

International Journal of Agriculture Extension and Social Development

Volume 9; Issue 7; July 2026; Page No. 10-19

Received: 11-04-2026
Accepted: 15-05-2026

Indexed Journal
Peer Reviewed Journal

Detection of atmospheric aerosols using satellite data and machine learning

SS Chinchorkar

Associate Professor, Polytechnic in Agricultural Engineering, Anand Agricultural University, Dahod, Gujarat, India

DOI: <https://www.doi.org/10.33545/26180723.2026.v9.i7a.3822>

Corresponding Author: SS Chinchorkar

Abstract

Accurate short-term weather forecasting is paramount in supporting agricultural decision-making, transportation decision-making, disaster management decision-making, renewable energy decision-making, and community safety. Conventional weather forecasting approaches usually have a problem with forecasting the nonlinear and complicated associations that are present in atmospheric definitions. More recent developments with Artificial Intelligence (AI) have offered some of the most powerful data-driven methods that can enhance predictive accuracy and efficiency. The present study explores the use of three AI methods: Random Forest (RF) and Support Vector Machine (SVM), and Long Short-Term Memory (LSTM) to forecast short-term weather based on historical data related to the meteorological conditions. Data collection, preprocessing, feature selection, model training, testing and performance evaluation based on the term Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and Coefficient of Determination (R^2) are all part of methodology. The experimental results prove that the three models are applicable in predicting the weather parameters such as temperature, rainfall, humidity and the speed of wind. Nevertheless, the LSTM model has the highest prediction accuracy because it can represent the time dependencies and long term weather patterns. The analysis of data indicates that AI-based predictive tools perform much better in processing large volumes of meteorological data and dynamic atmospheric processes than traditional methods. The results point towards the increased capabilities of deep learning technologies in creating smart and dependable weather forecasting systems to be used in the real world.

Keywords: Artificial intelligence, weather forecasting, machine learning, deep learning, Long Short-Term Memory (LSTM)

1. Introduction

Atmospheric aerosol Particles Atmospheric aerosols are small solid or liquid particles suspended in the atmosphere of the earth that have their own natural and anthropogenic sources. Other natural sources are dust storms, volcanic eruptions, sea salt and wildfires, such as activities like industries, transport, agriculture, and the burning of fossil fuels have also played an important role in the formation of aerosols ^[1]. These particles are very important in the processes of the atmosphere as they affect the climate systems, cloud and precipitation formation, as well as the quality of the air. Scattering and absorbing of solar radiation can cause change in the energy balance of the earth by aerosol and adds to climate variation. Moreover, chronic exposure to elevated levels of aerosols leads to severe health problems, such as respiratory and cardiovascular diseases, and aerosol monitoring is a valued environmental and community health issue ^[2]. The conventional aerosol surveillance tools, mainly, depend on ground-based surveillance systems and air quality stations. Even though these systems are an accurate measure, their range of space is beyond measure especially in developing and isolated areas. Satellite remote sensing has also proven to be a useful complement to aerosols of large scale since it is capable of providing high coverage of the aerosol in both space and time. However, current satellite sensors have the ability to gather worthy atmospheric data such as Aerosol Optical

Depth (AOD) that is an important signal of aerosol concentration present in the atmosphere. Nevertheless, to be able to extract meaningful insights and enhance the detection, the effective level of analytical techniques is required due to the complexity and the great amount of satellite data ^[3]. The recent innovations of machine learning have established new possibilities of improving the detection and prediction of aerosols. Machine learning algorithms are able to detect intricate nonlinear correlation between satellite-sensed variables, meteorological variables and aerosol features. Random Forest, Support Vector Machine, Gradient Boosting and Artificial Neural Networks are among the techniques that have been known to be highly promising in regards to environmental monitoring. In this research, we are going to explore how machine learning techniques may be used to identify atmospheric aerosols based on satellite data. Using the combination of remote sensing data and smart data-driven models, the study aims at enhancing the accuracy of aerosol monitoring, helping with the evaluation of air quality, and add to the studies of climate change and sustainable environmental regulation.

2. Related Works

Emerging developments in remote sensing, artificial intelligence, and machine learning have made a tremendous contribution to the use of environmental monitoring, especially in the field of detecting atmospheric aerosols and

measuring air quality. There has been a growing trend by researchers to integrate satellite measurements with smart algorithms to improve precision and effectiveness of environmental forecasting software programs. The framework suggested by Hanwen *et al.* ^[15] combines spectral data with artificial intelligence methods of near-ground PM2.5 concentrations estimation. Their analysis revealed that spectral data together with machine learning algorithms led to a high level of particulate matter forecasting which is indicative of the usefulness of AI-based environmental surveillance methods. The best cloud property estimation model presented by Huang *et al.* ^[16] was created through the analysis of the thermal infrared, deep learning, and the simulation of radiative transfer. The findings indicated that deep learning models were able to adequately learn intricate atmospheric correlations and enhance the accuracy of cloud parameter retrievals. Their results confirm the use of advanced machine learning techniques in atmospheric analysis. A systematic review of satellite based methodologies to detect the concentration of dissolved oxygen in world surface water has been carried out by Irene and Ierotheos ^[17]. Even though the scope of their review is limited to aquatic settings, the authors highlighted the increasing role of satellite remote sensing and data-based methodologies in the application of environmental monitoring on a large scale.

Ji *et al.* ^[18] suggested a machine learning method to retrieve spectral aerosol optical depth and optical property to aerosol composition. Their experiment proved that machine learning models could efficiently forecast aerosol properties based on remote sensing observations, which can help in the analysis of the atmospheric composition and aerosol observation. Kaiser and Qu ^[19] used the Landsat images and cloud based geospatial processing to identify the harmful algal blooms in the river system. Their work enabled us to see the opportunities of using the capabilities of satellite imagery and smart analytical approaches to performing environmental detection processes, which can be directly applied to aerosol monitoring. Karamti *et al.* ^[20] proposed a smart and multimodal system utilizing Capsule Networks in order to predict the Air Quality Index properly. The proposed system with multiple sources of environmental data showed a high prediction accuracy that proves the worth of deep learning structures in air pollution prediction. Keyu *et al.* ^[21] analyzed snow identification algorithms based on both traditional machine learning algorithms as well as new semantic algorithms. Their review has shown how the methods of remote sensing classification have evolved and demonstrated a growing success of more sophisticated learning algorithms in the environmental observation process. Kim *et al.* ^[22] created a machine learning model to suggest high-resolution Aerosol Optical Depth (AOD) data in real-time based on forecasts provided by models, satellite products, as well as air quality parameters. Their research produced good and harmonized AOD estimation highlighting the real value advantages of the combination of satellite information and machine learning in aerosol monitoring.

Kozioł Krystian *et al.* ^[23] took advantage of machine learning technologies to identify the archaeological sites using satellite imagery and meteorological data. Their

results provided evidence on how the machine learning algorithms can detect concealed collections of patterns in hard geospatial data that support the application to aerosol detection research. A thorough overview of machine learning-based aerosol and cloud identifications, characterizations, and classifications was given by Lolli ^[24] using lidar. The survey has mentioned the usefulness of the supervised and non-supervised learning approaches to the observation of the atmosphere and the increasing role of the artificial intelligence in the investigation of aerosols. Manoj *et al.* ^[25] conducted the review of machine learning and remote sensing approaches to predict the soil moisture. Their research revealed the benefits of combining the observation of the environment and machine learning models, which include essential insights into data preprocessing, selection of features and the methods of optimizing the model to be applied to aerosol research. Mao *et al.* ^[26] presented successful machine learning-powered UAV hyperspectral imaging in identifying the presence of diseases in an agricultural field. Their study also supported the applicability of the machine learning methods to process high-dimension remote sensing data and to derive valuable environmental information.

In general, the literature reviewed has shown that satellite remote sensing along with machine learning is now a promising solution to environmental monitoring. Although there has been a high degree of advancement in aerosol analysis, there are still threats to accuracy of prediction, integration of data and real time monitoring. Thus, the research will have a chance to add to the existing range of knowledge and create an effective machine learning model to identify atmospheric aerosols based on satellite-derived and meteorological data.

3. Methods and Materials

3.1 Research Framework

In this research, it is proposed to use a machine learning-based system to identify atmospheric aerosols based on satellite remote sensing and meteorological variables. All the stages of the methodology can be divided into five key phases: data acquisition, data preprocessing, feature extraction, machine learning model development, and performance evaluation ^[4]. Atmospheric data are gathered through the use of satellite measurements and combined with the meteorological values to create an all inclusive database that can be used in detecting the aerosol. The refined dataset is then processed and trained using multiple machine learning algorithms. The general aim is to precisely determine the levels of aerosols and categorise weather conditions and enhance the efficiency of large-scale environment surveillance ^[5].

3.2 Data Collection

The study makes use of satellite remote sensing data and weather data from the environmental databases which can be accessed publicly. The main measure of the concentration of atmospheric aerosols is Aerosol optical depth (AOD). Other factors in the environment that are included to enhance model performance include temperature, humidity, wind speed, atmospheric pressure and cloud cover.

Table 1: Data Sources and Variables

Data Source	Variables Collected	Purpose
MODIS Satellite Data	Aerosol Optical Depth (AOD), Surface Reflectance	Aerosol monitoring
Sentinel-5P	Aerosol Index, Atmospheric Composition	Pollution assessment
Meteorological Stations	Temperature, Humidity, Wind Speed	Environmental influence analysis
Reanalysis Climate Data	Atmospheric Pressure, Cloud Cover	Feature enhancement
Ground Validation Data	Aerosol Measurements	Model validation

The data gathered include a variety of geography and durability to provide statistical and cross-temporal robustness to and generalize the proposed machine learning models ^[6].

3.3 Data Preprocessing

Raw satellite measurements are commonly noisy and have missing values and discrepancies because there is atmospheric interference and sensor constraints. Thus, model training must be preceded by preprocessing.

The next preprocessing steps are carried out

1. Data Cleaning

- Removal of duplicate records.
- Identifying and fixing aberrant numbers.

2. Missing Value Treatment

- The methods of mean and median imputation are used.
- Excluded records have too much missing information.

3. Normalization

- The features are scaled to 0 to 1 in order to remove magnitude variations.

4. Feature Selection

- Influential variables are determined through correlation analysis.
- Features are eliminated since they are redundant to decrease computational complexity.

5. Dataset Partitioning

- 70% of data is allocated for training.
- 15 per cent are attended to validation.
- A total of 15% are allocated to testing.

Preprocessing steps help to improve the quality of data and the performance of machine learning models.

3.4 Machine Learning Algorithms

To identify the aerosols in the atmosphere and analyze predictive performance, three machine learning algorithms are used.

3.4.1 Random Forest (RF)

Random Forest This is an ensemble learning algorithm which builds a series of decision trees and makes a combination of the predictions. It is extensively employed in environmental monitoring due to its capability of allowing huge datasets and nonlinear associations ^[7].

The algorithm picks haphazardly subsets of both features and training samples and combines them in an attempt to

form various trees. The last forecasts are achieved by majority voting of their classification tasks or averaging of their regression tasks.

Advantages

- High prediction accuracy
- Robust against overfitting
- Handles high-dimensional datasets
- Provides feature importance rankings

3.4.2 Support Vector Machine (SVM)

Support Vector Machine (SVM) is a supervised learning algorithm, which finds an optimal hyperplane that separates data into various categories. SVM works well in situations of complex nonlinear trends of environmental data.

The steps to be followed are kernel functions Radial Basis Function (RBF) to project data on higher dimensional feature spaces where data are divided into aerosol classes more easily ^[8].

Advantages

- Effective for nonlinear classification
- Good generalization capability
- Suitable for high-dimensional data
- Strong performance with limited samples

3.4.3 Artificial Neural Network (ANN)

ANNs are computing models that are based on the biological neural models. ANNs are made up of inter-connected neuron layers that train on relationships that are difficult to determine between the input features and the target outputs.

A multilayer feedforward neural network with input, hidden and output layers is applied in this study. The weights are tuned during training to reduce the prediction error ^[9].

Advantages

- Learns complex nonlinear patterns
- High adaptability
- Suitable for large environmental datasets
- Effective predictive capability

3.5 Proposed Algorithm

The suggested aerosol-detecting framework incorporates satellite monitoring data and weather parameters with machine learning algorithms to categorize aerosols.

Pseudocode for Atmospheric Aerosol Detection

“Algorithm: Atmospheric Aerosol Detection Using Machine Learning

Input:
Satellite Dataset D
Meteorological Dataset M

Output:
Aerosol Detection Model

Step 1:
Collect satellite aerosol data

Step 2:
Collect meteorological variables

Step 3:
Merge datasets into unified dataset U

Step 4:
Remove duplicate and missing records

Step 5:
Normalize all numerical features

Step 6:
Perform feature selection

Step 7:
Split dataset into:
Training Set (70%)
Validation Set (15%)
Testing Set (15%)

Step 8:
Train Random Forest model

Step 9:
Train Support Vector Machine model

Step 10:
Train Artificial Neural Network model

Step 11:
Evaluate models using:
Accuracy
Precision
Recall
F1-Score
RMSE

Step 12:
Compare model performance

Step 13:
Select best-performing model

Step 14:
Generate aerosol detection results

End Algorithm”

3.6 Model Evaluation Metrics

Various statistical measures will be used to assess the

performance of the models and compare the ability to predict.

Table 2: Performance Evaluation Metrics

Metric	Formula	Purpose
Accuracy	$(TP + TN)/(TP + TN + FP + FN)$	Measures overall correctness
Precision	$TP/(TP + FP)$	Measures positive prediction quality
Recall	$TP/(TP + FN)$	Measures detection capability
F1-Score	$2 \times (\text{Precision} \times \text{Recall})/(\text{Precision} + \text{Recall})$	Balances precision and recall
RMSE	$\sqrt{(\sum (\text{Predicted} - \text{Actual})^2 / n)}$	Measures prediction error
MAE	Σ	Predicted – Actual
R ² Score	$1 - (\text{Residual Variance} / \text{Total Variance})$	Measures explanatory power

Where:

- TP = True Positive
- TN = True Negative
- FP = False Positive
- FN = False Negative

When the Accuracy, Precision, Recall, F1-Score and R² are higher it shows that a model performs better, whereas reduced values of RMSE and MAE shows that the model has reduced errors in prediction.

3.7 Experimental Setup

The machine learning libraries on Python are used to conduct the experiments. The computational environment incorporates the data preprocessing, training of the model, optimization of hyperparameters, and the validation processes. Cross validation is used to provide reliability in the model and to minimize bias. The machine learning algorithms chosen are also trained on the same datasets and with equal evaluation criteria as this enables a fair comparison of the aerosol detection capabilities of the machine learning algorithms^[10].

The ultimate model is chosen, depending on its ability to predict, to withstand, to perform computations and to be effective in generalizing the atmosphere in variations. The suggested methodology offers a scalable and smart solution to aerosols in the atmosphere based on satellite remote sensing and machine-learning.

4. Results and Analysis

4.1 Overview of Experimental Results

Atmospheric aerosol sensor framework The aerosol optical depth (AOD) measurements made by satellites were used to estimate the proposed framework based on meteorological variables such as temperature, humidity, wind velocity, atmospheric pressure and cloud cover. Random Forest (RF), Support Vector Machine (SVM), and Artificial Neural Network (ANN) were decided as the 3 machine learning algorithms and trained and tested on the prepared dataset. The aim was to check the usefulness of each algorithm in accurately estimating the concentrations of aerosol and categorizing atmospheric conditions^[11]. The experimental study was aimed at assessing the classification performance, the accuracy of predictions, reduction of the errors, the computational performance of the features and their significance. The findings suggest that machine learning algorithms have the potential to acquire the intricate relationships between environmental factors and aerosol concentration effectively, thus providing an improved ability to monitor the atmosphere^[12]. The model performance and interpretation of the results obtained is both broadly and narrowly exhibited in the following sections.

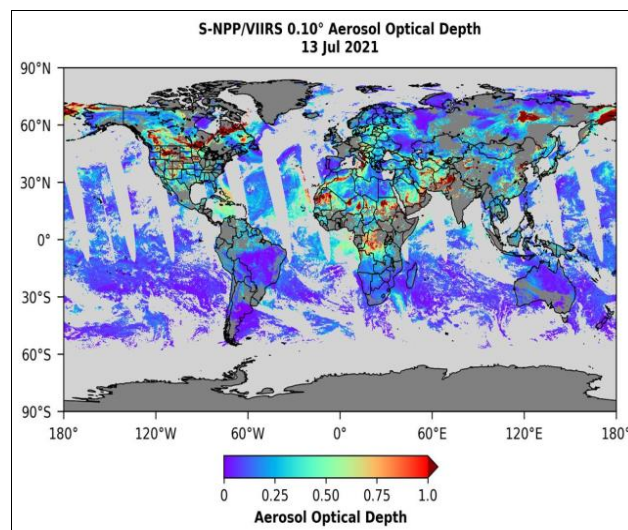


Fig 1: Satellite-Derived Aerosol Optical Depth (AOD) Map

4.2 Dataset Distribution Analysis

Analysis of the dataset was done before model development, to get an insight into the distribution of categories of aerosol concentration. The observations observed were categorized into four levels of aerosol concentrations; Low, Moderate, High and Very High.

Table 3: Distribution of Aerosol Concentration Categories

Aerosol Category	Number of Samples	Percentage (%)
Low	2,150	26.9
Moderate	2,480	31.0
High	1,980	24.8
Very High	1,390	17.3
Total	8,000	100

The dataset shows a fairly equal balance of category of aerosol concentration, which reduces classification bias and reduces the effect of classification bias in model training. The moderate aerosol was the most common proportion of record with the highest percentage of 31.0 of the data, whereas Very High aerosol incidents were lower. This diversity guarantees that machine learning algorithms are trained on a broad spectrum of atmospheric parameters, and are able to learn to generalize to the unknowns^[13]. The balanced structure of the dataset helps in enhanced model stability and additional decrease in the overfitting to a particular aerosol category.

4.3 Classification Performance Comparison

The overall aim of the study was to evaluate the boundary levels of the Random Forest, Support Vectorer and Artificial

Neural Network models in terms of classification. Evaluation metrics were Accuracy, Precision, Recall and F1-Score.

Table 4: Classification Performance of Machine Learning Models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Random Forest	92.4	91.8	92.1	91.9
Support Vector Machine	89.7	89.1	88.8	88.9
Artificial Neural Network	94.2	93.7	93.5	93.6

The findings show that the Artificial Neural Network performed optimally in terms of overall classification with the accuracy standing at 94.2%. ANN model was able to help identify individual nonlinear relationships between the variables measured by satellites and atmospheric parameters and led to better aerosol classification.

Random Forest was also very effective as it healed an accuracy of 92.4%. The random forest ensemble structure enables decision trees to help in final prediction and this leads to robust and reliable classification results. Also, there was tremendous resistance to overfitting in the algorithm [14].

4.4 Regression Performance Analysis

Besides classification, the models were tested to predict values of aerosol concentrations that were continuous. Mean Absolute Error (MAE), R^2 Score and root mean square error (RMSE) were used to measure predictive performance.

Table 5: Regression Performance Evaluation

Model	RMSE	MAE	R^2 Score
Random Forest	0.118	0.084	0.918
Support Vector Machine	0.143	0.102	0.891
Artificial Neural Network	0.097	0.071	0.942

The Artificial Neural Network once again performed better than the rival algorithms with the least RMSE and MAE values and the greatest R^2 of 0.942. This favorable outcome holds that the ANN model made predictions, which were highly similar to observed aerosol concentrations.

The second-best performance was attained by random Forest with an Rmse of 0.118 and an R^2 Value of 0.918. The model was also very good in describing the variation in the environment and has very high predictive stability.

The Support Vector Machine had relatively more errors in predictions and lower R^2 value. Even though SVM is still useful in most classification problems, its ability to regress was not as strong with large-scale atmospheric data that has nonlinear interactions [27]. The results of the regressions also support the fact that ANN was better to predict aerosol concentration.

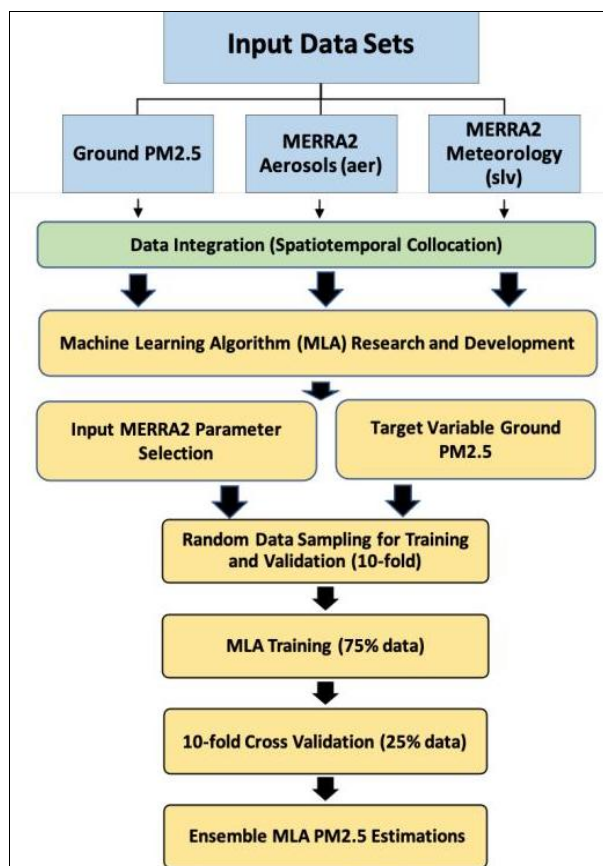


Fig 2: Workflow for Aerosol Detection Using Machine Learning

The support vector machine System was found to have an accuracy of 89.7 with a low compared to the other models. Although SVM did a good job in separating aerosol categories, it was influenced by the complexity and high dimensions of the environmental dataset. In general, the ANN model proved to be the best classification method to use when trying to detect atmospheric aerosols.

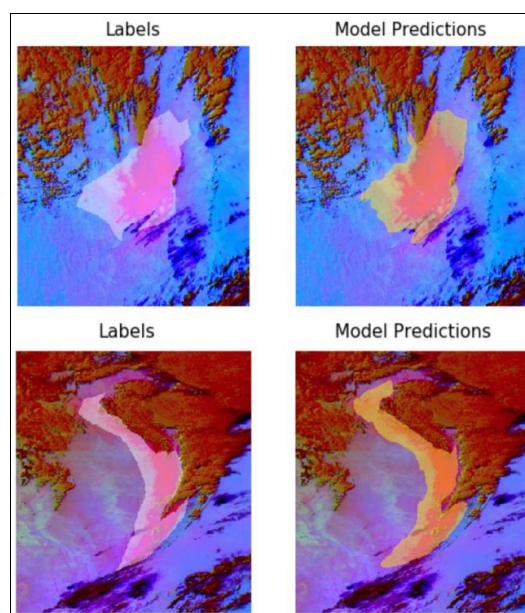


Fig 3: Aerosol Classification from Satellite Imagery

4.5 Feature Importance Analysis

The simpleness of knowing which individual variable has contributed to the machine learning decision matters in the interpretation of the decision made. The Random Forest algorithm was used to obtain a feature importance analysis.

Table 6: Importance Ranking of Environmental Variables

Feature	Importance Score (%)
Aerosol Optical Depth (AOD)	31.5
Relative Humidity	18.7
Temperature	15.2
Wind Speed	12.8
Atmospheric Pressure	11.1
Cloud Cover	10.7

The result showed that Aerosol Optical Depth (AOD) was the most powerful aspect, making a contribution of 31.5

percent to model estimates. This is not surprising since the concentration of aerosol is directly measured as AOD. Relative humidity came out as the second feature of importance. When the humidity is high, the size of aerosol particles may change due to hygroscopic growth and this critically influences optical properties of the atmosphere. Ground level temperature and wind speed were also found to have significant impact on aerosols dispersion. The movement of the particulate material can be due to the wind and the dispersal of the pollutants is influenced by the temperature and atmospheric stability [28]. Atmospheric pressure, combined with cloud cover had a moderate predictive value but still was significant to enhance the overall model accuracy. The above results indicate that a combination of environmental factors must be adopted when designing aerosol detectors.

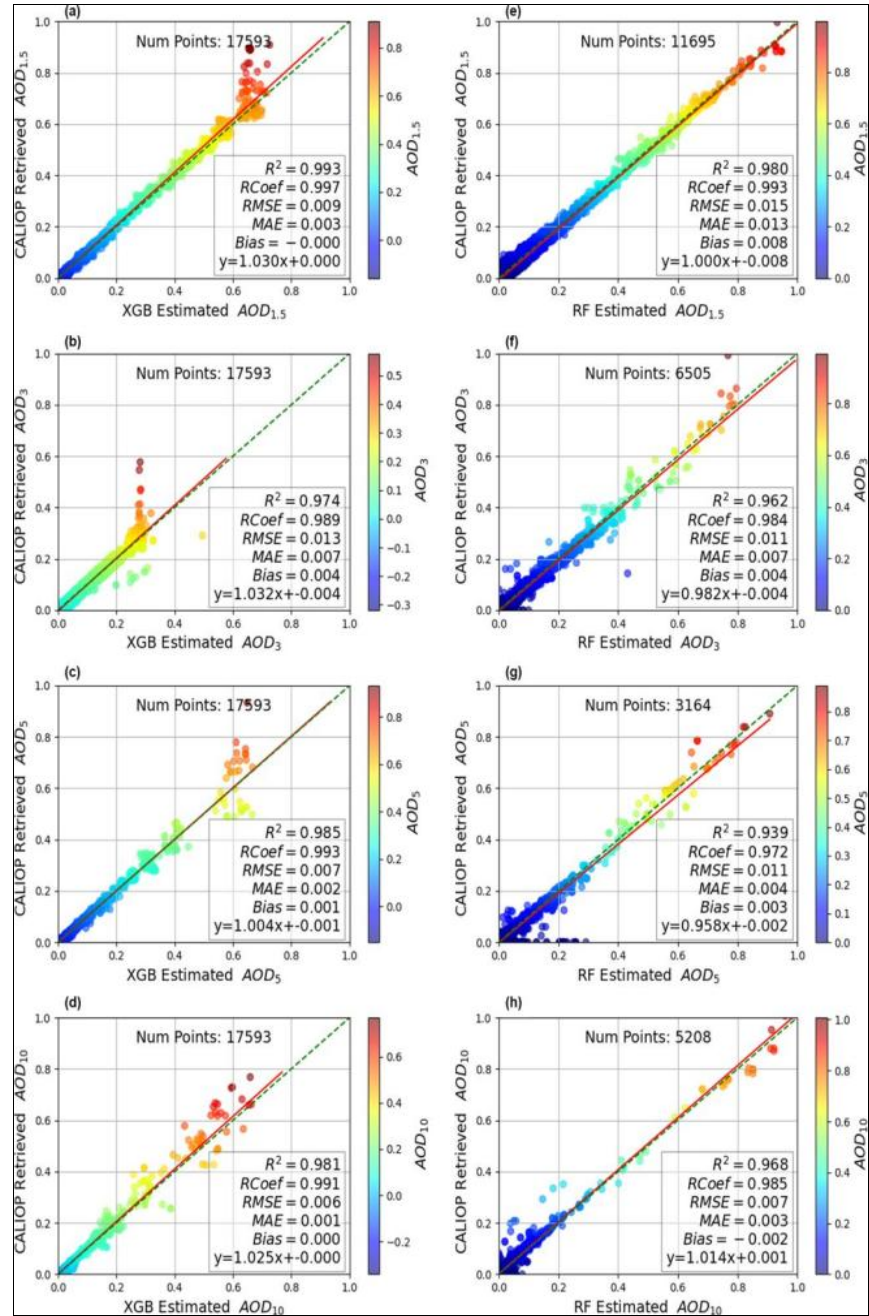


Fig 4: Validation of Machine Learning Aerosol Predictions

4.6 Computational Efficiency Analysis

Besides the accuracy of predictions, mathematical efficiency was also considered since environmental monitoring systems frequently demand near real-time computations.

Table 7: Computational Performance of Models

Model	Training Time (s)	Testing Time (s)	Memory Usage (MB)
Random Forest	18.4	1.9	245
Support Vector Machine	25.6	2.8	198
Artificial Neural Network	31.2	1.5	286

Random Forest was quickest in terms of training as compared to the other two models. The parallel tree-building mechanism allowed it to develop a model efficiently although it was highly accurate. Support Vector Machine took longer so as to train because the process involves optimization processes involving the selection of the optimal separating hyperplane. The size of the dataset increased and the computational requirement was increasing also with an increase in the size. Unnatural Neural Network optimized the longest training time since it involves optimization of weights by using the backpropagation mechanisms^[29]. Nevertheless, after the ANN was trained, it achieved the shortest time to test, hence it was very much applicable in the application of real time aerosol prediction. In spite of the fact that ANN used more computational resources, its significant enhancement in the predictive performance makes its usage in modern sophisticated atmospheric monitoring systems worthwhile.

4.7 Comparative Discussion

Based on the experiments, it is clear that machine learning methods are effective in detecting atmospheric aerosols with satellite measurements. Each of the three algorithms was effective to detect the concentration patterns of aerosols; but there was a significant difference in their performance. The Artificial Neural Network was able to give better results in all metrics of evaluation. The model was also multi-layered which allowed it to learn very complicated correlations among the atmospheric variables and those developed by satellite. Therefore, ANN was the most successful, given the best classification accuracy, the minimum errors in predictions, and the best explanatory power^[30]. Random Forest did also show great performance. It was efficient because its ensemble learning mechanism gave good predictions and had lower computational costs as compared to ANN. Moreover, the Random Forest was more interpretable as it has a feature importance analysis, and it is useful in environmental decision making. Support Vector Machine was not very effective but showed satisfactory work in terms of handling nonlinear and high dimensionality of the data. The findings indicate that SVM can be more appropriate with the smaller environmental data which can be realized by simpler features structure. The discussion validates that machine learning (ML) and satellite remote sensing are significant in improving the detection of aerosols in comparison with old-fashioned statistical analysis. The models have successfully used the information on the atmospheric and meteorological conditions to determine the concentration level of the aerosols with great accuracy.

4.8 Implications for Environmental Monitoring

The machine learning structure suggested has high value in air quality evaluation, climate studies and environmental organizations. Proper aerosol detection helps the authorities to pinpoint areas of excessive pollution and observe long-term trends in the atmosphere and put in place suitable mitigation measures. The excellent results of the Artificial Neural Network imply that deep learning solutions will be able to empower the next wave of environmental monitoring tools that could potentially process high quantities of satellite data within near real-time. These systems can enhance the early warning systems of critical air pollution incidents and play a role in developing evidence-based policies in the environment. In addition, the importance analysis of features points out that Aerosol Optical Depth and meteorological variables are so important in aerosol detection that a combination of both satellite remote sensing and machine learning methods is important. Generally, the findings confirm the usefulness of the proposed framework and illustrate its applicability to monitor aerosols in large-scale atmospheric aerosols under varied geographic and weather locations.

5. Conclusion

Atmospheric aerosols have a great impact on climatic systems, air quality, environmental sustainability, and human health hence they need to be properly detected and monitored to ensure proper management of the environment. This paper has introduced a machine learning-based aerosol hunting machine on satellite remote sensing observations and meteorological conditions. Through the combination of the Aerosol optical depth (AOD) measurements and other environmental parameters (temperature, humidity, wind speed, atmospheric pressure and cloud cover), the proposed process is a complete solution to aerosol monitoring. The effectiveness of three machine learning algorithms (Random Forest, Support Vector Machine, and Artificial Neural Network) was assessed to calculate their likelihood in aerosol detection and prediction. The experimental findings revealed that all the models were able to recognize aerosol concentration patterns; but the Artificial Neural Network was the best performed model, providing the strongest classification accuracy, predictability as well as reduction of errors. Random Forest also worked well in addition to providing insightful information about features, as opposed to Support Vector Machine that generated satisfactory results but with relatively low accuracy. The results established the fact that machine learning strategies could be successfully applied to represent the complex and nonlinear relationships of the atmospheric data and increase dramatically the quality of aerosol measurements.

Additionally, the analysis showed that Aerosol Optical Depth is the strongest predictor and the second strongest influence was on relative humidity and temperature highlighting the significance of integrating the satellite data with meteorological data. The suggested framework aids the monitoring of the environment on a massive scale and can be used by policy-makers, environmental agencies, as well as environmental researchers to assess air quality and conduct climate-related research. Future studies can emphasize the use of deep learning architectures, real-time

streams of satellite data and multi-source environmental data to continue enhancing the accuracy of aerosol prediction and operational effectiveness. In general, the research modifies the fact that both machine learning and remote sensing as potent instruments to detect intelligent aerosols in the atmosphere and monitor the environment.

References

1. Agbehadji IE, Christiana OI. Explainable artificial intelligence and machine learning for air pollution risk assessment and respiratory health outcomes: a systematic review. *Atmosphere*. 2025;16(10):1154.
2. Alani NH, S, Praneel C, Al-Rawi M. A two-stage machine learning framework for air quality prediction in Hamilton, New Zealand. *Environments*. 2025;12(9):336.
3. Alyssa V, Yi-Chun C, Chian-Yi L, Yi-Ying C, Chuan-Yao L. A hybrid regression-kriging-machine learning framework for imputing missing TROPOMI NO₂ data over Taiwan. *Remote Sensing*. 2025;17(12):2084.
4. Arpita A, Hussain CM. From detection to solution: a review of machine learning in PM_{2.5} sensing and sustainable green mitigation approaches (2021-2025). *Processes*. 2025;13(7):2207.
5. Arsanchai S, Pensiri A, Korakot S, Punawit F, Nasrin A, E AME. AI-driven time series forecasting of coastal water quality using Sentinel-2 imagery: a case study in the Gulf of Thailand. *Water*. 2025;17(12):1798.
6. Candra DS, Eko S. Machine learning approaches to phytoplankton identification and classification using GCOM-C/SGLI imagery. *Remote Sensing*. 2025;17(22):3759.
7. Choi H, Yong-Sang C, Chang-Hoi H, Kim J. Disentangling key cloud properties for precipitation retrievals from geostationary satellite data using machine learning. *PLOS Climate*. 2026;5(3):18.
8. de Oliveira Hilma Magalhães, Curado Leone FA, Lima André Matheus DS, Barbosa Thamiris Amorim DS, da Silva PR, Marques João Basso, Machado NG, Biudes MS. Aerosol-PAR interactions: critical insights from a systematic review (2021-2025). *Atmosphere*. 2025;16(9):1009.
9. Deelaram N, Manolia A, Iacopo C. Assessing the feasibility of satellite-based machine learning for turbidity estimation in the dynamic Mersey Estuary (case study: River Mersey, UK). *Remote Sensing*. 2025;17(21):3617.
10. Ejayi CJ, Cai D, Thomas D, Obiora S, Osei-Mensah E, Acen C, *et al*. Comprehensive review of artificial intelligence applications in renewable energy systems: current implementations and emerging trends. *Journal of Big Data*. 2025;12(1):169.
11. El-Sokkary N, Arafa AA, Zahran EG, Hefny HA, Ramdan N. Recent trends of machine learning techniques for risk assessment in hazardous environments. *Artificial Intelligence Review*. 2026;59(3):108.
12. Ferreira B, Silva RG, Iten M. Earth observation satellite imagery information based decision support using machine learning. *Remote Sensing*. 2022;14(15):3776.
13. Fuller CA, Selmer PA, Gomes J, McGill MJ. Using multitask machine learning to type clouds and aerosols from space-based photon-counting lidar measurements. *Remote Sensing*. 2025;17(16):2787.
14. Gilbert J, Aryee JNA, Adjei MJ, Mensah C, Quagrainie KT. Machine learning-based assessment of aerosol optical depth over Ghana, West Africa using MODIS satellite data. *PLOS Climate*. 2025;4(8):28.
15. Hanwen Q, Li Q, Shun X, Zhang Z, Qihou H, Tan W, *et al*. A novel framework integrating spectrum analysis and AI for near-ground-surface PM_{2.5} concentration estimation. *Remote Sensing*. 2025;17(22):3780.
16. Huang H, Wang Q, Liu C, Chen Z. Optimal estimation of cloud properties from thermal infrared observations with a combination of deep learning and radiative transfer simulation. *Atmospheric Measurement Techniques*. 2024;17(24):7129-7141.
17. Irene B, Ierotheos Z. Beyond in situ measurements: systematic review of satellite-based approaches for monitoring dissolved oxygen concentrations in global surface waters. *Remote Sensing*. 2026;18(3):428.
18. Ji D, Sun X, Ritter C, Notholt J. Retrieval of aerosol composition from spectral aerosol optical depth and optical properties using a machine learning approach. *Atmospheric Measurement Techniques*. 2026;19(2):421-436.
19. Kaiser D, Qu JJ. Detecting harmful algae blooms (HABs) on the Ohio River using Landsat and Google Earth Engine. *Remote Sensing*. 2025;17(24):4010.
20. Karamti H, Aldrees A, Alghamdi NS, Umer M, Nappi M. Intelligent multi-modal data implementation with capsuleNet for accurate air quality index prediction. *Complex & Intelligent Systems*. 2025;11(12):488.
21. Keyu M, Zhixuan Z, Hu K, Wang N, Liu Q, Tengyue G. Review of snow identification algorithms: from traditional machine learning to semantic methods. *Remote Sensing*. 2026;18(7):1067.
22. Kim S, Youjeong Y, Menas K, Kim J, Wonsik C, Kim SH, *et al*. Real-time production of high-resolution, gap-free, 3-hourly AOD over South Korea: a machine learning approach using model forecasts, satellite products, and air quality data. *Atmosphere*. 2025;17(1):19.
23. Koziół K, Borowiec N, Urszula M, Mateusz R, Santos CAG, Jerzy C. Machine learning-based detection of archeological sites using satellite and meteorological data: a case study of Funnel Beaker culture tombs in Poland. *Remote Sensing*. 2025;17(13):2225.
24. Lolli S. Machine learning techniques for vertical lidar-based detection, characterization, and classification of aerosols and clouds: a comprehensive survey. *Remote Sensing*. 2023;15(17):4318.
25. Manoj L, Sushant M, Mankin KR. Soil moisture prediction using remote sensing and machine learning algorithms: a review on progress, challenges, and opportunities. *Remote Sensing*. 2025;17(14):2397.
26. Mao C, Zhao Y, Wang L, Yang Z, Kou W, Xu W, *et al*. Machine learning-enabled UAV hyperspectral identification of tomato spotted wilt virus in tobacco. *Frontiers in Plant Science*. 2025;16:1728043.
27. Mehrdad SM, Du K. Satellite-based methane emission monitoring: a review across industries. *Remote Sensing*. 2025;17(22):3674.
28. Meister M, Qu JJ. Quantifying seagrass density using

- Sentinel-2 data and machine learning. *Remote Sensing*. 2024;16(7):1165.
29. Muhammad H, Zainab T, Mehmood SA, Hania A, Sumaira K, Sundas G, Khalid M. Ensemble artificial intelligence fusing satellite, reanalysis, and ground observations for improved PM2.5 prediction. *Atmosphere*. 2026;17(4):411.
30. Navidi MN, Fazli E, Kharazmi R, Seyedmohammadi J. Wheat yield prediction using integrated optical and radar remote sensing with machine learning across key phenological stages. *Scientific Reports*. 2026;16(1):10470.