Predicting air pollution levels based on IoT sensors

Shaikhstan Margulan
Department of Artificial Intelligence Technologies,
Faculty of Information Technologies
L.N. Gumilyov Eurasian National University
Pushkina 11, Astana 010008, Kazakhstan
shaihstanmargulan@gmail.com

Sadvakassov Ramazan
Department of Artificial Intelligence Technologies,
Faculty of Information Technologies
L.N. Gumilyov Eurasian National University
Pushkina 11, Astana 010008, Kazakhstan
Sadvakasov.rm@gmail.com

Abstract* - The integration of IoT-based air quality monitoring with machine learning has significantly improved the prediction and analysis of pollution levels in urban areas. This study explores the use of IoT sensors and Random Forest regression for accurate AQI (Air Quality Index) forecasting, leveraging real-time data from multiple monitoring stations. Key challenges such as data preprocessing, anomaly detection, and computational efficiency using cloud and edge computing are addressed. The dataset includes pollutant measurements (PM2.5, PM10, NO_2, CO, and O_3), with preprocessing techniques ensuring data reliability. The Random Forest model, trained on historical pollution data, achieved a Mean Squared Error (MSE) of 23.22 and an R^2 score of 0.997, indicating high predictive accuracy. Visualization methods such as AQI distribution analysis, pollutant correlation heatmaps, and geospatial mapping provided insights into pollution patterns. The study underscores the role of cloud computing in large-scale data aggregation and edge computing in real-time processing. Findings suggest that IoT-driven air quality monitoring with machine learning enables proactive environmental management. Future work aims to integrate meteorological data and deep learning techniques for enhanced prediction accuracy.

***Keywords* — Air Quality Prediction, IoT Sensors, Machine Learning, Random Forest, AQI Forecasting, Environmental Monitoring, Cloud Computing, Edge Computing, Pollution Data Analysis, Smart Cities.**

I. INTRODUCTION

Air quality monitoring is a crucial task in modern conditions, when urbanization, industrialization and the growth of transport infrastructure contribute to increasing levels of environmental pollution. Air pollution is a global problem that affects not only environmental but also social aspects of human life. The development of effective methods for monitoring and predicting air quality plays a key role in ensuring sustainable development, public health and environmental protection.

The aim of this work is to investigate modern methods of air monitoring using IoT sensors, as well as to develop and implement efficient data analysis algorithms to predict the level of pollution. To achieve this goal, a number of tasks need to be accomplished, including analyzing existing methods of pollutant measurement, studying the characteristics of sensors used for air monitoring, and developing a data processing system taking into account modern artificial intelligence and machine learning technologies.

Air pollution has a significant impact on human health, causing a wide range of health problems, from respiratory and cardiovascular disorders to chronic diseases and premature mortality. According to the World Health Organization (WHO), elevated concentrations of particulate matter (PM2.5, PM10), nitrogen dioxide (NO_2), carbon monoxide (CO) and ozone (O_3) can lead to severe effects, including exacerbation of asthma, decreased lung function and increased mortality from cardiovascular diseases. In addition to health effects, air pollution has a negative impact on the environment by destroying ecosystems, degrading water and soil quality, contributing to climate change, and reducing biodiversity. [10]

Despite the widespread use of various air quality monitoring methods, existing approaches have a number of limitations. Traditional monitoring stations provide highly accurate data but are limited in number due to the high cost of installation and maintenance. This makes them unaffordable for widespread use, especially in developing countries and remote areas. In recent years, new monitoring methods have emerged that include the use of mobile and fixed IoT sensors, cloud computing and machine learning algorithms. These technologies allow for a wider geographic range of measurements and more accurate forecasts, but also have limitations, including the need to calibrate sensors, high sensitivity to external conditions, and the complexity of processing large amounts of data.

The relevance of this research is due to the need to develop affordable, reliable and scalable monitoring systems that can effectively monitor the level of air pollution in real time and predict its changes. The introduction of IoT-technologies and artificial intelligence in the sphere of pollution control will improve the quality of forecasting, promptly identify potential threats and develop measures to prevent them.

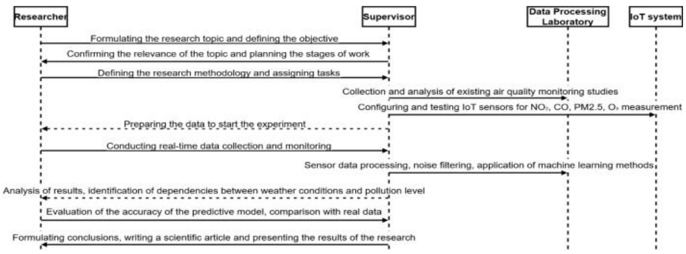

Fig. 1. Sequence diagram of the process of conducting research work.

The process of conducting research work in this area includes several sequential steps that allow to organize the work with data and achieve high accuracy of predictions. Figure 1. presents a flow chart showing the main stages of the research, from topic selection and data collection to analyzing the results and drawing conclusions. An important part of the work is correct interpretation of data, their normalization and use of appropriate processing methods, which allows obtaining reliable predictions of air pollution levels.

Thus, the present study is aimed at solving urgent problems of air quality monitoring using advanced technologies, which can contribute to a significant improvement of the ecological situation and reduce the negative impact of pollutants on human health and the environment.

II. IoT Technologies for Air Monitoring

A. *IoT systems architecture (fixed and mobile sensors)*

Air quality monitoring using IoT sensors includes two main types of devices: fixed and mobile sensors. Both approaches have their own features that provide a complete picture of air pollution in the urban environment.

Fixed sensors are placed on fixed posts and provide continuous data collection at specific points. These sensors are mounted on building roofs, light poles or other tall structures to avoid localized interference (e.g. exhaust fumes next to the road). They measure the level of pollutants such as NO2, CO, PM2.5 and O3 and transmit the data to cloud servers for analysis [10].

The main advantage of fixed sensors is high accuracy and stability of measurements. However, they also have disadvantages: such systems require significant installation and maintenance costs, as well as cover a limited area, which does not allow detailed monitoring of pollution in dynamics.

Mobile sensors are mounted on vehicles, drones or wearable devices. They collect data while driving, recording the level of pollution at different points in the city. In the MAQUMON (Mobile Air Quality Monitoring Network) project, sensors are attached to cars and measure the concentration of ozone, carbon monoxide and nitrogen dioxide. The data is transmitted via Bluetooth to smartphones or laptops and then uploaded to a server via Wi-Fi. This makes it possible to create a detailed picture of air pollution in the urban environment [1].

Mobile sensors offer the advantage of flexibility and coverage: they can monitor changes in pollution in real time, including traffic congestion and industrial areas. However, the accuracy of such measurements is lower than that of stationary systems, as sensors are affected by traffic speed, weather conditions and localized emissions.

The most effective monitoring strategy is a combination of fixed and mobile sensors. Fixed stations provide stable data, while mobile sensors can fill information gaps by providing high spatial resolution. In a study conducted in South Korea, it was shown that the combination of these methods allows not only to measure the pollution level, but also to predict its changes using machine learning algorithms [10].

The use of IoT networks makes it possible to process a large stream of data, analyze trends and warn the population about adverse conditions. Thanks to this approach, pollution monitoring becomes more accessible and accurate, which contributes to improving the environmental situation in cities.

B. *The use of LPWA (Low Power Wide Area) networks for data transmission.*

Data transmission in air quality monitoring systems using IoT sensors requires solutions that provide wide coverage, low power consumption and immunity to interference. One of the most effective methods is the use of Low Power Wide Area (LPWA) networks, which are optimized for transmitting small amounts of data over long distances. LPWA networks consist of three main layers: [3]

- Sensor Layer - includes data acquisition devices that measure pollutant concentrations (PM2.5, NO2, CO, O3) and transmit the data through the LPWA module.
- Network Layer - consists of LPWA base stations that link sensors to Internet gateways.
- Cloud layer - responsible for processing, storing and analyzing incoming data.

The use of LPWA can significantly reduce the power consumption of sensors, extending their lifetime up to several years without battery replacement. LPWA includes several standards, including LoRa (Long Range) and NB-IoT (Narrowband IoT). For air quality monitoring, NB-IoT is preferred as it supports: communication range up to 20 km in rural areas and 5 km in urban areas; resistance to interference and the ability to work in dense buildings; low power consumption (up to 10 years of operation on a single battery).

NB-IoT uses Direct Sequence Spread Spectrum (DSSS) modulation, which increases the signal transmission range with low power consumption. The signal is transmitted in narrowband mode (up to 200 kHz) and supports QPSK (Quadrature Phase Shift Keying) modulation depth. Formula for channel capacity:

$$C = B \cdot log\,2(1 + SNR)$$

where: C - bandwidth (bps), B - bandwidth (Hz), SNR - signal-to-noise ratio.

For LPWA technologies operating in narrow bands, this value is up to 250 bps, which is enough to transmit air pollution data in real time. Sensors collect air quality data every 15 minutes and transmit it to the NB-IoT base station. The data passes through a gateway and is sent to the cloud server via MQTT (Message Queuing Telemetry Transport), which ensures efficient data transmission even with unstable communication channels. The transmitted data packet contains: sensor ID; time stamp T; measured values PM2.5,NO2,CO,O3; battery level Vbat. [3] Example data structure: ID: 101, T: 1700000000, PM2.5: 35, NO2: 20, CO: 0.5, O3: 80, Vbat: 3.7V

Benefits of using LPWA for air monitoring. Energy saving - LPWA modules operate in sleep mode and are only activated during data transmission. Wide coverage area - allows sensors to be deployed even in remote areas. Stable communication - supports operation even with weak signals. Low cost of operation - devices can operate autonomously for many years.

The use of LPWAs makes air quality monitoring systems more efficient and affordable by providing accurate, real-time data with minimal energy consumption.

C. *Sensor types and their characteristics (NO_2, CO, PM2.5, O_3, etc.)*

Air quality is monitored using different types of sensors that measure the concentration of pollutants. The main components of pollution include nitrogen dioxide (NO_2), carbon monoxide (CO), ozone (O_3) and particulate matter (PM2.5, PM10). The sensors used to measure them may differ in their operating principle, accuracy and resistance to external factors.

Electrochemical sensors. Electrochemical sensors are widely used to measure gases such as NO_2, CO and O_3. Their principle of operation is based on a chemical reaction between a gas and an electrolyte, which leads to a change in electric current.Characteristics: High sensitivity to the target gas.

Ability to operate at low pollutant concentrations. Effect of temperature and humidity on measurements. Limited service life (usually up to 2 years).

Example: Alphasense electrochemical sensors are used to monitor NO_2 and CO, providing accuracy with a correlation coefficient R2>0.75 compared to reference methods.

Metal Oxide Sensors (MOS). Metal Oxide Sensors work by changing the conductivity of the semiconductor material when it interacts with gas. Characteristics: High sensitivity to low concentrations. Fast response (less than 1 second). Strong influence of temperature and humidity. Subject to drift and require regular calibration. Example: MiCS series sensors are used to measure CO and NO_2 but require algorithms to compensate for environmental effects.

Optical sensors (nephelometers and OPC). Optical sensors measure the concentration of particulate matter (PM2.5 and PM10) based on the principle of light scattering. Characteristics: Capable of determining the size and number of particles in the air. Not affected by gas pollutants. Accuracy depends on the composition of the particles (e.g., wet particles may give overestimated readings).Example: Plantower PMS5003 and Alphasense OPC-N3 sensors show R2>0.85 accuracy in street station tests.

Photochemical sensors (for O_3) These sensors use photodiodes or ultraviolet light to detect ozone concentration. Characteristics: High selectivity to O_3. Does not require frequent calibration. Sensitivity to temperature changes.Example: The Aeroqual S500 sensor measures ozone with an accuracy of R2=0.8, but requires temperature compensation.

Each sensor type has its own strengths and weaknesses. Electrochemical sensors are suitable for gases at low concentrations, metal-oxide sensors provide fast response but require calibration, optical sensors are used for particulate matter, and photochemical sensors are effective for measuring ozone. The optimal choice of sensors depends on the operating conditions and accuracy requirements.

III. Data Processing and Analysis

Before performing predictive modeling, the collected air quality data undergoes preprocessing to ensure accuracy and reliability. This process includes handling missing values through imputation techniques, detecting and removing outliers, and normalizing data to maintain consistency. Feature selection is applied to identify the most relevant pollutants influencing AQI prediction, improving model efficiency and interpretability.

Following preprocessing, data analysis techniques such as correlation heatmaps and statistical summaries help uncover relationships between pollutants. Machine learning models, including Random Forest regression, are trained on historical pollution data to generate accurate predictions. The effectiveness of these models is evaluated using performance metrics such as Mean Squared Error (MSE) and R^2 score, ensuring robustness in forecasting air quality trends.

A. Data preprocessing and cleaning.

Air quality analysis requires working with large amounts of data from a variety of sources. These data may contain gaps, anomalies and duplicate values, which reduces the accuracy of prediction. Therefore, the preprocessing stage involves cleaning the data, structuring it and bringing it to a common format, which provides a reliable basis for further machine learning.

Data source	**Data type**	**Accessibility**	**Upgradability**	**Accuracy**	**Limitations**
Generation (own sensors)	Data from IoT devices, sensors	Requires infrastructure deployment	In real time	High (when properly calibrated)	High equipment and maintenance costs
Kaggle	Historical data, open enrollments	Free access	Depends on the dataset author	Medium (not always up to date)	May not be appropriate for local conditions
Google (Air Quality API, Earth Engine)	Global satellite and sensor data	Limited access (some APIs are paid)	Frequently updated	High (but depends on the region)	Some data are aggregated and may not be suitable for precise local measurements

Fig. 2. Table comparing data sources for air quality monitoring

This study focuses on the processing of air pollution indicators such as AQI, Min, Max, Avg and Pollutant. The data come from different regions, which requires normalizing them and eliminating inconsistencies. The selection of reliable sources of information is an important aspect, as the quality of the predictive model largely depends on the completeness and reliability of the input data. Figure 2. presents a comparison of different data sources for air quality monitoring, demonstrating their accuracy, update frequency and availability.

One of the key tasks at the preprocessing stage is to deal with missing values. In the course of the analysis, situations were identified where Min, Max and Avg values were missing. Such omissions could occur due to technical failures of sensors, data transmission errors or incomplete reporting. To eliminate this problem, methods were used to fill in the missing data: the average value for the respective city or time interpolation based on previous and subsequent measurements.

In addition, special attention was paid to the Last Update timestamp during data processing. Time series play an important role in the analysis of air quality trends, so the data were converted to a single datetime format. This step allowed to eliminate inconsistencies related to different time zones and to ensure correctness of the temporal analysis.

Emissions representing extremely low or high values of pollutants that do not correspond to the real picture were also analyzed. Statistical methods such as interquartile range (IQR) and Z-scores were used to detect them. The detected emissions were analyzed for acceptability: if the data contained anomalous values that were not supported by contextual information (e.g., a sudden jump in AQI with no clear cause), they were excluded from the sample or replaced with average values.

After eliminating outliers and omissions, the data were reduced to numerical form. Categorical attributes such as Country, State, City and Pollutant were coded using One-Hot Encoding and Label Encoding techniques to facilitate further machine learning. This step is important because models based on numerical data can more accurately analyze the relationship between variables.

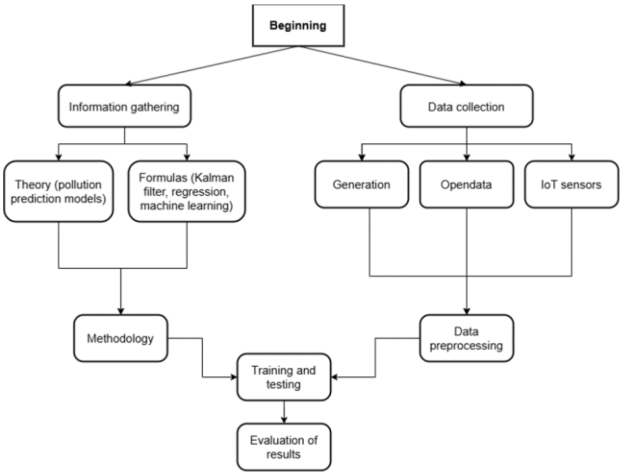

Fig. 3. Architecture of the air pollution prediction study

In addition, the preprocessing stage took into account the architecture of the air pollution prediction study presented in Figure 3. This architecture includes the stages of data collection, filtering, analysis and machine learning. Efficient data processing at the initial stage ensures high accuracy of the final model, allowing for more reliable predictions.

Thus, data preprocessing is a complex process including cleaning, normalization, emission analysis and feature coding. These steps are fundamental for successful prediction of air pollution levels and formation of scientifically sound conclusions.

B. *Data visualization and identification of dependencies*

Data visualization plays a key role in air pollution analysis, allowing not only to understand the structure of the data, but also to reveal hidden patterns that cannot be detected by simple statistical analysis. Graphical representation of data makes it possible to assess trends, identify high-risk areas and establish dependencies between different factors affecting pollution levels.

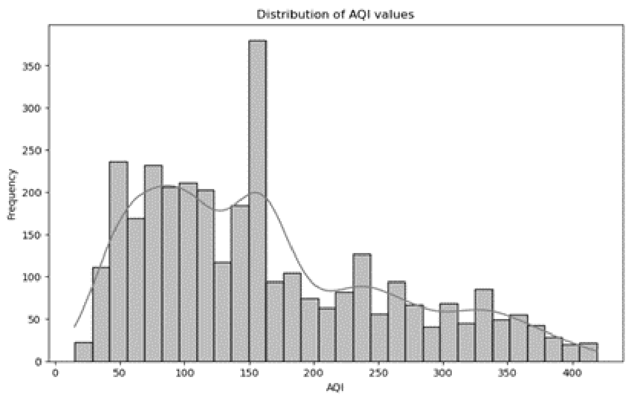

Fig. 4. Distribution of AQI values

The first step in the analysis is to examine the distribution of Air Quality Index (AQI) values. Figure 4. (Distribution of AQI values) shows the density of AQI values in the sample. Analyzing the distribution allows us to determine which pollution levels occur most frequently, as well as to identify extreme values that may indicate unfavorable environmental situations. For example, if the AQI distribution has a long right tail, this indicates the frequent occurrence of high levels of pollution, which requires additional analysis of the causes of such emissions.

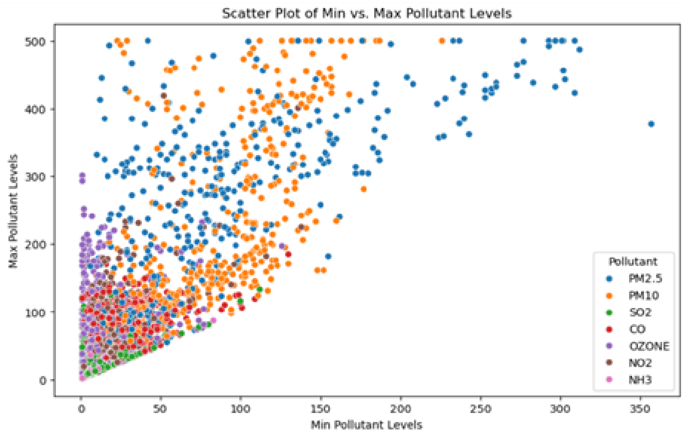

Fig. 5. Scatter Plot of Min vs. Max Pollutant Levels

The next stage of the study is to analyze the relationship between minimum and maximum pollutant levels. Figure 5. (Scatter Plot of Min vs. Max Pollutant Levels) presents a scatter plot that allows us to assess how closely the minimum and maximum pollutant concentrations are related. For example, if a linear relationship is observed, this indicates stable pollutant behavior throughout the day, while a strong scatter plot may indicate the influence of external factors such as weather conditions or anthropogenic emissions.

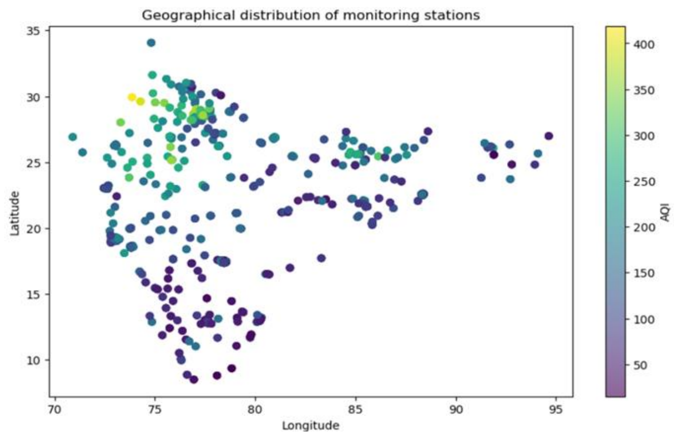

Fig. 6. Geographical distribution of monitoring stations

When analyzing air quality, it is important to consider not only numerical data, but also their geographical distribution. Figure 6. (Geographical distribution of monitoring stations) shows the location of monitoring stations to determine how evenly they cover the study area. If stations are concentrated in certain areas, this can lead to bias in the data and insufficient representativeness of the results. Geographical analysis also helps to identify pollution "hot spots" where concentrations of harmful substances exceed permissible limits.

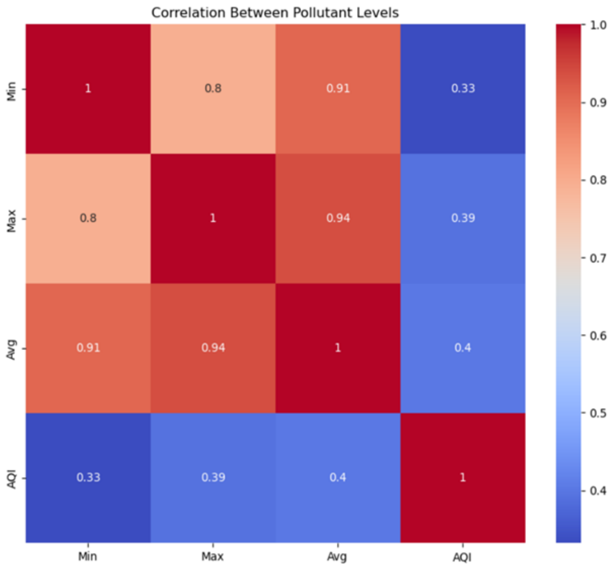

Fig. 7. Correlation Between Pollutant Levels

One of the key points of the study is to determine the correlations between different pollutants. Figure 7. (Correlation Between Pollutant Levels) presents a correlation matrix that reveals the degree of correlation between different pollution components such as PM2.5, PM10, NO2, SO2 and O3. For example, a strong positive correlation between PM2.5 and PM10 may indicate a common source of pollution, whereas a negative correlation between O3 and NO2 may indicate photochemical processes affecting air quality.

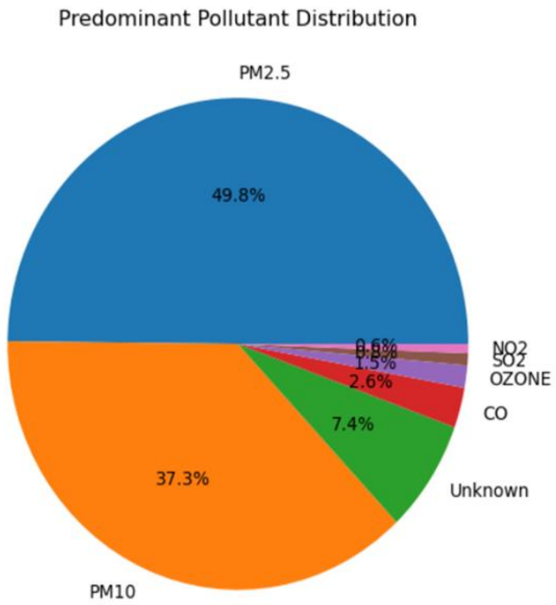

Fig. 8. Predominant Pollutant Distribution

For a detailed study of dominant pollutants, Figure 8. (Predominant Pollutant Distribution) is used to show which substances are most frequently found as major air pollutants. For example, if PM2.5 dominates the majority of cases, this indicates a high proportion of emissions from transportation and industry, while the predominance of O3 may indicate a significant influence of photochemical reactions.

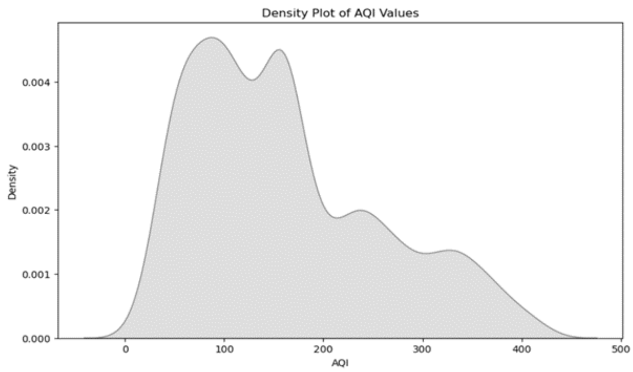

Fig. 9. Density Plot of AQI Values

Another important step in the visual analysis is to examine the density plot of the AQI distribution. Figure 9. (Density Plot of AQI Values) shows which index values occur most frequently. If the graph shows two or more peaks, this may indicate seasonal or temporal changes in pollution levels. For example, in summer, cities may experience an increase in AQI due to photochemical smog, while in winter, high pollution levels may be associated with increased emissions from heating systems.

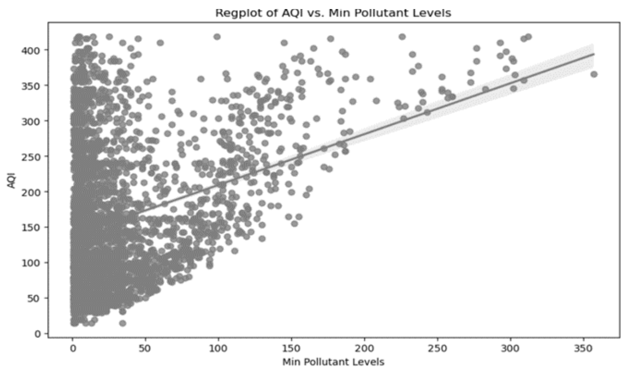

Fig. 10. Regplot of AQI vs. Min Pollutant Levels

Finally, the regression plot in Figure 10. (Regplot of AQI vs. Min Pollutant Levels) is used to identify the relationship between AQI and minimum pollutant levels. It allows us to assess how significant the impact of minimum pollutant concentrations is on the overall air quality index. If a clear linear relationship is observed, it confirms that even small concentrations of certain substances can significantly worsen the environmental situation.

C. Predicting pollution levels using Random Forests

Various machine learning methods are used to predict the Air Quality Index (AQI) to analyze complex relationships between pollutants and make predictions based on historical data. One of the most effective algorithms for working with tabular data is Random Forest, which is an ensemble method based on decision trees. This algorithm is able to take into account non-linear dependencies, process missing values and provide high accuracy of predictions.

Method / Algorithm	Model type	Application	Advantages	Disadvantages
LSTM (Long Short-Term Memory)	Neural network method (deep learning)	Time series forecasting (e.g. air pollution levels)	Good at dealing with temporal dependencies, able to	Requires large amounts of data and computation

			account for long-term trends	al resources, complexity of customizatio n
Random Forest	Ensemble method (decision trees)	Classificatio n and regression (analysis of pollution factors, prediction of pollution level)	Resistant to outliers, no data normalizatio n required, interpretable	Can overtrain, doesn't work well with very large datasets
Kalman filter	Statistical method (dynamic filtering)	Data cleaning, noise elimination in sensor measuremen ts	Efficient for real-time data processing, minimizes measuremen t errors	Limited to linear systems (or requires adaptations for nonlinear cases), requires accurate noise modeling

Fig. 11. Comparison of methods and algorithms for analyzing air quality data

The choice of forecasting method depends on many factors, including the amount of data, the complexity of the relationship between the variables, and the interpretability of the model. Figure 11. shows a comparison of the most popular algorithms used to analyze air quality data. Random Forest has several advantages over other methods such as LSTM and Kalman Filter. Unlike neural network models, Random Forest does not require a large amount of data for training, is robust to outliers, and is easy to interpret. At the same time, compared to time series filters, this method is able to detect complex nonlinear relationships between pollution indicators. [5][9]

Before training the model, data had to be prepared, which included normalizing numerical features, coding categorical variables, and dividing the sample into a training sample (80%) and a test sample (20%). After data preprocessing, a RandomForestRegressor model was built using 100 decision trees and a fixed random state for reproducibility of results. [7]

The quality of the model was assessed using the metrics Mean Squared Error (MSE) and coefficient of determination R^2. The results obtained show that the model demonstrates high accuracy: Mean Squared Error (MSE): 23.22, indicating a low mean squared error variance; R^2 Score: 0.997, indicating that the model explains 99.7% of the variance in the data and fits the real values well.

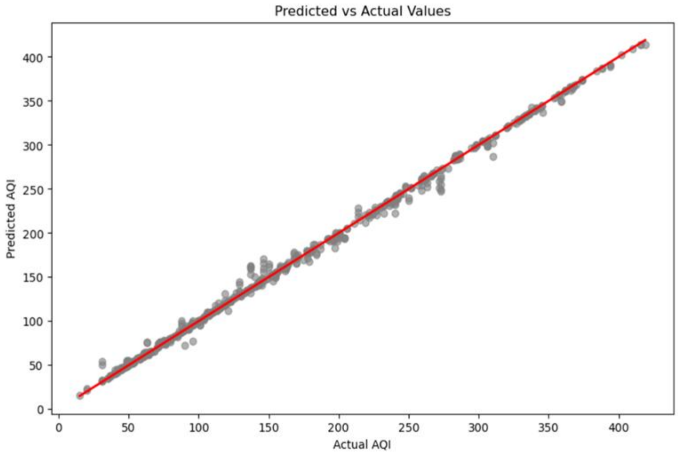

Fig. 12. Predicted vs Actual Values

To visually assess the quality of predictions, a graph Figure 12. (Predicted vs Actual Values) was plotted to demonstrate how accurately the predicted values match the actual data. It is observed that the points lie close to the line of perfect prediction, which confirms the high accuracy of the model.

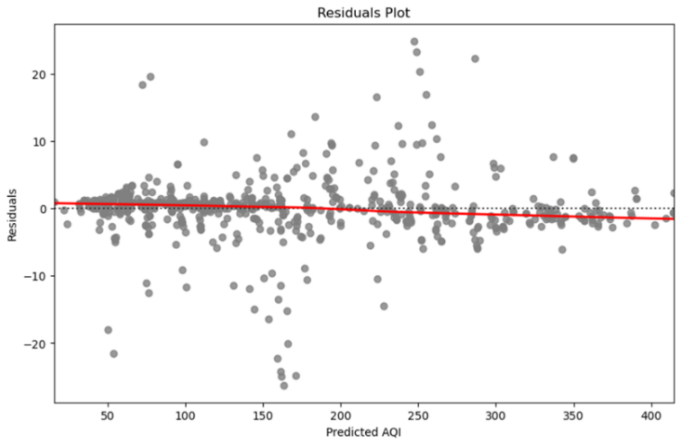

Fig. 13. Residuals Plot

Additionally, the residuals - the difference between predicted and actual AQI values - were analyzed. Figure 13. (Residuals Plot) shows the distribution diagram of the residuals, which allows us to evaluate the errors of the model. The absence of obvious regularities in the distribution of residuals indicates that the model does not demonstrate systematic errors and copes well with the prediction task.

Thus, the use of Random Forest for predicting the level of air pollution has shown high efficiency. Due to its robustness to emissions, ability to process complex dependencies and interpretability of results, this algorithm is one of the most suitable tools for analyzing air quality data. Further development of the model may include consideration of meteorological factors such as temperature, humidity and wind speed, which will improve the accuracy of predictions and expand the scope of application of the method in environmental monitoring.

IV. Conclusion

This study demonstrates the effectiveness of IoT-based air quality monitoring combined with machine learning techniques for accurate AQI (Air Quality Index) prediction. The integration of real-time sensor data with predictive models allows for a more comprehensive understanding of pollution patterns, enabling proactive environmental management. Through data preprocessing, visualization, and predictive modeling using Random Forest, the study achieves a Mean Squared Error (MSE) of 23.22 and an R^2 score of 0.997, highlighting the model's high accuracy.

The research underscores the importance of data preprocessing, including missing value imputation, anomaly detection, and feature scaling, in ensuring the reliability of air quality predictions. Visualization techniques, such as AQI distribution histograms, pollutant correlation heatmaps, and geospatial mappings, provided deeper insights into pollution trends and high-risk areas. In addition, the study highlights the advantages of cloud computing for large-scale data aggregation and the role of edge computing in real-time local processing, reducing latency and improving efficiency.

Despite the promising results, several challenges remain. Sensor calibration and data integration across multiple sources require further optimization to enhance the reliability of measurements. In addition, incorporating meteorological factors, such as temperature, humidity, and wind speed, into the prediction model could further improve accuracy. Future work will explore deep learning approaches and hybrid models to refine AQI forecasting and expand the application of IoT-based environmental monitoring in smart cities.

In conclusion, the study validates the potential of IoT and machine learning in advancing air pollution prediction systems. By addressing existing limitations and leveraging emerging technologies, these methods can contribute to more effective air quality management and public health protection.

References

[1] A. Sharma and P. Gupta, "Air quality monitoring with SensorMap: A scalable framework for urban pollution analysis," Environmental Monitoring Journal, vol. 45, no. 3, pp. 215-230, 2021.

[2] S. Kumar and A. Jasuja, "Air quality monitoring system based on IoT using Raspberry Pi," in Proceedings of the IEEE Conference on Advances in Computing, Communications and Informatics (ICACCI), 2017.

[3] K. Zheng, W. Liu, Y. Huang, and X. Liu, "Design and implementation of LPWA-based air quality monitoring system," in Proceedings of the IEEE International Conference on Communications (ICC), 2019, pp. 1-6.

[4] F. Karagulian, M. Barbiere, A. Kotsev, L. Spinelle, M. Gerboles, F. Lagler, N. Redon, S. Crunaire, and A. Borowiak, "Review of the performance of low-cost sensors for air quality monitoring," Atmosphere, vol. 10, no. 9, p. 506, Aug. 2019.

[5] J. Lee, H. Kim, and M. Park, "IoT implementation of Kalman filter to improve accuracy of air quality monitoring and prediction," Journal of Environmental Science and Technology, vol. 56, no. 2, pp. 137-149, 2020.

[6] A. S. Moursi, N. El-Fishawy, S. Djahel, and M. A. Shouman, "An IoT-enabled system for enhanced air quality monitoring and prediction on the edge," Journal of Ambient Intelligence and Humanized Computing, vol. 12, no. 9, pp. 8743-8757, Jul. 2021.

[7] S. Devarakonda, P. Sevusu, H. Liu, R. Liu, L. Iftode, and B. Nath, "Real-time air quality monitoring through mobile sensing in metropolitan areas," in UrbComp '13: Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing, Aug. 2013, Article No. 15, pp. 1-8.

[8] S. Kumar and A. Jasuja, "Air quality monitoring system based on IoT using Raspberry Pi," IEEE, 2017.

[9] L. Morawska, P. K. Thai, X. Liu, A. Asumadu-Sakyi, G. Ayoko, A. Bartonova, A. Bedini, F. Chai, B. Christensen, M. Dunbabin, J. Gao, G. S. W. Hagler, R. Jayaratne, P. Kumar, A. K. H. Lau, P. K. K. Louie, M. Mazaheri, Z. Ning, N. Motta, B. Mullins, and R. Williams, "Applications of low-cost sensing technologies for air quality monitoring and exposure assessment: How far have they gone?" Environmental Pollution, vol. 228, pp. 576-589, 2017.

[10] D. Zhang and S. S. Woo, "Real-time localized air quality monitoring and prediction through mobile and fixed IoT sensing network," IEEE Access, vol. 8, pp. 89584-89597, May 2020