Signal, Image and Video Processing (2026) 20:170
https://doi.org/10.1007/s11760-026-05144-5

ORIGINAL PAPER

An intelligent hybrid machine learning framework for low-visibility reconstruction for airport transportation

Fang-wei Liu[1] · **Ren-jie Wu**[2,3]

Received: 20 August 2025 / Revised: 5 January 2026 / Accepted: 19 January 2026 / Published online: 10 March 2026

Abstract

Accurate low-visibility prediction and real-time fog detection remain challenging, especially for sudden localized events where current methods often respond slowly, offer limited spatial resolution, and produce frequent false alarms. This study presents a hybrid machine learning framework that integrates video-based fog density estimation (using MVG modeling) with real-time atmospheric observations to bridge spatial and temporal data gaps. The framework includes a data preparation module and a training module combining LSTM and XGBoost. Experimentally, it achieves strong reconstruction performance with a test RMSE of 121.48 m and an R2 of 0.935, improving R2 by 5.67% over other LSTM hybrids. The optimized LSTM–XGBoost model also outperforms both baseline models and unoptimized variants. These results confirm that the framework effectively utilizes video-derived fog density to dynamically calibrate visibility and deliver fast, accurate fog impact reconstruction.

Keywords Low-visibility · Atmospheric factors · LSTM network · Deep learning

1 Introduction

Visibility is crucial for atmospheric monitoring, highway traffic management, and aviation operations. Once visibility drops below safe thresholds, measures such as road closures and flight adjustments are typically implemented to ensure safety. However, conventional monitoring systems like lidar and transmissometers are often costly, provide limited spatial coverage, and struggle to accurately detect sudden or localized fog events [1]. These limitations highlight the need for more reliable and scalable visibility prediction methods to improve transport safety.

Numerous factors such as air pollutants, humidity, temperature, solar radiation, and time of day affect visibility. Recent progress in artificial intelligence and machine learning has spurred new methods for fog and visibility prediction. Early approaches using Bayesian networks [2], Support Vector Regression (SVR), and Extreme Learning Machines (ELM) [3] achieved limited success. Later studies treated the task as regression or classification [4], with evolving neural networks [5] and ordinal classifiers [6] showing potential in specific cases. More recent techniques include LSTMs for capturing temporal patterns, ensemble methods for reliability, and neural networks combined with ELM or ensemble learning using atmospheric data. Emerging approaches like Bayesian ensemble graph attention networks [7] and physics-informed graph transformers [8] also offer insights for modeling complex spatiotemporal systems in visibility prediction. Despite these advances, a major limitation remains: most methods overlook highly localized and transient microclimatic variations, such as in wind and stability, leading to poor generalization in extreme or uneven fog conditions [9]. Video-based methods using dark channel priors complement these approaches but often ignore nonlinear interactions between atmospheric factors and visual obscuration [10]. While hybrid models combining image and atmospheric data show improvement [11], they still cannot adequately capture rapid fog evolution and spatial heterogeneity, especially in patchy fog [12]. Recent data-driven dynamics identification studies [13] further emphasize the need to incorporate physical mechanisms to improve generalization in complex, non-stationary fog environments.

✉ Ren-jie Wu
renji.wu@polyu.edu.hk

[1] Department of Art and Design, Dongguan Polytechnic, Dongguan, China

[2] Shenzhen Research Institute, The Hong Kong Polytechnic University, Hung Hom, Hong Kong, China

[3] College of Civil Engineering and Architecture, Zhejiang University, Hangzhou, China

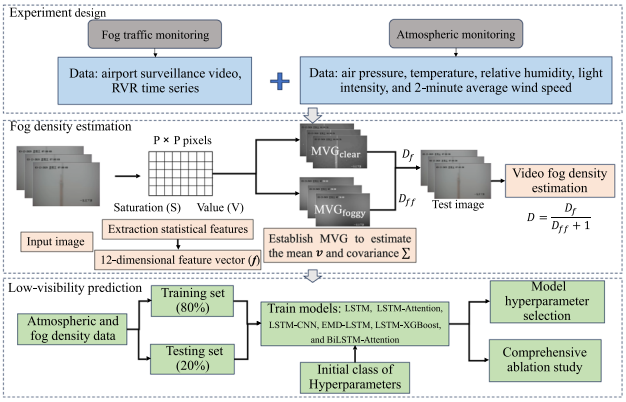

Fig. 1 Flow chart of the proposed intelligent hybrid machine learning framework

To address the aforementioned challenges, this paper proposes a novel two-stage framework. The core innovation lies in a hybrid modeling methodology which systematically combines Bayesian optimization, multi-source data fusion, and an interpretable analytical structure. Methodologically, this study makes three specific advances in Fig. 1. First, it introduces a multi-level data fusion strategy that synergistically merges locally estimated fog density from video sequences with atmospheric time-series data, overcoming the limitations of single-source inputs. Second, instead of a straightforward model combination, the architecture employs Bayesian optimization for coordinated hyperparameter tuning, effectively leveraging the complementary strengths of LSTM for temporal modeling and XGBoost for nonlinear fitting. Third, the framework embeds SHAP-based interpretability to quantitatively assess feature contributions, including those derived from images, thereby enhancing predictive transparency under low-visibility conditions. To validate the contribution of each component, ablation experiments are conducted using six state-of-the-art LSTM-based benchmarks, confirming that the proposed framework improves both predictive accuracy and computational efficiency.

2 Methodology

2.1 Fog density estimation via MVG modeling

The fog density assessment involves three key stages: (1) extracting fog-aware statistical features from the input image, (2) comparing them with pre-trained MVG models of clear and foggy scenes using a distance metric, and (3) fusing the distances into a final fog density index. Image degradation under fog is characterized by reduced color saturation and shifted intensity values. To capture this, saturation and value components are extracted in HSV space and averaged within non-overlapping patches. Multi-scale patching yields a 12-dimensional feature vector, which is log-transformed to stabilize variance and better fit the Gaussian model used in later analysis. The method reduces the influence of uniform lighting changes by relying on relative saturation-value statistics and log normalization. Its robustness under extreme lighting, however, depends on training data that covers varied illumination conditions.

The Multivariate Gaussian (MVG) modeling is used to establish statistical baselines for both clear and foggy conditions. This offline phase uses large, carefully curated datasets to ensure model robustness and generality. Two separate reference MVG models are constructed. The first model, the Fog-Free Image Model ($\mathrm{MVG_{clear}}$), is derived from a corpus of 500 natural fog-free images. Each image in this corpus is processed through the feature extraction pipeline to generate its respective 12-dimensional feature vector. A multivariate Gaussian model is then fitted to this entire collection of feature vectors using conventional maximum likelihood estimation, resulting in a model fully characterized by its mean vector $\boldsymbol{v}_{\mathrm{clear}}$ and its covariance matrix $\sum_{\mathrm{clear}}$. Similarly, a second model, the Foggy Image Model ($\mathrm{MVG_{foggy}}$), is built from a corpus of 500 natural foggy images, yielding a model parameterized by its mean vector $\boldsymbol{v}_{\mathrm{foggy}}$ and covariance matrix $\sum_{\mathrm{foggy}}$. These two models capture the essential statistical characteristics of clear and foggy scenes in the 12-dimensional feature space, enabling effective comparative analysis. The MVG probability density in this space is:

$$MVG(\boldsymbol{f}) = \frac{1}{(2\pi)^{12/2}\left|\sum\right|^{1/2}} \exp[-\frac{1}{2}(\boldsymbol{f}-\boldsymbol{v})^{t}\sum^{-1}(\boldsymbol{f}-\boldsymbol{v})] \tag{1}$$

where $\boldsymbol{v}$ and $\sum$ represent the mean and covariance matrix, respectively, and $\left|\sum\right|$ and $\sum^{-1}$ represent the determinant and inverse of the covariance matrix of the MVG model, and $\boldsymbol{f}$ is the collection of fog-aware statistical characteristics.

The fog density of a test image is estimated through a comparative analysis that measures the statistical deviation of its features from the pre-established offline models. The process is initiated by fitting the test image's own feature distribution; the extracted 12-dimensional feature vector $\boldsymbol{f}$ is used to calculate its sample mean vector $\boldsymbol{v}_{\mathrm{test}}$ and sample covariance matrix $\sum_{\mathrm{test}}$. The heart of the estimation involves calculating a Mahalanobis-like distance metric between this fitted test distribution and the two reference models. The first distance, termed the Fog Density Indicator (D_f), is computed between the test image's feature distribution and the fog-free model ($\mathrm{MVG_{clear}}$).

$$D_f = \sqrt{(\boldsymbol{v}_{\mathrm{test}}-\boldsymbol{v}_{\mathrm{clear}})^{t}(\frac{\sum_{\mathrm{test}}+\sum_{\mathrm{clear}}}{2})^{-1}(\boldsymbol{v}_{\mathrm{test}}-\boldsymbol{v}_{\mathrm{clear}})} \tag{2}$$

where $\sum_{\text{test}}$, $\sum_{\text{clear}}$, and $\boldsymbol{v}_{\text{test}}$, $\boldsymbol{v}_{\text{clear}}$ are the covariance matrices and mean vectors of the MVG fit of the test image and the MVG model of the fog-free corpus, respectively. A larger value of D_f signifies a greater statistical deviation from the characteristics of clear images, thereby suggesting a higher concentration of fog. The second distance, termed the Fog-Free Level Indicator (D_{ff}), is computed between the test image's feature distribution and the foggy image model ($\text{MVG}_{\text{foggy}}$). A larger value of D_{ff} indicates that the image is statistically atypical of dense fog, suggesting a lower fog density or a clearer scene.

$$D_{ff} = \sqrt{(\boldsymbol{v}_{\text{test}} - \boldsymbol{v}_{\text{foggy}})^t \left(\frac{\sum_{\text{test}} + \sum_{\text{foggy}}}{2}\right)^{-1} (\boldsymbol{v}_{\text{test}} - \boldsymbol{v}_{\text{foggy}})} \tag{3}$$

Finally, these two complementary distance measures are fused into a single, normalized fog density score D as follows:

$$D = \frac{D_f}{D_{ff} + 1} \tag{4}$$

Adding the stabilization constant '1' in the denominator ensures numerical robustness by preventing division by a negligible value. Consequently, a lower final value of D indicates a clearer scene. A higher value indicates a scene with denser fog, providing an intuitive and reliable metric for fog density assessment.

2.2 LSTM-based prediction for low-visibility

The accurate prediction of atmospheric visibility presents fundamental challenges arising from nonlinear and nonstationary dynamic interactions among multiple meteorological variables including fog density, humidity, and temperature. To address these complexities, this study develops a comprehensive modeling framework comprising standard LSTM networks [14] and six hybrid architectures: LSTM-Attention, LSTM-CNN [15], EMD-LSTM [16], LSTM-XGBoost [17], and BiLSTM-Attention [18]. These models are based on advanced computational methods with established effectiveness in meteorological forecasting. Each architecture targets specific prediction challenges: attention mechanisms improve severe weather forecasting through temporal weighting, CNN-LSTM hybrids excel in precipitation nowcasting, EMD-LSTM handles nonstationary wind speed prediction, XGBoost-LSTM enhances multi-step visibility accuracy, bidirectional attention networks advance typhoon track prediction, and Bayesian optimization aids parameter tuning for extreme weather forecasting. This study systematically details the design and implementation of these models for airport low-visibility prediction, beginning with the collection of multivariate atmospheric data including temperature, humidity, wind speed, pressure, and estimated fog density. These temporal sequences are segmented into fixed-length samples using a sliding window approach. Each sample is normalized to the [0, 1] range via min–max normalization:

$$x^* = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{5}$$

where x^* is the normalized value, x is the original input, and $x_{\min}$ and $x_{\max}$ denote the minimum and maximum values in the sample data, respectively.

Figure 2 illustrates six neural network architectures for low-visibility modeling. (a) The LSTM model captures temporal dependencies through three gating mechanisms. (b) LSTM-Attention adds a learnable weighting mechanism to highlight meteorologically significant historical states. (c) LSTM-CNN combines convolutional layers for spatial features with LSTM for temporal modeling. (d) EMD-LSTM decomposes data into intrinsic mode functions for separate prediction and reconstruction. (e) LSTM-XGBoost integrates LSTM with gradient boosting (100 cycles, rate 0.05) to enhance regression performance. (f) BiLSTM-Attention employs bidirectional LSTM, multiple fully-connected layers, and dropout, trained with Adam over 50 epochs. These architectures represent diverse approaches to modeling spatiotemporal features in visibility prediction. Evaluating the prediction model's effectiveness and predictive ability is crucial before using it for visibility analysis. The standard metrics for comparing actual and predicted values are the mean absolute error (MAE), variance accounted for (VAF), root mean square error (RMSE), weighted mean absolute percentage error (WMAPE), and coefficient of determination (R^2).

3 Data collection and preprocessing

The datasets used in this study, comprising airport security footage and atmospheric observations. Sourced from the 2020 China Post-Graduate Mathematical Contest in Modeling (https://cpipc.acge.org.cn/cw/hp/4), the atmospheric data was recorded at minute-resolution by an Automated Meteorological Observing System. The data's co-location with visibility measurements from the Rilan Expressway in Shandong Province logically identifies Jining Qufu Airport as the source. The experimental setup employs MATLAB-R 2024a on a Windows 11 system with an Intel Core i7-9750H CPU. The key metric, Runway Visual Range (RVR), is defined as the maximum distance a pilot can identify runway markings or lights, and its calibration adheres to ICAO or CAAC

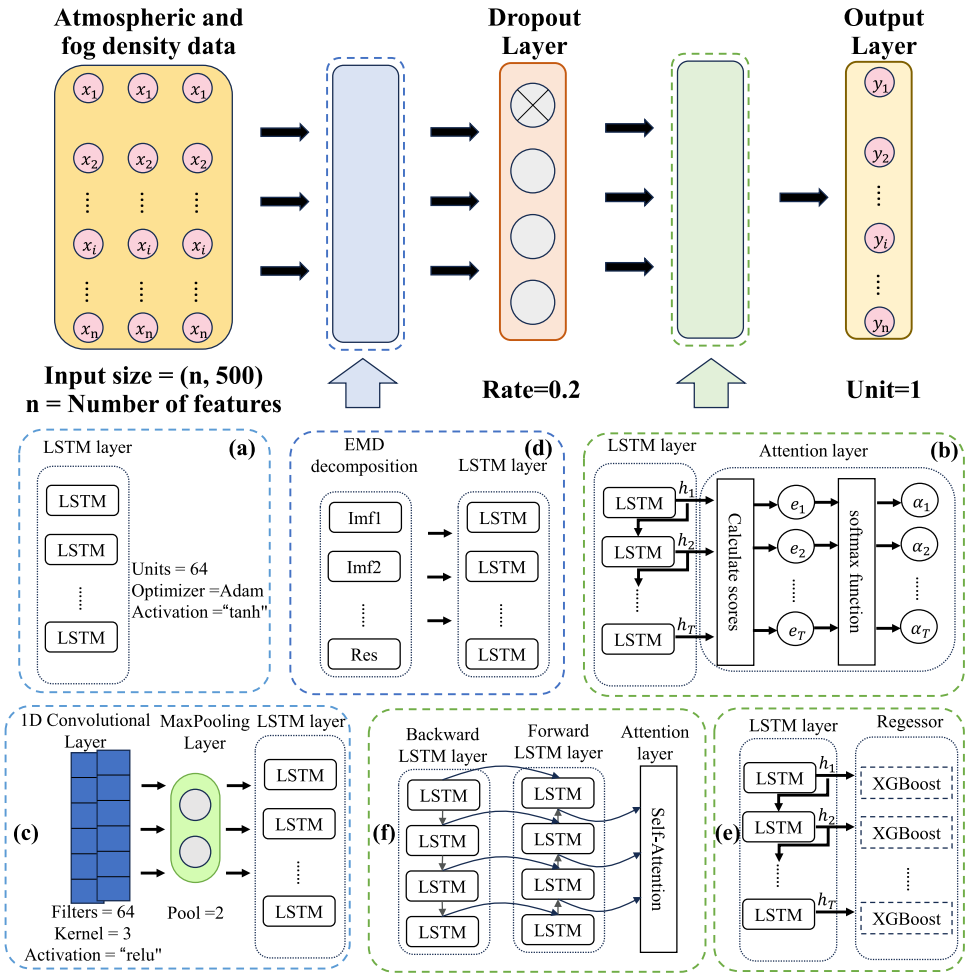

Fig. 2 The network framework diagram for **a** LSTM, **b** LSTM-Attention, **c** LSTM-CNN, **d** EMD-LSTM, **e** LSTM-XGBoost, and **f** BiLSTM-Attention architectures

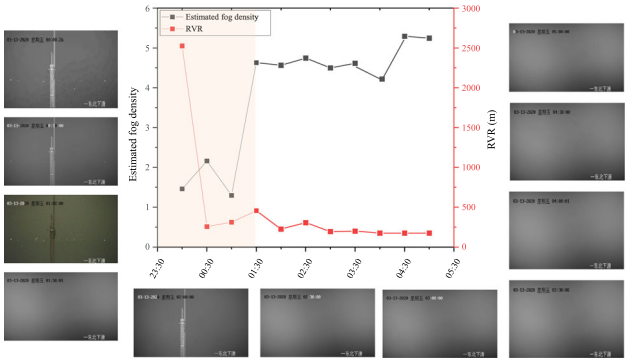

Fig. 3 The comparison results on fog density estimation, its corresponding RVR

standards. These standards govern the conversion of Meteorological Optical Range to RVR, incorporating background light intensity. The video's eleven essential frames are displayed in Fig. 3. The video (25 fps, 1280 × 720 resolution) was sampled at 30-min intervals, yielding 11 key frames for analysis. This uniform sampling captured representative variations in visual attributes such as fog density, text color, and lighting over time. The footage was recorded by a fixed-angle rotating camera at the airport from 00:00:26 to 11:47:48 on March 13, 2020. Analysis focused on the period from 00:00 to 08:00, when temperatures ranged from 8 to 13 °C under foggy conditions. The 500 fog-free and foggy images used for MVG training were confirmed to originate from the same location as the atmospheric data, thereby ensuring spatial consistency within the multimodal dataset. To meet the Gaussian assumption essential for the MVG classifier, a logarithmic transformation was applied to the 12-dimensional feature vector. The efficacy of this normalization was quantitatively verified using the Shapiro–Wilk test, which confirmed that the transformed feature distribution does not significantly deviate from normality ($p > 0.05$), thereby validating the use of the parametric MVG classifier. The estimated fog density values from the corresponding time points were plotted and compared with the RVR visibility trend. The results indicated that the fog density curve generally corresponds to the visibility changes: from 00:00 to 01:30, no fog patches were present, leading to low fog density and high RVR visibility; from 01:30 to 05:00, fog patches formed, resulting in increased fog density and significantly

Table 1 Airport atmospheric observation data

Notation	Description
PAINS (hPa)	Air pressure
TEMP (°C)	Temperature
RH (%)	Relative humidity
LIGHT (cd)	Light intensity
WS2A (MPs)	2-min average wind speed
CW2A (MPs)	2-min average vertical wind speed
WD2A (°)	2-min average wind direction
RVR	Visibility ($\leq$ 3000 m)

reduced visibility. During this period, both curves exhibited noticeable fluctuations, likely due to the influence of other meteorological factors on fog density estimation and RVR visibility measurements. This analysis confirms the correlation between fog density and visibility while highlighting that external weather conditions may affect measurement stability.

4 Results and discussion

4.1 Model verification

In this study, six deep learning models are established for low-visibility prediction based on historical data analysis. The aim is to reconstruct the relationship between known visibility, atmospheric environmental data, and image-based fog concentration. To ensure a fair and rigorous comparison, all models underwent systematic hyperparameter tuning via Bayesian optimization. The dataset was randomly split into a training set (70%) and a test set (30%), with Min-Max normalization applied to mitigate the influence of varying scales and units on prediction performance. For consistency, the architectural parameters of the LSTM components in each baseline model were kept identical to those in the standard LSTM structure. Each model was then trained with its respective set of optimized hyperparameters and evaluated using the same performance metrics, allowing for a direct comparison of their predictive capabilities under equivalently tuned conditions. Table 1 presents all atmospheric observation data, with the wind direction data (WD2A) having undergone a sine transformation. As shown in Fig. 4, all LSTM models converged stably after 500 training iterations. According to Table 2, the hybrid LSTM-XGBoost model performed best, outperforming both individual models. The standalone LSTM achieved an RMSE of 160.82 m and R^2 of 0.89, capturing temporal patterns but lacking feature refinement. The XGBoost-only model attained an RMSE of 143.68 m and R^2 of 0.90, demonstrating strong feature

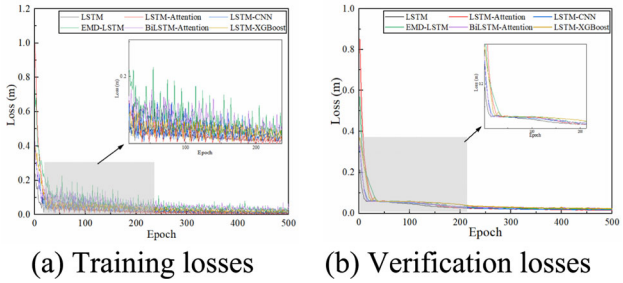

Fig. 4 The comparison of loss curves of deep learning models

optimization without utilizing temporal dependencies. By combining the two methods, the hybrid model achieved an RMSE of 139.36 m and R^2 of 0.92, confirming that their integration enhances both expressiveness and generalization. In addition to superior accuracy, the hybrid model exhibited smoother training with minimal loss fluctuations and faster convergence in just 223 iterations. Under standardized conditions (Intel i7-10750H, RTX 2060, 479 samples, batch size 32), the LSTM-XGBoost model achieved the best performance with an RMSE of 139.36 m and an R^2 of 0.92, outperforming both the standalone LSTM (RMSE 160.82 m) and XGBoost (RMSE 143.68 m) models. Although its inference time of 0.21 s is longer than that of LSTM only (0.04 s), it offers a favorable balance of accuracy and training efficiency, requiring only 2.6 s to train. In contrast, the LSTM-Attention model showed higher error (RMSE 175.08 m, R^2 0.87) and a significantly longer training time of 10.4 s, further validating the practicality of the LSTM-XGBoost hybrid approach.

For an intuitive comparison of predictive performance across models, we selected 75 consecutive samples from the test set (covering 02:00–04:00 on March 13, 2020) for visualization analysis, as shown in Fig. 5. Experimental results demonstrate significant performance differences among the evaluated models. While the baseline LSTM model captures general visibility trends, it shows substantial errors during volatile periods like rapid fog formation and dissipation. Contrary to theoretical expectations, several hybrid architectures (LSTM-CNN, EMD-LSTM, and LSTM-Attention) underperform compared to the basic LSTM. The LSTM-Attention mechanism's limitations stem from its single-layer context vector and lack of nonlinear transformation, restricting its ability to develop sophisticated weighting schemes. The absence of attention weight regularization further increases overfitting risks. LSTM-CNN suffers from incompatible feature representations between temporal and spatial modules, while EMD-LSTM accumulates errors through multiple decomposition stages. In contrast, LSTM-XGBoost emerges as the superior performer due to its complementary design. Three key factors contribute to its success: hierarchical feature transformation between components systematically refines temporal features; XGBoost's ensemble mechanism

Table 2 Model performance evaluation on the testing set

Models	Evaluation metrics						
	RMSE (m)	MAE (m)	VAF	WMAPE (m)	R^2	Training time (s)	Prediction time (s)
LSTM	160.82	94.10	89.10	26.97	0.89	3.5	0.04
XGBoost	143.68	55.16	91.80	15.81	0.90	5.1	0.11
LSTM-Attention	175.08	107.45	87.26	30.80	0.87	10.4	0.14
LSTM-CNN [16]	213.60	177.90	84.14	50.99	0.81	17.6	0.19
EMD-LSTM [17]	225.80	132.54	78.49	37.99	0.78	27.6	0.26
LSTM-XGBoost [18]	121.48	44.69	94.43	12.28	0.94	2.6	0.21
BiLSTM-Attention	183.91	117.85	85.74	33.78	0.86	24.3	0.21
Transformer	192.20	103.62	85.03	29.70	0.84	9.0	0.17
Random Forest	183.66	78.08	86.44	21.79	0.86	2.26	0.12
GRU	248.47	131.81	69.92	39.35	0.72	1.79	0.03

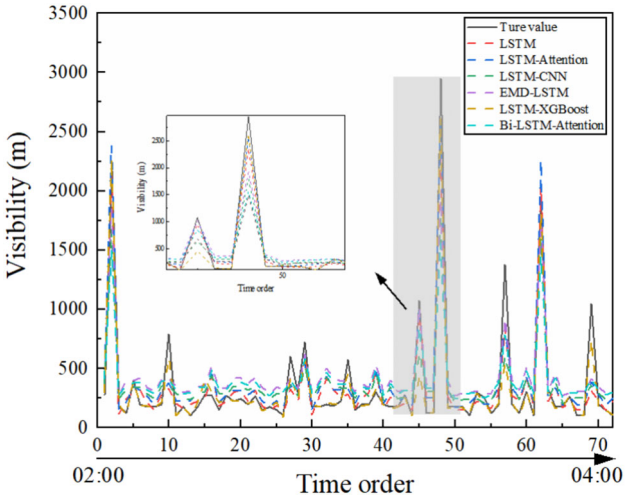

Fig. 5 Comparison of time series of the experimental data and the prediction data

naturally regularizes predictions through multiple decision trees; and Bayesian optimization ensures optimal parameter tuning for both components. The model's advantage is particularly evident in handling volatile visibility patterns, where its dual architecture combines temporal sensitivity with robust regression capability. LSTM-XGBoost demonstrates the closest alignment with ground truth values, especially in capturing local details like peaks and troughs, with Bayesian optimization significantly enhancing the network's sensitivity to data variations. The model performance is benchmarked using a radar chart in Fig. 6, where closer proximity to the center indicates superior predictive capability. Among all models, LSTM-XGBoost performs the best, achieving the highest R^2 (0.921) and the lowest MAE and WMAPE (54.87 m and 15.73 m, respectively), which clearly demonstrates its forecasting advantage.

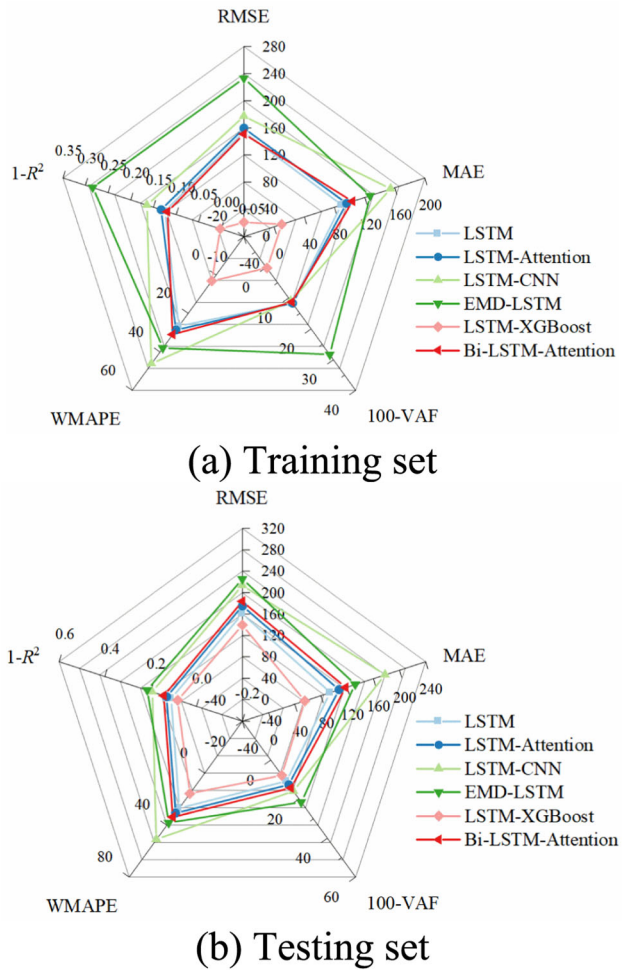

Fig. 6 Comparison of evaluation metrics

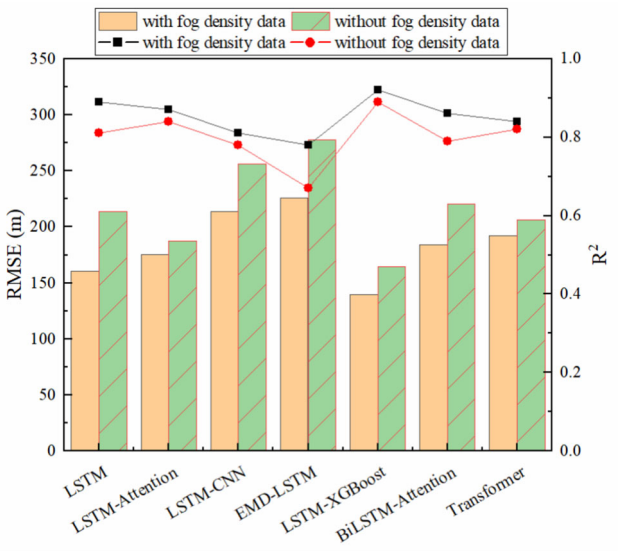

Fig. 7 Ablation experiments for fog density data

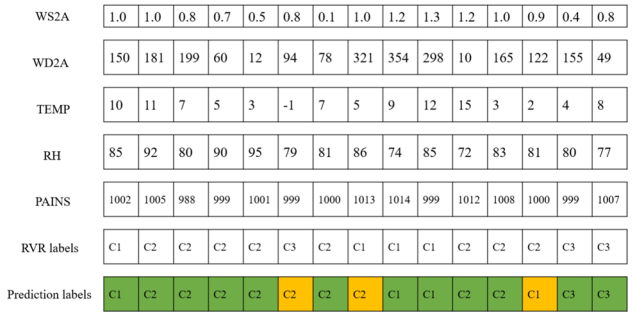

Fig. 8 Validation of model using data from different climate-region airports

To assess the individual contributions of key architectural components, we performed ablation studies on three structurally related models: a baseline LSTM model, a simplified LSTM-XGBoost model, and a comprehensively Bayesian-optimized LSTM-XGBoost model (Fig. 7). Ablation studies on three related models confirm that an effective hybrid framework requires complementary learning pathways: LSTMs capture temporal dependencies, while XGBoost models complex feature interactions. Bayesian optimization converged after 25 iterations, identifying the optimal configuration at iteration 14. The final model uses 56 and 51 units in the two LSTM layers, a 16-dimensional feature space, and an XGBoost ensemble with 121 trees, a depth of 9, and a learning rate of 0.0872. These results demonstrate that Bayesian optimization efficiently navigates complex parameter spaces to achieve strong model configurations. Based on evaluation using an independent test set, the optimized LSTM-XGBoost method achieved a root mean square error (RMSE) of 121.48 m. This error corresponds to approximately 18.3% of the average visibility in the test set (about 664 m), and exhibits a relative error lower than that of the standalone LSTM, XGBoost, and the unoptimized LSTM-XGBoost models. The results indicate that the integrated architecture, which combines the sequential modeling capability of LSTM, the feature selection mechanism of tree-based models, and the efficiency of Bayesian hyperparameter optimization, can effectively learn temporal patterns and feature interactions for visibility prediction. To ensure the reliability of the evaluation, the train/test split was performed randomly to enable the model to learn from critical fog transition phases across the entire dataset. Despite this choice, rigorous safeguards were implemented to prevent temporal data leakage: only historically available features were used at each prediction step to avoid look-ahead bias; time-series cross-validation was employed during model development; and the final model was evaluated exclusively on a held-out test set that never participated in training or tuning. These measures collectively mitigate the risk of data leakage and support the credibility of the reported performance metrics. Moreover, the experiment systematically compares LSTM variants with non-LSTM models such as Transformer. Results show our LSTM-XGBoost achieves the lowest RMSE, outperforming all competitors. Ablation studies confirm that incorporating video fog density data consistently improves accuracy across all models. Notably, our hybrid model maintains robust performance even without fog density inputs, demonstrating its ability to effectively leverage other meteorological parameters. To assess the generalizability of the LSTM-XGBoost model, we conducted validation using data from Valladolid Airport in Spain, a region characterized by distinct meteorological conditions that foster frequent radiation fog events (Fig. 8) [6].

The validation utilized an eight-year meteorological dataset from November 2009 to December 2016, specifically focusing on the November to February period when radiation fog is most frequent. The model employed only standard meteorological variables measured at airports, with no need for specialized parameters or architectural adjustments. Visibility was categorized into three classes based on hourly Runway Visual Range: C1 for low visibility (0 to 1000 m), C2 for moderate visibility (1000 to 1990 m), and C3 for normal conditions (1990 m and above). For the critical C1 low visibility events, the model achieved a Probability of Detection of 100%, a False Alarm Rate of 5.07%, a Critical Success Index of 94.93%, and an overall classification accuracy of 86.6%. This performance demonstrates high operational reliability, balancing safety-critical detection with minimal unnecessary disruptions, and confirms the model's robustness and cross-regional applicability under constrained data conditions. The model maintains a consistent architecture but can adapt its input features based on

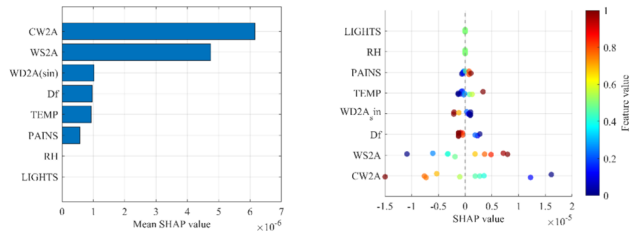

Fig. 9 Feature importance analysis of the best prediction model

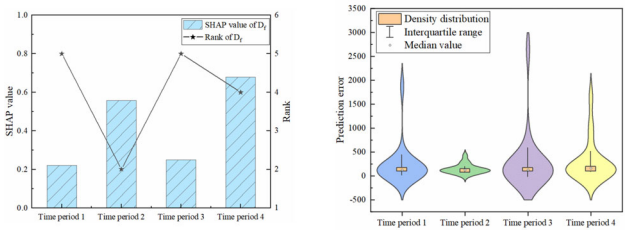

Fig. 10 Time-segmented SHAP analysis and error-case analysis

practical data availability. Regarding the role of fog density, the model exhibits dual capability: while integrating video-derived fog density enhances precision for continuous visibility estimation, standard meteorological parameters alone provide sufficient information for reliable categorical classification. This explains the model's strong performance on the Spain dataset without fog density data. The framework can leverage specialized data to refine predictions when available, yet maintains high reliability using only conventional inputs for operational decision support, highlighting its practical adaptability.

4.2 Analysis of relative influence and partial dependence analysis

The SHAP value ϕ_i for a feature i, which explains the model prediction f for a given sample x, is formally defined by the following equation:

$$\phi_i = \sum\nolimits_{S \in F_i} P(F)[f(S \cup \{i\}) - f(S)] \tag{6}$$

where F represents the entire set of features, while S is any subset of F that does not include feature i. The terms $f(S)$ and $f(S \cup \{i\})$ correspond to the model's output using the subset S and the subset augmented with feature i, respectively. The SHAP interaction value $\phi_{i,j}$ between features i and j is introduced to measure their synergistic effect. It captures the portion of the model's output that cannot be explained by the sum of each feature's individual contribution, as follows:

$$\phi_{i,j} = \sum\nolimits_{S \in F_{i,j}} P(F)[f(S \cup \{i, j\}) - f(S \cup \{i\}) - f(S \cup \{j\}) + f(S)] \tag{7}$$

The SHAP interaction values quantify feature interdependencies by comparing a model's output with both features present against the sum of their individual outputs across all contextual subsets S. As quantified by SHAP in Fig. 9, the LSTM-XGBoost model identifies CW2A and WS2A as the dominant predictive features, followed by WD2A and D_f, while other variables exhibit limited influence. The accompanying scatter plots further visualize how feature values positively or negatively impact predictions. To enhance model interpretability, this study conducted time-segmented SHAP analysis and error-case analysis, as shown in Fig. 10. The period from 00:00:26 to 11:47:48 was divided into four phases corresponding to key stages of fog evolution: formation through radiation cooling, a mature stage with peak density and stability, an early dissipation stage driven by solar heating, and a late dissipation stage dominated by turbulent mixing and rapid evaporation. SHAP analysis indicated that the importance of the fog density metric D_f peaked during the late dissipation stage. However, the highest prediction errors occurred in both the early and late dissipation stages. This divergence suggests that while the model relied heavily on D_f, its relationship with visibility likely became nonlinear or unstable under the complex, transitioning meteorological conditions of the dissipation phases, rendering predictions less reliable. Therefore, moving beyond static interpretability is necessary. This study emphasizes adopting time-aware modeling strategies, including optimizing feature weights specifically for dissipation stages, dynamic importance monitoring, and early concept drift detection, to systematically improve model robustness and interpretability in real-world complex environments.

Analysis of SHAP interaction values clarifies the meteorological roles of key features in fog evolution (Fig. 11). The SHAP interaction analysis quantifies the dynamic contributions of key meteorological features during the studied fog event. During the dense fog phase observed in this case (01:30–05:00), CW2A and WS2A exhibit the strongest interaction effects on visibility improvement, which is consistent with a physical mechanism whereby wind speed may accelerate fog dissipation through mixing or advection. In contrast, during the preceding low fog density phase (00:00–01:30), PAINS and TEMP show more pronounced influences, reflecting conditions often associated with high pressure and radiative cooling that can favor fog formation. Although WD2A ranks high in overall feature importance, its stage-specific interactive effect here appears secondary, suggesting its role may be more context-dependent. Notably, RH and LIGHTS exhibits minimal interaction with D_f, likely due to its limited variation or collinearity with other variables. These interpretations are derived from a single case study; thus, their generalizability requires validation

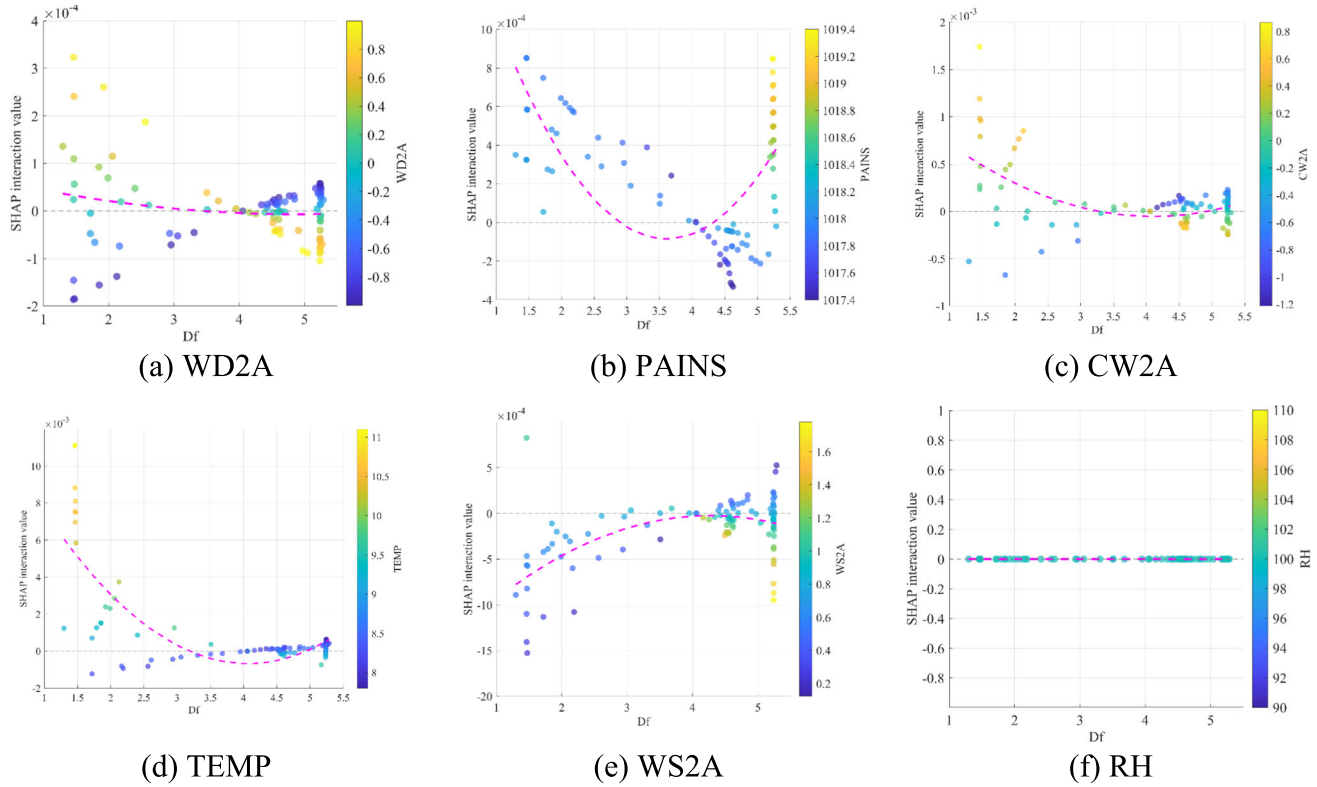

Fig. 11 The interaction between fog density and atmospheric variables

across multiple fog events under varied meteorological settings. These results demonstrate that fog density significantly modulates the impact of meteorological factors on visibility, reflecting the complex interactions observed during the event. A two-tier edge-center architecture deploys our hybrid ML framework for low-latency visibility monitoring. Edge units employ lightweight deep learning models to process 16 channels of 1080P video in real-time, estimating fog density with under 100 ms latency. Meanwhile, a central fusion center aggregates edge and meteorological data, running an optimized LSTM–XGBoost model to predict runway visual range and detect fog clusters with end-to-end latency under 3 s. This design balances edge efficiency with central precision for reliable airport operation in low-visibility conditions.

5 Conclusions

(1) A multi-source data fusion framework was developed to overcome the limitations of conventional atmospheric and video-based visibility estimation. By integrating video-inverted high-precision fog density based on the MVG model and real-time atmospheric observations, this approach effectively compensates for missing data and spatial discontinuities. The fused data strongly correlate with ground-truth RVR measurements, enabling real-time, high-resolution, low-visibility prediction at airport runways.

(2) The Bayesian optimized model demonstrates superior performance compared to both individual models and unoptimized hybrid variants. Systematic hyperparameter optimization through Bayesian methods proved essential for model performance. The optimized architecture achieved a test RMSE of 121.48 m and R^2 of 0.935, representing significant improvement over baseline models. The hybrid design successfully leverages the complementary strengths of both components.

(3) Interpretable analysis confirms that wind direction and fog density are the most influential variables, engaging in complex nonlinear interactions with other meteorological factors. The effects of pressure, temperature, and wind speed are context-dependent, emphasizing the need for scenario-aware modeling. SHAP and partial dependence analyses enhance operational interpretability, providing scientifically grounded insights for airport visibility forecasting.

6 Limitations and future work

The proposed framework exhibits certain limitations that also point to future research directions. Currently, its generalizability is constrained by the reliance on video-derived fog concentration data from a single environment. Furthermore, the approach lacks explicit modeling of atmospheric physical processes and does not fully capture dynamic feature interactions. Most notably, the current model remains diagnostic in nature and has not yet achieved the lead-time forecasting capability required for operational applications. Future work will focus on improving robustness through multi-source, multi-scale data fusion, and explore coupling with numerical weather prediction models to enable end-to-end, lead-time visibility forecasting for airport operations. Lightweight or physics-informed architectures will be developed to enhance adaptability, while integrating explainable AI and uncertainty quantification methods will support more reliable meteorological decision-making.

Acknowledgements This work was supported by the Shenzhen Science and Technology Program (JCYJ20240813162003005), Songshan Lake Science and Technology Special Agent Project (2023441201KCJG), Science and Technology Special Agent Project for Dongguan Polytechnic (KJTP202409).

Author contribution Fang-wei Liu: Writing—review & editing, Funding acquisition, Supervision. Ren-jie WU: Data analysis, Writing—review & editing.

Funding Open access funding provided by The Hong Kong Polytechnic University.

Data availability No datasets were generated or analysed during the current study.

Declarations

Competing interests The authors declare no competing interests.

References

1. Li, M., Wang, G.: A method combining mobile transmissometer and lidar for high precision measurement of visibility. IEEE Photonics J. **16**(4), 161–10 (2024)
2. Boneh, T., Weymouth, G.T., Newham, P., Potts, R., Bally, J., Nicholson, A.E., et al.: Fog forecasting for Melbourne airport using a bayesian decision network. Weather Forecast. **30**, 1218–1233 (2015)
3. Cornejo-Bueno, L., Casanova-Mateo, C., Sanz-Justo, J., Cerro-Prada, E., Salcedo-Sanz, S.: Efficient prediction of low-visibility events at airports using machine-learning regression. Bound. Layer Meteorol. **165**, 349–370 (2017)
4. Castillo-Botón, C., Casillas-Pérez, D., Casanova-Mateo, C., Ghimire, S., Cerro-Prada, E., Gutierrez, P.A., et al.: Machine learning regression and classification methods for fog events prediction. Atmos. Res. **272**, 106157 (2022)
5. Durán-Rosal, A.M., Fernández, J.C., Casanova-Mateo, C., Sanz-Justo, J., Salcedo-Sanz, S., Hervás-Martínez, C.: Efficient fog prediction with multi-objective evolutionary neural networks. Appl. Soft Comput. **70**, 347–358 (2018)
6. Guijo-Rubio, D., Gutiérrez, P.A., Casanova-Mateo, C., Sanz-Justo, J., Salcedo-Sanz, S., Hervás-Martínez, C.: Prediction of low-visibility events due to fog using ordinal classification. Atmos. Res. **214**, 64–73 (2018)
7. Xu, Q., Pang, Y., Liu, Y.: Air traffic density prediction using Bayesian ensemble graph attention network (BEGAN). Transp. Res. Part C, Emerg. Technol. **153**, 104225 (2023)
8. X, Q., P, Y., Z, X., L, Y.: PIGAT: physics-informed graph attention transformer for air traffic state prediction. IEEE Trans. Intell. Transp. **25**, 12561–12577 (2024)
9. Guo, H., Wang, X., Li, H.: Density estimation of fog in image based on dark channel prior. Atmosphere (2022). https://doi.org/10.3390/atmos13050710
10. Dai, M., Li, G., Shi, W.: Fog density analysis based on the alignment of an airport video and visibility data. Sensors (Basel) (2024). https://doi.org/10.3390/s24185930
11. Shan, Y., Zhang, R., Gultepe, I., Zhang, Y., Li, M., Wang, Y.: Gridded visibility products over marine environments based on artificial neural network analysis. Appl. Sci. (2019). https://doi.org/10.3390/app9214487
12. Tapiador, F.J., Sánchez, J., García-Ortega, E.: Empirical values and assumptions in the microphysics of numerical models. Atmos. Res. **215**, 214–238 (2019)
13. Xu, Q., Pang, Y., Zhang, Z., Liu, Y.: Data-driven governing equation identification of near terminal air traffic flow dynamics. J. Air Transp. Manag. **129**, 102871 (2025)
14. Wu, R., Wang, Y., Swe Hein, K., Xia, J.: An enhancement method for chloride diffusion coefficient long-term prediction based on hilbert dynamic probabilistic interpolation and BO-LSTM. Measurement **247**, 116820 (2025)
15. Portal-Porras, K., Fernandez-Gamiz, U., Zulueta, E., Irigaray, O., Garcia-Fernandez, R.: Hybrid LSTM+CNN architecture for unsteady flow prediction. Mater. Today Commun. **35**, 106281 (2023)
16. Hao, W., Sun, X., Wang, C., Chen, H., Huang, L.: A hybrid EMD-LSTM model for non-stationary wave prediction in offshore China. Ocean Eng. **246**, 110566 (2022)
17. Yang, Z., Wang, Y., Li, J., Liu, L., Ma, J., Zhong, Y.: Airport arrival flow prediction considering meteorological factors based on deep-learning methods. Complexity **2020**, 6309272 (2020)
18. Ni, C., Fam, P.S., Marsani, M.F.: A data-driven method and hybrid deep learning model for flood risk prediction. Int. J. Intell. Syst. **2024**, 3562709 (2024)