future internet

MDPI

Article

Machine Learning-Based Short-Term Visibility Classification for Wireless Optical Communication Systems Using METAR and Microwave-Link Features at Bangkok Airports

Sabai Phuchortham and Hakilo Sabit *

School of Engineering, Computer and Mathematical Sciences, Auckland University of Technology, 6 Saint Paul Street, Auckland Central, Auckland 1010, New Zealand; sabai.phuchortham@autuni.ac.nz
* Correspondence: hakilo.sabit@aut.ac.nz

Abstract

Rapid growth in connected devices, artificial intelligence applications, and the Internet of Things (IoT) is driving demand for ultra-high data rates, low latency, and energy-efficient communication infrastructure. Wireless optical communication (WOC), including free-space optical (FSO), is recognized as a disruptive technology for 6G and future-generation networks. However, atmospheric visibility critically affects WOC/FSO link availability, capacity, and reliability. This study proposes a machine learning (ML)-based low-visibility classification model that integrates Meteorological Aerodrome Reports (METARs) with microwave-link received-signal (Rx) features. Visibility below 6000 m is predicted at the 1 h, 3 h, and 6 h horizons using 18 months of data from Suvarnabhumi Airport (VTBS) and Don Mueang Airport (VTBD) in Bangkok, Thailand. Four ML algorithms, namely logistic regression, random forest, extreme gradient boosting, and light gradient boosting machine (LGBM), are evaluated against persistence and Terminal Aerodrome Forecast (TAF) baselines. In a 100-round block-bootstrap evaluation, LGBM with METAR-Rx achieved the highest mean F1 scores at the 1 h and 3 h horizons, outperforming TAF by 28 and 20 percentage points at the 1 h horizon for VTBS and VTBD, respectively. SHAP and ablation analyses suggested that current visibility is the dominant predictor, while Rx features provide complementary information and improve F1 performance by approximately 1–4 percentage points. Seasonal analysis shows stronger cool-season performance, while rainy-season prediction remains challenging. Adding visibility-trend features further improves performance, with the best combined model achieving 1 h F1 scores of 0.7253 for VTBS and 0.6495 for VTBD. These findings indicate that integrating the METAR-Rx feature set can support short-term low-visibility classification.

Keywords: machine learning; atmospheric visibility classification; wireless optical communication; 6G; METAR; microwave link; future-generation networks

Academic Editors: Evangelia Filiopoulou, Mara Nikolaidou, Teta Stamati, Cleopatra Bardaki, Georgia Dede, Alexandra Gasparinatou and Elena Politi

Received: 26 June 2026
Revised: 17 July 2026
Accepted: 22 July 2026
Published: 25 July 2026

1. Introduction

The rapid development of connected devices, artificial intelligence applications, IoT, and high-capacity wireless communication networks is accelerating digital transformation across smart cities, intelligent transportation systems, and airport infrastructure. In these environments, reliable, low-latency connectivity is essential for supporting data-driven operations, automation, safety monitoring, and resilient communication services. WOC is a complementary high-capacity technology for 6G networks and a key component of future convergent network architectures that can address these requirements [1,2]. WOC uses optical waves to transmit data through free space, providing fiber-like data rates without requiring physical transmission lines. It can support space, aerial, terrestrial, and underwater communication [3–6].

However, the performance of WOC systems can be affected by environmental conditions, particularly in outdoor wireless and sensing applications. In particular, atmospheric visibility is a critical factor because it directly affects WOC systems, including FSO links, as well as LiDAR-based autonomous-mobility systems [7]. Reduced visibility can degrade WOC link availability, disrupt aviation and transport services, and reduce the reliability of camera and LiDAR sensing systems. Therefore, short-term visibility awareness can support more resilient decision-making across digital communication, aviation, and intelligent transportation systems.

More specifically, in WOC systems, visibility directly affects atmospheric attenuation and signal reliability under adverse weather conditions, and it is a key variable in visibility-dependent path-loss models [4]. In aviation, visibility is essential for operational safety, flight planning, and runway visual range assessment [8]. In road transportation, reduced visibility due to fog, rain, or haze can increase the risk of crashes and reduce drivers' ability to detect hazards. It also affects camera- and LiDAR-based systems used in advanced driver-assistance systems and automated vehicles [9]. Therefore, low visibility can degrade communication-link performance, reduce operational efficiency, and increase risks across aviation and transportation systems.

To mitigate these effects, short-term visibility prediction is desirable because it can provide early warnings of low-visibility conditions and support timely decision-making across communication, aviation, and transportation systems. Visibility is typically predicted spatiotemporally using standard approaches, namely numerical weather prediction (NWP) [10] and terminal aerodrome forecasts (TAFs). However, these approaches tend to exhibit limited predictive performance [11–13]. Consequently, many researchers have investigated alternative approaches, including statistical models [14–16] and multiple ML algorithms [17–21].

Artificial intelligence is also transforming WOC research, especially through ML approaches for noise mitigation, channel estimation, system optimisation, performance monitoring, and network management [22]. Recent studies have used ML algorithms to improve WOC performance through modulation adaptation [23–25], signal detection [26], and power adaptation [27]. These studies indicate the growing integration of ML into WOC systems.

ML-based approaches require high-quality meteorological data for model training. Meteorological Aerodrome Reports (METARs) can provide such data and are commonly used as a standard source of airport weather observations. METARs routinely report information on visibility, temperature, dew point, wind conditions, pressure, cloud cover, and present weather conditions. Several studies have therefore used METAR observations with ML algorithms for short-term fog and visibility prediction, including studies on current visibility estimation during the dry season at VTBD [28], fog forecasting at airports in North India [29], visibility prediction using a random forest (RF) algorithm [30], comparisons of multiple ML algorithms for fog prediction at Spanish airports [31], solar irradiance forecasting [32], and short-term fog forecasting [33]. These publications demonstrate the potential of METAR-based ML algorithms for improving short-term low-visibility prediction performance.

Similarly, wireless telecommunications and sensing technologies, such as cellular links, radar, and microwave links, have proven useful for providing weather information. For instance, rainfall intensity has been estimated using cellular links [34,35], microwave links [36], and satellite radars [37]. In addition, wireless telecommunications data have been used to predict weather-related conditions, such as rainfall [38,39], fog [40], sandstorms [41], and dusty conditions [42]. These publications show that variations in wireless signals are associated with atmospheric conditions, particularly rain- and fog-related attenuation.

The combination of METARs and microwave-link received-signal (Rx) features represents an opportunity to improve short-term low-visibility prediction, especially in airport areas where microwave links are commonly part of the communication infrastructure. However, a limited number of studies have investigated visibility-related applications using weather observations and Rx features together. For example, Ref. [42] used weather services and microwave-link Rx to estimate visibility in dusty conditions, while Ref. [43] used weather-station data and microwave-link Rx to monitor fog. Consequently, the direct integration of METAR observations with Rx features for short-term low-visibility prediction remains limited in the existing literature.

Existing ML-based visibility studies have mainly used meteorological observations, numerical weather prediction outputs, weather-station measurements, satellite data, and camera observations to estimate or forecast visibility. Although wireless and microwave-link attenuation data have also been investigated for monitoring rain, fog, dust, and visibility-related conditions, the combined use of METAR observations and microwave-link Rx features for multi-horizon airport low-visibility classification remains limited. In parallel, ML-based WOC studies have largely focused on physical-layer functions, such as modulation adaptation, signal detection, channel estimation, system optimisation, and transmit-power control, rather than predicting atmospheric visibility in advance of optical-link degradation. More specifically, published studies using real-world environmental data to predict visibility for proactive WOC link operation remain scarce. Therefore, a technical gap remains between ML-based environmental visibility prediction and proactive WOC link operation.

In this paper, a short-term low-visibility classification approach that integrates METARs and Rx features is proposed. To the best of our knowledge, this is the first study to use METAR observations and microwave-link Rx features together for multi-horizon low-visibility classification at tropical airports and to interpret the resulting predictions for proactive WOC link operation. The main contributions are as follows:

- Real-world multi-source low-visibility classification: This study integrates METAR and Rx features for short-term low-visibility classification below 6000 m at the 1 h, 3 h, and 6 h prediction horizons using 18 months of data from two airports in Bangkok.
- Comparison of ML algorithms and operational baselines: Four ML algorithms, namely LR, RF, XGB, and LGBM, are evaluated using a chronological training–test split and a 100-round, 7-day block-bootstrap method. Their performance is compared with persistence and TAF baselines under highly imbalanced real-world conditions.
- Microwave-link received signal levels as complementary predictors: Feature importance, SHAP beeswarm, and feature ablation methods are used to verify the contribution of Rx features to the classification model. The addition of Rx features improves F1 performance by approximately 1–4 percentage points.
- Seasonal and visibility-trend evaluation: The model is further examined using cool- and rainy-season subsets and visibility-trend features.
- Connection to proactive WOC operation: Visibility-dependent atmospheric attenuation, path loss, and remaining link margin are estimated using the Kim model to demonstrate possible proactive WOC link-management actions.

The remainder of this paper is organised as follows. Section 2 describes the data sources and preprocessing steps. The machine-learning setup and evaluation metrics are outlined in Section 3. Section 4 reports experimental results, and Section 5 discusses the findings. The study is summarised in Section 6.

2. Data and Preprocessing

This section describes the data sources and preprocessing procedures used in this study. The dataset includes METAR observations, TAF forecasts, and Rx levels from two airport sites, namely VTBS and VTBD. The data were collected from 4 November 2024 to 5 May 2026, covering an 18-month observation period. The software environment comprised Python 3.12.2, NumPy 2.2.3, pandas 2.2.3, scikit-learn 1.8.0, XGBoost 3.1.2, LightGBM 4.6.0, PyMySQL 1.1.2, Pysnmp 4.4.12, requests 2.32.5, shap 0.51.0, MySQL-Front 6.0, MATLAB v25.2 and Visual Studio Code 1.116.0.

VTBS and VTBD are two major airports in Thailand. VTBS is located at latitude 13.68110 and longitude 100.74700, while VTBD is located at latitude 13.91260 and longitude 100.60700 [44], as shown in Figure 1. These airports represent tropical airport environments in Southeast Asia, where visibility is commonly affected by high humidity, fog, heavy rain, and thunderstorms. Such weather conditions can reduce visibility and cause short-term visibility fluctuations.

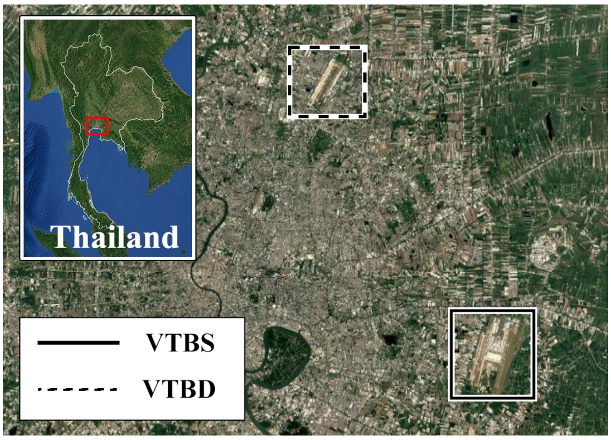


Figure 1. Geographic locations of the two selected airport sites in Thailand: VTBS and VTBD.

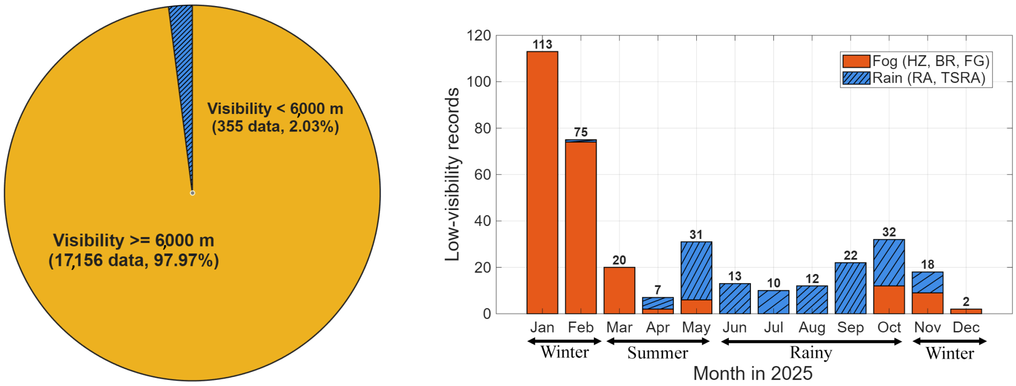
To obtain high-quality data, the process began with data acquisition, collection, and cleaning to ensure reliable input. The data were then transformed into ML features, which were used to train the selected ML algorithms.

2.1. Meteorological Aerodrome Report

METAR is a standard routine aviation weather report issued every 30 min [30]. It contains plain-text data, including visibility, temperature, dew point, wind variables, weather conditions, sky coverage, and observed time. For example, the row "METAR VTBS 272130Z 19009KT 9999 FEW020 30/25 Q1007 NOSIG" indicates that Suvarnabhumi Airport on the 27th day at 21:30 UTC had wind from 190 degrees at 09 Knots, visibility of 10 km or more, FEW clouds at 02000 feet, temperature of 30 °C, dew points 25 °C, pressure 1007 hPa, and no significant weather change expected soon.

METAR observations were obtained from the aviation weather centre website through an automated Python-based data collection process. The REQUESTS package was used to retrieve the data, the METAR package was used to parse and clean the observations, and the data was then stored in a local MySQL database.

Figure 2a shows the percentage distribution of visibility records in 2025. Low-visibility cases below 6000 m are rare at VTBS, accounting for only 355 of 17,511 records (2.03%). This indicates an extremely imbalanced dataset. Such an imbalance can cause ML algorithms to become biased in favour of the majority class, resulting in poor detection or misclassification of the minority class [45,46].



(**a**) (**b**)

Figure 2. Visibility characteristics at VTBS in 2025: (**a**) visibility distribution using the 6000 m threshold; (**b**) monthly weather-code distribution for visibility below 6000 m.

Figure 2b shows the monthly distribution of low-visibility cases by weather-code category at VTBS in 2025. The results indicate that low-visibility events are mainly associated with adverse atmospheric conditions, particularly fog, rain, and thunderstorms. During the winter season, fog-related conditions are more frequently associated with low-visibility events. In contrast, during the rainy season, rain and thunderstorms are the main causes of low-visibility events. This point is important because many visibility prediction studies focus primarily on fog events, while rain-related visibility degradation has received comparatively less attention [20,47].

2.2. Terminal Aerodrome Forecast

TAF is a traditional aviation weather forecast product, typically updated every 6 h. It provides plain-text information on expected weather conditions at an airport for a specified forecast period [48]. Unlike METAR, TAF describes forecast conditions and includes temporal and probabilistic indicators [49]. For example, PROB30 (probability) indicates a 30% probability of specific weather conditions. BECMG (becoming) indicates a gradual change expected within a specified time window, while TEMPO (temporary) indicates temporary weather conditions during a defined period. Further details on TAF coding and interpretation can be found in Ref. [13].

In this study, a Harris-style TAF interpretation [50] was used to convert TAF codes into event probabilities for binary low-visibility evaluation. This method was selected because the F1 score requires binary forecast labels, whereas the TAF codes contain temporal and probabilistic indicators. Although the Gerrity Skill Score (GSS) has been adopted by the Met Office since April 2020, it is not used to derive the F1 score because GSS is a categorical forecast-verification score rather than a binary label. In the Harris-style interpretation, TAF codes are converted into probabilistic values and then transformed into binary forecasts [0, 1] using a selected probability threshold. For example, assuming a probability threshold of 0.50, a binary forecast of 0 represents a high-visibility forecast. PROB30 and TEMPO are converted to a probability of 0.30. Therefore, PROB30 and TEMPO are classified as a binary forecast of 0 (high-visibility forecast) because their assigned probability is below the threshold of 0.50. The resulting binary forecast is then compared with the corresponding METAR observation to evaluate TAF performance.

To examine the sensitivity of TAF performance to the probability cutoff, thresholds from 0.10 to 0.90 were evaluated, as reported in Appendix A. The 0.50 threshold was used as a common reference operating point. The sensitivity analysis showed that it produced the highest or near-highest F1 score across the evaluated cases.

2.3. Microwave-Link Received-Signal Data

Microwave-link Rx-level data from two airport sites, VTBS and VTBD, were monitored. The equipment operates in the C-band (7.5 GHz) and is installed on rooftops to provide line-of-sight communication within the airport area, as shown in Figure 3.

(**a**) (**b**)

Figure 3. Microwave link equipment within the airport area: (**a**) VTBS; (**b**) VTBD.

The minimum and maximum Rx-level measurements were collected every 15 min via an automated Python-based data-collection process using the PySNMP package [51], and the data were stored in a local MySQL database. The equipment internally measured the Rx level every minute and reported the minimum and maximum values observed within each 15-min interval. To synchronise the Rx data with METAR reports, which were issued every 30 min, only Rx records timestamped at hh:00 and hh:30 were stored. The transmit power at the transmitter sites was fixed throughout the monitoring period. Therefore, variations in the received signal level primarily reflect propagation-related changes along the microwave links. Detailed link parameters, including link distance, antenna height, polarisation, exact site location, and full equipment configuration, are omitted for operational confidentiality.

2.4. Feature Selection

All datasets were stored and/or reprocessed at 30-min intervals. METAR timestamps (UTC) were converted to local time (UTC+7) to synchronise with Rx datasets. Table 1 shows the variables used for model training and evaluation.

Cyclic variables, including hour and month, were converted into sine and cosine components. This provides the models with a more meaningful representation of cyclical variables than raw numerical values. In addition, text-based weather codes were converted into binary weather flags [0, 1] to indicate the severe weather conditions associated with low-visibility, such as fog, haze, rain, and thunderstorms, as shown in Figure 2b.

Missing data were handled using both exclusion and imputation. Records were excluded when the METAR observation or matching Rx record was unavailable, ensuring that only valid observations were used. For the remaining records, missing values in the predictor variables, including temperature, dew point, wind, weather flag, cloud base, and cloud ceiling, were replaced with the median value from the training set.

A visibility threshold of 6 km was used as a proxy for low-visibility classification. This threshold is relevant to WOC assessment because visibility is a key parameter in atmospheric attenuation modelling. According to the Kim atmospheric attenuation model for WOC [52], visibility below 6000 m is important for estimating optical signal attenuation. In aviation, visibility is also essential for operational safety and for runway visual range (RVR) assessment. For instance, the Civil Aviation Authority of New Zealand specifies a minimum flight visibility of 5 km for VFR operations at aerodromes within a control zone [53].

Table 1. Variables used as predictors and targets for model training and assessment.

Variable	Description	Data Type	SI Unit
Predictor variables			
Visibility	Horizontal visibility observed from METAR	Numerical	m
Temperature	Air temperature	Numerical	°C
Dew point	Dew-point temperature	Numerical	°C
Dew-point depression	Temperature–dew-point difference	Numerical	°C
Wind speed	Wind speed from METAR	Numerical	kt
Wind gust	Maximum wind gust speed	Numerical	kt
Month	Month of observation encoded using sine and cosine components	Cyclic numerical	–
Time	Local time (UTC+7) encoded using sine and cosine components	Cyclic numerical	–
Weather flag	Binary flag indicating weather conditions related to visibility reduction	Binary	–
Cloud base	Lowest cloud-base height	Numerical	ft
Cloud ceiling	Cloud ceiling height	Numerical	ft
Rx received level	Minimum and maximum received signal levels from the microwave link	Numerical	dBm
Target variables			
Visibility 1 h	Binary low-visibility target 1 h ahead, where visibility < 6000 m	Binary	–
Visibility 3 h	Binary low-visibility target 3 h ahead, where visibility < 6000 m	Binary	–
Visibility 6 h	Binary low-visibility target 6 h ahead, where visibility < 6000 m	Binary	–

The prediction targets are visibility at 1, 3, and 6 h ahead. For each horizon, visibility below 6000 m was labelled 1, while visibility at or above 6000 m was labelled 0, representing low- and high-visibility conditions, respectively. These target variables enable the models to determine the likelihood of low-visibility conditions occurring within the specified prediction horizons.

3. Machine Learning Setup

3.1. Machine Learning Algorithms

To forecast visibility below 6000 m, a persistence baseline (PB) and four ML prediction algorithms were evaluated for performance: LGBM, Extreme Gradient Boosting (XGB), logistic regression (LR), and RF. The evaluated algorithms are explained below. These ML algorithms were selected because they are commonly used for supervised classification tasks and have been widely applied to feature-based time-series prediction applications.

3.1.1. Persistence Baseline Algorithm

PB is the simplest baseline algorithm, assuming the future visibility state is identical to the present state. If visibility is low at the current time, PB forecasts low visibility for the next prediction period. This model can achieve strong performance when environmental conditions do not change frequently, especially for very short-term forecasts. It can also reduce computational cost because no training is required. However, PB does not learn from other variables and fails in dynamic environments where visibility changes rapidly.

3.1.2. Light Gradient-Boosting Machine

LGBM is a decision-tree-based gradient boosting algorithm used for both classification and regression tasks [54]. LGBM adopts a leaf-wise growth strategy, enabling it to handle large datasets effectively while reducing computation time and memory usage. It has been applied to solar irradiance [32] and sea ice prediction [55]. These studies reported high predictive performance for LGBM.

3.1.3. Extreme Gradient Boosting

XGB is another decision-tree-algorithm that uses a level-wise tree-growth strategy for forecasting and handles missing values [56]. XGB includes regularisation mechanisms that can help control model complexity and overfitting. Many applications have adopted XGB. For example, XGB has been applied to visibility prediction using meteorological variables [57], radiation-fog visibility forecasting using meteorological and satellite data [20], and visibility prediction using WRF atmospheric chemistry during heavy-pollution conditions [58]. These findings highlight the XGBoost performance.

3.1.4. Logistic Regression

LR is one of the simplest ML algorithms used for binary classification. It estimates the likelihood of an event occurring using maximum-likelihood estimation and a sigmoid function, yielding a probability that can be converted into a binary output with a decision threshold. LR has been used to classify visibility levels at Vienna Airport [15,59].

3.1.5. Random Forest

RF builds multiple decision trees using the bootstrap method. Each tree is trained on a subsample of the dataset, and the final prediction is obtained by combining the outputs of all trees. RF is suitable for both regression and classification. It has been used to estimate visibility in South Korea [18], at Sofia Airport [19], and under dusty conditions in West Africa [42]. These studies suggest that RF can provide good predictive performance for visibility-related applications.

3.2. Training and Evaluation

3.2.1. Training

The entire dataset was split chronologically into 80% for training and 20% for testing in a time-series setting. To prevent temporal leakage, the earliest 80% of the data were used for training, while the most recent 20% were held out for testing. Therefore, no future test data were included during model training. The first 80% of the dataset comprised 20,945 and 20,875 records for VTBS and VTBD, respectively, while the final 20% comprised 5237 and 5219 records for VTBS and VTBD, respectively. However, due to the limited number of low-visibility cases, a separate validation set was not used. Instead, a fixed decision threshold of 0.50 was used as a common reference operating point for all models. Threshold sensitivity was additionally evaluated using thresholds from 0.05 to 0.95 across 100 bootstrap rounds for each algorithm, airport, and prediction horizon, as presented in Appendix A.

Due to the extreme class imbalance in the dataset, as shown in Figure 2a, class-weighting configurations were applied to the ML algorithms. This allowed the algorithms to pay more attention to low-visibility events during training.

To assess the statistical stability of the model performance, bootstrapping was implemented [60]. The test data were resampled over 100 bootstrap rounds for each evaluation. A 7-day overlapping block-bootstrap approach was used to preserve short-term atmospheric dependencies within the data. The 1 h, 3 h, and 6 h visibility lead-time prediction models were trained separately. The resulting distributions of the evaluation metrics across the 100 bootstrap rounds were used to estimate the mean performance and 95% confidence intervals for each dataset variant. The overall methodological framework is shown in Figure 4.

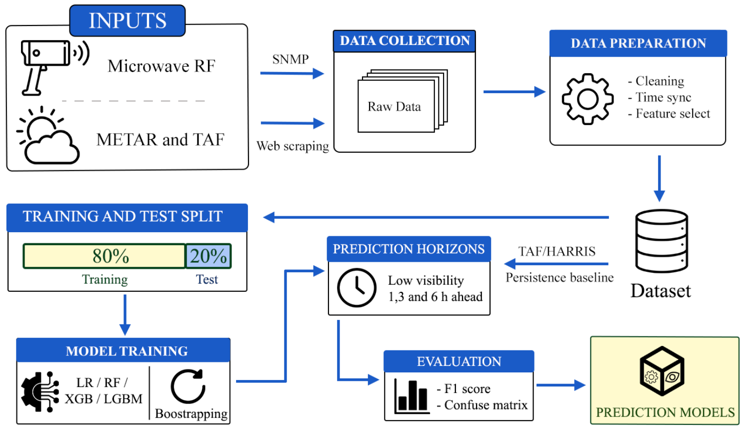


Figure 4. Overall workflow of the proposed low-visibility prediction models.

3.2.2. Evaluation

For each prediction horizon, the algorithms were trained and tested separately. Each prediction case was matched to an observation case, yielding the confusion matrix. The confusion matrix contains true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). In this study, TP denotes a correct prediction of low visibility, while FN denotes an actual low-visibility event that the model predicted as high visibility.

Accordingly, recall, specificity, precision, negative predictive value (NPV), and accuracy can be calculated from the confusion matrix. The model outputs were evaluated using a confusion matrix, as shown in Figure 5.

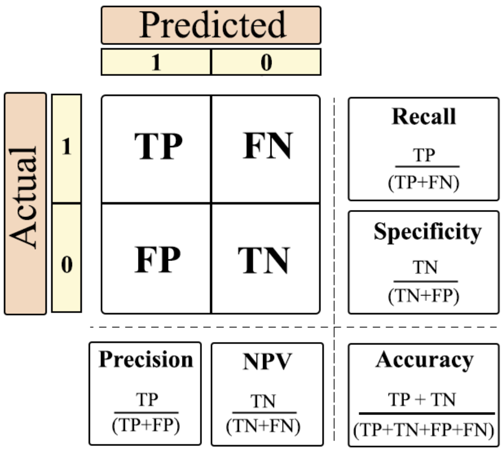


Figure 5. Confusion matrix and classification performance metrics.

In this study, the classification performance was evaluated using accuracy, precision, recall, and F1 score. Accuracy measures the overall proportion of correct predictions. Precision indicates the proportion of predicted low-visibility events that were correct. Recall measures the model's ability to detect actual low-visibility events. The F1 score was used because it balances precision and recall, making it suitable for imbalanced classification problems. The F1 score is expressed as follows:

$$\text{F1} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{1}$$

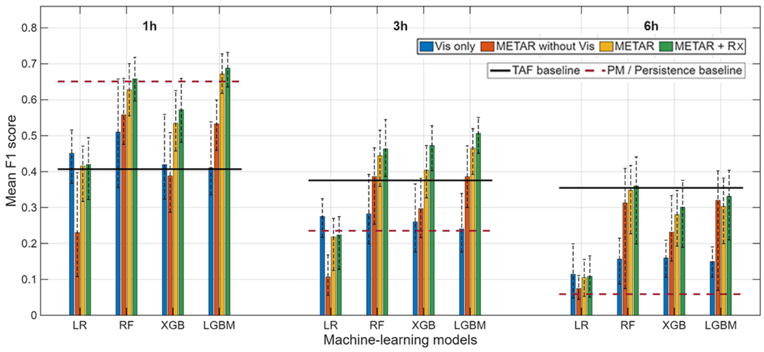
4. Experimental Results

This section presents the classification performance of the selected ML algorithms across the entire dataset. The evaluation results are reported in Section 4.1. The feature and subgroup analyses are presented in Section 4.2, while the event-based analysis and visibility-trend results are demonstrated in Section 4.3.

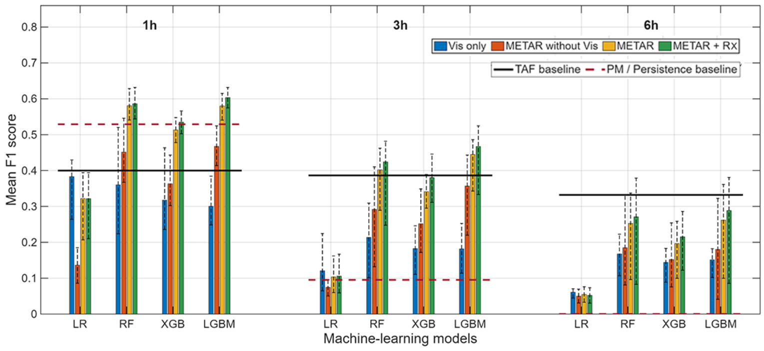
4.1. Classification Results

4.1.1. F1 Score

The F1 scores for VTBS and VTBD visibility prediction at 1, 3, and 6 h ahead using the four selected ML algorithms and feature sets are shown in Figure 6.



(**a**)



(**b**)

Figure 6. F1 score comparison of low-visibility prediction models using different feature sets: (**a**) VTBS; (**b**) VTBD.

According to Figure 6, the algorithms using the METAR-Rx feature set (green bars) generally achieved the highest F1 scores, except for the LR algorithm. For the 1 h prediction horizon at VTBS, as shown in Figure 6a, the LGBM, RF, XGB, LR, PB, and TAF baselines achieved mean F1 scores of 0.69, 0.66, 0.57, 0.42, 0.65, and 0.41, respectively. The vertical dashed lines show the minimum and maximum F1 scores obtained from the 100 bootstrap runs. Among the proposed ML algorithms, LGBM with the METAR-Rx feature set achieved the highest F1 score and outperformed both the PB and TAF baselines across the 1 h and 3 h prediction horizons. However, none of the proposed ML models exceeded the TAF baseline at the 6 h horizon. These results suggest that LGBM and RF are the most suitable ML models for short-term visibility classification at the 1 h and 3 h horizons in this setup.

Moreover, Figure 6 shows a comparison of mean F1 scores using different feature sets. In Figure 6a, which represents the VTBS dataset, the METAR-Rx feature set achieved the highest F1 scores among the tested feature sets for the LGBM, XGB, and RF algorithms. The only exception was LR, where the visibility-only feature set achieved the highest F1 score. For LGBM at the 1 h prediction horizon, the METAR-only feature set achieved an F1 score of 0.67, while the METAR-Rx feature set improved the score to 0.69. Similarly, at the 3 h prediction horizon, the METAR-Rx feature set achieved a higher F1 score than the METAR-only feature set, increasing from approximately 0.46 to 0.51. These results indicate that adding Rx features can improve the F1 score by approximately 2–5 percentage points for the VTBS dataset. Figure 6b shows a similar trend for the VTBD dataset, where the METAR-Rx feature set also provided modest improvements, typically around 2–3 percentage points, particularly for LGBM, XGB, and RF. These results suggest that the Rx feature can be used with the METAR feature to improve visibility classification for the LGBM, XGB, and RF algorithms.

For the PB (horizontal red dashed lines), the 1 h F1 score was moderate at both airports. This indicates that visibility conditions often remain stable over short time periods. However, as the prediction horizon increased to 6 h, the persistence baseline dropped sharply because visibility conditions were more likely to change over longer periods. In contrast, the TAF baseline (black lines) provided more stable F1 scores across the 1-, 3-, and 6-h prediction horizons, with values generally in the range of 0.33–0.41. This suggests that TAF remains a valid operational baseline for practical forecasting, particularly at longer prediction horizons.

The METAR-Rx feature set (green bars) was also compared with the PB and TAF baselines. At the 1 h prediction horizon, LGBM with the METAR-Rx feature set increased the F1 score by 28 and 20 percentage points compared with the TAF baseline for VTBS and VTBD, respectively. When compared with the PB baseline, the F1 score increased by 4 and 7 percentage points for VTBS and VTBD, respectively. Likewise, at the 3 h prediction horizon, the LGBM model achieved higher F1 scores than TAF, with improvements of 13.1 and 8.1 percentage points for VTBS and VTBD, respectively. When compared with PB, the LGBM model achieved higher F1 scores, with improvements of 27.1 and 37.7 percentage points for VTBS and VTBD, respectively. However, at the 6 h prediction horizon, all the tested ML models underperformed the TAF baseline but still outperformed PB. These results suggest that LGBM with the METAR-Rx feature set achieved higher F1 scores than the PB and TAF baselines at the 1- and 3-h prediction horizons.

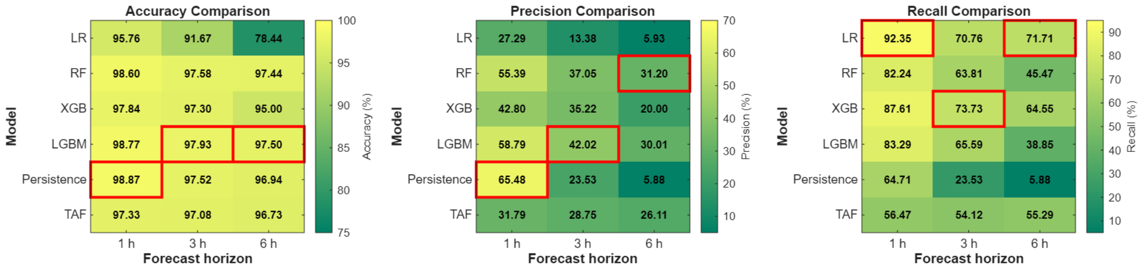
Comparing Figure 6a and Figure 6b, which represent VTBS and VTBD, respectively, differences in prediction performance can be observed even though the two airports are located relatively close to each other, approximately 30 km apart. The results show that the VTBS models achieved higher F1 scores than the VTBD models across all prediction horizons. The F1 scores of the ML algorithms at VTBS were often more than 0.10 higher than those at VTBD, whereas the TAF baselines for both airports showed relatively similar performance. This may indicate that visibility conditions at VTBS were more stable or changed more slowly than those at VTBD. Therefore, even airports located in the same metropolitan area can experience different visibility-prediction behaviour.

4.1.2. Accuracy, Recall, and Precision

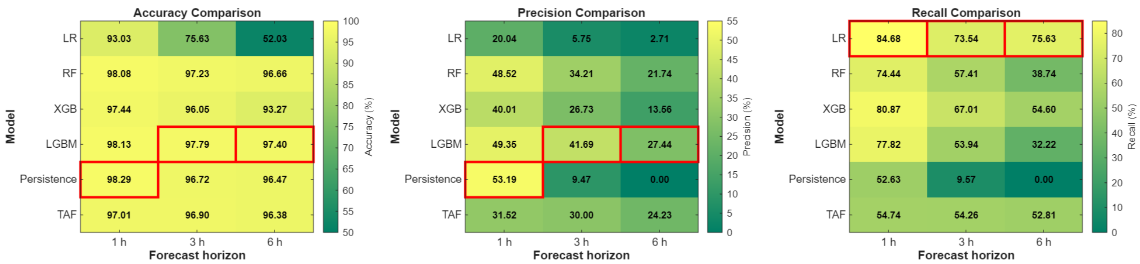
In this section, the METAR-Rx feature set was used for the ML algorithms, as it generally achieved the highest F1 score. Figure 7 shows the accuracy, recall, and precision scores for VTBS and VTBD at 1-, 3-, and 6-h prediction horizons using different ML algorithms and the TAF baseline. These scores provide complementary information to the

F1 score by evaluating model performance from different perspectives. The results are discussed separately for each metric in the following paragraphs.

As shown in Figure 7, lighter green represents higher scores, while the red boxes highlight the highest score for each performance category at each prediction horizon.



(**a**)



(**b**)

Figure 7. Classification-score comparison of low-visibility prediction models using the METAR-Rx feature set and operational baselines: (**a**) VTBS; (**b**) VTBD. Red boxes mark the highest score for each metric at each prediction horizon.

Accuracy is the least informative metric in this study due to the extremely imbalanced dataset. All prediction algorithms achieved more than 95% accuracy, except for the LR algorithm at some prediction horizons. PB achieved the highest scores at both airports, with a 1-h prediction accuracy ranging from 98.29% to 98.87%, while LGBM achieved the highest scores for the 3- and 6-h predictions, ranging from 97.79% to 97.93% and from 97.40% to 97.50%, respectively. The scores show that even when the PB algorithm assumes the current visibility condition will persist, it can still achieve high scores due to the dataset's imbalance. Therefore, accuracy is not a reliable standalone metric in this study.

The precision metric shows mixed results. Depending on the airport and prediction horizon, PB, LGBM, and RF achieved the highest precision scores, with values ranging from 27.44% to 65.48%. In the VTBS scenario (Figure 7a), the PB, LGBM, and RF algorithms achieved the highest scores of 65.48%, 42.02%, and 31.20% for the 1 h, 3 h, and 6 h predictions, respectively. For the VTBD scenario (Figure 7b), the PB achieved the highest score of 53.19%, while the LGBM algorithm achieved the highest scores of 41.69% and 27.44% for the 3- and 6-h predictions, respectively. The results suggest that LGBM tended to achieve higher precision than the other ML algorithms at longer prediction horizons, potentially reducing false alarms compared with lower-precision models.

The recall metric shows that LR achieved the highest scores among the algorithms for most prediction horizons, except for the 3 h prediction at VTBD, where XGB achieved the best score. However, the recall score of LR was only slightly higher than that of other algorithms. For example, for the 1 h prediction at VTBD, LR, RF, XGB, and LGBM achieved recall scores of 84.68%, 74.44%, 80.87%, and 77.82%, respectively. The differences in recall scores ranged from 3.81% to 10.20%, whereas the differences in precision scores ranged from 4.67% to 33.15%. These results indicate that LR performed best in terms of recall among the ML algorithms, while XGB and LGBM also achieved competitive recall scores.

From these results, LGBM with the METAR-Rx feature set achieved the most balanced performance across the three metrics, consistent with its highest F1 score. Although other models achieved the best score in some specific cases, such as PB for the 1 h prediction horizon and LR for recall, their performance was less consistent across all metrics and prediction horizons. Therefore, LGBM can be considered the most suitable ML algorithm for this study.

4.1.3. Computational Cost

The computational cost of the four machine-learning algorithms was evaluated for the models used in the 100-round block-bootstrap analysis. Experiments were conducted on an ASUS laptop equipped with an Intel Core i5-11400H CPU (6 cores, 12 threads, 2.70 GHz) and 16 GB of RAM. All models were trained and evaluated using the CPU without GPU acceleration. The software environment included Python 3.12.2, scikit-learn 1.8.0, XGBoost 3.1.2, and LightGBM 4.6.0.

For each airport and prediction horizon, the training time, single-sample inference latency, and serialized model size were recorded. Training time included data preprocessing and model fitting, while inference latency included preprocessing and prediction.

Table 2 summarises the results across the two airports and three prediction horizons. LR had the lowest training time and serialized model size, while XGB was the most computationally efficient among the evaluated tree-based algorithms. The LGBM model yielded an average training time of 1.83 s and an inference latency of 8.14 ms. Although RF had the highest inference latency, its latency remained below 53 ms. The evaluated models required less than 3 s for training. These results suggest that the models are computationally feasible for near-real-time deployment.

Table 2. Computational cost of the evaluated ML algorithms using the METAR–Rx feature set.

Model	Training Time (s)	Inference Latency (ms)	Model Size (MB)
LR	0.049	3.68	0.0017
RF	1.149	52.96	2.387
XGB	0.358	5.61	0.220
LGBM	1.825	8.14	2.631

4.2. Feature Analysis

This section focuses on the contributions of the predictor features across all test observations. The LGBM algorithm with the METAR–Rx feature set was selected because it achieved the highest F1 score in the previous section. It was therefore used to analyse feature effectiveness.

4.2.1. Bootstrap-Based Evaluation

To further assess the contribution of the Rx features, the F1 scores obtained using the METAR–Rx and METAR-only feature sets were compared across the same 100 bootstrap rounds. The results are shown in Figure 8.

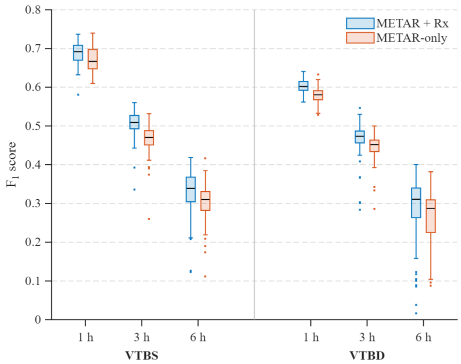


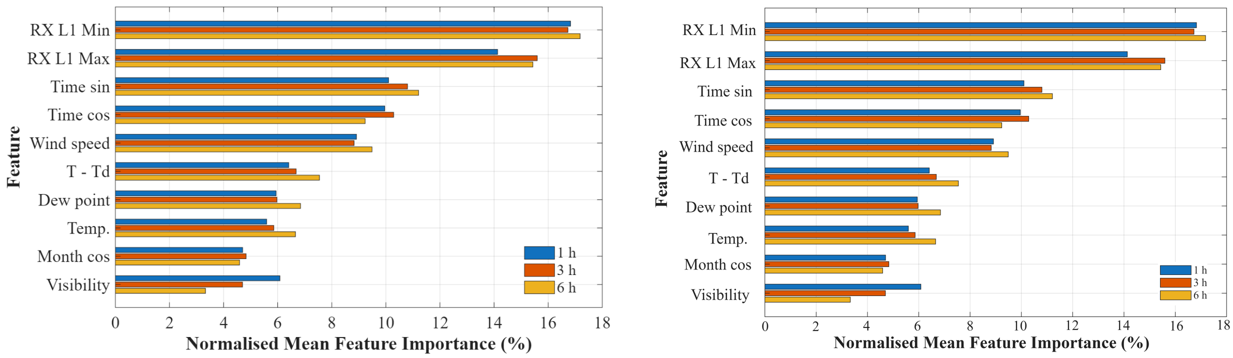
Figure 8. LGBM F1-score distributions across 100 bootstrap rounds at VTBS and VTBD.

The box plots show that the METAR-Rx feature set distributions were generally shifted towards higher F1 scores than the METAR-only distributions. Although the spread of the bootstrap values increased in some cases after adding the Rx features, the mean F1 score improved in all evaluated cases.

Combining the Rx features with METAR observations increased the mean F1 score across all airport–horizon combinations. For VTBS, the paired mean F1 improvements were 1.6, 4.1, and 2.8 percentage points at the 1 h, 3 h, and 6 h horizons, respectively, with corresponding 95% confidence intervals of −2.56 to 6.27, −0.93 to 8.99, and −1.95 to 8.64 percentage points. The METAR-Rx feature set achieved a higher F1 score than the METAR-only feature set in 78%, 96%, and 83% of the bootstrap rounds, respectively. For VTBD, the paired mean improvements were 2.3, 2.2, and 2.7 percentage points, with corresponding 95% confidence intervals of −0.50 to 5.27, −3.34 to 7.48, and −4.02 to 9.27 percentage points. The METAR-RX feature set was favoured in 93%, 85%, and 88% of the bootstrap rounds, respectively. Although the mean improvement was positive in every evaluated case, all confidence intervals included zero. Therefore, the Rx contribution should be interpreted as modest and generally favourable across the resampled periods, rather than as statistically conclusive.

4.2.2. Feature Importance

Figure 9 shows the feature importance of the predictor variables in the LGBM algorithm. The Rx variables are the most important features for both airports, indicating that received signal levels provide useful information for distinguishing between low- and high-visibility conditions. Also, local-time features are the second-most important predictors, suggesting that time-of-day features are informative for classifying low and high visibility. Likewise, the month-related features are important, indicating that seasonal effects contribute to low-visibility classification. These time and month features suggest that low visibility is more likely to occur during specific periods. For example, according to Figure 2b, January and February experience more low-visibility events than other months, while fog, which reduces visibility, is more likely to occur at night due to higher humidity and lower temperatures [47].



(**a**) (**b**)

Figure 9. Relative importance of features in the LGBM model using the METAR-Rx feature set at different prediction horizons: (**a**) VTBS; (**b**) VTBD.

4.2.3. SHAP-Based Predictor Importance

The Shapley Additive Explanations (SHAP) method is a mathematical technique for explaining each feature's contribution. While the previous section evaluates the overall ranking of predictors, SHAP provides a more detailed explanation of how each feature contributes to individual model outputs. In this study, the SHAP beeswarm plot [61] was used to analyse the feature importance of the trained LGBM model.

Accordingly, in Figure 10, visibility has the largest SHAP impact on low-visibility prediction, as shown by the positive SHAP values on the horizontal axis, because current visibility is directly related to future visibility. Figure 10a,b show that low current visibility values, represented by many blue dots, strongly influence the model output at the 1 h prediction horizon. This also explains why PB is a strong baseline at the 1 h horizon. As the prediction horizon increases, PB performance drops sharply (Figure 6), which is consistent with the SHAP results (Figure 10c–f). This is because some high-visibility values (red dots) appear on the positive SHAP side, indicating that current visibility becomes less reliable as a standalone predictor at longer prediction horizons.

The weather flag is another interesting feature. At the 1 h prediction horizon (Figure 10a,b), a clear separation is shown between low and high weather-flag values. The high values (red dots) represent the presence of weather codes such as HZ, FG, RA, and TSRA (Figure 2b), which indicate visibility-reducing atmospheric conditions. These red dots are mainly located on the positive SHAP side, meaning that the presence of these weather phenomena increases the predicted probability of low visibility. In contrast, low weather-flag values (blue dots) are mostly clustered near zero, suggesting a weaker contribution when no significant weather condition is reported. However, at the 3- and 6-h predictions (Figure 10c–f), some of the red dots appear on the negative SHAP side, indicating that the same weather codes do not always lead to future low-visibility predictions at longer horizons. As a result, the predictive effect of the weather flag becomes less consistent as the forecast horizon increases. Therefore, the weather flag provides useful categorical information for detecting low-visibility events, especially for short-term prediction at the 1 h horizon.

The Rx max and Rx min features have smaller SHAP spreads than visibility does, but they still provide useful information for the predictions. Low Rx max values show a strong relationship with low-visibility prediction for the 1 h and 3 h prediction horizons at both airports (Figure 10a–d). However, the Rx min feature shows a more mixed pattern, suggesting that its contribution is less consistent across prediction horizons and sites. The SHAP results indicate that Rx levels can serve as useful indicators for visibility classification, particularly the Rx max feature.

Other features also contributed to the predictions. For example, month-related features, humidity-related indicators, and cloud-base features tended to show different distributions of SHAP results.

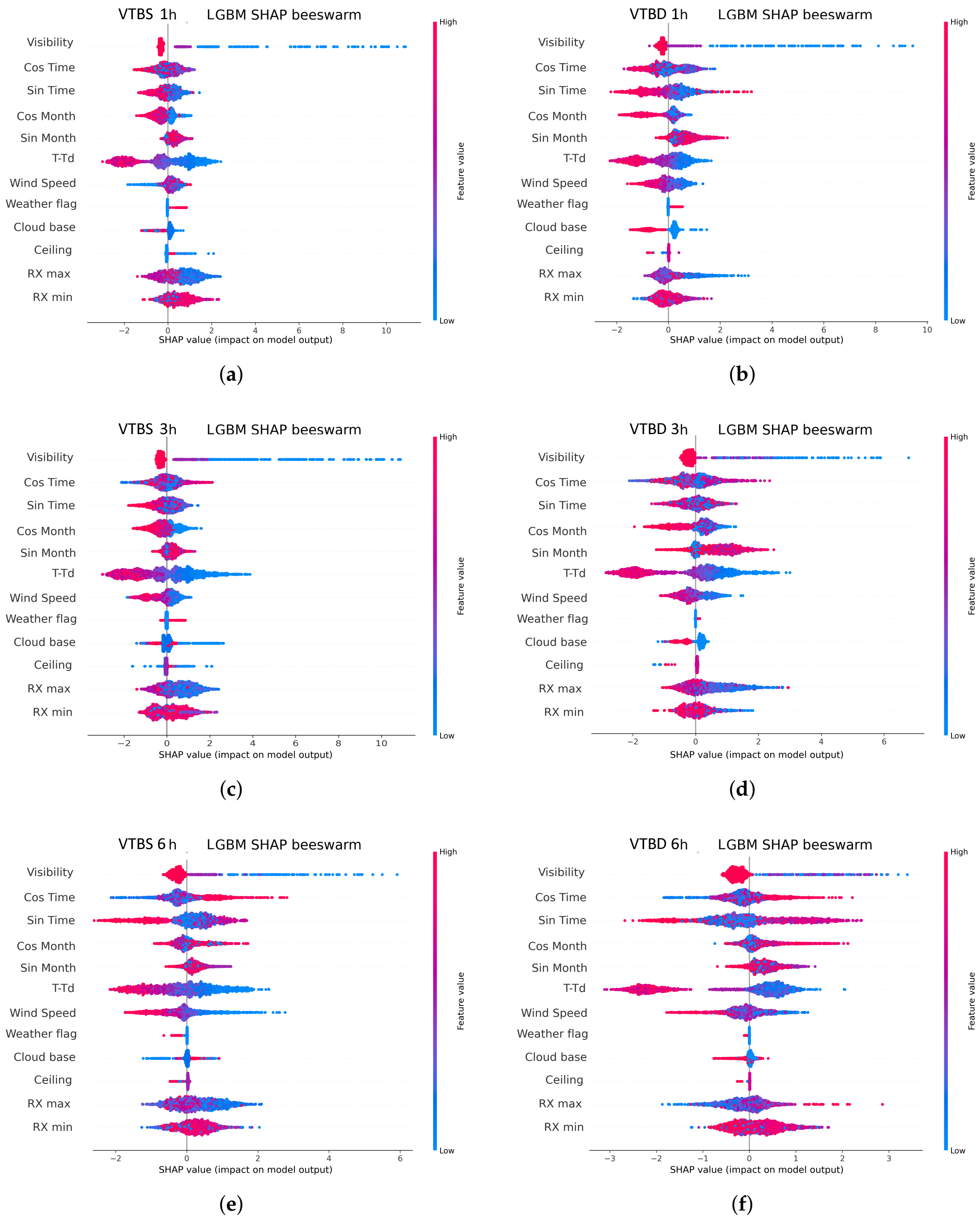


Figure 10. SHAP-value beeswarm distributions for the low-visibility prediction class in the binary classification models: (**a**) VTBS 1 h; (**b**) VTBD 1 h; (**c**) VTBS 3 h; (**d**) VTBD 3 h; (**e**) VTBS 6 h; (**f**) VTBD 6 h.

4.2.4. Feature Group Ablation Analysis

To confirm the feature contributions, a feature-group ablation analysis was conducted by removing selected feature groups from the METAR-Rx feature set and comparing the resulting model's performance with that of the full-feature model. The ablation analysis was conducted over 100 bootstrap rounds to provide a more robust comparison. The results are summarised in Table 3.

The feature ablation table shows that the Rx, time, weather, temperature, and visibility feature groups contribute differently to model performance. When these feature groups were removed, the F1 score generally decreased, indicating that these variables provide useful information for low-visibility prediction. The strongest degradation was observed after removing the visibility feature group. For example, at VTBS, removing the visibility feature reduced the 1 h F1 score from 0.6879 to 0.5331. A similar pattern was observed at VTBD, where the 1 h F1 score decreased from 0.6034 to 0.4674 after removing the visibility feature. These results are consistent with the SHAP analysis in Section 4.2.3, which identified recent visibility information as an important predictor of short-term low-visibility classification.

In some cases, however, the classification metrics increased after removing certain feature groups. For example, removing the Rx group increased recall at several horizons, including VTBS at 3 h (from 0.6559 to 0.6889) and VTBD at 3 h (from 0.5394 to 0.6340). Similar behaviour was also observed when the visibility group was removed at some prediction horizons, such as VTBS at 6 h, where recall increased from 0.3885 to 0.5579. This suggests that removing some features may make the model more sensitive to low-visibility events, but this often occurs at the expense of precision and F1 score. For example, for VTBD at 3 h, removing the Rx group increased recall but reduced precision from 0.4169 to 0.3458 and reduced the F1 score from 0.4670 to 0.4446.

Table 3. Feature-group ablation results for the LGBM-based low-visibility prediction model.

Site	Horizon	Metric	Full Features	Without Rx Feature Group	Without Time Group	Without Weather Flag	Without Temperature Group	Without Visibility Feature
VTBS	1 h	F1	0.6879	0.6720	0.6517	0.6756	0.6848	0.5331
		Accuracy	0.9877	0.9861	0.9862	0.9869	0.9881	0.9764
		Precision	0.5879	0.5505	0.5557	0.5686	0.6061	0.3995
		Recall	0.8329	0.8685	0.7922	0.8369	0.7898	0.8205
	3 h	F1	0.5067	0.4657	0.3889	0.5137	0.4758	0.3857
		Accuracy	0.9793	0.9743	0.9699	0.9797	0.9798	0.9631
		Precision	0.4202	0.3550	0.2956	0.4269	0.4160	0.2743
		Recall	0.6559	0.6889	0.5889	0.6594	0.5705	0.7066
	6 h	F1	0.3312	0.3032	0.2027	0.3286	0.2441	0.3201
		Accuracy	0.9750	0.9682	0.9582	0.9754	0.9731	0.9625
		Precision	0.3001	0.2393	0.1472	0.2989	0.2321	0.2322
		Recall	0.3885	0.4328	0.3444	0.3841	0.2705	0.5579
VTBD	1 h	F1	0.6034	0.5801	0.5439	0.5986	0.6029	0.4674
		Accuracy	0.9813	0.9793	0.9778	0.9809	0.9822	0.9702
		Precision	0.4935	0.4618	0.4354	0.4874	0.5099	0.3508
		Recall	0.7782	0.7823	0.7261	0.7775	0.7407	0.7120
	3 h	F1	0.4670	0.4446	0.3686	0.4657	0.3787	0.3569
		Accuracy	0.9779	0.9715	0.9704	0.9778	0.9742	0.9651
		Precision	0.4169	0.3458	0.3036	0.4158	0.3359	0.2740
		Recall	0.5394	0.6340	0.4837	0.5382	0.4426	0.5357
	6 h	F1	0.2885	0.2618	0.1039	0.2865	0.2420	0.1802
		Accuracy	0.9740	0.9652	0.9549	0.9738	0.9725	0.9506
		Precision	0.2744	0.2065	0.0781	0.2735	0.2347	0.1302
		Recall	0.3222	0.3766	0.1674	0.3203	0.2646	0.3255

Green shading indicates ablation results that are higher than the corresponding full-feature result for the same site, prediction horizon, and metric.

The Rx feature group had a smaller effect than the visibility feature group, but it still improved overall classification performance. For VTBS, adding the Rx feature group increased the F1 score by 1.59, 4.10, and 2.80 percentage points at the 1 h, 3 h, and 6 h horizons, respectively. For VTBD, the corresponding F1 score improvements were 2.33, 2.24, and 2.67 percentage points, respectively. The results indicate a trade-off between recall and precision. When the Rx features were included, precision generally increased, whereas recall decreased. However, the F1 and accuracy scores generally improved, suggesting that the Rx features enhanced the overall classification balance. This result should be interpreted carefully because the Rx contribution was not uniform across all metrics. Nevertheless, the Rx features can be seamlessly integrated with the METAR dataset.

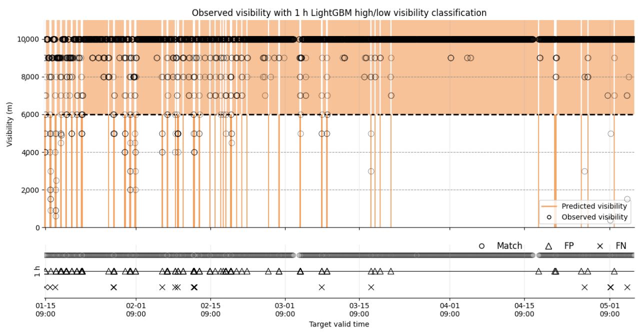
To summarise the table, the full-feature LGBM model achieved the best F1 performance in most cases. The ablation results show that the visibility feature group had a strong influence on the predictions, particularly at the 1 h and 3 h horizons, while the time-related feature group became more influential at the 6 h horizon. Removing the Rx feature group reduced the F1 score across all prediction horizons for both sites. These results suggest that Rx features provide complementary information to the METAR dataset and can improve the F1 performance of low-visibility classification.

4.3. Event-Based Analysis

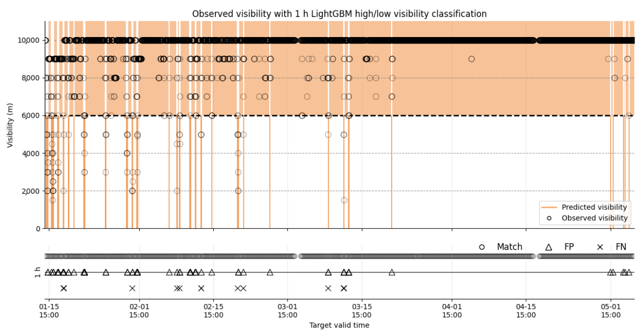
This section presents the event-based evaluation of the LGBM model for 1-h-ahead low-visibility classification. The model used the combined METAR-Rx feature set and was trained without bootstrapping, using observations from 4 November 2024 to 12 January 2026. The test period covers 13 January 2026 to 5 May 2026 for both airports. To provide a more detailed interpretation of the classification behaviour, the analysis further incorporates seasonal subset evaluation and visibility-trend-based analysis.

4.3.1. Full Test-Period Classification Prediction Performance

Figure 11 shows the full test period 1-h-ahead classification results. Observed METAR visibility is represented by circular markers, and the orange vertical lines indicate the predicted visibility cases. The dashed horizontal line denotes the 6000 m visibility classification threshold. The lower panel shows the corresponding classification outcomes, with circles representing correct classifications, triangles representing FP cases, and crosses representing FN cases.



(**a**)



(**b**)

Figure 11. Full test-period result of the 1 h ahead low-visibility classification using LGBM: (**a**) VTBS; (**b**) VTBD.

In Figure 11a, at VTBS, the model correctly detected 73 low-visibility events (TP = 73) and missed 12 events (FN = 12), yielding a recall of 85.88%. Although the model produced 58 FP cases, the precision was 55.73%. This result indicates that the model effectively captured most low-visibility events, although some false alarms were still generated. The F1 score was 0.6759, outperforming both the PB (0.65) and the TAF baseline (0.41).

In Figure 11b, at VTBD, the model obtained TP = 78, FP = 96, and FN = 17, yielding a recall of 82.11%, precision of 44.83%, and F1 score of 0.58. Although these scores were lower than those of VTBS, the F1 score remained higher than both the PB baseline (0.53) and the TAF baseline (0.40).

These two full-test results correspond to the 1 h horizon in the bootstrap results shown in Table 3, where the model achieved the highest F1 score. The model improved 1 h-ahead low-visibility classification performance compared with the operational baselines at both airports. Table 4 summarises the full-test-period classification performance of the model for the 1 h, 3 h, and 6 h prediction horizons.

The full test results showed that the LGBM with the METAR-Rx feature set achieved the best performance at the 1 h horizon for both airports. The model achieved high recall values above 82% at both sites, while the precision remained moderate, ranging from 44.83% to 55.73%. These results suggest that false alarms remain a key limitation under the highly imbalanced visibility conditions in this experimental setup.

Table 4. Test-period performance of 1 h, 3 h, and 6 h low-visibility classification using LGBM.

Airport	Horizon	Accuracy	Precision	Recall	F1
VTBS	1 h	0.9866	0.5573	0.8588	0.6759
	3 h	0.9776	0.3933	0.6941	0.5021
	6 h	0.9726	0.3040	0.5294	0.3862
VTBD	1 h	0.9783	0.4483	0.8211	0.5799
	3 h	0.9741	0.3624	0.5744	0.4444
	6 h	0.9705	0.2695	0.4270	0.3304

As the prediction horizon increased, the classification performance decreased. At the 3 h horizon, the F1 score dropped to 0.5021 for VTBS and 0.4444 for VTBD. The recall also decreased to 0.6941 and 0.5744, respectively. Nevertheless, the model still outperformed both the PB and TAF baselines at this horizon. At the 6 h horizon, however, performance further decreased, with F1 scores of 0.3862 for VTBS and 0.3304 for VTBD. These scores were lower than the TAF baselines, indicating that the model became less reliable as the prediction lead time increased.

This trend suggests that the model's predictive advantage is strongest for short-term forecasting, particularly at the 1 h and 3 h horizons, whereas the 6 h horizon remains more challenging in this setup.

4.3.2. Seasonal Prediction Performance

Seasonal subset evaluation was conducted because visibility may decrease under different atmospheric conditions throughout the year (Figure 2b). In the cool season (November–February), low visibility is more likely to be associated with fog and mist, whereas in the rainy season, it is more commonly related to rainfall. Therefore, the model was tested separately on seasonal subsets to further examine its classification behaviour under different atmospheric conditions.

(a) Cool-Season Subset Evaluation (November–February)

The cool-season subset was evaluated from November to February, when low-visibility events are more likely to involve fog, mist, and haze. For VTBS, the subset contained

11,365 samples, comprising 9092 training and 2273 test samples. The low-visibility base rate in the test set was 3.56%. For VTBD, the subset contained 11,323 samples, consisting of 9058 training and 2265 test samples, with a test-set low-visibility base rate of 3.88%. The results are summarised in Table 5.

Table 5. Cool-season low-visibility classification performance using LGBM.

Airport	Horizon	Accuracy	Precision	Recall	F1
VTBS	1 h	0.9722	0.5692	0.9135	0.7014
	3 h	0.9555	0.4295	0.7530	0.5471
	6 h	0.9414	0.2968	0.4691	0.3636
VTBD	1 h	0.9611	0.5754	0.8636	0.6333
	3 h	0.9417	0.3764	0.7613	0.5037
	6 h	0.9245	0.2935	0.6705	0.4083

Green shading indicates scores higher than the full test-period results in Table 4.

The cool-season subset evaluation mostly yielded higher scores than the full-test-period evaluation for the 1 h and 3 h prediction horizons at both airports (Green highlight). At the 1 h horizon, the F1 score increased from 0.6759 to 0.7014 for VTBS and from 0.5799 to 0.6333 for VTBD. Similarly, at the 3 h horizon, the F1 score increased from 0.5021 to 0.5471 for VTBS and from 0.4444 to 0.5037 for VTBD. These results indicate that the model performed better under cool-season conditions, particularly for short-term low-visibility classification.

However, the improvement was not consistent at the 6 h horizon. For VTBS, the F1 score decreased from 0.3862 in the full-test-period evaluation to 0.3636. In contrast, VTBD improved from 0.3304 to 0.4083, with higher precision, recall, and F1 score. This suggests that the benefit of seasonal filtering was site-dependent at longer prediction horizons.

These results suggest that the predictor features were more meteorologically consistent during this period. Fog-related visibility reduction is associated with temperature, dew point, relative humidity, wind conditions, cloud base, and received signal levels. Therefore, the model can distinguish low-visibility cases more effectively. In addition, the cool-season test set had a higher low-visibility base rate than the full test-period data, thereby increasing the representation of low-visibility cases. Nevertheless, the 6 h horizon remained challenging because visibility becomes more uncertain as the prediction lead time increases.

(b) Rainy-Season Subset Evaluation (May–September)

The rainy-season subset was evaluated from May to September, when visibility reductions are more likely to be associated with rainfall. For VTBS, the subset contained 7581 samples, comprising 6064 training and 1517 test samples, with a test-set low-visibility base rate of 1.38%. For VTBD, the subset also contained 7581 samples, comprising 6064 training and 1517 test samples, with a test-set low-visibility base rate of 1.12%. The results are summarised in Table 6.

Table 6. Rainy-season low-visibility classification performance using LGBM.

Airport	Horizon	Accuracy	Precision	Recall	F1
VTBS	1 h	0.9703	0.0714	0.0952	0.0816
	3 h	0.9775	0.0000	0.0000	0.0000
	6 h	0.9802	0.0000	0.0000	0.0000
VTBD	1 h	0.9854	0.1428	0.0588	0.0833
	3 h	0.9835	0.1667	0.1176	0.1379
	6 h	0.9795	0.0625	0.0588	0.0606

Green shading indicates scores higher than the full test-period results in Table 4.

The rainy-season subset showed extremely weak classification performance at both airports. Although the accuracy remained high, this was mainly due to the dominance of normal-visibility samples rather than to successful low-visibility detection. The precision, recall, and F1 scores were very low, indicating that the model showed limited ability to detect low-visibility events. This weak performance may be explained by the rarity of low-visibility cases during the rainy season and the less consistent relationship between the available predictors and the reduction in rain-related visibility. Unlike cool-season fog-related visibility reduction, which is more closely associated with temperature, dew point, and stable wind conditions, rain-related visibility reduction can occur more suddenly and irregularly. Moreover, the small number of low-visibility cases limited the model's ability to learn reliable low-visibility patterns during this season.

4.3.3. Visibility Trend and Event-Level Prediction Analysis

Visibility trend information was further considered as an additional temporal predictor to be integrated into the METAR-Rx feature set. Motivated by Ref. [62], which showed that smoothed temporal trend information improved the F1 score of an XGB-based fog nowcasting model, this study integrated visibility-trend features into the LGBM model. The aim was to examine whether the short-term visibility trend could further improve low-visibility classification performance.

Table 7 presents the effect of visibility-trend features on the classification performance for VTBS and VTBD. The results show that adding visibility-trend information generally improved accuracy, precision, and F1 score across most prediction horizons, although recall tended to decrease. The 1 h visibility-trend window improved the F1 score for all prediction horizons at both airports compared with the original full-test-period results. It also achieved the highest recall among the visibility-trend windows. However, the 6 h visibility-trend windows tended to yield the highest accuracy and precision scores. From these results, adding the visibility-trend feature can improve the F1 score. The 1 h visibility-trend window is recommended for integration with the LGBM model using the METAR-Rx feature set, as it requires less additional computation while still providing an acceptable improvement in low-visibility classification performance.

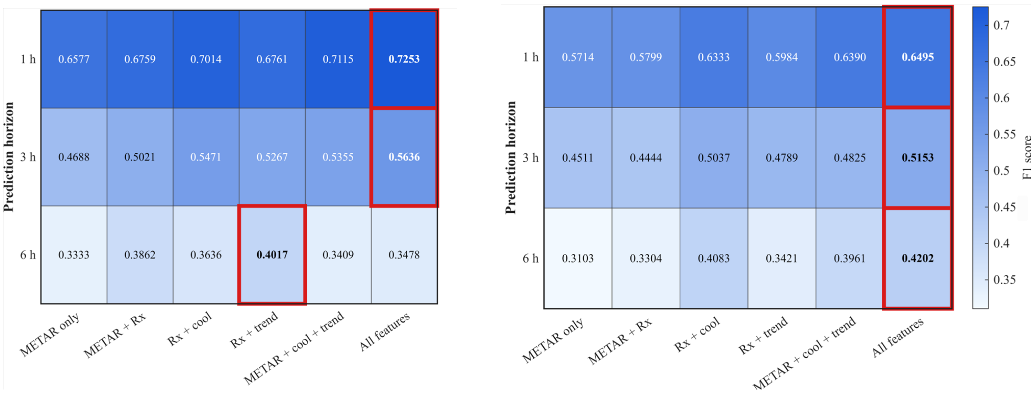
Table 7. Classification performance using visibility-trend features with different trend windows.

Airport	Prediction Horizon	Visibility-Trend Window	Accuracy	Precision	Recall	F1
VTBS	1 h	1 h	**0.9870**	**0.5680**	**0.8352**	**0.6761**
		3 h	0.9866	0.5600	0.8235	0.6667
		6 h	0.9862	0.5511	0.8235	0.6603
	3 h	1 h	0.9797	0.4244	**0.6941**	0.5267
		3 h	0.9799	0.4264	0.6823	0.5248
		6 h	**0.9807**	**0.4393**	0.6823	**0.5345**
	6 h	1 h	0.9738	0.3194	**0.5411**	0.4017
		3 h	0.9759	0.3410	0.5176	**0.4112**
		6 h	**0.9772**	**0.3534**	0.4823	0.4079
VTBD	1 h	1 h	0.9804	0.4779	**0.8000**	0.5984
		3 h	**0.9817**	**0.5000**	0.7894	**0.6122**
		6 h	0.9816	0.4966	0.7789	0.6065
	3 h	1 h	0.9762	0.3958	**0.6063**	0.4789
		3 h	0.9762	0.3928	0.5851	0.4701
		6 h	**0.9777**	**0.4179**	0.5957	**0.4912**
	6 h	1 h	0.9712	0.2805	**0.4382**	0.3421
		3 h	0.9747	0.3193	0.4269	**0.3653**
		6 h	**0.9752**	**0.3245**	0.4157	0.3645

Green shading indicates scores higher than the corresponding full-feature LGBM results in Table 4; bold values indicate the highest score within each airport–horizon group.

4.3.4. Combination of Cool-Season and Visibility-Trend Evaluation

To examine whether the effects of Rx features, seasonal filtering, and visibility-trend information were complementary, a combined evaluation was conducted. Figure 12 compares the F1 scores obtained using different feature and subset strategies for both airports. The comparison includes the METAR-only model, the METAR-Rx model, the cool-season subset model, the visibility-trend model, and the combined model.



(**a**) (**b**)

Figure 12. Heatmap comparison of F1 scores of different feature and subset strategies: (**a**) VTBS; (**b**) VTBD. Red boxes mark the highest F1 score for each prediction horizon.

The results show that combining the METAR-Rx feature set, cool-season filtering, and the 1 h visibility-trend feature produced the highest F1 scores, particularly at the 1 h and 3 h horizons for both airports. Although the Rx features alone provided only a modest improvement, its contribution became more meaningful when combined with season-specific filtering and short-term visibility-trend information. These findings support the use of Rx signal levels as complementary predictors for short-term low-visibility classification.

4.4. Implications of Visibility Prediction for WOC Link Operation

A sufficient link margin is vital for WOC link availability. However, atmospheric attenuation reduces the received optical power, thereby reducing the remaining link margin. The Kim attenuation model can estimate visibility-dependent atmospheric attenuation, which can then be used to determine the remaining link margin. This section demonstrates how short-term visibility prediction can provide an early indication of atmospheric loss in a WOC link.

Assuming a 10 km WOC link operating at a wavelength of 1550 nm with an initial link margin of 2 dB, atmospheric attenuation can be estimated from visibility using the Kim model. The corresponding visibility-dependent path loss and remaining link margin are summarised in Table 8.

As shown in Table 8, atmospheric path loss increases nonlinearly as visibility decreases, particularly below 6 km. For example, the path loss increases from 1.69 dB at 6 km visibility to 9.87 dB at 2 km and 23.29 dB at 1 km. At 5 km visibility, the path loss reaches 2.40 dB and exceeds the assumed 2 dB link margin, causing the link to become unavailable. If the additional atmospheric loss exceeds the available link margin, the WOC link may become unavailable. Longer transmission distances would further increase the total visibility-dependent loss.

According to Figure 12a, the proposed model achieved F1 scores of 0.7253 and 0.5636 for the 1 h and 3 h prediction horizons at VTBS, respectively. Therefore, short-term visibility prediction could provide advance warning of increased WOC attenuation. Such predictions could support proactive link-management actions, including reducing the modulation order, switching to a more robust transmission mode, activating an RF backup link, or repositioning a movable relay.

Table 8. Visibility-dependent atmospheric attenuation and remaining link margin for a 10 km WOC link operating at 1550 nm with an initial link margin of 2 dB.

Visibility (km)	Kim Parameter	Attenuation (dB/km)	Path Loss (dB)	Remaining Margin (dB)	Link State
10	1.30	0.102	1.02	0.98	Available
9	1.30	0.113	1.13	0.87	Available
8	1.30	0.127	1.27	0.73	Available
7	1.30	0.145	1.45	0.55	Available
6	1.30	0.169	1.69	0.31	Available
5	1.14	0.240	2.40	−0.40	Outage
4	0.98	0.354	3.54	−1.54	Outage
3	0.82	0.557	5.57	−3.57	Outage
2	0.66	0.987	9.87	−7.87	Outage
1	0.50	2.329	23.29	−21.29	Outage

Under the assumed link conditions, the link remains available at 6 km visibility, with a remaining margin of 0.31 dB, but becomes unavailable below 6 km. According to Figure 2a, low-visibility observations below 6 km account for approximately 2% of the 2025 VTBS observations. These observations represent potential outage-risk periods in which protective link-management actions could be initiated. Therefore, if all observations below 6 km are treated as visibility-induced outage periods, the visibility-only link availability would be approximately 98%. Additional channel impairments would reduce the overall link availability further. Advance warning from the proposed model, which achieved an F1 score of 0.7253 at the 1 h prediction horizon, could enable the system to take proactive link-management actions and potentially maintain service continuity.

5. Discussion

A short-term prediction framework for classifying low-visibility below 6000 m for Bangkok airports, using the METAR-Rx feature set and tested with multiple ML algorithms, was proposed. The PB and TAF were used as operational baselines for performance comparison, as measured by the confusion matrix and F1 score.

The data were gathered from VTBS and VTBD over 18 months, from 4 November 2024 to 5 May 2026. The Python packages requests and PySNMP were used to automate the data collection. Missing predictor values were replaced with the median value from the training set. The dataset was extremely imbalanced because low-visibility events occurred at a rate of around 2%. The data were split into 80% for training and 20% for testing the ML algorithms, including LR, RF, XGB, and LGBM. The performance was evaluated using F1 score, precision, and recall.

The fixed 0.50 decision threshold used for the ML algorithms could instead be selected using a time-ordered validation procedure. For example, the data could be split chronologically into 64% for training, 16% for validation, and 20% for testing, with the most appropriate threshold selected using the validation set. However, because the dataset covers only 18 months, a single validation period may not represent the full range of seasonal visibility conditions. For example, if the validation period were dominated by rainy-season observations, the selected threshold might not generalise well to other seasonal conditions, particularly given the weak rainy-season performance reported in Section 4.3.2. This validation effect cannot be confirmed from the current dataset and would require a larger multi-year dataset. Therefore, the 0.50 threshold was retained as a common reference operating point. As shown in Appendix A, the F1 response varies across thresholds, ML algorithms, airports, and prediction horizons. Therefore, the operational threshold should be tuned separately for each case.

In the 100-round bootstrap evaluation, the LGBM algorithm with the METAR-Rx feature set achieved the highest F1 scores among the tested ML algorithms at the 1 h and 3 h prediction horizons. This finding is consistent with Ref. [28], which reported that LGBM outperformed other ML algorithms in a Bangkok airport scenario. However, at the 6 h horizon for VTBS, RF achieved a higher F1 score than LGBM, whereas at the 6 h horizon for VTBD, LGBM still produced the highest F1 score. Therefore, LGBM with the METAR-Rx feature set is considered the most suitable ML algorithm in this study for short-term (1 h and 3 h) low-visibility classification.

Moreover, the computational-cost analysis shows that the proposed framework is unlikely to be limited by ML model execution time. The evaluated models required less than 3 s for training, while the single-sample inference latency remained below 53 ms in the tested laptop environment. Since the METAR–Rx inputs are updated at 30-min intervals, these processing times are small enough for near-real-time low-visibility warnings. Therefore, the practical deployment limitations are more likely to come from data acquisition, database access, and system integration rather than from the ML model computation itself.

At the 6 h prediction horizon, the selected ML algorithms underperformed compared with the TAF baseline. This suggests that the selected features are more suitable for short-term prediction than for longer-horizon prediction. As the prediction horizon increases, the relationship between current observations and future visibility becomes weaker. In contrast, TAF may perform better because it includes operational meteorological knowledge and human interpretation of expected weather transitions. Future work could address this limitation by using sequence-based algorithms, such as LSTM, GRU, or Transformer models, which may better capture environmental dependencies over longer time horizons.

The precision–recall performance trade-off is also important. According to Figure 7, the ML algorithms achieved high recall at the 1 h horizon (approximately 74–92%) but only moderate precision (approximately 20–59%). This means that most low-visibility events were detected, but false alarms remained common. For safety-critical applications, this trade-off may still be acceptable because missing a true low-visibility event can be more serious than generating a false alarm.

The METAR-Rx feature set was investigated using bootstrap-based evaluation, feature-importance and feature-contribution analyses. Based on the LGBM feature-importance results, the predictor ranking showed that Rx features were the most important. However, the SHAP beeswarm results showed a different ranking of feature influence. In the SHAP analysis, current visibility was the most significant feature. The difference is because LGBM feature importance reflects how frequently features are used in tree splitting, whereas SHAP explains how each feature contributes to individual predictions. Therefore, these two methods should be interpreted as complementary rather than contradictory.

The feature ablation results confirmed that the current visibility feature had the strongest influence on the prediction model, consistent with the SHAP beeswarm results. When current visibility was removed, the F1 score decreased by approximately 10–15 percentage points in most cases.

Although Rx features improved the classification performance, they should be interpreted with caution. As the microwave link operates in the C-band, the signal is not strongly affected by fog. Therefore, the Rx features may not be directly related to low-visibility events caused by fog. However, the received signal power may decrease during rain because rainfall can affect microwave-link propagation, depending on rainfall intensity. Therefore, Rx is recommended as a complementary propagation-related predictor rather than a direct visibility sensor. This may explain why adding the Rx features provided only a modest improvement, while stronger classification performance was achieved when the Rx features were combined with the METAR variables.

The test results also showed different performance between the two airports. At VTBS, the prediction model achieved better performance than at VTBD. For the LGBM model with the METAR-Rx feature set, the F1 scores at the 1 h prediction horizon were 0.6759 and 0.5799 for VTBS and VTBD, respectively. For the 3 h prediction horizon, the F1 scores were 0.5021 and 0.4444, while for the 6 h prediction horizon, the F1 scores were 0.3862 and 0.3304 for VTBS and VTBD, respectively. Although the two airports are located in the same metropolitan region, approximately 30 km apart, the model performance still differed between the two sites. This difference may be related to local environmental and site-specific conditions. For example, VTBS is located farther from dense urban areas, whereas VTBD is located closer to the city centre, where urban aerosols or PM2.5-related visibility degradation may occur more frequently. These local effects may not be fully captured by the selected METAR and Rx features, leading to lower predictive performance at VTBD. However, this interpretation requires further validation using additional air-quality or aerosol data, such as PM2.5 observations.

Seasonal subset evaluation was conducted to examine the model performance. The results showed that the cool-season subset improved most classification scores, whereas the rainy-season subset performed poorly. This is likely because the predictors were more directly related to fog- and mist-related reductions in visibility, including temperature, dew point, relative humidity, and stable wind conditions. In contrast, during the rainy season, the near-zero F1 score indicates that the model was unreliable. Rain-related visibility reductions can occur suddenly and irregularly, making them difficult to capture with METAR data. In addition, the low-visibility base rate during the rainy season is very low, at around 1%, creating a severe class imbalance. METAR weather descriptors also do not provide detailed rainfall-intensity information. Therefore, future work should include higher-resolution data, such as radar reflectivity, rain rate, disdrometer measurements, or satellite nowcasting features.

The evaluation of visibility-trend features was motivated by Ref. [62]. The results showed that the 1 h visibility-trend window was the most suitable option in this experiment, as it improved the F1 score across prediction horizons for both airports. This suggests that recent visibility evolution provides useful additional information for short-term low-visibility classification.

Overall, the findings indicate that the LGBM with the METAR-Rx feature set is promising for short-term low-visibility classification at Bangkok airports. The model performed best at the 1 h and 3 h prediction horizons. When the 1 h visibility-trend feature and cool-season filtering were added, the F1 scores increased to 0.7253 for VTBS and 0.6495 for VTBD at the 1 h horizon. These results outperformed the traditional TAF baseline, indicating that the proposed approach can serve as a useful complementary tool for operational short-term low-visibility classification.

The current study is limited by the 18-month observation period and the small number of low-visibility events, especially during the rainy season. The analysis also relied primarily on the METAR and Rx features, even though METAR reports may contain observational or coding uncertainty [13]. In addition, the TAF was evaluated using only a Harris-style binary conversion, whereas other operational verification methods, such as the Gerrity Skill Score, could also be considered in future work.

The model's generalisability is limited. The model was trained and tested on Bangkok airport data, which represent a hot, humid tropical climate. Other regions may have different visibility-reduction mechanisms, such as dust storms in desert regions, fog in temperate climates, or snow and ice fog in cold regions. Therefore, the model should be retrained or validated with local data before being applied to other visibility-sensitive applications.

Future research should extend this work by using a longer, multi-year dataset, thereby improving the robustness of conclusions about seasonal and rare-event patterns. Additional datasets, including air quality, PM2.5, liquid water content, ERA5, WRF outputs, satellite data, or closed-circuit television (CCTV) imagery [63], could also be included. Although Rx represents an additional telecommunications-based data source, other airport-based systems, such as radar, AeroMACS, higher-frequency microwave links, and free-space optical links, could be explored to provide additional visibility-related information. Other algorithms, including stacking, ensemble learning, LSTM, GRU, time-series foundation models [64], and Transformer-based models, could also be investigated.

6. Conclusions

Since visibility is essential for WOC reliability, aviation safety, and LiDAR-based sensing performance, short-term visibility prediction is important for mitigating the effects of visibility degradation. This study proposed a short-term low-visibility classification framework for Bangkok airports using METAR and Rx features. Visibility below 6000 m was predicted at the 1 h, 3 h, and 6 h horizons using LR, RF, XGB, and LGBM, with comparisons against the PB and TAF baselines. The results showed that LGBM with the METAR-Rx feature set achieved the highest F1 score, particularly at the 1 h and 3 h horizons. Compared with the PB and TAF baselines, the model performed better at the 1 h and 3 h horizons, whereas the TAF baseline outperformed the ML models at the 6 h horizon. Feature analysis confirmed that current visibility was the dominant predictor, while Rx features provided useful complementary information. Integrating Rx features with METAR observations improved the F1 score by approximately 1–4 percentage points under the 100-round bootstrap evaluation. Seasonal evaluation showed better performance during the cool season, when low-visibility events were more often associated with fog and mist. Adding the 1 h visibility-trend feature further improved the results, with the best combined model achieving F1 scores of 0.7253 for VTBS and 0.6495 for VTBD at the 1 h horizon.

Author Contributions: Conceptualization, S.P.; Validation, S.P. and H.S.; Resources, H.S.; Writing original draft, S.P.; Writing review and editing, S.P. and H.S.; Visualization, S.P.; Supervision, H.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: METAR data are publicly available from the Aviation Weather Center (https://aviationweather.gov/data/metar/ accessed on 24 July 2026). The Rx data, Python v3.12.2 scripts, and MATLAB v25.2 code used in this study are available from the authors upon reasonable request, subject to operational confidentiality constraints.

Acknowledgments: We sincerely thank Supanut Rattanasenee for his contribution to this research project by preparing Python scripts. This research would not be possible without these valuable datasets.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

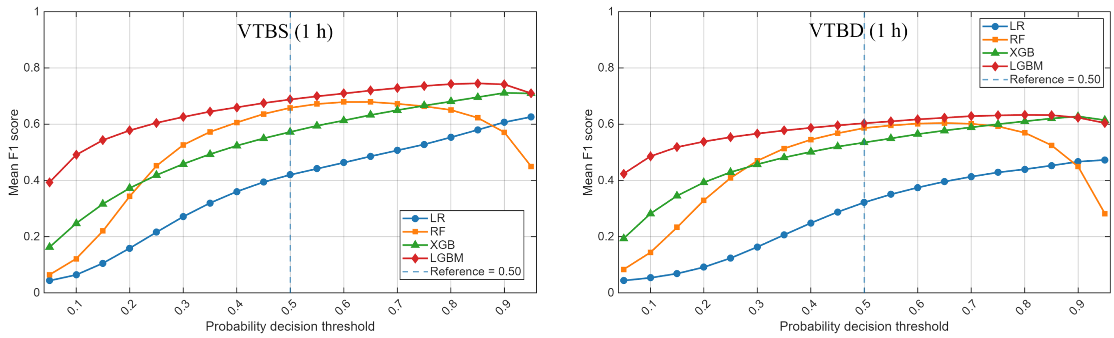
BECMG	Becoming
FN	False negative
FP	False positive
FSO	Free-space optical

GSS	Gerrity Skill Score
IoT	Internet of Things
LGBM	Light gradient-boosting machine
LiDAR	Light detection and ranging
LR	Logistic regression
METAR	Meteorological Aerodrome Report
ML	Machine learning
NOSIG	No significant change
NPV	Negative predictive value
NWP	Numerical weather prediction
PB	Persistence baseline
PROB	probability
RF	Random forest
RVR	Runway visual range
Rx	Received signal
SHAP	Shapley Additive Explanations
SNMP	Simple Network Management Protocol
TAF	Terminal Aerodrome Forecast
TEMPO	Temporary
TN	True negative
TP	True positive
UTC	Coordinated Universal Time
VFR	Visual flight rules
VTBD	Don Mueang Airport
VTBS	Suvarnabhumi Airport
WOC	Wireless optical communication
XGB	Extreme gradient boosting

Appendix A. Threshold Sensitivity Analysis

Appendix A.1. Threshold Sensitivity Report

Figure A1 presents the threshold-sensitivity results for the four ML algorithms across probability thresholds from 0.05 to 0.95, based on 100 bootstrap rounds for each algorithm.



(**a**) (**b**)

Figure A1. *Cont.*

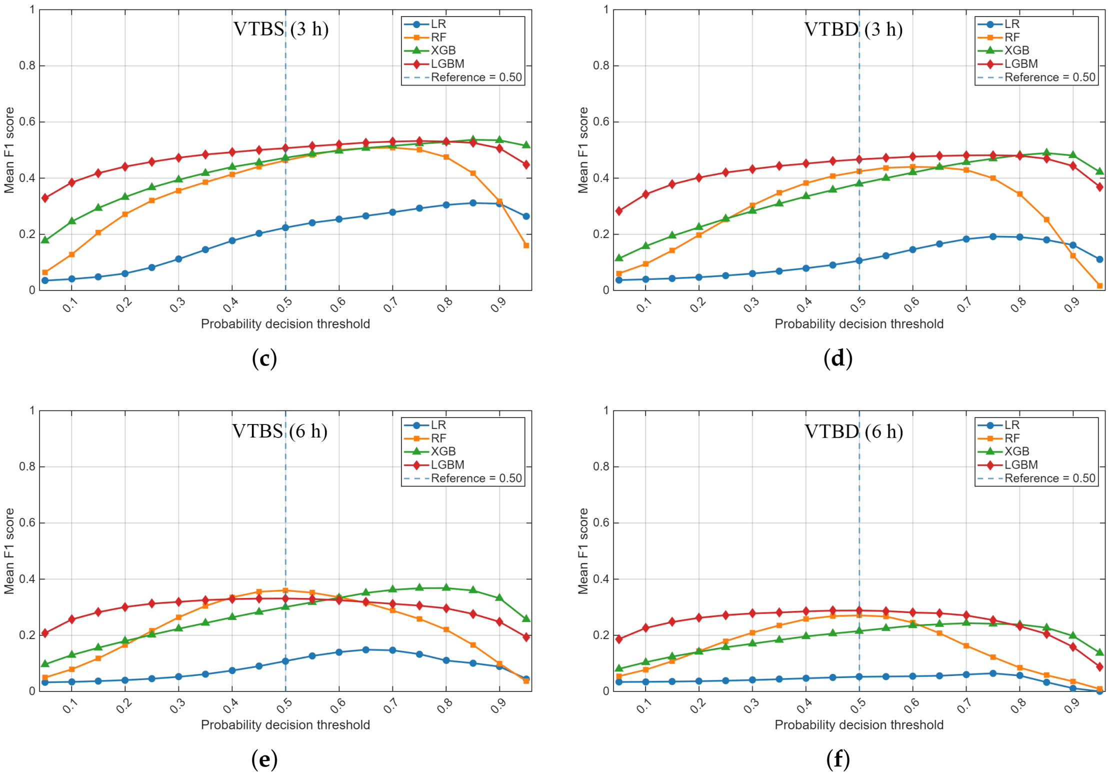


Figure A1. Probability-threshold sensitivity of the low-visibility classification models using the METAR–Rx feature set: (**a**) VTBS at 1 h; (**b**) VTBD at 1 h; (**c**) VTBS at 3 h; (**d**) VTBD at 3 h; (**e**) VTBS at 6 h; and (**f**) VTBD at 6 h.

Appendix A.2. TAF Threshold Sensitivity Results

Table A1 presents the TAF F1 scores across probability thresholds from 0.1 to 0.9 for both airports and all prediction horizons. The 0.40 and 0.50 thresholds show the highest or near-highest F1 scores across all airport–horizon combinations.

Table A1. TAF threshold-sensitivity analysis based on the F1 score.

Site	Horizon	TAF Threshold								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
VTBS	1 h	0.193	0.205	0.199	0.406	0.406	0.411	0.411	0.395	0.395
	3 h	0.177	0.186	0.180	0.375	0.375	0.373	0.373	0.348	0.348
	6 h	0.183	0.189	0.183	0.354	0.354	0.348	0.348	0.329	0.329
VTBD	1 h	0.193	0.211	0.214	0.400	0.400	0.400	0.400	0.378	0.378
	3 h	0.193	0.208	0.211	0.386	0.386	0.384	0.384	0.362	0.362
	6 h	0.175	0.185	0.196	0.332	0.332	0.338	0.338	0.330	0.330

Green shading indicates the highest F1 score among the evaluated TAF thresholds for each site and prediction horizon.

References

1. Nguyen, H.; Al-Imran.; Jang, Y.M. Survey of next-generation optical wireless communication technologies for 6G and Beyond 6G. *ICT Express* **2025**, *11*, 576–589. https://doi.org/10.1016/j.icte.2025.04.006.
2. Krishnamoorthy, A.; Safi, H.; Younus, O.; Kazemi, H.; Osahon, I.N.O.; Liu, M.; Liu, Y.; Babadi, S.; Ahmad, R.; Ihsan, A.; et al. Optical Wireless Communications: Enabling the Next-Generation Network of Networks. *IEEE Veh. Technol. Mag.* **2025**, *20*, 20–39. https://doi.org/10.1109/MVT.2025.3555366.
3. Tan, H.; Fang, C.; Shen, J.; Bhuiyan, Z.A.; Wu, Q.M.J. Cross-Domain Heterogeneous Data Aggregation With Dynamic Group Key Agreement for Hybrid Satellite Networks. *IEEE Trans. Dependable Secur. Comput.* **2026**, *23*, 4830–4844. https://doi.org/10.1109/TDSC.2025.3649054.
4. Govinda Waduge, T.; Yang, Y.; Seet, B.C. A Review of Reconfigurable Intelligent Surfaces in Underwater Wireless Communication: Challenges and Future Directions. *J. Sens. Actuator Netw.* **2025**, *14*, 97. https://doi.org/10.3390/jsan14050097.

5. McCann, A.; Walsh, S.; Maron, N.; Gozzard, D.; McSorley, S.; Schediwy, S. Rapid tactical deployment capability of a transportable optical ground station. *Sci. Rep.* **2025**, *15*, 12026.
6. Sen, M.A.; Al-Rubaye, S.; Tsourdos, A. Securing UAV Flying Ad Hoc Wireless Networks: Authentication Development for Robust Communications. *Sensors* **2025**, *25*, 1194. https://doi.org/10.3390/s25041194.
7. Rattenbury, N.J.; Ashby, J.; Bennet, F.; Birch, M.; Cater, J.E.; Ferguson, K.; Giggenbach, D.; Grant, K.; Knopp, A.; Knopp, M.T.; et al. Update on the German and Australasian Optical Ground Station Networks. *Int. J. Satell. Commun. Netw.* **2025**, *43*, 147–163. https://doi.org/10.1002/sat.1564.
8. Civil Aviation Authority of New Zealand. *Civil Aviation Rules Part 139: Aerodromes—Certification, Operation and Use*; CAA Consolidation: Wellington, New Zealand, 2025.
9. Capraru, R.; Lupu, E.; Wang, J.G.; Soong, B.H. Leveraging Adverse Weather for Enhanced LiDAR Spoofing in Autonomous Driving: Challenges and Opportunities. *IEEE Veh. Technol. Mag.* **2026**, *21*, 27–37. https://doi.org/10.1109/MVT.2025.3648068.
10. Koh, T.Y.; Ng, J.S. Improved diagnostics for NWP verification in the tropics. *J. Geophys. Res. Atmos.* **2009**, *114*, *D12102*. https://doi.org/10.1029/2008JD011179.
11. Kim, M.; Lee, K.; Lee, Y.H. Visibility Data Assimilation and Prediction Using an Observation Network in South Korea. *Pure Appl. Geophys.* **2020**, *177*, 1125.
12. Singh, A.; George, J.P.; Iyengar, G.R. Prediction of fog/visibility over India using NWP Model. *J. Earth Syst. Sci.* **2018**, *127*, 26.
13. Novotny, J.; Dejmal, K.; Repal, V.; Gera, M.; Sladek, D. Assessment of TAF, METAR, and SPECI Reports Based on ICAO ANNEX 3 Regulation. *Atmosphere* **2021**, *12*, 138. https://doi.org/10.3390/atmos12020138.
14. Miao, Y.; Potts, R.; Huang, X.; Elliott, G.; Rivett, R. A fuzzy logic fog forecasting model for Perth Airport. *Pure Appl. Geophys.* **2012**, *169*, 1107–1119.
15. Kneringer, P.; Dietz, S.J.; Mayr, G.J.; Zeileis, A. Probabilistic nowcasting of low-visibility procedure states at Vienna International Airport during cold season. *Pure Appl. Geophys.* **2019**, *176*, 2165–2177.
16. Boneh, T.; Weymouth, G.T.; Newham, P.; Potts, R.; Bally, J.; Nicholson, A.E.; Korb, K.B. Fog forecasting for Melbourne Airport using a Bayesian decision network. *Weather Forecast.* **2015**, *30*, 1218–1233.
17. Hwang, S.H.; Park, S.K.; Park, S.H.; Kwon, K.W.; Im, T.H. RDCP: A Real Time Sea Fog Intensity and Visibility Estimation Algorithm. *J. Mar. Sci. Eng.* **2024**, *12*, 53. https://doi.org/10.3390/jmse12010053.
18. Kim, B.Y.; Cha, J.W.; Chang, K.H.; Lee, C. Visibility Prediction over South Korea Based on Random Forest. *Atmosphere* **2021**, *12*, 552. https://doi.org/10.3390/atmos12050552.
19. Penov, N.; Guerova, G. Sofia Airport Visibility Estimation with Two Machine-Learning Techniques. *Remote Sens.* **2023**, *15*, 4799. https://doi.org/10.3390/rs15194799.
20. Schütz, M.; Schütz, A.; Bendix, J.; und Boris Thies, J.M. Evaluating station, satellite, and combined data for XGBoost-based visibility forecast. *Atmos. Res.* **2026**, *328*, 108395. https://doi.org/10.1016/j.atmosres.2025.108395.
21. Ortega, L.C.; Otero, L.D.; Solomon, M.; Otero, C.E.; Fabregas, A. Deep learning models for visibility forecasting using climatological data. *Int. J. Forecast.* **2023**, *39*, 992–1004. https://doi.org/10.1016/j.ijforecast.2022.03.009.
22. Refaai, A.; Abaza, M.; Abbas, M.A. Artificial Intelligence for Optical Wireless Communication: A Review. *SSRN Electron. J.* 2026, *preprint*. https://doi.org/10.2139/ssrn.6867577.
23. Tefera, M.A.; Yao, C.K.; Peng, C.H.; Chung, Y.J.; Lin, T.C.; Chen, H.M.; Hsu, W.Y.; Peng, P.C. Optimizing Free-Space Optical Communication Based Distributed Temperature Sensing Demodulation Using Meta-Learning. *J. Light. Technol.* **2025**, *43*, 10032–10039. https://doi.org/10.1109/JLT.2025.3605754.
24. Hasegawa, Y.; Kamoi, A.; Aizono, M.; Kato, T. Free-Space Optical Communication System with Wide-Steering Beam for Terrestrial Access Networks. In *Proceedings of the ICC 2024—IEEE International Conference on Communications*; IEEE: New York, NY, USA, 2024; pp. 55–60. https://doi.org/10.1109/ICC51166.2024.10623060.
25. Shao, J.; Liu, Y.; Du, X.; Xie, T. Adaptive Modulation Scheme for Soft-Switching Hybrid FSO/RF Links Based on Machine Learning. *Photonics* **2024**, *11*, 404. https://doi.org/10.3390/photonics11050404.
26. Obeed, M.; Jian, M. Deep-Learning Based Detectors for SISO and SIMO IM/DD FSO Systems. In *Proceedings of the 2024 IEEE Wireless Communications and Networking Conference (WCNC)*; IEEE: New York, NY, USA, 2024; pp. 1–6. https://doi.org/10.1109/WCNC57260.2024.10571014.
27. Azodi, M.; Avazkar, P.; Amirabadi, M.A. Reinforcement Learning-Based FSO Power Adaptation by Deep Computer Vision-Based Weather Classification. *IEEE Trans. Commun.* **2026**, *74*, 754–763. https://doi.org/10.1109/TCOMM.2025.3631646.
28. Aman, N.; Panyametheekul, S.; Sudhibrabha, S.; Pawarmart, I.; Xian, D.; Gao, L.; Tian, L.; Manomaiphiboon, K.; Wang, Y. Estimating visibility and understanding factors influencing its variations at Bangkok airport using machine learning and a game theory–based approach. *Environ. Sci. Pollut. Res.* **2025**, *32*, 31162–31179.
29. Sharma, S.; Bajaj, K.; Deshpande, P.; Bhattacharya, A.; Tripathi, S. Short-Term Fog Forecasting using Meteorological Observations at Airports in North India. In *Proceedings of the 7th Joint International Conference on Data Science & Management of Data (11th*

ACM IKDD CODS and 29th COMAD), CODS-COMAD '24; Association for Computing Machinery: New York, NY, USA, 2024; pp. 307–315. https://doi.org/10.1145/3632410.3632449.

30. Kharisma, A.; Fadhillah, M.; Haryanto, Y.D. Advancing Aviation Meteorology: Airport Visibility Prediction Using Random Forest Regressor On Integrated Metar Parameters. *JIIF J. Ilmu Inov. Fis.* **2025**, *9*, 62–72.
31. Salcedo-Sanz, S.; Guijo-Rubio, D.; Pérez-Aracil, J.; Peláez-Rodríguez, C.; Gomez-Orellana, A.M.; Gutiérrez-Peña, P.A. Artificial Intelligence-Based Methods and Algorithms in Fog and Atmospheric Low-Visibility Forecasting. *Atmosphere* **2025**, *16*, 1073. https://doi.org/10.3390/atmos16091073.
32. Roy, A.; Heinemann, D.; Schroedter-Homscheidt, M.; Lezaca Galeano, J.E. Data-driven combination of METAR observations and CAMS reanalysis aerosols to enhance satellite retrieval of surface solar irradiance. *Sci. Rep.* **2026**, *16*, 6716.
33. Castillo, M.C. Geographic Transferability of Machine Learning Models for Short-Term Airport Fog Forecasting. *arXiv* **2025**, arXiv:2510.21819.
34. Messer, H.; Zinevich, A.; Alpert, P. Environmental Monitoring by Wireless Communication Networks. *Science* **2006**, *312*, 713. https://doi.org/10.1126/science.1120034.
35. Liu, L.; Shu, L.; Xu, Y.; Li, K.; Han, R.; Su, Q.; Fang, J. Rain Detection in Solar Insecticidal Lamp IoTs Systems Based on Multivariate Wireless Signal Feature Learning. *Electronics* **2026**, *15*, 465. https://doi.org/10.3390/electronics15020465.
36. Polz, J.; Chwala, C.; Graf, M.; Kunstmann, H. Rain event detection in commercial microwave link attenuation data using convolutional neural networks. *Atmos. Meas. Tech.* **2020**, *13*, 3835–3853. https://doi.org/10.5194/amt-13-3835-2020.
37. Pudashine, J.; Velasco-Forero, C.; Vale, M.; Zhang, Y.; Frost, A.; Wang, Q.; Trewin, B. Real-time blending of radar, satellite and gauge rainfall data for hydrological applications using Kalman Filter. In *Proceedings of the 2024 Hydrology and Water Resources Symposium (HWRS 2024)*; National Committee on Water Engineering, Engineers: Melbourne, Australia, 2024; pp. 1057–1062.
38. Pudashine, J.; Guyot, A.; Petitjean, F.; Pauwels, V.R.N.; Uijlenhoet, R.; Seed, A.; Prakash, M.; Walker, J.P. Deep Learning for an Improved Prediction of Rainfall Retrievals From Commercial Microwave Links. *Water Resour. Res.* **2020**, *56*, e2019WR026255. https://doi.org/10.1029/2019WR026255.
39. Zheng, L.; Lin, J.; Huang, Z.; Lin, Y.; Zheng, Q.; Chen, Q.; Lin, L.; Chen, J. Cost-Sensitive Rainfall Intensity Prediction with High-Noise Commercial Microwave Link Data. *Sustainability* **2024**, *16*, 8067. https://doi.org/10.3390/su16188067.
40. Perelet, A.O.; Gultepe, I.; Hoch, S.W.; Pardyjak, E.R. Discriminating fog and rain at the kilometre scale using the extinction from collocated infrared and microwave scintillometers. *Bound. Layer Meteorol.* **2021**, *181*, 295–315.
41. Eltahir, E.I.; Elsheikh, E.A.A.; Islam, M.R.; Abdalla, A.H.; Hadi Habaebi, M.; Zyoud, A.; Hamdan, M. Sand and Dust Storm Attenuation Prediction Using Visibility and Humidity Measurements. *IEEE Access* **2024**, *12*, 79602–79612. https://doi.org/10.1109/ACCESS.2024.3409576.
42. Ouedraogo, W.Y.S.B.; Tiemounou, S.; Djibo, M.; Doumounia, A.; Sanou, S.R.; Sawadogo, M.; Guira, I.; Zougmore, F. Application of machine learning methods on climate data and commercial microwave link attenuations for estimating meteorological visibility in dusty condition. *Engineering* **2022**, *14*, 85–93.
43. van Grinsven, I.; Grootes, M.W.; Uijlenhoet, R.; Steeneveld, G.J. Fog Monitoring through Machine Learning of Signal Attenuation Data from Microwave Links from Cellular Communication Networks. *EGUsphere* **2025**, *2025*, 1–36. https://doi.org/10.5194/egusphere-2025-2634.
44. Google Maps. *Satellite Imagery of Bangkok Airports: Suvarnabhumi Airport and Don Mueang Airport*; Online Map; Google Maps: Mountain View, CA, USA, 2026.
45. He, H.; Garcia, E.A. Learning from Imbalanced Data. *IEEE Trans. Knowl. Data Eng.* **2009**, *21*, 1263–1284. https://doi.org/10.1109/TKDE.2008.239.
46. Niaz, N.U.; Shahariar, K.N.; Patwary, M.J.A. Class Imbalance Problems in Machine Learning: A Review of Methods And Future Challenges. In *Proceedings of the 2nd International Conference on Computing Advancements, ICCA '22*; Association for Computing Machinery: New York, NY, USA, 2022; pp. 485–490. https://doi.org/10.1145/3542954.3543024.
47. Kim, J.; Kim, S.H.; Seo, H.W.; Wang, Y.V.; Lee, Y.G. Meteorological characteristics of fog events in Korean smart cities and machine learning based visibility estimation. *Atmos. Res.* **2022**, *275*, 106239. https://doi.org/10.1016/j.atmosres.2022.106239.
48. Boucouvala, D.; McCooey, D. A Verification Procedure for Terminal Aerodrome Forecasts. *Environ. Earth Sci. Proc.* **2025**, *35*, 30. https://doi.org/10.3390/eesp2025035030.
49. Sharpe, M.A.; Bysouth, C.E.; Trueman, M. Towards an improved analysis of Terminal Aerodrome Forecasts. *Meteorol. Appl.* **2016**, *23*, 698–704. https://doi.org/10.1002/met.1593.
50. Harris, G.R. Comparison of different scoring methods for TAFs and other probabilistic forecasts. In Proceedings of the 15th AMS Conference on Probability and Statistics in the Atmospheric Sciences, Asheville, NC, USA, 8–12 May 2000.
51. PySNMP Contributors. *PySNMP: Python SNMP Library*; GitHub Repository: San Francisco, CA, USA, 2026.
52. Kim, I.I.; McArthur, B.; Korevaar, E.J. Comparison of laser beam propagation at 785 nm and 1550 nm in fog and haze for optical wireless communications. In *Proceedings of the Optical Wireless Communications III*; SPIE: Bellingham, WA, USA, 2001; Volume 4214, pp. 26–37.

53. Civil Aviation Authority of New Zealand. Part 91: Subpart D—Visual Flight Rules. 2026. Available online: https://aviation.govt.nz/rules/rule-part/part-91/subpart-d/ (accessed on 24 June 2026).
54. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.Y. Lightgbm: A highly efficient gradient boosting decision tree. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*; Curran Associates Inc.: Red Hook, NY, USA, 2017; Volume 30.
55. Tian, W.; Zhang, C.; Fu, S.; Zhu, F.; Hu, H. Predicting High-Resolution Gridded Sea Ice Concentration by Integrating LightGBM and Kriging Algorithms. *J. Mar. Sci. Eng.* **2026**, *14*, 1092. https://doi.org/10.3390/jmse14121092.
56. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16*; Association for Computing Machinery: New York, NY, USA, 2016; pp. 785–794. https://doi.org/10.1145/2939672.2939785.
57. Kim, B.Y.; Belorid, M.; Cha, J.W. Short-Term Visibility Prediction Using Tree-Based Machine Learning Algorithms and Numerical Weather Prediction Data. *Weather Forecast.* **2022**, *37*, 2263–2274. https://doi.org/10.1175/WAF-D-22-0053.1.
58. Zhang, X.; Wang, Y.; Zhuang, Z.; Liu, Y.; Yuan, C.; Su, L.; Shao, J.; Chan, P.W. Comparison of simulating visibility using XGBoost and IMPROVE method: A case study in East China. *Front. Environ. Sci.* **2025**, *12*, 1534113. https://doi.org/10.3389/fenvs.2024.1534113.
59. Dietz, S.J.; Kneringer, P.; Mayr, G.J.; Zeileis, A. Low-visibility forecasts for different flight planning horizons using tree-based boosting models. *Adv. Stat. Climatol. Meteorol. Oceanogr.* **2019**, *5*, 101–114. https://doi.org/10.5194/ascmo-5-101-2019.
60. Berkowitz, J.; Kilian, L. Recent developments in bootstrapping time series. *Econom. Rev.* **2000**, *19*, 1–48. https://doi.org/10.1080/07474930008800457.
61. Ponce-Bobadilla, A.V.; Schmitt, V.; Maier, C.S.; Mensing, S.; Stodtmann, S. Practical guide to SHAP analysis: Explaining supervised machine learning model predictions in drug development. *Clin. Transl. Sci.* **2024**, *17*, e70056.
62. Schütz, M.; Schütz, A.; Bendix, J.; Thies, B. Improving classification-based nowcasting of radiation fog with machine learning based on filtered and preprocessed temporal data. *Q. J. R. Meteorol. Soc.* **2024**, *150*, 577–596. https://doi.org/10.1002/qj.4619.
63. Bartok, J.; Šišan, P.; Ivica, L.; Bartoková, I.; Malkin Ondík, I.; Gaál, L. Machine Learning-Based Fog Nowcasting for Aviation with the Aid of Camera Observations. *Atmosphere* **2022**, *13*, 1684. https://doi.org/10.3390/atmos13101684.
64. Jetwiriyanon, J.; Susnjak, T.; Ranathunga, S. Generalisation Bounds of Zero-Shot Economic Forecasting Using Time Series Foundation Models. *Mach. Learn. Knowl. Extr.* **2025**, *7*, 135. https://doi.org/10.3390/make7040135.