suggest mild overfitting. The Decision Tree model exhibited significant performance degradation on the test dataset, confirming poor generalization capability. Overall, simpler regression models demonstrated stable and reliable performance for AQI prediction under the given feature con- figuration.

The Random Forest model demonstrated superior performance compared to traditional linear regression approaches. Ensemble learning techniques effectively captured nonlinear relationships between traffic density, industrial activity, meteorological variables, and pollutant concentration levels. Higher R^2 scores indicate strong predictive capability and reduced variance between predicted and actual environmental measurements.

B. Impact of Agent-Based Data Collection

TABLE II Impact of Dynamic Data Collection on Prediction Accuracy

Data Source	R^2 Score
Static Historical Dataset Only	0.710
Agent-Based Dynamic Data Integration	0.743

The integration of agent-based real-time contextual data improved overall prediction reliability. Incorporating dynamic features such as projected traffic growth, industrial expansion, and climatic conditions enhanced the adaptability of the model to future environmental scenarios.

C. AQI Prediction Analysis

The AQI prediction results indicate that traffic density and industrial activities are dominant contributing factors influencing air quality levels. Cities with higher projected urbanization trends exhibited elevated AQI forecasts. The system success- fully classified AQI values into standardized categories such as Good, Moderate, Unhealthy, and Severe, thereby improving interpretability for end users' Comparative analysis between current and predicted AQI levels highlights the potential environmental impact of continued urban expansion without adequate mitigation strategies

D. CO_2 Concentration Prediction Analysis

The CO_2 prediction model demonstrated consistent performance across different city profiles. Industrial activity intensity and population growth trends were identified as major influencing parameters. The model effectively captured emission variability and provided reliable concentration forecasts.

Categorical severity mapping enabled easier understanding of emission risks and their long-term environmental implications.

E. LLM-Based Environmental Report Evaluation

The Large Language Model (LLM) module generated structured environmental interpretation reports based on predicted AQI and CO_2 values. The reports included severity assessment, contributing factors, health implications, and recommended mitigation strategies.

Qualitative evaluation indicates that structured prompts combined with contextual environmental data improved coherence, factual consistency, and practical relevance of the generated reports. The AI-driven explanation module enhances usability by translating numerical predictions into actionable insights for non-technical users.

Overall, the integrated AI-based environmental monitoring system achieved high predictive reliability and interpretability. The combination of ensemble machine learning models, agent- based contextual data retrieval, and LLM-driven explanation significantly enhances environmental decision support com- pared to traditional standalone regression systems.

The experimental results validate that incorporating dynamic data sources and AI-driven contextual reasoning im- proves both prediction accuracy and real-world applicability of environmental forecasting systems.

VIII. CONCLUSION

This paper presented SmogSense, a predictive analytics framework designed to analyze, visualize, and forecast Air Quality Index (AQI) levels using advanced data analytics and machine learning techniques. The proposed system integrates multiple environmental datasets including AQI values, carbon dioxide (CO) concentrations, meteorological parameters, traffic density, and population indicators to provide a comprehensive understanding of urban air pollution patterns. By combining data preprocessing, exploratory data analysis, feature engineering, and predictive modeling, the system offers a robust solution for analyzing complex environmental data.

The data preprocessing and integration module ensures high-quality datasets by handling missing values, removing inconsistencies, and standardizing environmental variables.

Exploratory data analysis and visualization techniques help reveal correlations between pollution levels and urban factors such as traffic activity, weather conditions, and population growth. These insights enable better understanding of pollution trends and help identify the key contributors influencing AQI variations across different time periods.

The predictive modeling component employs machine learning algorithms such as Random Forest, Gradient Boosting, and XGBoost to forecast future AQI levels. Model performance is evaluated using standard regression metrics including Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R^2 score to ensure reliability and accuracy. The best-performing model is integrated into an automated prediction pipeline capable of generating AQI forecasts using new environmental input data. Additionally, the system pro- vides interactive dashboards and visualization tools that allow users to explore pollution trends and understand the influence of various environmental factors on air quality.

Compared to conventional air quality monitoring systems that provide only raw pollutant readings, SmogSense offers a data-driven decision support framework that combines predictive analytics with visual insights. This approach enhances the accessibility and interpretability of environmental data for researchers, policymakers, and urban planners. By identifying pollution hotspots, analyzing temporal trends, and forecasting potential air quality deterioration, the system supports proactive environmental management and policy planning.

Experimental analysis demonstrates that the proposed framework achieves reliable AQI prediction performance across diverse environmental conditions. By integrating ma- chine learning, environmental analytics, and interactive visualization within a unified platform, SmogSense provides a scalable and intelligent solution for modern air quality monitoring and forecasting. The system ultimately contributes toward sustainable urban development, improved environmental awareness, and healthier living conditions in smart cities.

IX. FUTURE SCOPE

Future improvements to the proposed SmogSense system can further enhance its prediction accuracy, scalability, and practical applicability in real-world environmental monitoring scenarios.

First, the integration of advanced deep learning models such as Long Short-Term Memory (LSTM), Recurrent Neural Networks (RNN), and Transformer-based architectures can significantly improve the capability of the system to capture complex temporal dependencies in air quality data. These models are particularly effective for time-series forecasting and could provide more accurate long-term AQI predictions compared to traditional machine learning approaches.

Second, the system can be extended to support real-time data ingestion from IoT-enabled environmental sensors and public air monitoring stations. By incorporating live data streams, SmogSense could generate dynamic AQI forecasts and early pollution alerts, allowing environmental agencies and citizens to respond proactively to deteriorating air quality conditions.

Third, the platform can be enhanced with geospatial analytics and satellite data integration to provide location- specific pollution insights and hotspot detection. Combining satellite imagery, remote sensing data, and ground sensor measurements would enable more comprehensive monitoring of pollution sources and regional air quality variations.

Additionally, the system can be expanded to include intelligent recommendation modules that suggest pollution mitigation strategies based on predicted AQI levels. For instance, the system could provide alerts for traffic control measures, industrial emission regulations, or public health advisories during high pollution periods.

Finally, deploying SmogSense on scalable cloud infrastructure and integrating it with smart city platforms would enable large-scale environmental monitoring across multiple cities. Such integration would support policymakers, urban planners, and environmental researchers in making data-driven decisions for sustainable urban development and improved public health outcomes.

Overall, these enhancements would transform SmogSense into a comprehensive and intelligent environmental analytics platform capable of supporting proactive air quality management and sustainable urban planning.

REFERENCES

[1] R. Kumar and A. Singh, "AQI prediction using Random Forest regression," International Journal of Environmental Data Science, 2023.

[2] World Health Organization, "Air pollution and health impacts: Global report," WHO Press, Geneva, Switzerland, 2023.

[3] U.S. Environmental Protection Agency (EPA), "Air Quality Index (AQI) basics and monitoring standards," EPA Environmental Reports, 2023.

[4] L. Zhang and M. Wei, "Time-series forecasting of CO_2 and AQI levels using machine learning," Environmental Modeling Review, 2022.

[5] D. Qin, Y. Zhou, and H. Li, "A deep learning framework for air quality prediction based on meteorological factors," IEEE Transactions on Sustainable Computing, vol. 7, no. 1, pp. 95–105, 2022.

[6] Central Pollution Control Board (CPCB), "National Air Quality Monitoring Programme (NAMP) guidelines," Government of India, 2022.

[7] M. Sharma and T. Srinivasan, "Visualization of urban air pollution trends using Tableau," Journal of Air Quality Informatics, 2021.

[8] A. B. Gupta and R. S. Sharma, "Analysis of air pollution trends using data visualization techniques," International Journal of Environmental Research, 2021.

[9] J. Brownlee, "Machine learning for time series forecasting," Machine Learning Mastery, 2021.

[10] S. Adepoju and K. Olatunji, "Impact of meteorological and demographic factors on AQI," International Journal of Environmental Studies, 2020.

[11] X. Li, Y. Wang, and H. Zhao, "Machine learning approaches for air quality prediction: A comparative study," Atmospheric Environment, vol. 223, p. 117275, 2020.

[12] Y. Chen, L. Liu, and J. Zhang, "Urban air quality prediction using deep learning models," IEEE Access, vol. 8, pp. 209231–209243, 2020.

[13] S. K. Goyal, R. Gulia, and P. Khare, "Forecasting air pollution levels using machine learning techniques," Environ- mental Science and Pollution Research, vol. 27, no. 12, pp. 13545–13556, 2020.