Check for updates

Meteorological Applications

RMetS Royal Meteorological Society

RESEARCH ARTICLE OPEN ACCESS

Hybrid Deep Learning-Based Networks for Deterministic and Probabilistic Wind Nowcasting Within a Wake Vortex Warning System for Vertiport Operations

Sooseok Lee[1,2,3] | Frank Holzäpfel[1] | Anton Stephan[1] | Dieter Moormann[3]

[1]Institute of Atmospheric Physics, German Aerospace Center (DLR), Oberpfaffenhofen, Germany | [2]Innovation Center for Small Aircraft Technologies (INK), German Aerospace Center (DLR), Würselen, Germany | [3]Institute of Flight System Dynamics, RWTH Aachen University, Aachen, Germany

Correspondence: Sooseok Lee (sooseok.lee@dlr.de)

Received: 28 August 2025 | **Revised:** 19 February 2026 | **Accepted:** 13 March 2026

Keywords: continuous wavelet transform | convolutional neural network | mixture density network | multivariate distribution | probabilistic prediction | recurrent neural network | wake vortex warning system | wind nowcasting

ABSTRACT

Accurate wind prediction is essential for forecasting the transport and decay of wake vortices that might cause an adverse impact to light aircraft operations. This paper proposes an innovative framework for deterministic and probabilistic wind nowcasting, designed to be integrated into a Wake Vortex Warning System for Vertiports (WVWS-V) near major airports. The study presents and tests different deterministic nowcasting models. The superior approach employs a hybrid concept combining complex Continuous Wavelet Transform (CWT), two-dimensional Convolutional Neural Networks (2D-CNN), Long Short-Term Memory (LSTM) networks, and Light Gradient Boosting Machine (LGBM), enabling it to capture both the general patterns and abrupt fluctuations of wind occurring intermittently and repeatedly. The proposed hybrid model achieved a Mean Absolute Error (MAE) of 0.75 m/s for wind speed and 9.1° for wind direction over a 20-min prediction horizon, outperforming all other evaluated models including the statistical persistence prediction method serving as baseline. The probabilistic nowcasting establishing 95% confidence intervals of wind speed and direction utilizes a multivariate-Mixture Density Network (MDN) with weighted Mahalanobis distance. The MDN-based model provides a narrower probability range than the statistical persistence baseline model, with average improvements by 0.91 m/s for wind speed and 12.05° for wind direction. The comparison of the statistical persistence baseline and other machine learning architectures demonstrates that the proposed models achieve superior performance, enabling the WVWS-V to ensure safe and reliable light aircraft operations at vertiports.

Abbreviations: 2D-CNN, Two-Dimensional Convolutional Neural Network; ANN, Artificial Neural Network; AVOSS, Aircraft Vortex Spacing System; BFU, German Federal Bureau of Aircraft Accident Investigation; CARA, Computer for Advanced Research in Aerospace; CI, Confidence Interval; CMOR, Complex Morlet; CNN, Convolutional Neural Network; COI, Cone of Influence; CR, Contamination Rate; CWT, Continuous Wavelet Transform; DC, Direct Current; DFS, German Air Navigation Service; DLR, German Aerospace Center (Deutsches Zentrum für Luft- und Raumfahrt); DWT, Discrete Wavelet Transform; ERF, Effective Receptive Field; EUROCONTROL, European Organisation for the Safety of Air Navigation; eVTOL, electric Vertical Take-Off and Landing; FAA, Federal Aviation Administration; FCO, Rome Fiumicino Airport; GMM, Gaussian Mixture Model; GRU, Gated Recurrent Unit; ICAO, International Civil Aviation Organization; ICWT, Inverse Continuous Wavelet Transform; IF, Isolation Forest; LGBM, Light Gradient Boosting Machine; LSTM, Long Short-Term Memory; MAE, Mean Absolute Error; MDN, Mixture Density Network; ML, Machine Learning; NTSB, National Transportation Safety Board; PDF, Probability Density Function; RECAT, Wake Turbulence Re-categorization (of Separation Minima); RMSE, Root Mean Square Error; RNN, Recurrent Neural Network; SP, Statistical Persistence; UAM, Urban Air Mobility; WD, Wind Direction; WS, Wind Speed; WSVBS, Wake Vortex Prediction and Advisory System; WSWS, Wake Vortex Warning System; WVWS, Wake Vortex Warning System; WVWS-V, Wake Vortex Warning System for Vertiports.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2026 The Author(s). *Meteorological Applications* published by John Wiley & Sons Ltd on behalf of Royal Meteorological Society.

1 | Introduction

The development of Urban Air Mobility (UAM) with the connected vertiport operations has brought about new challenges in predicting environmental hazards. One significant concern is related to wake vortices generated by large aircraft during landing or take-off, which can impact electrical Vertical Take-Off and Landing (eVTOL) aircraft like air taxis operating at low altitudes near vertiports located adjacent to the runways of major airports. The 2022 incident, in which an aircraft with a weight below one metric ton lost control due to a sudden uncommanded roll and impacted the ground (National Transportation Safety Board 2022), clearly illustrates how hazardous wake turbulence can be to small aircraft. In another 2022 accident, a flight at Aachen-Merzbrück Airport lost control in wake turbulence and crashed, resulting in fatal injuries to the pilots (Bundesstelle für Flugunfalluntersuchung (BFU) 2022). Given their relatively low mass and limited momentum compared to larger aircraft, eVTOL vehicles are more vulnerable to such disturbances (Thompson et al. 2022), which is aggravated by the little time and altitude available for recovery during encounters. Even before these concerns emerged, the International Civil Aviation Organization (ICAO) had suggested the Re-categorization of Wake Turbulence Separation Minimuma (RECAT) (EUROCONTROL 2024), introducing modified wake turbulence categories for increasing airport capacity by more precisely scheduling the minimum separation distances between aircraft. Based on the suggestion, FAA and EUROCONTROL have implemented the RECAT phase I and have improved separation criteria for smaller aircraft, enhancing safety. RECAT phase III shall eventually take into account also the atmospheric conditions to predict wake vortex transport and decay and the effects of potential wake vortex encounters. For eVTOL operations in ground proximity, wind speed and direction are the main environmental parameters controlling vortex transport and decay.

Consequently, this study focuses on wind nowcasting used in wake vortex warning systems to predict wake vortex behavior. As part of the Flight Safety of Small Aircraft under Adverse Meteorological Conditions (SMART-MET) project conducted at the German Aerospace Center (DLR), this research aims to support the safe operation of small aircraft. As there is no suitable globally unified formal standard for the minimum separation between eVTOLs and conventional large aircraft, the SMART-MET project has embarked on developing a Wake Vortex Warning System for Vertiports (WVWS-V) with an initial focus on the vertiport planned at Rome Fiumicino Airport (FCO) to mitigate risks that small aircraft can potentially face posed by wake vortices generated by large aircraft during landing. While the Wake Vortex Warning System (WSWS) developed by German Air Navigation Service (DFS) (Gurke and Lafferton 1997), the Aircraft Vortex Spacing System (AVOSS) developed by NASA (Hinton 2001), and the Wake Vortex Warning System (WSVBS) developed at DLR (Gerz et al. 2009; Holzäpfel et al. 2009, 2021) deal with conventional aircraft, the WVWS-V is particularly pertinent for the operations of air taxis, which shall connect between the city center of Rome and FCO. UrbanV plans to build a vertical takeoff and landing area adjacent to runways 16 L and 34R at FCO. UrbanV has evaluated the safety factors needed to integrate the vertiport into the airport environment, supporting the investigation of wake vortex impacts to verify the feasibility of the proposed location. Considering their proximity, it is possible that wake vortices generated by medium and large aircraft such as A320 and A333 are likely to reach the flight tracks to the vertiport under certain wind conditions, potentially impacting the safety of air taxi operations. Therefore, implementing more precise wind nowcasting algorithms based on both deterministic and probabilistic approaches will be essential to support safe and efficient air taxi operations at the airport. Furthermore, considering that aircraft such as the eVTOLs, which operate at the vertiport, will be significantly affected by the circulation strength generated by aircraft classified as heavy or medium, this study is an indispensable component of the WVWS-V.

This study primarily employs a recurrent neural network (RNN) structure (Elman 1990) for time series nowcasting for deterministic prediction and additionally implements various architectures based on long short-term memory (LSTM) (Hochreiter and Schmidhuber 1997). Building on this, the study extends the approach to combine continuous wavelet transform (CWT) (Grossmann and Morlet 1984), Two-dimensional convolutional neural network (2D-CNN) (LeCun et al. 1998), and LSTM in order to seize intermittent and recurrent patterns over extended periods in the wind profile. Moreover, probabilistic nowcasting using the mixture density network (MDN) (Bishop 1994) is implemented in accordance with deterministic prediction. Notably, instead of relying on normal distributions, the new approach integrates a multivariate distribution with the MDN to capture the dependent distributional changes in the u and v components of horizontal wind. This approach shall enable an accurate representation of the probabilistic distributions for both wind speed and direction. Likewise, this study encompasses the implementation of an established statistical persistence model serving as a benchmark Frech et al. (2002), as well as the development of machine learning-based models. The benchmarking of the investigated machine learning and statistical persistence approaches for both deterministic and probabilistic predictions yields the superior solutions that will be integrated into the WVWS-V.

2 | Previous Approaches

2.1 | Statistical Persistence Nowcasting for WVWS

The statistical persistence approach initially developed by (Frech et al. 2002) for the wake vortex warning system (WSWS) at Frankfurt airport (Gurke and Lafferton 1997) has demonstrated robust performance for closely-spaced departure operations at San Francisco airport. The model applies a probabilistic approach using a probability density function (PDF) to compute a range of likely wind conditions. The nowcasting model builds on the statistical analysis of 523 days of anemometer data to predict horizontal wind vectors over a 20-min forecast horizon. Unlike earlier systems that relied on only crosswind components, this model predicts both wind speed and direction using a combination of statistical persistence and fluctuation analysis. The model is implemented

14698080, 2026, 2, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70173 by Universitatsbibliothek RWTH Aachen University, Wiley Online Library on [08/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

by defining medium-range fluctuation parameters for wind speed and direction, which are derived from standard deviations of sequentially averaged wind components. These parameters are used to classify predefined wind classes, each representing a wind probability range. Within each class, PDFs define the maximum expected changes in wind speed and direction at a 95% confidence interval (CI).

For real-time predictions first the past 20-min wind speed and direction are averaged and then assigned to one of the predefined classes. Weighted averaging methods smoothen transitions between wind classes. This Statistical Persistence (SP) model was combined with a wake vortex prediction model, which used the predicted crosswind component to estimate the lateral displacement of wake vortices between the closely-separated runways.

This design not only enhances the stability and reliability of the WSWS but also supports increased landing capacity by safely reducing aircraft separation minima for over 80% of operational time. So, this algorithm can be considered a benchmark in modern airport meteorological systems, paving the way for future enhancements to improve operational safety and efficiency. Nevertheless, the WSWS never became operational because pilot associations argued that the consideration of wake vortex evolution in ground proximity may not be sufficient and called for an advisory system covering the complete glide path.

2.2 | Deterministic Time Series Analysis Using Machine Learning

The persistence-based prediction does not adequately capture dynamically shifting wind conditions within the deterministic and probabilistic framework (Theuer et al. 2022; Worsnop et al. 2018). To address this limitation, this study presents different time series machine learning models, developing more sophisticated architectures that have been widely used for wind nowcasting by analyzing temporal dependencies of time series data. The proposed methodology includes both deterministic and probabilistic nowcasting, which will be applied to better account for capturing uncertainties in wind conditions. For the largest part, the LSTM network has been utilized for wind prediction and the excellent performance of LSTM has been demonstrated until recently (Demirtop and Sevli 2024; Prabha et al. 2019). Moreover, various approaches have been merged with LSTM to enhance the performance of the models. A study introduced a composite of an artificial neural network (ANN) model that integrates a convolutional input layer, an LSTM hidden layer, and a transposed convolutional output layer, demonstrating superior performance compared to traditional models (Cao et al. 2017). Another approach is to use LSTM with an additional function with LSTM to find hyper parameters (Shao et al. 2021). In a different effort, a model proposed a multi-variable stacked LSTM model (MSLSTM) for short-term wind speed forecasting (Liang et al. 2018). This model incorporates multiple meteorological variables to improve prediction accuracy. One more approach proposed the unit designed for wind field prediction by analyzing the horizontal wind components u and v (Gu et al. 2024; Manero et al. 2022). This approach reconstructs wind speed and direction from these components, which emphasizes the importance to analyze wind by separating it into u and v components in order to avoid unsteady jumps across the 360° value of wind direction. There are cases where performance has been improved by integrating CNN to LSTM (Manero et al. 2022; San et al. 2023; Sari et al. 2022). In addition, there are also studies that add independent processes to CNN-LSTM beyond structural transformations and additions within LSTM itself (Tekin et al. 2021; Ren et al. 2022). One study augments the CNN–LSTM structure with attention, external memory, and K-shape clustering, and selects one of K outputs based on the assigned pattern (Yang and Zhang 2022). These studies suggest that performance can be improved in wind prediction by using components as dependent features in combination with improved LSTM architectures and CNN.

Despite these previous advancements, constraints have been found in the existing models. One example of this limitation is in anomaly detection, particularly pronounced in the periodically nonstationary patterns (Lindemann et al. 2021; Tran and Hogg 2022). For this reason, methods for predicting wind using time series machine learning models and mathematical techniques such as the discrete wavelet transform (DWT) (Berrezzek et al. 2019; Liu et al. 2019; Zhao et al. 2021) and the continuous wavelet transform (CWT) (Mohammed and Ahmed 2023), as well as models incorporating all of the LSTM, CNN, and DWT (Pei et al. 2024), demonstrate the potential for enhancing performance when mathematical methods are combined with various machine learning techniques. Hence, this study proposes new approaches using advanced methods to find hidden patterns that cannot be found in single-line patterns but can be seen when decomposed into multiple profiles with an in-depth analysis.

2.3 | Composite Stochastic Distribution

A major challenge for probabilistic wind nowcasting arises from the observation that the empirical distribution of the wind components deviates from a Gaussian distribution. To achieve predictions within a 95% confidence interval (CI) for wind speed and wind direction, it is crucial to accurately estimate the probability distribution based on historical data. One alternative approach to deal with the issue is to use an intermittency term (Bock et al. 2024; Castaing et al. 1990). This model aims to explain the observed probability distribution derived by velocity differences in a high Reynolds number turbulence regime from Gaussian distributions. In particular, distributions with heavy-tailed characteristics are addressed in the study. Although this model can be adjusted to match empirical distributions, it is constrained in predicting the changes in kurtosis of the time series analysis. The concept needs to be enhanced to provide time steps and predicted wind as independent variables for kurtosis, and to accommodate multiple instances of predicted wind in a single time step.

Predicting dynamically changing distribution shapes under varying conditions necessitates more advanced modeling techniques. Datasets including high volatility of variance between deterministic predicted wind and actual data bring

14698080, 2026, 2, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70173 by Universitatsbibliothek RWTH Aachen University, Wiley Online Library on [08/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

along significant challenges for probabilistic prediction by volatile kurtosis and thick distribution tails in this study. To address these challenges, the Mixture Density Network (MDN) is used for probabilistic prediction (Bishop 1994). This model combines the Gaussian mixture model (GMM) (Dempster et al. 1977) with neural networks, enabling both the implementation of complex data distributions and their use for predictive tasks. Although conventional MDN models often struggle to accurately capture such irregularities, a recent study introduced an enhanced MDN framework that implemented a concept of the Wasserstein distance in the loss function that encourages the model to generate more precise empirical distributions (Yang et al. 2022). This approach demonstrated that an appropriate modification of MDN can achieve precise performance in predicting probabilistic distributions. An improved generative probabilistic wind speed forecasting framework was proposed to better model the likelihood of the wind speed sequence and produce predictive distributions Zheng et al. (2022). However, despite these advancements, the method does not fully address the joint probabilistic prediction of wind speed and wind direction at the same time, and it does not account for correlations among features within the generative probabilistic distribution.

Therefore, this study aims to incorporate changes in the sequence of u and v as a covariance matrix and represent these dynamics using a multivariate distribution integrated into the MDN framework. Furthermore, the mixture weights extracted from MDN will be applied in computing the weighted Mahalanobis distances, which is addressed in this study for the first time. Mahalanobis distance measures distances between points and probability distributions while accounting for variable correlations and distribution shape. Since this is based on a single distribution formula and is difficult to apply to superposed distributions, the proposed model introduces an advanced framework capable of producing outstanding results for both wind speed and direction.

2.4 | Data Acquisition Framework and Preprocessing

For training, validation, and testing of the nowcasting methods, this study employs wind data captured by an ultrasonic anemometer installed 10 m above the ground close to the threshold of runway 34R of FCO. Figure 1 presents a windrose diagram of wind direction and wind speed measured at the runway threshold, with the runway alignment indicated by the dashed red line. The distribution provides an overview of prevailing directions and associated wind speed classes. The distribution is broadly consistent across 2018–2023 and exhibits similar seasonal and diurnal patterns. Methods to account for these biases are described in the subsequent section. The data was provided by UrbanV in Rome. The dataset is preprocessed with linear interpolation of the missing data points, while gaps exceeding a threshold of 5-min missing data were excluded for both training and testing. After preprocessing, the dataset was uniformly resampled at 1-min intervals, spanning the period from 2018 to 2023. For the purposes of performance testing, data from 2018 to 2022 were used for training, while the data from 2023 were reserved for testing to assess the predictive accuracy under real-time prediction conditions. Real-time recorded historical anemometer data is used as minimalistic data input which is readily available and keeps the prediction system simple, cost-efficient and robust to develop and test the nowcasting system consisting of mathematical and machine learning methods.

For a single time series analysis, the methodology follows a rolling-window approach to account for the sequential dataset

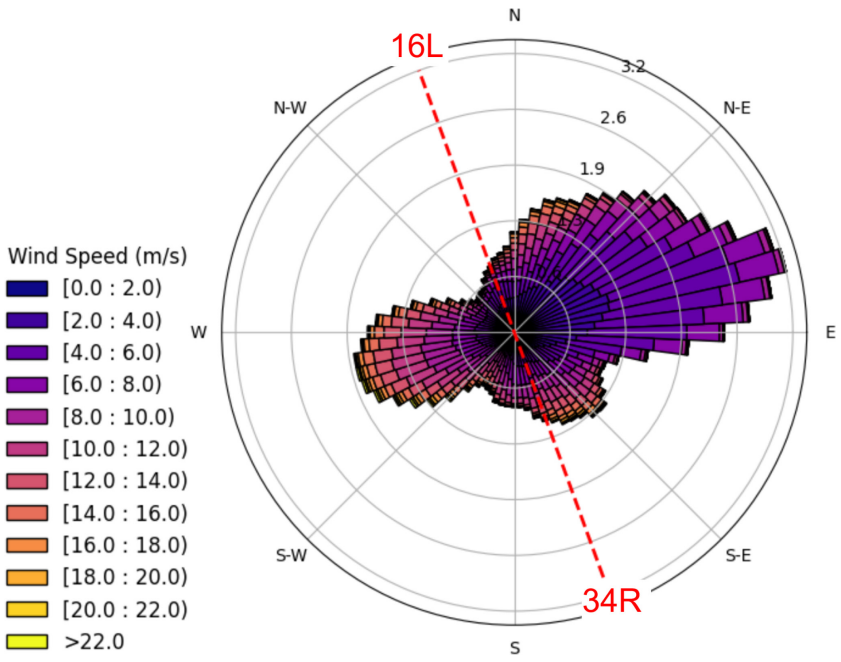


FIGURE 1 | Wind rose of wind speed (m/s) and wind direction, with wind speeds grouped into discrete bins from 2018 to 2023. The dashed red line denotes the runway alignment (runway 34R-16L).

14698080, 2026, 2, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70173 by Universitatsbibliothek RWTH Aachen University, Wiley Online Library on [08/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

except for windows including excluded data points. At each step, a 20-min period preceding the current time point was used as the sequence of input X for single-time series models, and the subsequent 20 min were designated as the target sequence. By shifting the prediction timestamp, the entire dataset was utilized efficiently to represent its temporal dependencies. In the hybrid approach proposed in this paper, the previous 160 min are used for training and the prediction time horizon is set to 20 min after the reference time, as before. The prediction horizon is adjusted to the maximum flight time of 20 min for air taxis considering some buffer for headwind conditions. Various tests using different training sequences in the single-time series analysis have been conducted and the same sequence is applied to all single-time series models.

As training data sets, 2,318,000 data points were generated using the data from 2018 to 2022. For testing, the trained model was applied to the dataset of 2023 while following the same rolling-window methodology, resulting in a total of 452,500 prediction points. Additionally, within the training dataset, every fifth day of data was set aside for probabilistic prediction training. The deterministic model developed from the 4-day subset was then used for generating the model of probabilistic prediction employing the separated fifth day data. This prevents overfitting and supports accurate assessment.

2.5 | Feature Selection

The model basically aims to predict the trends for the next minutes based on data from the past minutes. In this study, all models using machine learning express wind as dependent u and v components of horizontal wind for both deterministic and probabilistic predictions. The basic approach uses u and v vectors over the past minutes, with the target of u and v of the subsequent minutes. Historical data shows seasonal and temporal features, necessitating the inclusion of seasonal indication as input features. Given that the FCO is located in close proximity to the coast, the influence of sea breeze and land breeze should be reflected through features. Therefore, hourly and daily data are converted as cyclic factors to reflect the possibility of seasonal winds and land sea breezes. Since time and days have a cycle of 24 h and 365 days (366 days in the case of leap years), sine and cosine functions are used to convert the time information. Directly using numbers of hours (0–24) and days (1–365) as input values for training results in the model interpreting them as linearly increasing numbers, failing to capture the periodic patterns. Thus, hours are transformed with a 24-h cycle, and days are converted using sine and cosine based on a 365-day (366-day for leap years) cycle, resulting in four additional input features. Equations (1) exemplify the day information encoded using a 365-day cycle to reflect its periodic property. This approach effectively captures patterns regularly occurring by seasons or times of day of breezes. This study also compares the performance of models with and without these features.

$$day_{sin},\ day_{cos} = \sin\left(day \times \frac{2\pi}{365}\right),\ \cos\left(day \times \frac{2\pi}{365}\right) \tag{1}$$

2.6 | Statistical Persistence Model

The Statistical Persistence (SP) model introduced by Frech et al. (2002) is used as the baseline model in this study. The workflow of the SP model begins by dividing wind direction into four classes. Each of the four intervals is further classified into nine subclasses for wind speed. Additionally, medium-range fluctuations are represented by two classes, indicating the degree of atmospheric turbulence, and this process is performed for each time step. As a result, the total of 72 classes have their own probability ranges for wind direction and wind speed. This defines the ranges of the persistence approach within which wind speed and direction may vary within a 95% confidence interval. Figure 2 illustrates the workflow of the SP model, detailing the classification process for wind direction and speed. In the originally proposed SP model a smoothing process was applied to the values defined at each time step; the same approach was adopted in this study. Even with the applied smoothing, the intended 95% probabilities were consistently achieved. Unlike the SP model suggested by Frech et al. (2002) that used the average over the past 20 min as the persistence value forming the center of the 95% intervals, this study adopts the average over the past 2 min. This adaptation was made to achieve better results when comparing with the methods introduced later. The confidence intervals for both wind speed and wind direction are set to 95%

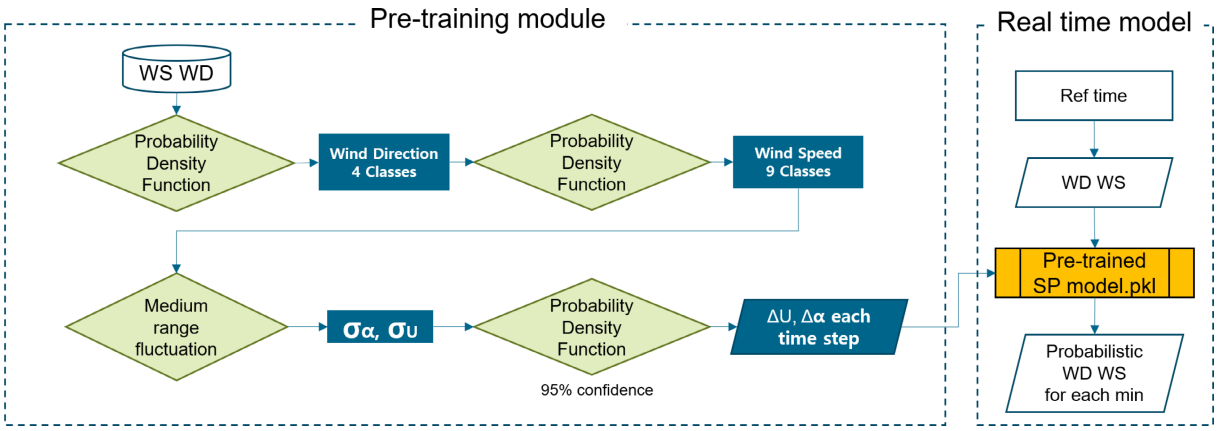


FIGURE 2 | Overview of the Statistical Persistence (SP) model workflow, including the classification of wind direction and speed into 72 probabilistic classes (Frech et al. 2002).

14698080, 2026, 2, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70173 by Universitatsbibliothek RWTH Aachen University, Wiley Online Library on [08/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

for the probabilistic predictions in this study. Accordingly, the overall confidence level for simultaneously satisfying both wind speed and wind direction conditions is defined as 90.25%, which results from the product of the two individual probabilities.

2.7 | Single-Time Series Models: RNN-Based Model

LSTM (Hochreiter and Schmidhuber 1997) is built based on an RNN (Elman 1990) that is a neural network architecture designed for handling time series analysis utilizing a cell state and a gating mechanism such that the three gates (Input Gate, Forget Gate, and Output Gate) selectively store or discard information. The Input Gate determines how much new information is added to the cell state, whereas the Forget Gate decides which part of the previous cell state is retained. Lastly, the Output Gate controls the final cell state reflected in the output. GRU (Cho et al. 2014) has also emerged from the RNN architecture similar to the LSTM network. Unlike LSTM, it regulates only two gates: the update gate and reset gate. The update gate determines what is to be retained with previous information while the reset gate controls whether the previous state's information is to be integrated with the current input, filtering out unnecessary information. A few studies have shown that GRU outperforms LSTM over shorter time steps (Kumar et al. 2021; Zhang et al. 2019). This may occur due to the relatively simpler structure of GRU, which allows it to retain more information from previous cells over shorter time intervals.

This study also includes a comparative evaluation of the two different cells to assess their respective performances. Furthermore, by designing the different architectures that constitute the LSTM or GRU cells, it is possible to achieve additional performance gains. The architectures implemented in this study contain the bidirectional sequence-to-sequence model (Graves et al. 2005; Schuster and Paliwal 1997), the encoder-decoder model with attention function (Bahdanau et al. 2015), and the basic encoder-decoder model (Sutskever et al. 2014). This paper focuses on comparing the performance of the bidirectional model using LSTM cells, the encoder-decoder model with attention function, and the basic encoder-decoder (sequence to sequence) model using GRU and LSTM cells. However, using the LSTM or GRU model alone makes it difficult to predict nonstationary patterns, making the model likely to focus on predicting general patterns. This is because the loss function tends to minimize the overall error so that rapidly changing intervals are less reflected. To address this limitation, this study proposes a hybrid model by integrating a mathematical method and additional machine learning techniques rather than using the RNN-based model alone. In order to prevent wake vortex encounters of air taxis, accurately predicting these intervals of gustiness is essential.

2.8 | Complex Continuous Wavelet Transform and Implementation Choices

The model proposed herein first adopts the wavelet transform technique. The wavelet transform is a signal processing technique that breaks down a signal into components at different frequency domains with preserving time information. Unlike the Fourier transform, which only gives frequency details, the wavelet transform allows for both time and frequency localization for analyzing nonstationary time series with periodic traits (Torrence and Compo 1998). The wavelet transform can be used as two types: *discrete wavelet transform* (DWT) (Daubechies 1988) and *continuous wavelet transform* (CWT) (Grossmann and Morlet 1984). Although using the DWT for wind analysis suggests that it is capable of capturing rapidly changing intervals by defining high and low frequencies (Barjasteh et al. 2024; Mannelli et al. 2021), the DWT approach still introduces uncertainty by severing the dependence and continuity between frequencies and thus failing to capture precisely local features and interactions between frequencies (Arts and van den Broek 2022; Selesnick et al. 2005). In contrast, the CWT preserves the temporal and frequency continuity of the signal across a continuous range of frequency and time, enabling a more detailed and smooth representation of localized variations in wind patterns. So, this study performs wind analysis and prediction using CWT.

Among the frequencies decomposed through CWT, low frequency domains reflect overall patterns occurring over a wide range, while high frequencies represent the volatility that appears in localized patterns. The outcomes of CWT vary depending on the form of the mother wavelet. While real-valued wavelets, such as the Haar (Haar 1910) or Daubechies wavelets (Daubechies 1988), are able to capture sharp transitions and analyze signal discontinuities, the identification of the exact directional transitions is still limited with the real-valued wavelet (Olhede and Metikas 2009; Torrence and Compo 1998; Unser et al. 2009). In contrast, complex-valued wavelets, such as the Morlet wavelet (CMOR) (Morlet et al. 1982), provide both amplitude and phase information, making them more suitable for analyzing the structure of time-frequency patterns. Accordingly, this study uses complex wavelet to perform the convolutional process during the CWT. Based on the variables of bandwidth parameter B and the center frequency C, this study sets the CMOR B–C configuration to CMOR 0.5–1.0 and CMOR 1.0–2.0. Both configurations are collectively used as input features for the regression to enhance the reconstruction process.

Theoretically, the range of integration when performing CWT and inverse continuous wavelet transform (ICWT) with the mother wavelet ranges from negative infinity to positive infinity. However, since this requires the padding region to cover the entire domain, the cone of influence (COI) is defined based on the point where the wavelet power decays to the e-folding time as e^{-2} of its maximum (Torrence and Compo 1998; De Moortel et al. 2004). In the range of the given COI, the boundary allows acceptable margins of regions affected by edge effects. In this study, the padding region is separately set for CMOR 0.5–1.0 and CMOR 1.0–2.0 to be less than 20 and 40 defined by the cone of influence time $t_{\mathrm{COI}} = a\sqrt{B}$ at the scale parameter a that adjusts the frequency resolution. This implies that there is a temporal gap between the prediction time and the reference time due to the padding region, so that only patterns that repeat or gradually emerge or disappear within this interval can be captured effectively. For the CMOR 0.5–1.0 and CMOR 1.0–2.0 wavelets, the respective maximum scales are respectively assigned as 28 and 40, making the frequency ranges as 0.07–1.0 Hz and 0.05–2.0 Hz. These ranges of each scale were divided into 100 uniform

14698080, 2026, 2, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70173 by Universitatsbibliothek RWTH Aachen University, Wiley Online Library on [08/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

intervals, corresponding to the number of pixels in the vertical direction. When scaling is limited in this way, the CWT and ICWT are performed without very low-frequency but relatively high-frequency ranges. Nevertheless, meaningful information still remains outside the COI boundary (Kralj et al. 2024), and prediction discrepancies due to padding regions inevitably arise when the reconstruction is performed. From this perspective, performing postprocessing with values extracted from a single LSTM model facilitates the advantage of compensating for this limitation.

The resulting CMOR-CWT convolutional process can be represented by amplitude and phase, yielding the following relations by real and imaginary parts,

$$W(a,b) = A(a,b)e^{i\phi(a,b)} \tag{2}$$

$$\phi(a,b) = \arctan\left(\frac{\mathrm{Im}(W(a,b))}{\mathrm{Re}(W(a,b))}\right) \tag{3}$$

$$\mathrm{Power}(a,b) = \ |W(a,b)|^2 \tag{4}$$

where b denotes the shift parameter. When these values are filled in for each pixel according to its values of scale and time step, they can be represented as power and phase values, as shown in Figure 3. The power image (top) reveals patterns that are not discernible in the given wind profile (bottom). This allows for a multi-dimensional representation of the univariate time series that shows excessive time dependency as a single profile. The phase difference (middle) is used to determine the local modulation, and temporal characteristics of a pattern. This type of image, which propagates from left to right, can be learned using a deep learning model that combines 2D-CNN and LSTM.

However, one problem that arises here is that two independent models are required to learn power and phase. In particular, the learning phase requires a more complicated model as there is a discrete transition from $-\pi$ to π. Even if a model is built to learn phase information, operating two independent models increases complexity and uncertainty in prediction. Consequently, instead of relying on utilizing power and phase components for time series prediction, the hybrid model proposes utilizing coefficients of real and imaginary terms that are directly extracted from Equation (2), and they are fed into inverse continuous wavelet transform (ICWT) (Grossmann and Morlet 1984; Torrence and Compo 1998), defined by

$$f(t) = \frac{1}{C_\psi}\int_0^\infty\int_{-\infty}^\infty W_\psi(a,b)\psi_{a,b}(t)\frac{da\,db}{a^2}, \tag{5}$$

Here, $W_\psi(a,b)$ is the CWT coefficient, and $\psi_{a,b}(t)$ is the corresponding scaled and translated mother wavelet. The constant C_ψ is the wavelet normalization constant that depends only on the mother wavelet. By following the postprocessing module introduced in the next section, the reconstruction process omits the term involving C_ψ, thereby reducing unnecessary computations as the value is constant by the defined mother

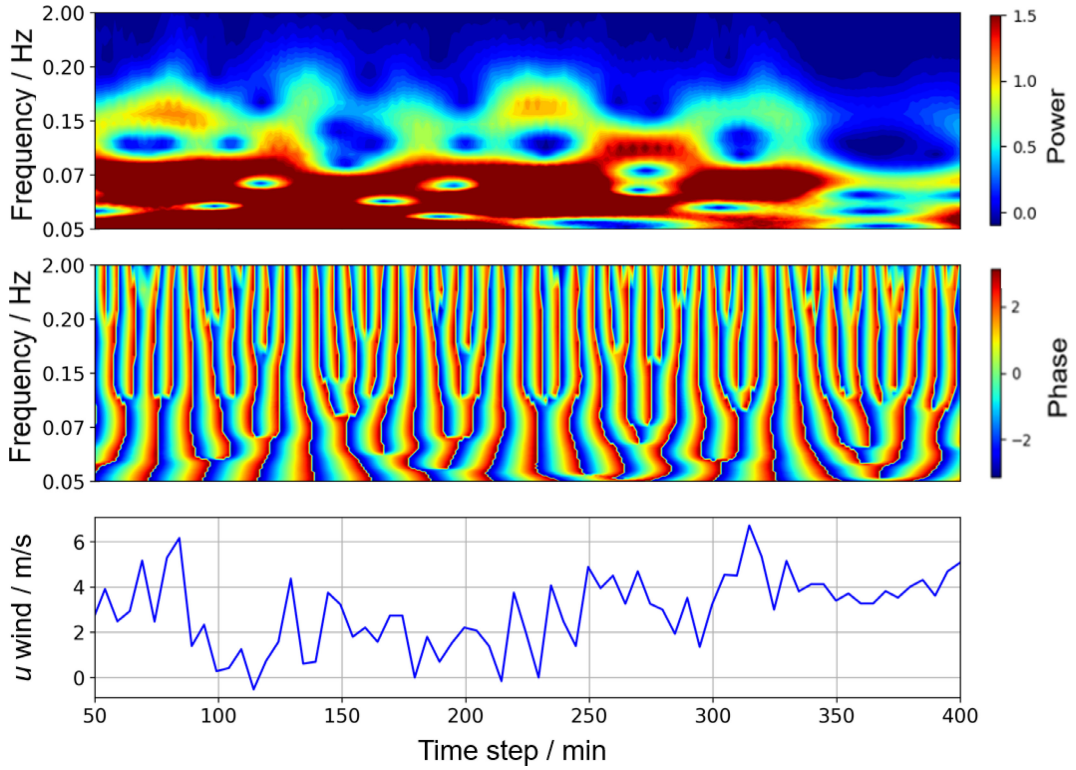
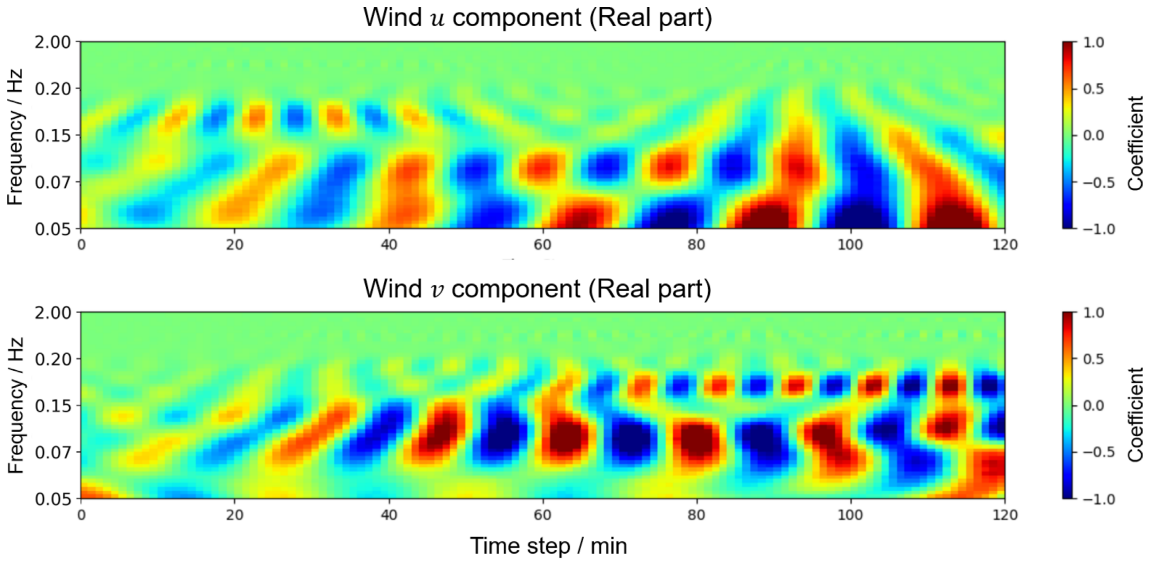
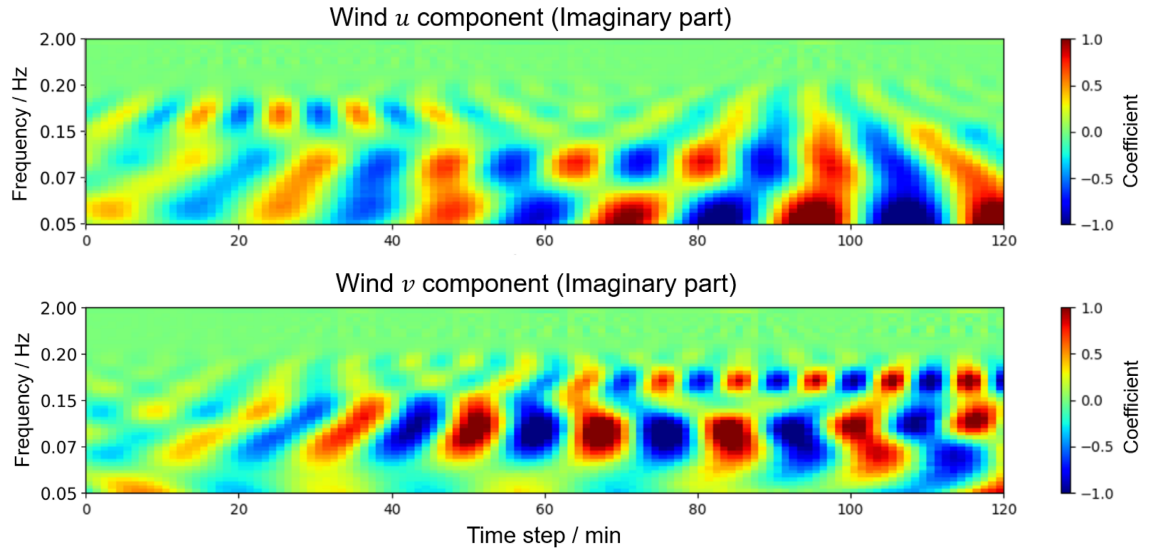


FIGURE 3 | The time–frequency analysis of the u-component wind speed using CMOR was conducted for data collected on January 25, 2022. The power (top) and phase (middle) images are extracted from the wind u component (bottom) over 350 min. The power image shows the wavelet power spectrum, where high-power regions (in red) indicate localized dominant fluctuation activity in both time and frequency. The phase in the CWT reveals that there are discrete transitions and discontinuous patterns.

14698080, 2026, 2, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70173 by Universitatsbibliothek RWTH Aachen University, Wiley Online Library on [08/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License



(a) Real part of CMOR-CWT coefficients for the u and v components.



(b) Imaginary part of CMOR-CWT coefficients for the u and v components.

FIGURE 4 | Time-frequency analysis of wind velocity components using CMOR-CWT: (a) real part of the CMOR-CWT coefficients and (b) imaginary part, with the u component shown on top and the v component on the bottom. The real and imaginary parts provide complementary information about oscillatory features and wave propagation. These representations capture the dominant oscillatory patterns in the signal over time and frequency. Phase shifts are observed between the real and imaginary parts, indicating temporal differences in oscillatory features.

wavelet over the processes. The CWT process using complex-mother wavelets is applied independently to both u and v components, and the resulting coefficients are used as inputs to the subsequent deep learning stage, which combines 2D-CNN and LSTM. Figure 4 presents the analysis results of the real and imaginary parts of the u and v components using the CMOR-CWT. The imaginary parts in Figure 4b exhibit similar structures as the real parts in Figure 4a, suggesting that a unified machine learning model for both imaginary and real components can be designed under the 2D-CNN–LSTM architecture. This approach yields images with more clearly defined periodic patterns, making them easier to train compared to power-phase pictures.

2.9 | Hybrid Deterministic Prediction Model

The hybrid model first combines CWT, 2D-CNN, and LSTM, see Figure 5 which is designed to enable a more precise analysis of more complicated wind behavior. The features of time-frequency extracted by the CMOR-CWT are mapped using 2D-CNN to capture localized spatial patterns. Two-dimensional feature extraction from input data is performed with applying two iterations of two-dimensional convolution layer (Conv-2D) and max pooling layers. This subsampling concept was introduced by LeCun et al. (1998) and addresses the fact that the subsampling concept can increase the size of the effective receptive field (ERF) (Luo et al. 2016). Thus,

14698080, 2026, 2, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70173 by Universitatsbibliothek RWTH Aachen University, Wiley Online Library on [08/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

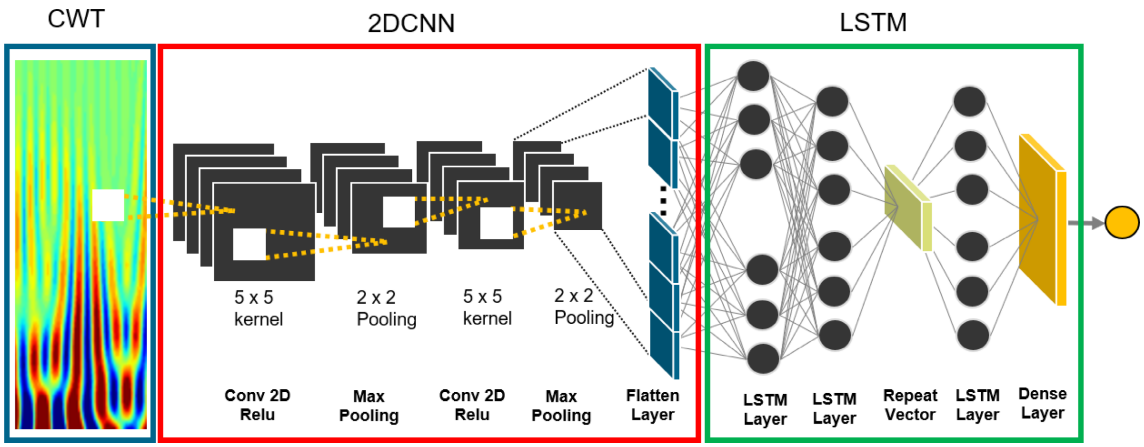


FIGURE 5 | A hybrid model that processes input time series data with 2D-CNN and LSTM. In the 2D-CNN block, feature mapping is conducted through 2D convolution with a 5×5 kernel, ReLU activation function, and max pooling layers as sub-sampling, and converted into vectors to be input into the flatten layer. The LSTM block uses multiple LSTM layers for sequence to sequence learning, and generates the final output via the last dense layer.

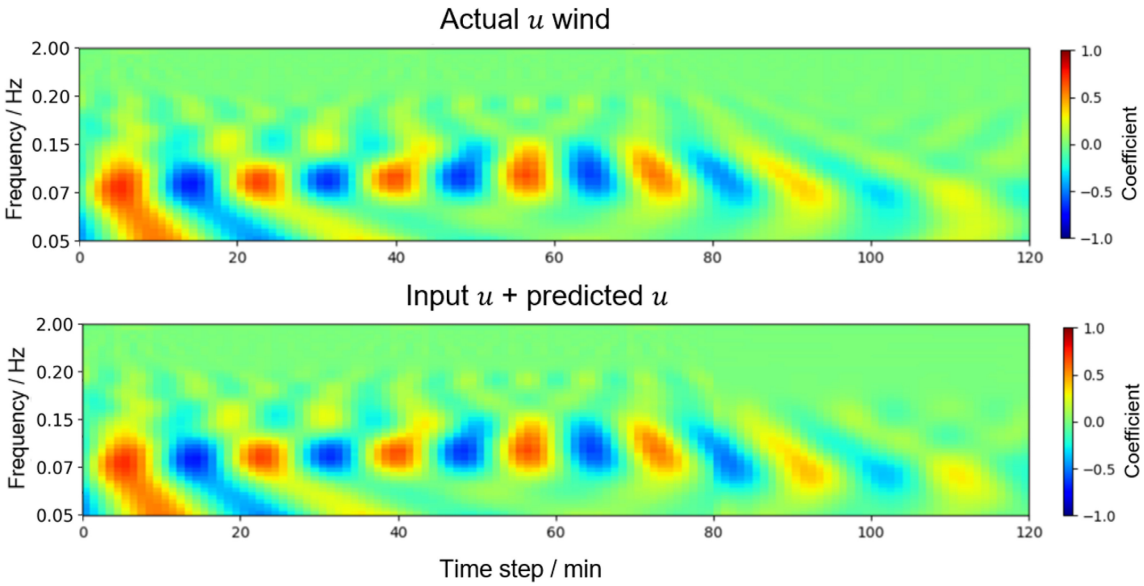


FIGURE 6 | Prediction of coefficient magnitudes using a 2D-CNN-LSTM model trained on CWT-transformed data. The model analyzes patterns up to time step 80 to predict the coefficient dynamics for time steps 80–120.

the 5×5 Conv-2D initially conducts the feature mapping that can capture relatively long-range spatial correlations. The 2×2 max pooling layer located after each convolutional layer improves the interpretation of the mapping by enlarging the effective receptive field and suppressing local noise so that it can mitigate the overfitting of unnecessary patterns (Luo et al. 2016). After hierarchically extracting features from input format with complex patterns in the time-frequency domain, this 2D-CNN model provides these spatial linkages to the subsequent module using LSTM. The resulting feature maps are flattened through the flatten layer and fed into the LSTM network to analyze temporal dependencies. As the sequence to sequence model (Sutskever et al. 2014) is incorporated in the layer, the layer replicates hidden states that implicate condensed information from the former layer and distributes it to the cells in the next layer as additional inputs to retain contextual information over the extended sequence. Figure 6 shows how the 2D-CNN-LSTM predicts the spatial and temporal propagation of the coefficients from left to right where the time steps between 80 to 120-min are a predicted section. From the figure and loss monitoring, it is demonstrated that subtle variability intermittently occurring at periodic intervals can be predicted by detecting appropriate contours at the corresponding time steps.

The proposed model is trained with the data points of the years 2018–2022 using the sliding window method where 2022 data is used for training of the machine learning regressor. The training during this period is intended to compensate for one inherent disadvantage of CMOR-CWT for wind analysis. In signal processing using CMOR-CWT, the term direct current (DC) component refers to the average component of a signal,

14698080, 2026, 2, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70173 by Universitatsbibliothek RWTH Aachen University, Wiley Online Library on [08/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

which does not vary with time. In the context of CWT, the admissibility condition is required to derive the mother wavelet as zero-mean (Foufoula-Georgiou and Kumar 1994; Torrence and Compo 1998), meaning that the wavelet cannot capture the constant DC component of the input wind pattern. Since the DC component is not captured during the CWT due to the zero-mean condition of the wavelet, this part is lost in the process. Consequently, the ICWT reconstruction reflects only the oscillatory part of the wind profile and does not include its original mean. Furthermore, this study confirms that even after compensating for the loss of the DC component by adding the mean of previous time steps, the reconstructed wind field may still appear shifted or sloped compared to the original profile. This suggests that the DC component in wind data is not always a constant offset, but may also contribute a linear trend over time into the very low-frequency domain. However, when the predicted coefficients are used for reconstruction, the nodes exhibiting contour changes of coefficients in the high frequencies are precisely represented by the CWT-ICWT process. Thus, the proposed model conducts a more robust recovery through the additional modules using regression, rather than performing reconstruction using ICWT alone. Moreover, the hybrid model combines models that detect these intervals of dynamic fluctuations using CMOR-CWT and 2D-CNN-LSTM with a single LSTM model that can show the general patterns. Both the single LSTM process and the hybrid model using the 2D-CNN-LSTM flow in the middle of the CWT-ICWT process for each of the u and v components are processed in a parallel architecture as a unified module. By using a machine learning-based regression model, the results from each model in the parallel process are used as input features. Additionally, to reflect DC information that is not a constant value, the average value of past data used as input in the CWT process and the average value of predicted wind over 20 min from the single LSTM are also included as input features for the regression stage, which leverages Light Gradient Boosting Machine (LGBM) to capture nonlinear relationships and maintain robustness against overfitting (Ke et al. 2017).

Figure 7 is a simple diagram of the entire process of the hybrid model, showing that LGBM uses the predicted wind of the single LSTM, the predicted values from the ICWT processes, and additional features defined above. The last additional feature in the regression model is the inclusion of time step indices ranging from 1 to 20. This allows the model to account for the aspect that in short time horizons, predictions tend to rely more on a persistence-likely concept so that the model should place greater weights on the output of the initial LSTM time step. Therefore, adding these indices enables the model to assign different weights to each time step accordingly. After the prediction process, the predicted u and v are converted into final wind speed and wind direction for the deterministic evaluation, while the u and v themselves are fed into the probabilistic nowcasting process.

Figure 8 illustrates predictions of the hybrid model including the full reconstruction process with the postprocessing. In the selected example the wind direction remained relatively steady. The u and v components are individually decomposed and analyzed, and the subsequent propagation of patterns is predicted using the CWT-2D-CNN-LSTM model (second and third rows in Figure 8). The final output obtained through the process described in Figure 7 is depicted in the first and the last rows of Figure 8. This is one of the example days where the LSTM-based model fails to detect dynamic changes as it produces outputs approximating a smoothed profile, which tends to minimize errors. In contrast, the hybrid model attempts to detect patterns dynamically so that it captures a sudden increase in wind speed around time step 100.

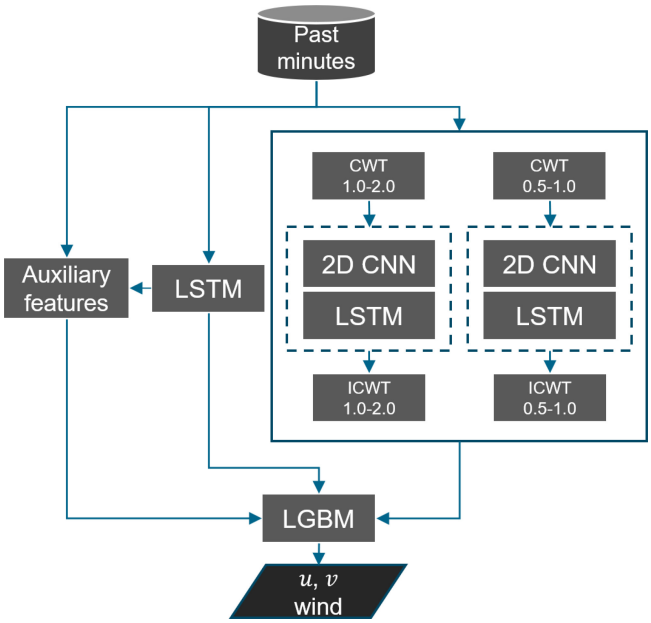


FIGURE 7 | A simplified schematic of the hybrid model is as follows: The single LSTM module and the module utilizing the CWT are trained in parallel. Within the CWT-based module, separate models are constructed for different bandwidths and center frequencies, and within each of these models, the analysis of the u and v components is conducted in parallel. The outputs from these components are then merged and subjected to regression analysis simultaneously using a LGBM model.

2.10 | Probabilistic Prediction Model: Mixture Density Network With Multivariate Distribution

For deterministic prediction, the models were trained using data of the first 4 days of each five-day period, whereas the fifth day in each period was reserved as training data for probabilistic prediction. During the training period, predicted wind is generated on every fifth day using the pretrained deterministic model. The differences between these predicted results and the actual wind are then used to construct the final probabilistic nowcasting model. The deviations and means obtained in this way cannot adequately be described by a Gaussian probability distribution, particularly when the empirical error distribution exhibits high kurtosis with heavy tails, see for example, Castaing et al. (1990). Such error distributions vary dynamically across different time steps depending on the deterministically predicted wind.

Figure 9 illustrates how the distribution of deviations between predicted wind and actual wind differs when the mean and variance from the empirical distribution (blue line) are directly used in the Gaussian distribution (red dotted line). The

14698080, 2026, 2, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70173 by Universitatsbibliothek RWTH Aachen University, Wiley Online Library on [08/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

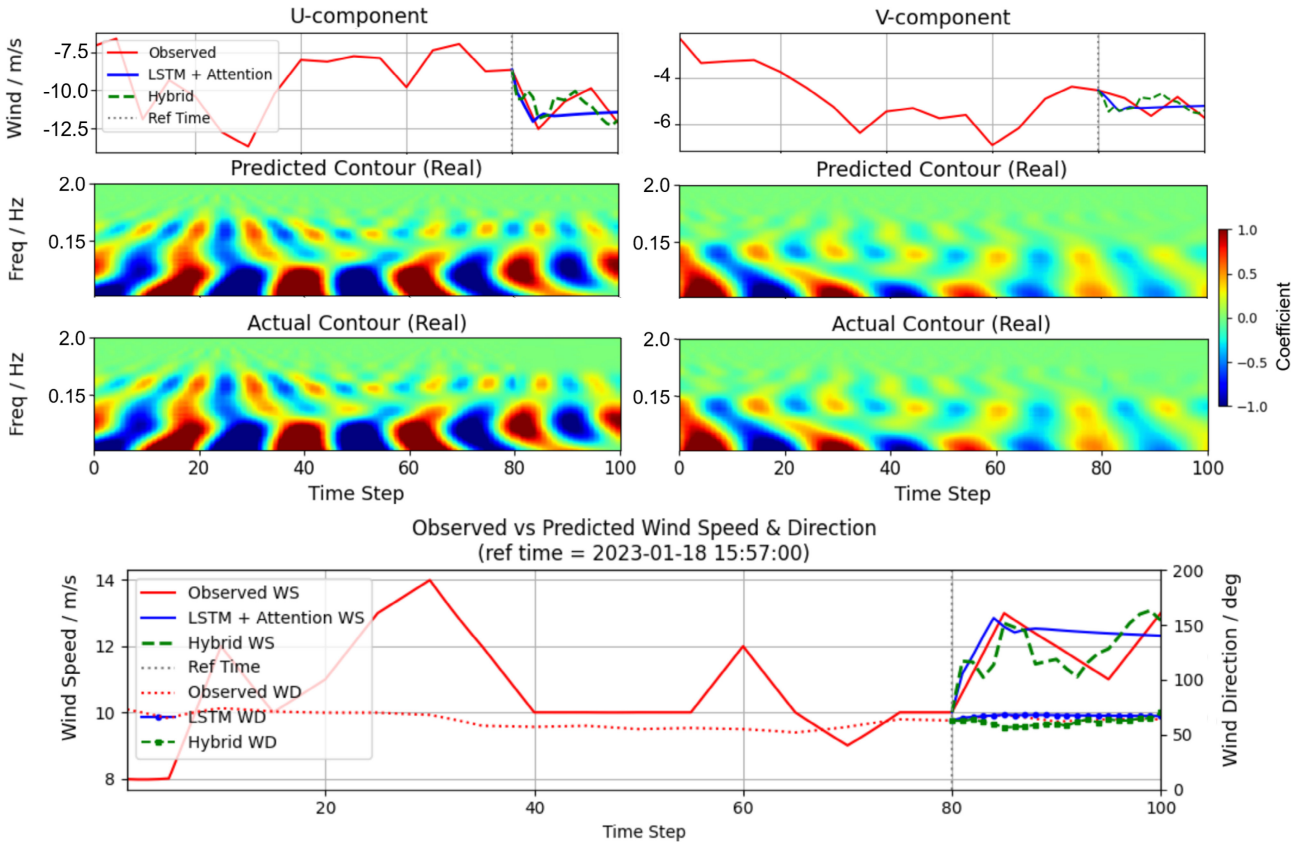


FIGURE 8 | A well-predicted case of the wind speed (WS) and wind direction (WD) over the following 20 min (time step 80–100). This figure focuses on wind speed changes rather than wind direction, as wind direction remained relatively stable during the selected period. The u (left) and v (right) components of the wind direction profile are transformed using CWT, and the real part of the actual contour and the predicted output are visualized. After applying regression-based postprocessing, the final reconstruction results are shown as the green dotted line in the comparison at the lowest graph.

variances of predicted wind directions as a log-scale density where the heavy tail is observed in sub-figure (a) of Figure 9 and the original density in sub-figure (b) of Figure 9 show that the empirical distributions of both time step 1 and 13 still exhibit higher kurtosis and heavier tails than the Gaussian distribution, indicating mode-seeking behavior of the time series machine learning model. The use of a Gaussian Mixture Model (GMM, green line) demonstrates that eight superposed weighted Gaussian distributions are capable of sufficiently representing the empirical distributions. The deviation distribution of the predicted wind speed also displayed characteristics of high kurtosis and heavy tails, which can be successfully fitted using the GMM. Nevertheless, the GMM fitting is not applied directly to wind speed and wind direction in the MDN, but rather to the distribution of the u and v components. Due to its circular nature, where wind direction errors cannot exceed the range of −180° to 180°, it is difficult to apply MDN directly to wind direction. For instance, when there is an absolute error exceeding 180° from the predicted wind, it cannot reliably distinguish the direction of the error relative to the 180-degree threshold, which contributes to the extremely high kurtosis observed in the wind direction deviation distribution. Therefore, the MDN is applied to the transformed u and v components using a multivariate GMM, enabling the reflection of the correct kurtosis variations across different time steps and wind conditions. The model provides the entire conditional probability distribution with GMM parameters for their likelihoods. The MDN learns by maximizing the log-likelihood of the observed target over the predicted distribution, which is described by

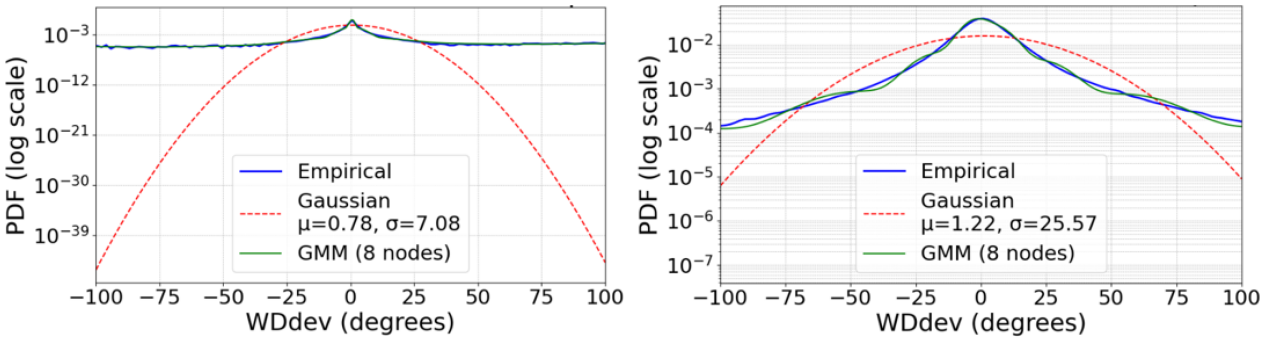
$$\text{Loss} = -\log\left(\sum_{k=1}^{K} \pi_k(x) \cdot \mathcal{N}\left(y \middle| \mu_k(x), \sigma_k(x)\right)\right). \tag{6}$$

Here, $\pi_k(x)$ denotes the mixture weight obtained via the Softmax function, while $\mu_k(x)$ and $\sigma(x)$ represent the mean and standard deviation of the k-th Gaussian component, which are functions depending on x, the deterministic wind prediction. As eight superposed distributions were used in this study, the model was designed with a setting of $K = 8$. Using the computed parameters, the basic distribution equation of GMM in MDN is constructed by substituting them into

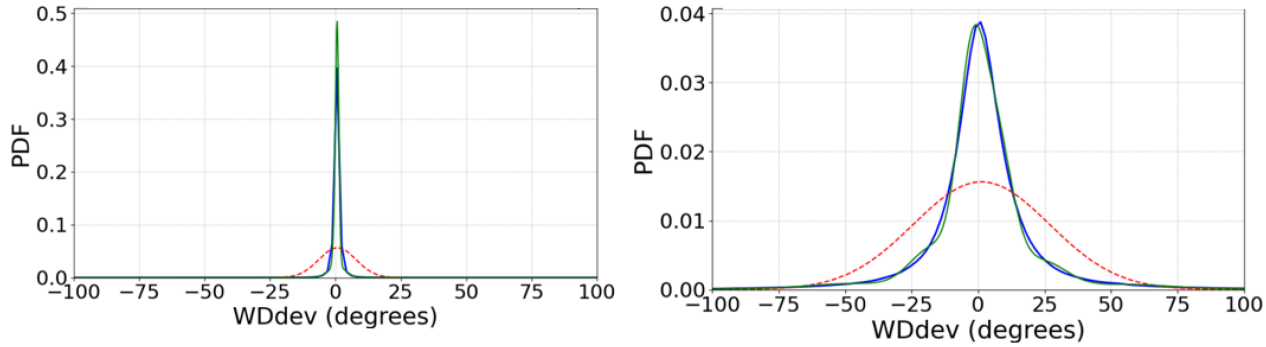
$$p(y|x) = \sum_{k=1}^{K} \pi_k(x) \cdot \frac{1}{\sqrt{2\pi\sigma_k^2(x)}} \exp\left(-\frac{\left(y-\mu_k(x)\right)^2}{2\sigma_k^2(x)}\right). \tag{7}$$

However, this formulation does not account for the covariance trait between u and v. Therefore, a modified GMM approach is adopted:

14698080, 2026, 2, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70173 by Universitatsbibliothek RWTH Aachen University, Wiley Online Library on [08/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License



(a) Log scale of the density distribution at time step 1 (left) and time step 13 (right)



(b) Original scale of the density distribution at time step 1 (left) and time step 13 (right)

FIGURE 9 | Comparison of the GMM fit and Gaussian fit with the distribution of empirical wind direction deviations (WDdev) extracted from the probabilistic prediction module between 2018 and 2022. The density of the logarithmic scale (a) from the original density (b) is presented to better capture the heavy tails.

$$p(y|x)=\sum_{k=1}^{K}\pi_k(x)\frac{1}{(2\pi)^{1/2}|\Sigma_k(x)|^{1/2}}\exp\Big(-\frac{1}{2}(y-\mu_k(x))^{\top}\Sigma_k(x)^{-1}(y-\mu_k(x))\Big),$$
$$\text{with }\Sigma=\begin{pmatrix}\sigma_u^2 & \mathrm{Cov}(v,u)\\ \mathrm{Cov}(u,v) & \sigma_v^2\end{pmatrix} \tag{8}$$

The MDN obtains the GMM based on mixture weights, means, and covariances (Σ) using neural networks. For the implementation, the covariance matrix Σ is flattened into a 1-dimensional list format, which is then used as a feature during MDN training, allowing the model to learn a multivariate distribution. Although the incorporation of multivariate distributions with covariance into the MDN framework is to achieve an optimal probabilistic distribution, an important aspect is to filter only the range that falls within the 90.25% confidence interval among scattered samples based on the predicted GMM equation, while the distributions over the u and v coordinates by the MDN inherently cover the probabilistic ranges of both wind speed and wind direction. When defining the confidence region in this study, the results are confirmed that a density-based approach leads to an incorrect confidence range in the uv plane. O'Hagan et al. (2015) attribute this difficulty to the fact that standard errors are not directly computed in GMMs. Although an alternative solution is to generate more than 100,000 samples from the computed GMM to achieve a more precise confidence region of the distribution, this approach increases computational demands, making it impractical for real-time probabilistic wind prediction. To measure the distance of deterministic prediction points x that can probabilistically occur from a distribution p, the Mahalanobis distance provides an appropriate metric:

$$D_M(x,p)=\sqrt{(x-\mu)^{\top}\Sigma^{-1}(x-\mu)} \tag{9}$$

Unlike the Euclidean distance, the Mahalanobis distance accounts for correlations between variables and scales of the distance based on the variance. It is formally defined as Equation (9), where x represents the point being measured, μ is the mean vector of the distribution. However, since this distance metric is applicable only to a single distribution, an extended approach is required when dealing with superposed distributions. Herein, a weighted average Mahalanobis Distance is proposed.

$$D_M(x,p)=\sqrt{\big(x-\mu_{\text{weighted}}\big)^{\top}\Sigma_{\text{weighted}}^{-1}\big(x-\mu_{\text{weighted}}\big)},$$
$$\text{with }\Sigma_{\text{weighted}}=\sum_{k=1}^{K}\pi_k\Sigma_k,\quad \mu_{\text{weighted}}=\sum_{k=1}^{K}\pi_k\mu_k, \tag{10}$$

where the mixture weights are applied to each covariance matrix and mean vector. The resulting distance metric is employed in the deterministic–probabilistic integrated algorithm to sample the corresponding wind range in the 90.25% confidence interval from 1000 sample points.

14698080, 2026, 2, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70173 by Universitatsbibliothek RWTH Aachen University, Wiley Online Library on [08/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

In addition, managing a numerical instability encountered in this workflow is incorporated in the proposed model. During the computation of mixed probability distributions, numerical failures due to singularities occur particularly when calculating the log-probability in the loss function by its high kurtosis. That phenomenon leads to determinants that are zero or negative (Rossi 2014). Given that such numerical failures are inevitable under the structure of the algorithm developed herein, an additional iteration loop is added into the model, where a small epsilon is progressively added to the covariance matrix whenever an error is encountered. In the iteration, outliers are removed by employing an Isolation Forest (IF) (Liu et al. 2008) as the contamination rate (CR) gradually increases by 0.5% at each iteration. Isolation Forest is an ensemble-based model for outlier detection that works by isolating data, detecting outliers based on the depth of randomly split trees. Moreover, by employing Cholesky decomposition in the middle of each iteration, the process supports the computational efficiency as well.

The aforementioned deterministic and probabilistic modules have been implemented as a unified module. In this design, the deterministic prediction is performed first and the predicted deterministic result is used as input to produce the final probabilistic result. The integrated structure of the wind nowcasting system is illustrated in Figure 10. The real-time module uses the pretrained models stored in files as a pickle format. At each reference time, the windspeed and wind direction of past 20 min are fed into the pretrained models, and the deterministic module produces predicted u and v of wind. These deterministic predictions are then used as inputs to compute probabilistic range of u and v values for each minute up to 20-min time horizon, and these values are subsequently converted into wind speed and wind direction. The results of this integrated algorithm described in Figure 11 are visualized as two-dimensional plots. The spot of the highest contour density corresponds to the deterministic prediction, from which only ranges that fall within the 90.25% confidence interval (red dotted lines) are extracted using the weighted Mahalanobis distance. In particular, at the earlier time step (left) in Figure 11, the density is higher concentrated near the center by the high kurtosis.

2.11 | Evaluation Metrics

To compare the performance of each model, metrics of the Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and standard deviation (Std) are used for deterministic predictions. The error is analyzed separately for each time step. Both MAE and RMSE are used to evaluate the accuracy of the predicted wind speed and direction, where MAE quantifies the magnitude of prediction errors, while RMSE gives greater weight to larger errors. In addition, skill scores are reported with respect to the SP baseline to quantify relative improvements. For ablation analysis, the performance of an LSTM-only model is compared against variants that combine the LSTM backbone with additional modules in order to assess the contribution of each component.

Extreme-event performance is evaluated on the basis of measured wind time series. A baseline trend is defined by a linear fit to the 20-min sequence in the actual wind profile. The wind change amplitude is then defined as the absolute deviation from this baseline. The extreme-event evaluation considers cases in which the amplitude exceeds the selected quantile threshold. For each threshold, categorical verification metrics are computed, and the MAE conditioned on event occurrences is also reported. Thresholds are set from the test set using quantiles as the 50th and 95th percentiles of the amplitude, which represent percentiles of the deviation distance from the baseline trend. A higher percentile means larger variability relative to the fitted baseline, and the goal is to evaluate prediction quality with error values under these highly variable conditions rather than simply counting threshold exceedances.

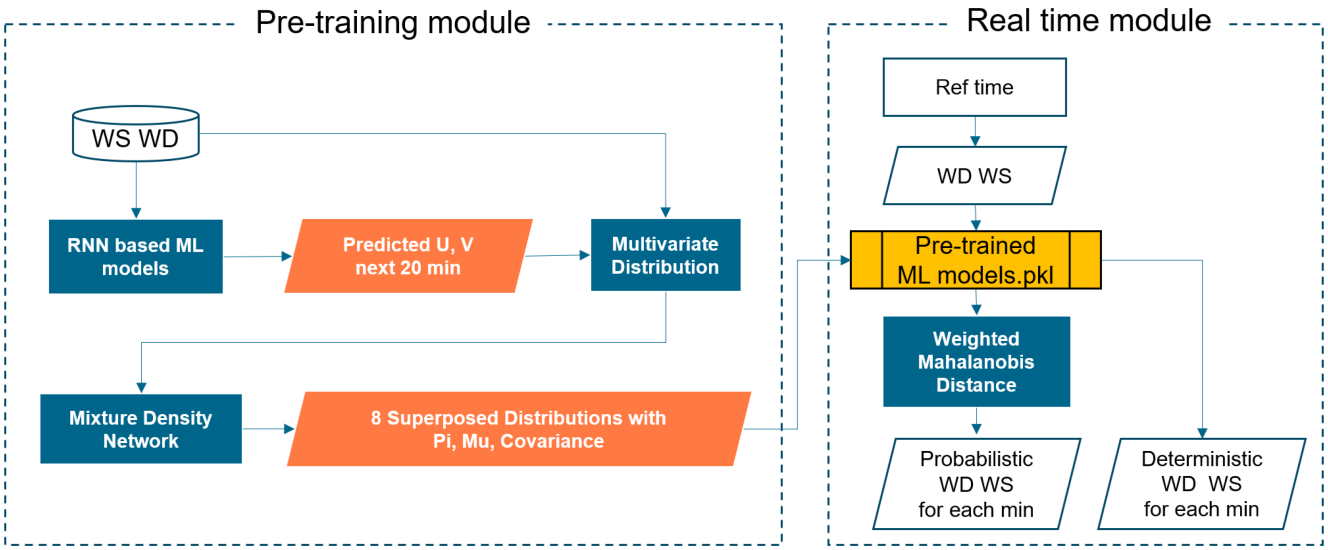


FIGURE 10 | Integrated module of deterministic nowcasting and probabilistic nowcasting. In a unified module, deterministic and probabilistic predictions are conducted in order. The deterministically predicted u and v are fed into multivariate MDN model, which generates superposed probabilities. The trained models are stored as a pickle format and used for real-time prediction. The predicted u and v are then converted as final wind speed (WS) and wind direction (WD).

14698080, 2026, 2, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70173 by Universitatsbibliothek RWTH Aachen University, Wiley Online Library on [08/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

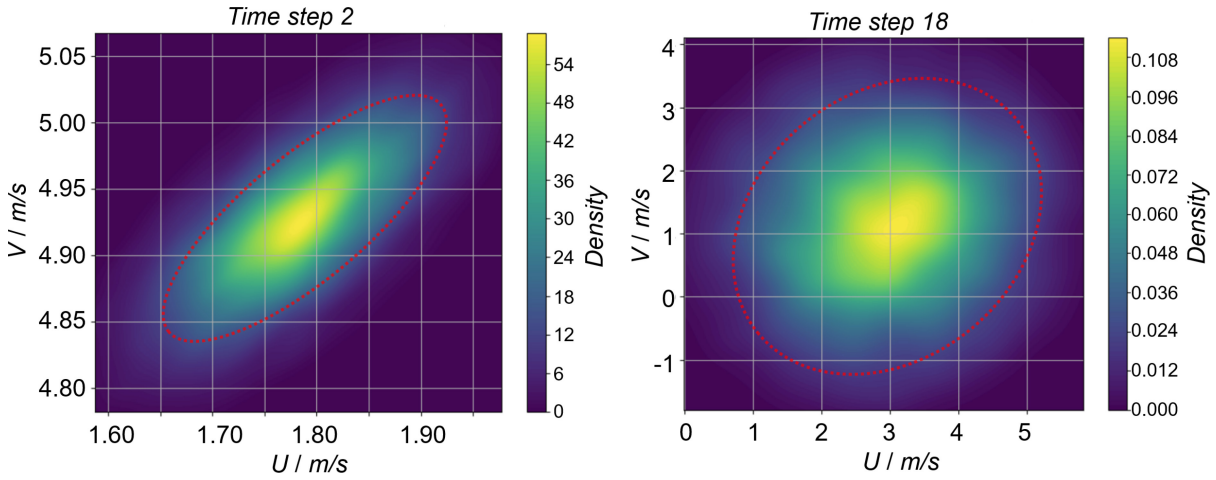


FIGURE 11 | Outcome format of the integrated module that conducts the deterministic and probabilistic nowcasting. The contour describes the probability density and the red dotted lines represent the 90.25% confidence interval. Density values denote probability density per unit area $(m/s)^{-2}$. Grid units are defined by dividing the u and v ranges between minimum and maximum into 100 intervals. Predicting from a reference time 16:20, 11th, January, 2023, the u and v probabilistic wind are generated from 1 to 20 min ahead while the figures depict the probabilistic wind at steps 2 (left) and 18 (right). The location where the contour density reaches its maximum corresponds to the wind from deterministic prediction.

The best performing deterministic model from the evaluation is integrated into the probabilistic wind prediction module. For probabilistic predictions, the SP model provides independent predictions of wind speed and direction, assigning a 95% confidence interval to each. So, it is verified whether both the deterministic wind speed and wind direction fall within the 95% confidence interval. The ratio of correct predictions to the total number of predictions is then calculated as the product of the probabilities, which should be at least 90.25%. Accuracies exceeding 90.25% indicate that the predicted confidence intervals are overly conservative, although this also suggests that the model performs well probabilistically. This implies that the probabilistic prediction ranges cannot be manually adjusted to align with specific research objectives or operational purposes. Under the condition that the predictions fall within the accuracy ranges defined by their respective confidence intervals, the widths of the prediction intervals for wind speed and wind direction are compared in terms of how narrow they are, as narrower intervals indicate greater precision and support uncertainty management. To ensure a consistent comparison of predictive sharpness at the common target joint coverage of 90.25%, the interval width of each model is rescaled by

$$\widetilde{w}_m = w_m \times \frac{90.25\%}{C_m}, \tag{11}$$

where w_m and C_m denote the interval width and the coverage of models m, respectively.

3 | Output Summary

The following test results are based on a model trained with data up to 2022 and evaluated using data from January 2023. Since the results for February through December exhibit similar patterns to those observed in January, only the representative findings from January are presented. Both deterministic and probabilistic nowcasting were trained and tested as a single integrated model under an environment with 80 CPUs and 512 GB of RAM while GPUs were not used. All models could alternatively be trained also with only 20 CPUs and less than 128 GB of RAM, except for the hybrid model. Only the hybrid model required higher memory capacity, as it simultaneously performed parallel processing of CMOR 1.0–2.0 and CMOR 0.5–1.0 in the CWT and ICWT procedure of the 2D-CNN-LSTM structure, as well as running a single LSTM model. However, this can be mitigated by implementing the process in series. The iteration of increasing epsilon and CR within the loop occurred four times between time step 1 and time step 4, indicating that singularities arose due to high kurtosis at the initial time steps. On the other hand, it is interpreted that the deterministic predictions during these time steps exhibit high accuracy.

For the SP model, the baseline was determined based on the persistence concept that yielded the best results when the persistence interval was set as the average of past data values over a certain period. The configuration using the average of the past 2 min outperformed other settings including the 20-min period suggested by Frech et al. (2002). The MAE for wind speed was 1.15 m/s with a 20-min average and 1.04 m/s with a 2-min average, calculated as the mean error over predicted 20 time steps. For wind direction, the 20-min average MAE was 14.54° and 13.21°, respectively. Therefore, the deterministic nowcasting baseline model was established as the SP model using the 2-min average for persistence, and the probabilistic nowcasting of the SP model was also based on 2-min averages.

For the single machine learning (ML) models, tests included evaluation of the impact of input sequence length. Both 20-min and 40-min sequences produced superior results compared to other sequence configurations. In contrast, increasing the sequence length more than 40-min sequences did not lead to further improvement. Considering computational efficiency, the 20-min input sequence was determined for the single MLs. The hybrid model adopted a 160-min past time sequence, which accounts for the padding region that is excluded from the valid

14698080, 2026, 2, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70173 by Universitatsbibliothek RWTH Aachen University, Wiley Online Library on [08/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

prediction interval. These configurations are chosen to ensure that each model is evaluated based on its optimal achievable performance.

TABLE 1 | Comparison of wind speed (WS) and wind direction (WD) MAE between models with and without periodic features over short (step 1–4) and full (step 1–20) prediction horizons.

	Mean MAE (Time step 1–4)		Mean MAE (Time step 1–20)	
	Wind speed (m/s)	**Wind direction (deg)**	**Wind speed (m/s)**	**Wind direction (deg)**
Without periodic features	0.23	**2.30**	0.88	9.92
With periodic features	0.24	2.35	**0.86**	**9.71**

Note: Bold values indicate the lowest mean MAE for the full prediction horizon (time step 1–20).
Abbreviations: MAE, Mean absolute error; WD, Wind direction; WS, Wind speed.

Skill score is computed as $\text{score}(\%) = 100 \cdot \left(1 - \text{metric} / \text{metric}_{\text{ref}}\right)$ for metric $\in$ {MAE, RMSE}.

The use of periodic features on performance was also evaluated. Although the model without periodic features exhibited better accuracy during the initial four time steps, the inclusion of periodic features led to consistently superior capability in subsequent time steps from time step 5 onward. As shown in Table 1, the average of MAE without periodic features from time step 1–4 is 0.23 m/s for wind speed and 2.30° for wind direction, whereas the MAE with periodic features is 0.24 m/s and 2.35°. The model with periodic features achieved lower mean MAE over 20 time steps (0.86 m/s for wind speed and 9.71° for wind direction) compared to the model without periodic features (0.88 m/s for wind speed and 9.92° for wind direction). Thus, the ML models configured with periodic features were used for comparison with the hybrid model.

The results present six methods: the statistical persistence model; the LSTM model; the LSTM-Attention model, which adopts the encoder-decoder structure with an attention mechanism; a bidirectional sequence-to-sequence model as LSTM-bidirectional; the GRU where LSTM cells are replaced with GRU cells; and the hybrid model. In the case of the LSTM-Attention model, only the

TABLE 2 | Mean wind speed (WS, m/s) and wind direction (WD, degree) prediction errors (MAE, RMSE, and Std) over 20-min time steps for different models, with MAE and RMSE based skill scores using SP and LSTM as reference models.

	SP	LSTM	LSTM-attention	LSTM-bidirectional	GRU	Hybrid
MAE						
WS (m/s)	1.04	0.86	0.86	0.90	0.87	0.75
WD (deg)	13.21	9.71	9.68	10.16	9.77	9.06
MAE-based skill score (%, ref = SP)						
WS	—	17.31	17.31	13.46	16.35	27.88
WD	—	26.50	26.72	23.09	26.04	31.42
MAE-based skill score (%, ref = LSTM)						
WS	−20.93	—	0.00	−4.65	−1.16	12.79
WD	−36.05	—	0.31	−4.63	−0.62	6.69
RMSE						
WS (m/s)	1.57	1.30	1.29	1.33	1.31	1.26
WD (deg)	25.74	19.10	19.05	19.64	19.20	19.17
RMSE-based skill score (%, ref = SP)						
WS	—	17.20	17.83	15.29	16.56	19.75
WD	—	25.80	25.99	23.70	25.41	25.52
RMSE-based skill score (%, ref = LSTM)						
WS	−20.77	—	0.77	−2.31	−0.77	3.08
WD	−34.76	—	0.26	−2.83	−0.52	−0.37
Std						
WS (m/s)	1.64	1.29	1.29	1.31	1.29	1.27
WD (deg)	25.33	18.65	18.53	19.01	18.95	18.83

Abbreviations: MAE, mean absolute error; RMSE, root mean square error; Std, standard deviation; WD, wind direction error (deg); WS, wind speed error (m/s).

14698080, 2026, 2, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70173 by Universitatsbibliothek RWTH Aachen University, Wiley Online Library on [08/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

attention mechanism was attached to the LSTM model while keeping all other structures identical. An ablation study compares the baseline LSTM with LSTM-Attention, LSTM-bidirectional, and the hybrid model to quantify the impact of architectural enhancements, and the results are presented as skill scores in Table 2.

Table 2 presents the mean MAE, RMSE, and Std values across all time steps for each wind speed (WS) and wind direction (WD). Overall, the RMSE exhibits a similar trend to the MAE, however, for the hybrid model in wind direction prediction, the MAE improved whereas the RMSE and Std increased. Despite this, the hybrid model still exhibits the largest proportion of improvement. The LSTM without the attention function led to a slight degradation in performance. For wind speed, the MAE remains unchanged at a meaningful decimal place, whereas for wind direction, the MAE decreases marginally from 9.71° (without attention) to 9.68° (with attention). The GRU model, which uses GRU cells in place of LSTM cells in the same architecture of the LSTM model, yielded a wind speed MAE of 0.87 m/s and a wind direction MAE of 9.72°. Compared to the LSTM model, the GRU model shows similar or slightly inferior error over 20-time steps. However, in time step-wise analysis, the GRU model outperformed the LSTM model from first time step to fourth time step. However, the improvements are not substantial enough to have a meaningful impact on operational WVWS-V as in situ applications. In case of the LSTM-bidirectional model, the performance of all time steps underperformed the other ML models. Compared to the SP baseline, the LSTM model reduces MAE by 17.31% (WS) and 26.50% (WD), while the Hybrid model reduces MAE by 27.88% (WS) and 31.42% (WD). In terms of RMSE, the LSTM model improves over the SP baseline by 17.20% (WS) and 25.80% (WD), whereas the Hybrid model improves over the SP baseline by 19.75% (WS) and 25.52% (WD), and relative to the LSTM model it reduces WS RMSE by 3.08% but increases WD RMSE by 0.37%.

Table 3 summarizes extreme event performance for the SP, LSTM, and Hybrid models using amplitude thresholds of 0.26/1.15 m/s for WS and 1.86/12.72° for WD at the 50th/95th percentiles. The test set was evaluated using F1, CSI, Precision, Recall, and conditional MAE. As the threshold percentile increases, the Hybrid model attains higher F1, CSI, and Precision than the SP and LSTM models but exhibits lower Recall, while the Hybrid model's conditional MAE decreases. While the Hybrid output fluctuates more, these oscillations are often fragmented in time, which makes the model less consistent in covering the full duration of extreme conditions. Consequently, Recall decreases because extreme intervals are only partially detected even when peak magnitudes are occasionally captured. At the 50th percentile thresholds, the Hybrid model shows lower F1 and CSI than the LSTM model, indicating that the LSTM model captures more moderate events. This suggests a more conservative behavior for the SP and LSTM baselines under extreme conditions.

Figure 12 shows the development of the MAE of the deterministic nowcasting models over the 20-min prediction range. The ML models exhibit superior performance at the initial time steps, but their errors tend to approach toward those of the SP model as time progresses. In contrast, the proposed hybrid model shows lower degradation of the performance over time, whereas it exhibits an inferior performance compared to the LSTM-based and GRU model up to time step 3. The hybrid model generates a profile that is less smooth over the time steps. This characteristic arises from the incorporation of the time step indices in the postprocessing stage.

The probabilistic predictions based on the MDN were evaluated with and without the use of a multivariate distribution. In Figure 13 (right), the univariate MDN, which models u and v

TABLE 3 | Classification performance MAE at the 50th and 95th percentile thresholds for wind speed and wind direction. 50th/95th thresholds are shown as the maximum threshold value among models for each percentile.

	Model	F1	CSI	Precision	Recall	MAE
WS (m/s)						
50th (thr = 0.26 m/s)	Hybrid	0.64	0.47	0.55	0.77	0.58
	LSTM	0.68	0.51	0.57	0.84	0.62
	SP	0.64	0.47	0.51	0.84	0.84
95th (thr = 1.15 m/s)	Hybrid	0.21	0.11	0.14	0.43	0.89
	LSTM	0.20	0.11	0.12	0.57	0.97
	SP	0.13	0.07	0.08	0.53	1.14
WD (deg)						
50th (thr = 1.86°)	Hybrid	0.67	0.51	0.57	0.82	9.50
	LSTM	0.70	0.53	0.58	0.88	9.52
	SP	0.67	0.50	0.53	0.89	14.98
95th (thr = 12.72°)	Hybrid	0.33	0.20	0.22	0.70	21.80
	LSTM	0.31	0.19	0.19	0.80	23.24
	SP	0.20	0.11	0.12	0.79	30.36

Abbreviation: CSI, Critical Success Index.

14698080, 2026, 2, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70173 by Universitatsbibliothek RWTH Aachen University, Wiley Online Library on [08/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

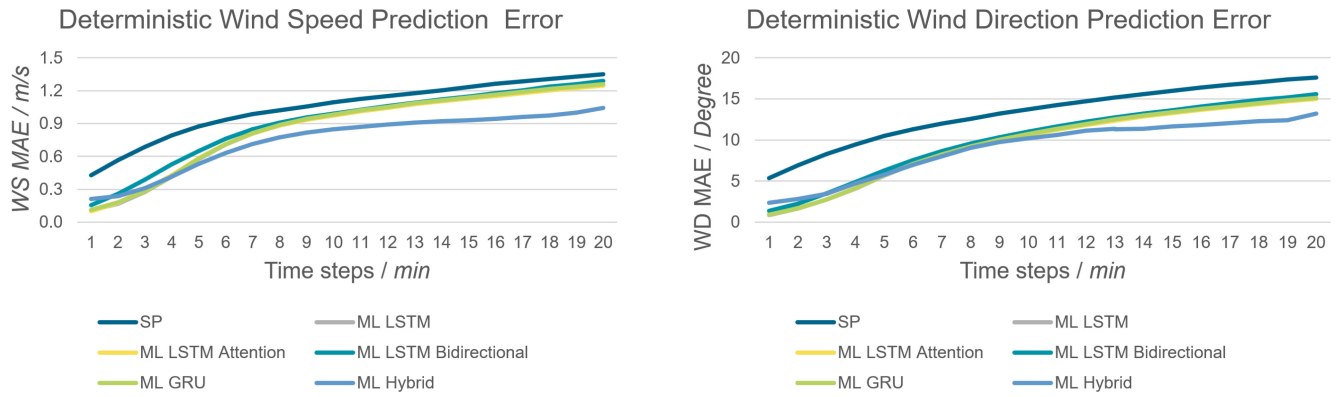


FIGURE 12 | Mean Absolute Error (MAE) comparison of deterministic nowcasting models. The MAE of wind direction (left) and wind speed (right) are computed by each comparison model.

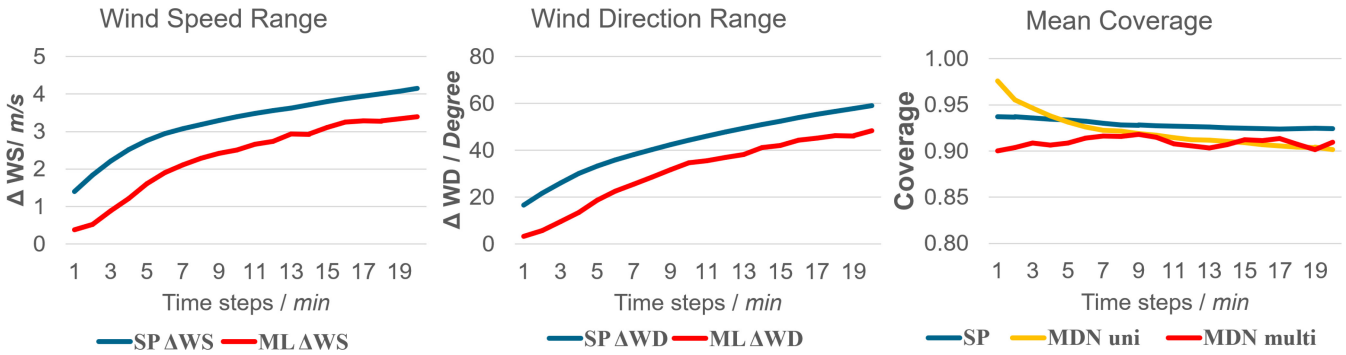


FIGURE 13 | Comparison of probabilistic prediction intervals, where narrower intervals demonstrate superior model performance (left and center). CI coverage is shown for the SP model and MDN variants with (MDN multi) and without (MDN uni) u–v wind feature dependence (right).

independently following Equation (7), produces conservative uncertainty estimates at the initial time steps. By contrast, the multivariate MDN, formulated with Equation (8), yields interval coverage consistent with the prescribed confidence level. The comparison focuses on the SP baseline and the MDN-based ML model trained using the hybrid model outputs that achieved the best deterministic nowcasting performance. The SP model exhibits conservative coverage across all time steps, achieving an average coverage of 92.88%. In contrast, the univariate MDN (MDN-uni) attains a comparable mean coverage of 92.17%, but produces overly conservative coverage at the earliest time steps. The multivariate MDN (MDN-multi) yields a lower mean coverage of 90.93%, which is closer to the target joint coverage of 90.25% for the coupled WS and WD uncertainty. The early time discrepancy between MDN-uni and MDN-multi is illustrated in Figure 11. The initial step uncertainty contours have a flatter elliptical shape, and their shape is governed by the u–v covariance. When covariance is neglected and a circular CI is used to represent the 90.25% coverage, the resulting area exceeds the given distributional area, thereby inflating the enclosed probability area and leading to overly conservative coverage. As lead time increases, the uncertainty ellipse becomes more circular, and the MDN-uni coverage correspondingly approaches the target CI level. Figure 13 (left and center) compares the predicted ranges of wind direction (ΔWD) and wind speed (ΔWS). The range is defined as the width extending symmetrically around the deterministic prediction, providing a measure to evaluate how narrow the prediction uncertainty is contained within the accuracy under the given confidence level. The analysis results show that the confidence intervals exhibit similar distribution characteristics in both positive and negative directions, with no observed significant bias toward one side. Compared with the SP baseline, the ML-based model reduces the mean interval width from 42.95° to 30.90° for ΔWD and from 3.24 to 2.33 m/s for ΔWS, corresponding to reductions of 12.05° and 0.91 m/s, respectively. The degrees of improvement are ΔWD 28.06% and ΔWS 28.09%. The ML-based model showed narrower ΔWD and ΔWS values over the whole prediction horizon compared to the SP-based model while guaranteeing the given coverage. However, the SP model already achieved the higher coverage, and therefore wider probabilistic ranges are naturally expected. After the coverage scaling, the ML-based model reduces the mean width from 41.73° to 30.67° for ΔWD and from 3.15 to 2.31 m/s for ΔWS, corresponding to 26.51% and 26.54% reduction in both metrics.

4 | Discussion and Conclusions

This study investigates both deterministic and probabilistic approaches for wind nowcasting, focusing on the prediction of wake vortex behavior to enhance the safety of small aircraft operations under adverse meteorological conditions. The best-performing model will be integrated into the Wake Vortex Warning System for Vertiports (WVWS-V) which is supposed to enable safe air taxi operations in the vicinity of a conventional runway.

14698080, 2026, 2, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70173 by Universitatsbibliothek RWTH Aachen University, Wiley Online Library on [08/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

For the single time series models, using 20 time steps as the input sequence yields the best performance, suggesting that longer sequences do not effectively capture additional historical patterns in univariate analysis. With the time sequence, the comparison of LSTM and GRU models indicates that the simpler gate structure of GRU provides slightly better performance in the early time steps, whereas LSTM outperforms with respect to the average MAE and RMSE of the entire 20 min time series. However, the methods using single time series models are unable to effectively capture repetitive ramp events blended in the general trend, which illustrate a common tendency of overlooking outliers to prioritize minimizing overall error. In particular, the potential impact of such rare phenomena is directly linked to safety when operating eVTOLs at vertiports; thus, incorporating the effects of disturbances such as gusts is essential for a reliable and safe WVWS-V.

In this regard, the hybrid model that combines complex mother wavelet-based CWT, 2D-CNN, LSTM, and LGBM as postprocess exhibits more accurate results, particularly in mitigating error growth over time. The model is capable of detecting intermittently periodic sparks occurring within the general pattern. Although the model can only predict patterns that exhibit regular behavior, such as contour patterns of CWT-derived coefficients that decay, grow, or stay constant over time, the hybrid model suggests better performance in analyzing and predicting longer time periods compared to using LSTM or GRU alone. This results in superior performance for the deterministic nowcasting of wind speed. However, while an improvement in MAE is observed for wind direction, the gain in RMSE is limited. This is attributed to the fact that the module from CWT to ICWT processes the u and v components independently. Additionally, this limitation is due to the characteristics of the LGBM model itself, which cannot produce two interdependent output dimensions. While the individual values of u and v are substantially corrected, the wind direction, represented by the ratio between these two components, exhibits only a relatively slight improvement. Moreover, due to the padding section set by the scale-based ranges, these repetitive hidden patterns must persist for at least 20 min to contribute to the effect on performance improvement.

The results of probabilistic nowcasting show the advantage of ML-based approaches using the multivariate MDN framework that achieves the predefined confidence levels. The ML models provide narrower 95% confidence intervals of wind speed and wind direction compared to the SP-based model. Notably, the confidence intervals are devoid of a substantial imbalance between positive and negative directions from the deterministic value, indicating the absence of significant bias. Simultaneously, the MDN functions in identifying subtle biases, contributing to the reduction of the probabilistic width. The results further indicate that using a multivariate MDN, rather than a univariate formulation, better captures the coupled uncertainty between the components u and v for the wind speed and wind direction through their joint dependency structure, leading to more reliable uncertainty quantification at the target confidence level. The singularities observed at the early time steps during the MDN training process reveal that the deterministic prediction results are highly consistent with the actual values. Although mathematical computation failures occurred in some cases due to these singularities, it was confirmed that gradually increasing the epsilon and the CR of Isolation Forest resolved the issue without any degradation in performance. In particular, by utilizing the weighted average Mahalanobis distance allows to extract the points in the generated sample that fall within the corresponding confidence interval. These findings demonstrate that combining the deterministic ML-based model with the multivariate GMM within the MDN outperforms the SP model in probabilistic wind nowcasting.

There remain areas that require further improvement. First, the data resolution as minute-level was generated by linearly interpolating the interval of missing data that is available at a several-minute resolution. This approach likely attenuated certain frequency components of the wind, resulting in their ineffective detection in time series analysis using wavelet transform. To address this limitation, it is necessary to utilize higher-resolution data, with the raw data having a temporal resolution of at least 1 min. In addition, instead of analyzing the u and v components separately during the CWT process, employing methods such as multivariate wavelet analysis or cross-wavelet analysis may serve as alternatives to better account for the interactions between these components. Updating the postprocessing stage of the hybrid model can be regarded as a possible approach to enhance the regression performance for wind direction prediction. However, herein incorporating regression methods that jointly learn the dependencies of u and v tended to degrade the performance of wind speed prediction. For this reason, such approaches were excluded from this comparison. Future work may explore methodologies that complement both aspects simultaneously or consider ensemble approaches that aggregate the two different types of regressions for wind speed and wind direction predictions.

The findings of this study suggest that employing the pipeline of the hybrid model enables deterministically accurate and probabilistically stable predictions, thereby contributing to the reliability of WVWS-V and the safe operation of eVTOLs at vertiports.

Author Contributions

Sooseok Lee was responsible for the design of the methodology, the simulation experiments, and the drafting of the original manuscript. Frank Holzäpfel, Anton Stephan, and Dieter Moormann provided supervision and contributed to manuscript revisions.

Acknowledgments

This research was funded by the German Aerospace Center (DLR). This work was conducted within the DLR SMART-MET project as part of a wind-nowcasting development effort for integration into the Wake Vortex Warning System for Vertiports (WVWS-V). Computational resources were provided by the high-performance computing (HPC) platforms: the TerraByte system at the Leibniz Supercomputing Centre (LRZ) and the Computer for Advanced Research in Aerospace (CARA) at DLR. Open Access funding enabled and organized by Projekt DEAL.

Funding

This work was supported by the German Aerospace Center (DLR) under internal research funding for the SMART-MET project.

14698080, 2026, 2, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70173 by Universitatsbibliothek RWTH Aachen University, Wiley Online Library on [08/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

All wind-anemometer data from Rome Fiumicino Airport (FCO) used in this study were provided to the authors by UrbanV. The official owner of the data is Ente Nazionale per l'Assistenza al Volo S.p.A. (ENAV S.p.A), and the data are locally managed by Aeroporti di Roma (AdR).

References

Arts, L. P. A., and E. L. van den Broek. 2022. "The Fast Continuous Wavelet Transformation (Fcwt) for Real-Time, High-Quality, Noise-Resistant Time-Frequency Analysis." *Nature Computational Science* 2: 47–58. https://doi.org/10.1038/s43588-021-00183-z.

Bahdanau, D., K. Cho, and Y. Bengio. 2015. "Neural Machine Translation by Jointly Learning to Align and Translate." *in Proceedings of the International Conference on Learning Representations (ICLR)*. https://arxiv.org/abs/1409.0473.

Barjasteh, A., S. H. Ghafouri, and M. Hashemi. 2024. "A Hybrid Model Based on Discrete Wavelet Transform and Bidirectional Recurrent Neural Networks for Wind Speed Prediction." *Engineering Applications of Artificial Intelligence* 127: 107340. https://doi.org/10.1016/j.engappai.2023.107340. https://www.sciencedirect.com/science/article/pii/S0952197623015245.

Berrezzek, F., K. Khelil, and T. Bouadjila. 2019. "Efficient Wind Speed Forecasting Using Discrete Wavelet Transform and Artificial Neural Networks." *Revue D'intelligence Artificielle* 33, no. 6: 447–452. https://doi.org/10.18280/ria.330607.

Bishop, C. M. 1994. *Mixture Density Networks*. Aston University. NCRG/4288. https://www.microsoft.com/en-us/research/publication/mixture-density-networks/.

Bock, M., K. Yassin, H. Kassem, J. Theron, L. J. Lukassen, and J. Peinke. 2024. "Intermittency, an Inevitable Feature for Faster Convergence of Large Eddy Simulations." *Physics of Fluids* 36, no. 5: 055149. https://doi.org/10.1063/5.0202514.

Bundesstelle für Flugunfalluntersuchung (BFU). 2022. *Zwischenbericht: Unfall Cessna 172 r und Robinson r 44 Raven i Bei Aachen-Merzbrück (Würselen)*, 28–37. Bundesstelle für Flugunfalluntersuchung. https://www.bfu-web.de/DE/Publikationen/Bulletins/2022/Bulletin2022-07.pdf.

Cao, L., Z. Zhang, S. Tang, X. Wang, and Y. Zhang. 2017. "Spatial-Temporal Wind Field Prediction by Artificial Neural Networks." https://arxiv.org/abs/1712.05293.

Castaing, B., Y. Gagne, and E. J. Hopfinger. 1990. "Velocity Probability Density Functions of High Reynolds Number Turbulence." *Physica D: Nonlinear Phenomena* 46, no. 2: 177–200. https://doi.org/10.1016/0167-2789(90)90035-N.

Cho, K., B. van Merriënboer, C. Gulcehre, et al. 2014. "Learning Phrase Representations Using RNN Encoder-Decoder for Statistical Machine Translation." https://arxiv.org/abs/1406.1078.

Daubechies, I. 1988. "Orthonormal Bases of Compactly Supported Wavelets." *Communications on Pure and Applied Mathematics* 41, no. 7: 909–996.

De Moortel, I., S. A. Munday, and A. W. Hood. 2004. "Wavelet Analysis: The Effect of Varying Basic Wavelet Parameters." *Solar Physics* 222, no. 2: 203–228. https://doi.org/10.1023/B:SOLA.0000043578.01201.2d.

Demirtop, A., and O. Sevli. 2024. "Wind Speed Prediction Using LSTM and ARIMA Time Series Analysis Models: A Case Study of Gelibolu." *Turkish Journal of Engineering* 8, no. 3: 524–536. https://doi.org/10.31127/tuje.1431629.

Dempster, A. P., N. M. Laird, and D. B. Rubin. 1977. "Maximum Likelihood From Incomplete Data via the EM Algorithm." *Journal of the Royal Statistical Society, Series B: Statistical Methodology* 39, no. 1: 1–38.

Elman, J. L. 1990. "Finding Structure in Time." *Cognitive Science* 14, no. 2: 179–211. https://doi.org/10.1207/s15516709cog1402_1.

EUROCONTROL. 2024. "European Wake Turbulence Categorisation and Separation Minima on Approach and Departure (RECAT-EU Edition 2.0)." EUROCONTROL. Edition 2.0. https://www.eurocontrol.int/publication/european-wake-turbulence-categorisation-and-separation-minima-approach-and-departure.

Foufoula-Georgiou, E., and P. Kumar. 1994. "Wavelet Analysis in Geophysics: An Introduction." In *Wavelets in Geophysics*, edited by E. Foufoula-Georgiou and P. Kumar, 1–43. San DiegoAcademic Press.

Frech, M., F. Holzäpfel, T. Gerz, and J. Konopka. 2002. "Short-Term Prediction of the Horizontal Wind Vector Within a Wake Vortex Warning System." *Meteorological Applications* 9, no. 1: 9–20. https://doi.org/10.1017/S1350482702001020.

Gerz, T., F. Holzäpfel, M. Frech, et al. 2009. "The Wake Vortex Prediction and Monitoring System WSVBS Part ii: Performance and Atc Integration at Frankfurt Airport." *Air Traffic Control Quarterly* 17, no. 4: 323–346. https://doi.org/10.2514/atcq.17.4.323.

Graves, A., S. Fernández, and J. Schmidhuber. 2005. "Framewise Phoneme Classification With Bidirectional LSTM and Other Neural Network Architectures." *Neural Networks* 18, no. 5–6: 602–610. https://doi.org/10.1016/j.neunet.2005.06.042.

Grossmann, A., and J. Morlet. 1984. "Decomposition of Hardy Functions Into Square Integrable Wavelets of Constant Shape." *SIAM Journal on Mathematical Analysis* 15, no. 4: 723–736. https://doi.org/10.1137/0515056.

Gu, Z. X., J. Wu, W. Jin, and G. J. Han. 2024. "MotionUVRNN: A Motion Capture Recurrent Neural Network for Wind Field Forecast Correction via U and V Components Modeling." *IAENG International Journal of Computer Science* 51, no. 7: 791–800. https://www.iaeng.org/IJCS/issues_v51/issue_7/IJCS_51_7_09.pdf.

Gurke, T., and H. Lafferton. 1997. "The Development of the Wake Vortices Warning System for Frankfurt Airport: Theory and Implementation." *Air Traffic Control Quarterly* 5, no. 1: 3–29.

Haar, A. 1910. "Zur Theorie Der Orthogonalen Funktionensysteme." *Mathematische Annalen* 69: 331–371.

Hinton, D. A. 2001. "Description of Selected Algorithms and Implementation Details of a Concept-Demonstration Aircraft Vortex Spacing System (Avoss)." NASA Langley Research Center. Technical Memorandum NASA/TM-2001-211027. https://ntrs.nasa.gov/api/citations/20010071526/downloads/20010071526.pdf.

Hochreiter, S., and J. Schmidhuber. 1997. "Long Short-Term Memory." *Neural Computation* 9, no. 8: 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735.

Holzäpfel, F., T. Gerz, M. Frech, et al. 2009. "The Wake Vortex Prediction and Monitoring System WSVBS Part i: Design." *Air Traffic Control Quarterly* 17, no. 4: 301–322. https://doi.org/10.2514/atcq.17.4.301.

Holzäpfel, F., L. Strauss, and C. Schwarz. 2021. "Assessment of Dynamic Pairwise Wake Vortex Separations for Approach and Landing at Vienna Airport." *Aerospace Science and Technology* 112: 106618. https://doi.org/10.1016/j.ast.2021.106618.

Ke, G., Q. Meng, T. Finley, et al. 2017. "LightGBM: A Highly Efficient Gradient Boosting Decision Tree." in Advances in Neural Information Processing Systems. 30. https://proceedings.neurips.cc/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf.

Kralj, L., M. Hultman, and H. Lenasi. 2024. "Wavelet Analysis and the Cone of Influence: Does the Cone of Influence Impact Wavelet Analysis

14698080, 2026, 2, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70173 by Universitatsbibliothek RWTH Aachen University, Wiley Online Library on [08/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Results?" *Applied Sciences* 14, no. 24: 11736. https://doi.org/10.3390/app142411736.

Kumar, V. B., V. M. Nookesh, B. S. Saketh, S. Syama, and J. Ramprabhakar. 2021. "Wind Speed Prediction Using Deep Learning-LSTM and GRU." in *Proceedings of the 2nd International Conference on Smart Electronics and Communication (ICOSEC 2021)* (Bengaluru, IEEE). 602–607.

LeCun, Y., L. Bottou, Y. Bengio, and P. Haffner. 1998. "Gradient-Based Learning Applied to Document Recognition." *Proceedings of the IEEE* 86, no. 11: 2278–2324. https://doi.org/10.1109/5.726791.

Liang, S., L. Nguyen, and F. Jin. 2018. "A Multi-Variable Stacked Long-Short Term Memory Network for Wind Speed Forecasting." https://arxiv.org/abs/1811.09735.

Lindemann, B., B. Maschler, N. Sahlab, and M. Weyrich. 2021. "A Survey on Anomaly Detection for Technical Systems Using LSTM Networks." *Computers in Industry* 131: 103498. https://arxiv.org/abs/2105.13810.

Liu, F. T., K. M. Ting, and Z. H. Zhou. 2008. "Isolation Forest." in *Proceedings of the 2008 Eighth IEEE International Conference on Data Mining (ICDM)* 413–422.

Liu, Y., L. Guan, C. Hou, et al. 2019. "Wind Power Short-Term Prediction Based on LSTM and Discrete Wavelet Transform." *Applied Sciences* 9, no. 6: 1108. https://doi.org/10.3390/app9061108.

Luo, W., Y. Li, R. Urtasun, and R. Zemel. 2016. "Understanding the Effective Receptive Field in Deep Convolutional Neural Networks." *Advances in Neural Information Processing Systems* 29: 4898–4906.

Manero, J., J. Béjar, and U. Cortés. 2022. "Wind Prediction Using Deep Learning and High Performance Computing." *High Performance Computing (CARLA 2021)* 1540: 193–207.

Mannelli, A., F. Papi, G. Pechlivanoglou, G. Ferrara, and A. Bianchini. 2021. "Discrete Wavelet Transform for the Real-Time Smoothing of Wind Turbine Power Using Li-Ion Batteries." *Energies* 14, no. 8: 2184. https://doi.org/10.3390/en14082184. https://www.mdpi.com/1996-1073/14/8/2184.

Mohammed, M. A., and L. A. Ahmed. 2023. "Forecasting Wind Speed Using the Proposed Wavelet Neural Network." *Discrete Dynamics in Nature and Society* 2023: 9940038. https://doi.org/10.1155/2023/9940038.

Morlet, J., G. Arens, E. Fourgeau, and D. Giard. 1982. "Wave Propagation and Sampling Theory—Part I: Complex Signal and Scattering in Multilayered Media." *Geophysics* 47, no. 2: 203–221. https://doi.org/10.1190/1.1441328.

National Transportation Safety Board. 2022. "Aviation Investigation Final Report: Cessna 120, n76646, Upland, California." National Transportation Safety Board. Accident date: January 3, 2022 (12:05 local). Original publish date: June 22, 2022. Aviation Investigation Final Report WPR22LA072. https://data.ntsb.gov/carol-repgen/api/Aviation/ReportMain/GenerateNewestReport/104480/pdf.

O'Hagan, A., K. Burke, and I. C. Gormley. 2015. "Investigation of Parameter Uncertainty in Clustering Using a Gaussian Mixture Model via Jackknife, Bootstrap and Weighted Likelihood Bootstrap." https://arxiv.org/abs/1510.00551.

Olhede, S. C., and G. Metikas. 2009. "The Monogenic Wavelet Transform." *IEEE Transactions on Signal Processing* 57, no. 9: 3426–3441. https://doi.org/10.1109/TSP.2009.2023397.

Pei, C., Y. Bao, X. Zhang, X. Cheng, and J. Feng. 2024. "A CNN-LSTM Model for Predicting Wind Speed in Non-Stationary Wind Fields in Mountainous Areas Based on Wavelet Transform and Adaptive Programming." *AIP Advances* 14, no. 11: 115009. https://doi.org/10.1063/5.0230026.

Prabha, P. P., V. Vanitha, and R. Resmi. 2019. "Wind Speed Forecasting Using Long Short Term Memory Networks." *in Proceedings of the 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT)* (IEEE). 1310–1314.

Ren, J., Z. Yu, G. Gao, G. Yu, and J. Yu. 2022. "A CNN-LSTM-LightGBM Based Short-Term Wind Power Prediction Method Based on Attention Mechanism." *Energy Reports* 8: 437–443. https://doi.org/10.1016/j.egyr.2022.02.206.

Rossi, P. E. 2014. *Bayesian Non- and Semi-Parametric Methods and Applications*. Princeton University Press. https://assets.press.princeton.edu/chapters/s10259.pdf.

San, K. M., J. G. Singh, and K. P. Natarajan. 2023. "Wind Speed Forecasting Using Hybrid Model of CNN and LSTM With Wavelets." *in Proceedings of the 2023 International Conference in Advances in Power, Signal, and Information Technology (APSIT)* 297–301.

Sari, A. P., H. Suzuki, T. Kitajima, T. Yasuno, D. A. Prasetya, and R. Arifuddin. 2022. "Short-Term Wind Speed and Direction Forecasting by 3DCNN and Deep Convolutional LSTM." *IEEJ Transactions on Electrical and Electronic Engineering* 17, no. 11: 1620–1628. https://doi.org/10.1002/tee.23669.

Schuster, M., and K. K. Paliwal. 1997. "Bidirectional Recurrent Neural Networks." *IEEE Transactions on Signal Processing* 45, no. 11: 2673–2681. https://doi.org/10.1109/78.650093.

Selesnick, I. W., R. G. Baraniuk, and N. G. Kingsbury. 2005. "The Dual-Tree Complex Wavelet Transform." *IEEE Signal Processing Magazine* 22: 123–151. https://doi.org/10.1109/MSP.2005.1550194.

Shao, B., D. Song, G. Bian, and Y. Zhao. 2021. "Wind Speed Forecast Based on the LSTM Neural Network Optimized by the Firework Algorithm." *Advances in Materials Science and Engineering* 2021: 1–13. https://doi.org/10.1155/2021/4874757.

Sutskever, I., O. Vinyals, and Q. V. Le. 2014. "Sequence to Sequence Learning With Neural Networks." *Advances in Neural Information Processing Systems* 27: 3104–3112. https://arxiv.org/abs/1409.3215.

Tekin, S. F., A. Fazla, and S. S. Kozat. 2021. "Numerical Weather Forecasting Using Convolutional-LSTM With Attention and Context Matcher Mechanisms." https://doi.org/10.48550/arXiv.2102.00696.

Theuer, F., A. Rott, J. Schneemann, L. von Bremen, and M. Kühn. 2022. "Observer-Based Power Forecast of Individual and Aggregated Offshore Wind Turbines." *Wind Energy Science* 7: 2099–2116. https://wes.copernicus.org/articles/7/2099/2022/. https://doi.org/10.5194/wes-7-2099-2022.

Thompson, E. L., A. G. Taye, W. Guo, et al. 2022. "A Survey of eVTOL Aircraft and AAM Operation Hazards." *in AIAA Aviation* 2022 Forum.

Torrence, C., and G. P. Compo. 1998. "A Practical Guide to Wavelet Analysis." *Bulletin of the American Meteorological Society* 79, no. 1: 61–78.

Tran, H. T. M., and D. Hogg. 2022. "Anomaly Detection Using Prediction Error With Spatio-Temporal Convolutional LSTM." https://arxiv.org/abs/2205.08812.

Unser, M., D. Sage, and D. Van De Ville. 2009. "Multiresolution Monogenic Signal Analysis Using the Riesz-Laplace Wavelet Transform." *IEEE Transactions on Image Processing* 18: 2402–2418. https://doi.org/10.1109/TIP.2009.2027628.

Worsnop, R. P., M. Scheuerer, T. M. Hamill, and J. K. Lundquist. 2018. "Generating Wind Power Scenarios for Probabilistic Ramp Event Prediction Using Multivariate Statistical Post-Processing." *Wind Energy Science* 3, no. 2: 371–393. https://wes.copernicus.org/articles/3/371/2018/. https://doi.org/10.5194/wes-3-371-2018.

Yang, L., and Z. Zhang. 2022. "A Deep Attention Convolutional Recurrent Network Assisted by k-Shape Clustering and Enhanced Memory for Short Term Wind Speed Predictions." *IEEE Transactions on Sustainable Energy* 13, no. 2: 856–867. https://doi.org/10.1109/TSTE.2021.3135278.

14698080, 2026, 2, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70173 by Universitatsbibliothek RWTH Aachen University, Wiley Online Library on [08/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Yang, L., Z. Zheng, and Z. Zhang. 2022. "An Improved Mixture Density Network via Wasserstein Distance-Based Adversarial Learning for Probabilistic Wind Speed Predictions." *IEEE Transactions on Sustainable Energy* 13, no. 2: 755–766. https://doi.org/10.1109/TSTE.2021.3131522.

Zhang, X., Y. Wang, Y. Chen, X. Cheng, and L. Gao. 2019. "Short-Term Wind Speed Prediction Based on GRU." *in 2019 IEEE Sustainable Power and Energy Conference (iSPEC)*. 882–887.

Zhao, R., J. Yang, D. Song, L. Wang, J. Liu, and Y. Xiao. 2021. "Short-Term Wind Speed Prediction Based on Data Preprocessing With Discrete Wavelet Transform-Mutual Information and Neural Network." *in Proceedings of the IEEE International Conference on Industrial Electronics and Applications (ICIEA)* (Changsha, China).

Zheng, Z., L. Wang, L. Yang, and Z. Zhang. 2022. "Generative Probabilistic Wind Speed Forecasting: A Variational Recurrent Autoencoder Based Method." *IEEE Transactions on Power Systems* 37, no. 2: 1387–1397. https://doi.org/10.1109/TPWRS.2021.3105101.

14698080, 2026, 2, Downloaded from https://rmets.onlinelibrary.wiley.com/doi/10.1002/met.70173 by Universitatsbibliothek RWTH Aachen University, Wiley Online Library on [08/05/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License