Techno.Com, Vol. 25 No. 3, August 2026: 959 - 971

Comparative Performance of Tree-Based Machine Learning Algorithms for Horizontal Visibility Classification at Soekarno–Hatta International Airport Using Meteorological and $PM_{2.5}$ Data

Muhammad Aulia Zikri[1*], Aries Kristianto[2]
[1]Undergraduate Meteorology Program, Indonesia State College of Meteorology Climatology and Geophysics, Tangerang, Indonesia, 15211
[2]Meteorology Department, Indonesia State College of Meteorology Climatology and Geophysics, Tangerang, Indonesia, 15211
Email: auliazikri@stmkg.ac.id[1]; aries.kristianto@stmkg.ac.id[2]*
Received 9 July 2026; Revised 24 July 2026; Accepted 25 July 2026

Abstract – Low horizontal visibility phenomena are a major threat to aviation safety and operational efficiency, particularly at Soekarno–Hatta International Airport (WIII), which is vulnerable to fog and haze arising from complex interactions between meteorological conditions and air quality. This study develops and comparatively evaluates four tree-based machine learning algorithms as independent base learners, Random Forest, Extra Trees, XGBoost, and LightGBM, for horizontal visibility classification, without employing any ensemble learning scheme. The dataset comprises AWOS meteorological observations integrated with PM2.5 concentrations from 2021 to 2025. Results show that XGBoost achieved the best performance, with 97.62% accuracy, 97.65% F1-score, and an AUC-ROC of 0.9926, followed by LightGBM, Random Forest, and Extra Trees. Feature importance analysis identified PM2.5 as the most dominant predictor, followed by wind-related variables. These findings indicate that a well-optimized standalone base learner can achieve competitive predictive performance while offering greater interpretability and computational efficiency for operational implementation.

Keywords: horizontal visibility, machine learning, $PM_{2.5}$, XGBoost, aviation meteorology

1. INTRODUCTION

Horizontal visibility is one of the most critical meteorological parameters governing aviation safety, particularly during approach, landing, and take-off. Adverse weather remains a leading cause of operational disruption, flight delays, and economic losses [1], [2], and low visibility further reduces pilot situational awareness and the time available for reactive decision-making [3]. Visibility degradation rarely results from a single factor; it arises from complex interactions among atmospheric processes such as advection or radiation fog, precipitation, smoke, and haze, often intensified by elevated particulate matter concentrations [4], [5]. Local and large-scale synoptic conditions, such as nocturnal cooling under clear skies, are primary drivers of localized fog formation that remains difficult to predict using standard approaches [6], and such disruptions often force temporary airport closures, causing cascading delays across flight networks [7].

Visibility prediction has long been dominated by Numerical Weather Prediction (NWP) models, which simulate atmospheric evolution through numerical integration of fluid-dynamic and thermodynamic equations. Although NWP remains the backbone of global weather forecasting, its performance often deteriorates when representing microscale atmospheric fluctuations around airports due to coarse grid resolution and

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i3.17336

uncertainties in cloud microphysics and aerosol–cloud parameterizations [8], [9], [10]. Aerosol-meteorology interactions, particularly in high-emission urban airport environments, also remain a persistent source of systemic uncertainty in visibility forecasting [11], [12].

These structural limitations have driven a shift toward more computationally adaptive, data-driven approaches. Unlike NWP, Machine Learning (ML) algorithms can automatically extract non-linear patterns from observational data without requiring rigid physical assumptions [13], [14], reducing computational complexity while maintaining predictive performance competitive with deterministic models [15]. In aviation meteorology specifically, ML approaches have demonstrated strong potential in capturing rapid transitions of low-visibility events at the airport scale [16], including the dynamics of low-visibility events across various regions [17].

Physically, visibility is governed by complex radiative transfer and light-scattering processes. According to the WMO [18], horizontal visibility represents the maximum distance at which an object remains distinguishable from the horizon background, physically determined by the atmospheric extinction coefficient, as formulated in Koschmieder's law, an inverse relationship between visibility and extinction [19]. Aerosol loading such as PM2.5 accelerates visibility degradation through Mie scattering, particularly in industrial and urban areas [20]. Integrating satellite-based data with surface sensors can further refine visibility estimation [21], while relative humidity, dew point depression, and wind speed modulate aerosol hygroscopic growth and pollutant dispersion [22], [23].

Various single-model ML algorithms, such as LightGBM, LSTM, and CNN, have shown promising early results in visibility prediction [24], [25]. Nevertheless, single-model architectures are inherently susceptible to the bias–variance trade-off, often suffering from limited generalization and overfitting when trained on data with high temporal autocorrelation or few transitional events [26]; proper feature selection is therefore crucial to prevent overfitting on high-dimensional pollutant data [27].

Among tree-based algorithms, Random Forest has proven effective in modeling non-linear atmospheric relationships on high-dimensional data [28], while Extra Trees offers greater split-point randomization, making it well-suited for noisy meteorological data [29]. Several studies have compared such algorithms individually, including approaches combining ML with game-theoretic interpretation to understand factors influencing visibility variation in tropical airport environments [30].

Nevertheless, most prior studies direct attention toward combining models through joint learning schemes, leaving the standalone performance of tree-based base learners, particularly with $PM_{2.5}$ integration in a tropical airport environment, relatively underexplored. Evaluating single-model performance directly carries distinct practical value: higher interpretability, lighter computational requirements, and easier real-time implementation compared with joint learning schemes. Empirical studies in tropical airport environments, shaped by humidity, urban heat island effects, and monsoon circulation, also remain relatively scarce.

Based on this gap, this study aims to develop and comparatively evaluate four tree-based machine learning algorithms as independent base learners: Random Forest, Extra Trees, XGBoost, and LightGBM, for horizontal visibility classification at Soekarno–Hatta International Airport (WIII), using AWOS observations integrated with PM2.5 data. Unlike joint learning approaches, this study evaluates each model independently to identify the algorithm offering the optimal balance of accuracy, computational efficiency, and interpretability for visibility-based operational decision-making in aviation.

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i3.17336

2. DATA AND METHOD

A. Study Area and Period

This study was conducted at Soekarno–Hatta International Airport (ICAO: WIII; IATA: CGK), Tangerang, Banten Province, Indonesia (6.12°S, 106.65°E), Indonesia's busiest airport and a major Southeast Asian aviation hub highly susceptible to low-visibility events. Situated on a low-lying coastal plain near the Java Sea, the airport experiences elevated aerosol concentrations from rapid urbanization, transportation, and industrial activity, which interact with the Asian–Australian monsoon system to produce significant seasonal variability in rainfall, humidity, wind, and cloud cover [31]. Land–sea breeze interactions further influence boundary-layer dynamics and pollutant dispersion, reinforcing haze, mist, and fog formation under high humidity and weak wind. This combination of characteristics makes the study area highly representative for evaluating tree-based machine learning algorithms in horizontal visibility prediction [30].

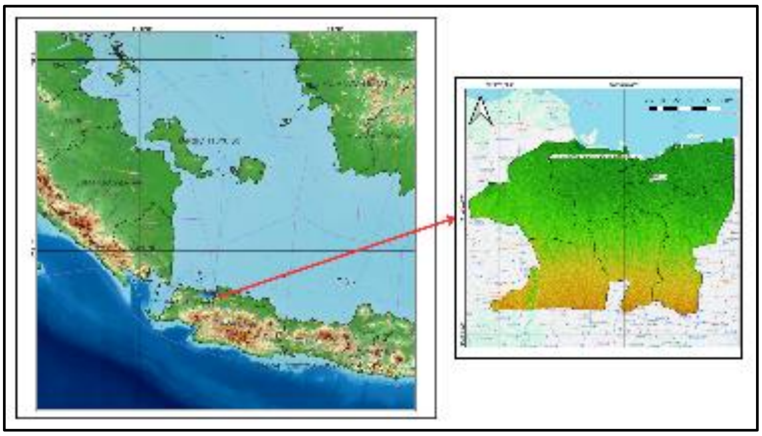

Figure 1. Spatial extent of the study area: (left) Western Indonesia region showing Banten Province and surrounding areas; (right) Soekarno–Hatta International Airport (WIII) location and adjacent BMKG meteorological observation stations.

B. Research Data

This study uses minute-resolution meteorological and air-quality data collected at Soekarno–Hatta International Airport from January 2021 to December 2025, obtained from the Automated Weather Observing System (AWOS) and the BMKG PM2.5 monitoring network. The datasets were synchronized by timestamp to create an integrated database representing diverse atmospheric conditions, including seasonal variability, extreme weather, and low-visibility events. Predictor variables comprise air temperature, dew point temperature, dew point depression, relative humidity, surface air pressure, wind speed, wind direction, precipitation, and PM2.5 concentration. Temporal features encoded using sine–cosine transformations of hour, minute, and month, together with a day/night indicator, were incorporated to capture diurnal and seasonal patterns. Horizontal visibility was defined as the target classification variable.

C. Data Pre-processing

Data Cleaning and Missing Value Handling

Meteorological and PM2.5 data were synchronized based on timestamps and subjected to quality-control procedures. Duplicate and inconsistent observations were identified and removed to preserve data integrity. Outlier detection was performed using the Z-score method, in which observations exceeding ±3 standard deviations from the mean were

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i3.17336

considered anomalous and adjusted accordingly. The standardized score for each observation was calculated as:

$$z = \frac{x-\mu}{\sigma} \quad (1)$$

Where x is the observed value, μ is the variable mean, and σ is the standard deviation [32]. Missing values resulting from sensor malfunction or data transmission interruptions were handled using a combination of linear interpolation, forward filling, and backward filling to preserve temporal continuity and minimize information loss in the time-series data.

Feature Selection

Visibility degradation is influenced by complex interactions among meteorological and air-quality variables; feature selection was therefore performed to identify the most relevant predictors while reducing data redundancy. Spearman rank correlation analysis was used to evaluate the strength of monotonic relationships between predictor variables and horizontal visibility. Unlike Pearson correlation, Spearman correlation captures non-linear monotonic relationships commonly found in atmospheric data [33], [34]:

$$\rho = 1 - \frac{6\sum {d_i}^2}{n(n^2-1)} \quad (2)$$

Where ρ is the Spearman correlation coefficient, d_i is the rank difference of the *i-th* paired observation, and nn n is the number of observations. In addition to correlation analysis, Random Forest–based feature importance was used as a supplementary exploratory technique to evaluate the relative contribution of each predictor to visibility variability.

Target Class Definition

Horizontal visibility was categorized into several classes based on visibility and relative humidity thresholds, referring to WMO criteria and BMKG regulations on aviation meteorological observation and reporting [35]. This classification scheme represents fog/dense fog, mist/haze, low visibility, and normal visibility conditions, allowing the resulting classification model to represent the operationally relevant spectrum of visibility conditions for aviation safety.

Data Splitting and Cross-Validation

To objectively evaluate model performance, the dataset was divided into training and testing subsets using a 70:30 ratio. The training subset was used for model development and hyperparameter optimization, while the testing subset was used for independent performance assessment. This strategy allows evaluation of the models' predictive capability on previously unseen data and minimizes the risk of performance overestimation.

To further enhance model reliability and reduce sampling bias, a 10-fold cross-validation procedure was applied during the training stage. The training dataset was divided into ten equally sized subsets; in each iteration, nine subsets were used for training while the remaining subset was used for validation. This procedure was repeated ten times, and the final performance was obtained by averaging the results across all folds.

D. Machine Learning Algorithms (Base Learners)

This study specifically evaluates four tree-based machine learning algorithms as standalone base learners, without combining them through any ensemble learning scheme. All four algorithms were selected for their ability to capture complex non-linear

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i3.17336

relationships among meteorological and air-quality variables while maintaining computational efficiency [36].

Random Forest (RF)

Random Forest, developed by Breiman [28], constructs multiple decision trees using bootstrap samples drawn from the original dataset and aggregates their outputs through majority voting for classification. Each node split is selected based on the reduction in Gini impurity, defined as:

$$Gini(t) = 1 - \sum_{k=1}^{K} p_k{}^2 \quad (3)$$

Where p_k is the proportion of class k at node t, and K is the number of visibility classes.

Extra Trees (ET)

Extra Trees (Extremely Randomized Trees) introduces a higher degree of randomization into tree construction [29]. Unlike Random Forest, which searches for the optimal split threshold, ET selects candidate split points randomly and then determines the best split among these candidates using the same Gini impurity criterion as in Equations (3)–(4). This additional randomization reduces variance and computational cost while maintaining predictive capacity on noisy meteorological data.

Extreme Gradient Boosting (XGBoost)

Extreme Gradient Boosting (XGBoost) is a tree-based boosting algorithm that constructs weak learners sequentially, where each subsequent model aims to correct the prediction errors produced by previous models. The algorithm combines gradient boosting optimization with regularization mechanisms to enhance predictive performance while mitigating the possibility of overfitting [36]. Compared with traditional boosting techniques, XGBoost offers several advantages, including improved computational efficiency, automatic processing of missing values, and greater model adaptability. These advantages make XGBoost highly applicable for atmospheric datasets, which commonly exhibit nonlinear relationships and complex variations in environmental conditions.

Light Gradient Boosting Machine (LightGBM)

Light Gradient Boosting Machine (LightGBM) is an optimized gradient boosting framework developed to achieve faster computation and more efficient memory usage when handling large-scale datasets. The algorithm utilizes histogram-based learning and a leaf-wise tree growth strategy, enabling accelerated model training compared with conventional boosting algorithms while maintaining competitive predictive accuracy [36]. Several studies have demonstrated the effectiveness of LightGBM in visibility forecasting and environmental prediction applications due to its capability to identify complex nonlinear patterns and interactions among meteorological variables [24].

E. Model Performance Evaluation

Confusion Matrix

Model predictions were evaluated using a multiclass confusion matrix, which summarizes the relationship between actual and predicted visibility classes. For a multiclass classification problem with K classes (K = 4, corresponding to Class 0: Fog, Class 1: Mist, Class 2: Haze, and Class 3: Normal), the confusion matrix components for each specific class k {0, 1, 2, 3} are defined on a One vs Rest (OvR) basis: True Positives (TP_k), True Negatives (TN_k), False Positives (FP_k), and False Negatives (FN_k).

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i3.17336

Performance Metrics

To accommodate the multiclass nature of the visibility classification task and address potential class imbalance, evaluation metrics were formulated both on a per-class basis and through weighted aggregation across all classes. For each class k, the class-specific Accuracy, Precision, Recall, and F1-score are calculated as follows:

$$Accuracy_K = \frac{TP_k + TP_N}{TP_k + TN_k + FP_k + FN_k} \tag{4}$$

$$Precision_K = \frac{TP_k}{TP_k + FP_k} \tag{5}$$

$$Recall_K = \frac{TP_k}{TP_k + FN_k} \tag{6}$$

$$F1score_K = 2 \times \frac{Precision_k \times Recall_K}{Precision_k + Recall_K} \tag{7}$$

Overall model accuracy across all classes is defined as the proportion of correctly classified instances out of the total dataset size N:

$$Overall\ accuracy = \frac{\sum_{k=0}^{K-1} TP_k}{N} \tag{8}$$

To account for class imbalance across the target categories (where Class 3 represents 86.1% of observations while Class 0 represents 0.2%), performance was aggregated using the weighted F1-score ($F1_{Weighted}$). The $F1_{Weighted}$ calculates the mean of all per-class F1-scores weighted by the support n_k (the number of true instances in class k) relative to the total number of samples N:

$$F1_{Weighted} = \sum_{k=0}^{K-1} \left(\frac{n_k}{N}\right) \times F1_{score} \tag{9}$$

Where K= 4 is the total number of visibility classes, n_k is the number of samples belonging to class k, $\sum_{k=0}^{K-1} n_k$ is the total number of evaluated samples.

3. RESULT

A. Dataset Characteristics

The research dataset consists of synchronized AWOS meteorological observations and PM2.5 measurements collected over the period January 2021–December 2025. Descriptive statistics of the meteorological and air-quality variables are presented in Table 2.

Table 1. Descriptive Statistics of Meteorological and Air Quality Variables

Variable	Unit	Min	Max	Mean	Median	Std. Dev.
Temperature	°C	-7.8	39.9	28.03	27.5	2.62
Dew Point	°C	-15.2	30.3	23.94	24.0	1.39
Relative Humidity	%	14.8	100	79.07	80.6	12.80
Pressure	hPa	1003.0	1015.9	1009.75	1009.8	1.82
Wind Speed	kt	0.1	30.2	6.43	5.8	3.44
Wind Direction	°	0	360	194.17	214	98.21
Precipitation	mm	0	148	0.01	0	0.59
PM2.5	µg m^{-3}	0.3	769.1	38.34	34.0	24.84

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i3.17336

The distribution of visibility classes derived from the WMO classification criteria is shown in Table 3.

Table 2. Distribution of visibility classes

Class	Category	Percentage
0	Fog/ dense fog	0.2%
1	Mist	3.3%
2	Haze	10.3%
3	normal	86.1%

B. Correlation Analysis Among Predictor Variables

The Spearman correlation analysis reveals a near-perfect negative correlation between relative humidity (RH) and dew point depression/DPD ($\rho = -1.00$). Air temperature shows a strong positive correlation with DPD ($\rho = 0.86$) and a strong negative correlation with RH ($\rho = -0.86$). Wind speed shows moderate correlations with RH ($\rho = -0.40$) and DPD ($\rho = 0.40$). Air pressure, wind direction, and precipitation show relatively weak correlations with most other predictor variables. PM2.5 shows weak correlations with air temperature ($\rho = -0.02$), RH ($\rho = -0.10$), and DPD ($\rho = 0.10$).

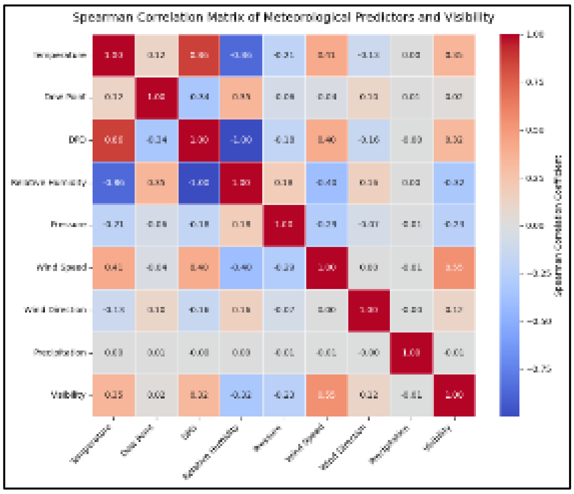

Figure 2. Spearman correlation matrix

C. Base Learner Performance on Test Data

Four tree-based machine learning models, Random Forest (RF), Extra Trees (ET), XGBoost (XGB), and LightGBM (LGBM), were evaluated using an independent test dataset. The comparative performance results are presented in Table 4.

Table 3. Performance Comparison of base learners

Model	Accuracy	F1 Weighted	Recall Fog	Recall Haze	AUC
RF	0.9249	0.9295	0.9123	0.9613	0.9765
ET	0.8990	0.9080	0.9325	0.9809	0.9716
XGB	**0.9762**	**0.9765**	**0.9469**	0.9647	**0.9926**
LGBM	0.9288	0.9343	0.9406	**0.9822**	0.9859

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i3.17336

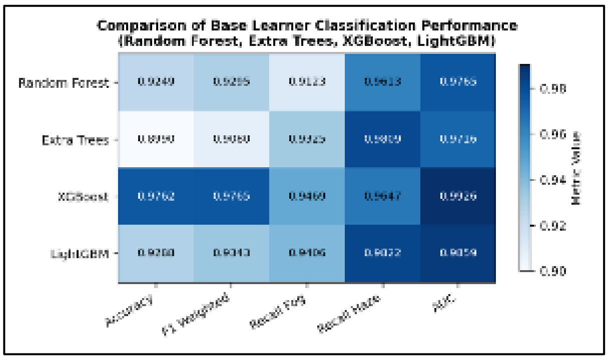

Figure 3. Heatmap comparison of classification performance metrics

D. Cross-Validation and Model Robustness

A 10-fold cross-validation procedure was performed to assess model stability and generalization. The cross-validation results are presented in Table 5.

Table 4. Cross validation

Model	CV F1 Mean	CV F1 Std	CV Accuracy Mean	Train F1	Gap
Random Forest	0.8989	0.0014	0.8931	0.9157	0.0168
Extra Trees	0.8729	0.0018	0.8594	0.8863	0.0134
XGBoost	**0.9556**	**0.0008**	**0.9555**	1.0000	0.0444
LightGBM	0.9352	0.0016	0.9319	0.9767	0.0415

E. Sensitivity Analysis of Low-Visibility Detection

A dedicated sensitivity analysis for low-visibility events (Fog and Mist) is presented in Table 6.

Table 5. Performa comparion between models

Model	Precision	Recall	F1-score	FNR	Precision
Random Forest	0.6506	0.9123	0.7595	0.0447	0.6506
Extra Trees	0.7026	0.9325	0.8014	0.0346	0.7026
XGBoost	0.9585	0.9469	0.9527	0.0448	0.9585
LightGBM	**0.9668**	0.9406	**0.9535**	0.0523	**0.9668**

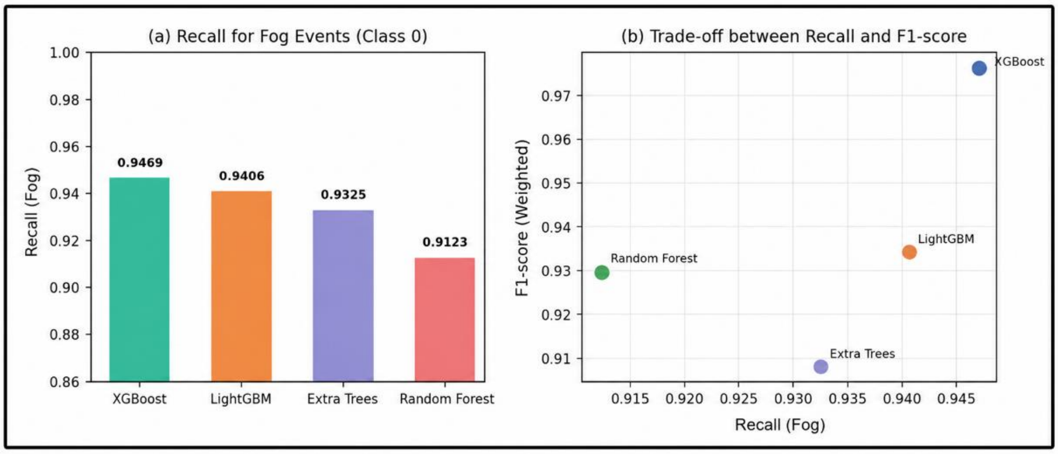

Figure 4. Comparison of low-visibility detection performance among evaluated models. (a) Recall values for Fog events (Class 0). (b) Trade-off between Recall and F1-score for Fog event detection.

A dedicated sensitivity analysis was conducted to evaluate model capability in detecting critical low-visibility conditions, specifically focusing on Fog events (Class 0), which represent the most severe hazard to aviation safety. As presented in Table 6 and

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i3.17336

Figure 4, model evaluation highlights noticeable performance variations across the algorithms for minority class detection.

Based on Figure 4(a), XGBoost achieved the highest recall for Fog events (0.9469), followed closely by LightGBM (0.9406), Extra Trees (0.9325), and Random Forest (0.9123). Figure 4(b) illustrates the trade-off between Recall and F1-score for Fog detection, showing that XGBoost occupies the most optimal balance with high recall and the highest weighted F1-score (0.9765) among all evaluated base learners.

F. Feature Importance

The feature importance analysis shows that PM2.5 is the most influential predictor, followed by wind direction, wind speed, air pressure, and dew point temperature. Temporal variables such as month and the night-time indicator contribute relatively little to the prediction process.

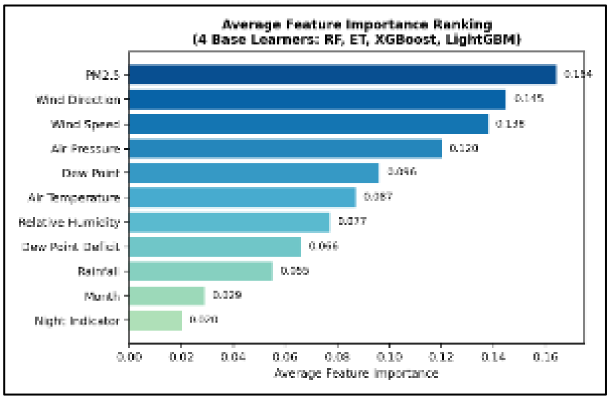

Figure 5. Average feature importance ranking derived from the machine learning

Furthermore, Spearman rank correlation analysis revealed a perfect inverse monotonic relationship (ρ= -1.00) between Relative Humidity (RH) and Dew Point Depression (DPD). Thermodynamically, DPD represents the moisture deficit (T - T_d), making it functionally redundant with RH in capturing atmospheric saturation. While tree-based algorithms (such as Random Forest and XGBoost) are inherently robust against multicollinearity during split determination, retaining collinear features can artificially split their feature importance scores. In this study, both variables were initially retained to observe split preference across ensemble architectures; however, their combined influence should be interpreted as a single thermodynamic saturation signal.

4. DISCUSSION

The results of this study show that all four base learners, Random Forest, Extra Trees, XGBoost, and LightGBM, were able to classify horizontal visibility at Soekarno–Hatta International Airport with accuracies above 89%, with XGBoost consistently outperforming the others across nearly all metrics (97.62% accuracy; 97.65% weighted F1-score; 0.9926 AUC-ROC). This superiority is consistent with the fundamental characteristics of gradient boosting algorithms, which sequentially focus learning on observations misclassified in previous iterations, allowing them to capture complex non-linear interactions between meteorological variables and PM2.5 [36]. This finding aligns with the study by Yu [24], which reported that histogram-based boosting algorithms such as LightGBM are effective in predicting atmospheric conditions with high computational efficiency, as well as with Chen [25], who emphasized the advantages of tree-based learning approaches for environmental data with complex non-linear interactions.

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i3.17336

The AUC-ROC value of 0.9926 achieved by XGBoost in this study is also comparable to the high performance reported by Shankar and Sahana [17] in predicting the dynamics of low-visibility events, indicating that tree-boosting algorithms generally possess strong discriminative capability for classifying fog and haze conditions, even without being combined through any joint learning scheme.

Meanwhile, the relatively lower overall accuracy of Extra Trees (89.90%) but high recall for haze detection (98.09%) reinforces the argument of Geurts [29] that the additional randomization in split-point selection makes ET more tolerant of noise in the data, albeit potentially at the expense of precision on the majority class. A similar pattern is observed for Random Forest, which reaffirms the fundamental characteristics of bagging algorithms as described by Breiman [28]: low variance and stable generalization, reflected in the small train–validation gap (0.0168 for RF and 0.0134 for ET) shown in Table 5, even though their peak performance does not match that of the boosting algorithms.

The finding that PM2.5 emerged as the most influential predictor (Figure 7) reinforces several previous studies on the role of aerosols in visibility degradation through Mie scattering processes [20]. This result is also consistent with Koschmieder's law[19], which states that visibility is inversely proportional to the atmospheric extinction coefficient, to which particulate loading contributes significantly. Interestingly, although relative humidity and dew point depression showed a very strong statistical correlation ($\rho = -1.00$ in Figure 4) and have long been associated with fog formation [6], neither variable ranked among the top features in the feature importance analysis. This suggests that aerosol transport processes, represented by wind direction and wind speed as the second- and third-ranked predictors, provide complementary predictive information not fully captured by conventional humidity variables, consistent with the findings of Dewita and Alfiandy [22] and Prasetyo et al. [23] regarding the modulating role of wind on pollutant dispersion and haze formation.

The feature importance analysis (Figure 5) identified $PM_{2.5}$ as the primary contributor to horizontal visibility variation (0.164), followed by wind direction (0.145) and wind speed (0.138). Interestingly, although Relative Humidity (RH) and Dew Point Depression (DPD) exhibited a perfect negative correlation ($\rho = -1.00$, Figure 2), their individual feature importance scores were split into 0.077 and 0.066, respectively. Summing their contributions yields a combined importance of 0.143, placing atmospheric moisture saturation as the second most vital physical driver alongside wind dynamics. This confirms that while multicollinearity does not impair the predictive accuracy of tree-based base learners, it causes importance score dilution. Future deployment pipelines can safely drop DPD to reduce feature dimensionality without losing predictive power.

From an aviation operational perspective, the sensitivity analysis results (Table 6) show that XGBoost and LightGBM offer the best balance between precision and recall for low-visibility event detection (F1-scores of 0.9527 and 0.9535, respectively), whereas Random Forest and Extra Trees produced substantially lower precision (0.6506 and 0.7026) despite their high recall. This pattern indicates that both bagging-based models tend to generate more false alarms for the minority class, a limitation also identified by Li et al. [16] in the context of airport-scale visibility prediction, where models with high sensitivity to minority classes often sacrifice precision due to extreme class imbalance.

Overall, these findings reinforce the notion that superior performance does not necessarily require complex joint learning architectures. A properly optimized single base learner model, in this case, XGBoost, proved capable of delivering highly competitive visibility classification performance, consistent with the views of Alpaydin [14] and Patel et al. [15] that the computational efficiency and interpretability of standalone models remain important considerations in the design of operational prediction systems,

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i3.17336

particularly for real-time implementation at airports such as WIII, as also suggested by the interpretability-based evaluation framework applied by Aman et al. [30] at Bangkok Airport.

5. CONCLUSION

This study developed and comparatively evaluated four tree-based machine learning algorithms, Random Forest, Extra Trees, XGBoost, and LightGBM, as independent base learners for horizontal visibility classification at Soekarno–Hatta International Airport, using AWOS meteorological observations integrated with PM2.5 data from 2021–2025. All four models achieved good classification performance (accuracy above 89%) without relying on any ensemble learning scheme, with XGBoost consistently outperforming the others (97.62% accuracy, 97.65% weighted F1-score, and 0.9926 AUC-ROC), followed by LightGBM, Random Forest, and Extra Trees. PM2.5 emerged as the most influential predictor of visibility variability, followed by wind direction and wind speed, confirming the significant role of aerosol loading and its transport processes in visibility degradation at a tropical urban airport.

These findings demonstrate that a well-optimized standalone base learner can deliver highly competitive visibility prediction performance while offering greater interpretability and lighter computational requirements than more complex joint learning approaches, making it a practical candidate for real-time operational decision-support systems at airports such as WIII. Future research is encouraged to explore deep learning architectures, extended prediction lead times, improved class-imbalance handling, and cross-location validation to further strengthen the generalizability of these results.

ACKNOWLEDGEMENT

The author would like to express sincere gratitude to Dr. Aries Kristianto, S.T., M.Si., for his invaluable guidance, supervision, and constructive feedback throughout the course of this research. Appreciation is also extended to the Applied Meteorology Study Program, Sekolah Tinggi Meteorologi Klimatologi dan Geofisika (STMKG), for its academic support. The author further extends heartfelt thanks to Stasiun Meteorologi Kelas I Soekarno–Hatta and the Meteorology, Climatology, and Geophysics Agency (BMKG) for providing the AWOS meteorological and PM2.5 air-quality data that served as the foundation of this study.

REFERENCES

[1] FAA, “United States 2021 Aviation Climate Action Plan,” 2021.

[2] B. Rochmat and S. Martha, “Pengaruh Faktor Geografis Terhadap Keselamatan Penerbangan di Indonesia,” *Jurnal Lemhannas RI*, 2021, [Online]. Available: https://api.semanticscholar.org/CorpusID:271022346

[3] A. D. Saputra, S. Priyanto, and I. Muthohar, “Faktor Faktor yang Mempengaruhi Kinerja Pilot dan Kecelakaan Pesawat Terbang dengan Pendekatan Partial Least Square (PLS),” *Warta Penelitian Perhubungan*, vol. 28, no. 2, pp. 71–91, 2016.

[4] W.-S. Won *et al.*, “Impact of Fine Particulate Matter on Visibility at Incheon International Airport, South Korea,” *Aerosol Air Qual. Res.*, vol. 20, no. 5, pp. 1048–1061, 2020.

[5] D. Alves, M. Belo-Pereira, F. Mendonça, and F. Morgado-Dias, “Intelligent Visibility Forecasting at Airports: a Systematic Review,” *Environ. Res. Commun.*, 2025.

[6] D. Bari and T. Bergot, “Influence of Environmental Conditions on Forecasting of an Advection-Radiation Fog: A Case Study from the Casablanca Region,

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i3.17336

Morocco," *Aerosol Air Qual. Res.*, vol. 18, no. 1, pp. 62–78, 2018, doi: 10.4209/aaqr.2016.11.0520.
[7] A. Kwasiborska, M. Grabowski, A. N. Sedláčková, and A. Novák, "The Influence of Visibility on the Opportunity to Perform Flight Operations with Various Categories of the Instrument Landing System," *Sensors*, vol. 23, no. 18, pp. 1–22, 2023.
[8] W.-K. Wong, C.-S. Lau, and P.-W. Chan, "Aviation Model: A Fine-Scale Numerical Weather Prediction System for Aviation Applications at the Hong Kong International Airport," *Advances in Meteorology*, vol. 2013, no. 1, pp. 1–11, 2013.
[9] I. R. Velde, G. J. Steeneveld, B. G. J. W. Schreur, and A. A. M. Holtslag, "Modeling and Forecasting the Onset and Duration of Severe Radiation Fog Under Frost Conditions," *Mon. Weather Rev.*, vol. 138, no. 11, pp. 4237–4253, 2010.
[10] S. Sen Roy *et al.*, "A New Paradigm for Short-Range Forecasting of Severe Weather Over the Indian Region," *Meteorology and Atmospheric Physics*, vol. 133, no. 4, pp. 989–1008, 2021.
[11] S. S. Lee *et al.*, "Effects of Model Resolution and Parameterizations on the Simulations of Clouds, Precipitation, and Their Interactions with Aerosols," *Atmos. Chem. Phys.*, vol. 18, no. 1, pp. 13–29, 2018.
[12] T. Beucler, M. Pritchard, P. Gentine, and S. Rasp, "Towards Physically-Consistent, Data-Driven Models of Convection," in *IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium*, IEEE, Sep. 2020, pp. 3987–3990. doi: 10.1109/IGARSS39084.2020.9324569.
[13] T. M. Mitchell, *Machine Learning*. Redmond: McGraw-Hill Science/Engineering/Math, 1997.
[14] E. Alpaydin, *Introduction to Machine Learning*, Third edition. London, England: The MIT Press, 2014.
[15] S. K. Patel, S. R. Panchal, and S. S. Panchal, "Advanced Learning Techniques for Short-Term Atmospheric Predictions: an Overview," in *2023 2nd International Conference on Automation, Computing and Renewable Systems (ICACRS)*, 2023, pp. 838–843.
[16] J. Li, F. Wang, K. Xue, and Y. Hao, "Research on Machine Learning-Based Nowcasting Method for Low Visibility Weather in Urumqi Airport," in *2022 4th International Conference on Advances in Computer Technology, Information Science and Communications (CTISC)*, 2022, pp. 1–6.
[17] A. Shankar and B. C. Sahana, "Machine Learning for Predicting Low-Visibility Event Dynamics: An Ensemble Approach," *Model. Earth Syst. Environ.*, vol. 11, no. 3, Jun. 2025, doi: 10.1007/s40808-025-02384-1.
[18] WMO, "Guide to Instruments and Methods of Observation Volume I - Measurement of Meteorological Variables," 2024.
[19] H. Koschmieder, *Theorie der Horizontalen Sichtweite*. Keim & Nemnich, 1925.
[20] S. K. Sahni, "The Visibility Test: How Particulate Matter Changes the Way We See Our City," *International Journal of Advanced Research in Science, Communication and Technology International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal*, vol. 10, no. 1, 2021, [Online]. Available: www.ijarsct.co.in
[21] M. I. Shahzad, J. E. Nichol, J. Wang, J. R. Campbell, and P. W. Chan, "Estimating surface visibility at Hong Kong from ground-based LIDAR, sun photometer and operational MODIS products," *J. Air Waste Manage. Assoc.*, vol. 63, no. 9, pp. 1098–1110, Sep. 2013, doi: 10.1080/10962247.2013.801372.

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i3.17336

[22] A. Dewita and S. Alfiandy, "Kajian Awal: Prediksi Kabut Berdasarkan Data Observasi di Wilayah Bandara Zainuddin Abdul Madjid," *Buletin GAW Bariri (BGB)*, vol. 5, no. 2, pp. 60–69, 2024.

[23] T. F. Prasetyo, A. Gani, and Z. Mufid, "Pengaruh Suhu, Kelembapan dan Angin terhadap Polusi Udara: Studi Kasus Dataset Air Quality," *INSANtek*, vol. 6, no. 1, pp. 43–51, 2025.

[24] D. Yu, W. Zhao, K. Nie, and G. Zhang, "Visibility forecast model based on LightGBM algorithm," *J. Comput. Appl*, vol. 41, no. 04, pp. 1035–1041, 2021.

[25] C.-J. Chen, C.-N. Huang, and S.-M. Yang, "Aviation visibility forecasting by integrating Convolutional Neural Network and long short-term memory network," *Journal of Intelligent & Fuzzy Systems*, vol. 45, no. 3, pp. 5007–5020, 2023.

[26] S. Shalev-Shwartz and S. Ben-David, *Understanding machine learning: From theory to algorithms*. Cambridge university press, 2014.

[27] Y. Zhang, Y. Du, M. Feng, and S. Hu, "Long-lasting marine heatwaves instigated by ocean planetary waves in the tropical Indian Ocean during 2015–2016 and 2019–2020," *Geophys. Res. Lett.*, vol. 48, no. 21, 2021, doi: 10.1029/2021GL095350.

[28] L. Breiman, "Random Forests," *Kluwer Academic*, vol. 45, pp. 5–32, 2001.

[29] P. Geurts, D. Ernst, and L. Wehenkel, "Extremely Randomized Trees," *Mach. Learn.*, p. 3, 2006, doi: 10.1007/s10994-006-6226-1.

[30] N. Aman *et al.*, "Estimating visibility and understanding factors influencing its variations at Bangkok airport using machine learning and a game theory–based approach," *Environmental Science and Pollution Research*, vol. 32, no. 59, pp. 31162–31179, Aug. 2024, doi: 10.1007/s11356-024-34548-4.

[31] L. D. Rotstayn *et al.*, "Have Australian rainfall and cloudiness increased due to the remote effects of Asian anthropogenic aerosols?," *Journal of Geophysical Research: Atmospheres*, vol. 112, no. D9, May 2007, doi: 10.1029/2006JD007712.

[32] P. Venkataanusha*, Ch. Anuradha, Dr. P. S. R. Chandra Murty, and Dr. S. K. Chebrolu, "Detecting Outliers in High Dimensional Data Sets Using Z-Score Methodology," *International Journal of Innovative Technology and Exploring Engineering*, vol. 9, no. 1, pp. 48–53, Nov. 2019, doi: 10.35940/ijitee.A3910.119119.

[33] E. van den Heuvel and Z. Zhan, "Myths About Linear and Monotonic Associations: Pearson's r , Spearman's ρ , and Kendall's τ," *Am. Stat.*, vol. 76, no. 1, pp. 44–52, Jan. 2022, doi: 10.1080/00031305.2021.2004922.

[34] C. Spearman, "The Proof and Measurement of Association Between Two Things," *American Journal of Psychology*, vol. 15, no. 1, pp. 72–101, 1904, doi: 10.2307/1412159.

[35] BMKG, *Peraturan Deputi Bidang Meteorologi BMKG Nomor 1 Tahun 2015 tentang Tata Cara Pengamatan dan Pelaporan Local Routine Report (Met Report) dan Local Special Report (Special) untuk Pelayanan Informasi Meteorologi Penerbangan di Lingkungan Badan Meteorologi Klimatologi dan Geofisika*. 2015.

[36] I. H. Sarker, "Machine learning: Algorithms, real-world applications and research directions," *SN Comput. Sci.*, vol. 2, no. 3, p. 160, 2021.

[37] D. M. W. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," *ArXiv*, vol. abs/2010.16061, 2011, [Online]. Available: https://api.semanticscholar.org/CorpusID:3770261

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI : 10.62411/tc.v25i3.17336