https://doi.org/10.5194/egusphere-2026-1438
Preprint. Discussion started: 21 April 2026

Predicting Forecast Errors with Diffusion Model for Uncertainty Quantification in Wind Speed Nowcasting

Yanwei Zhu[1], Aitor Atencia[2], Markus Dabernig[2], Yong Wang[1,3], Shuyan Zhou[4]

[1]School of Atmospheric Science, Nanjing University of Information Science and Technology, Nanjing, China
[2]GeoSphere Austria, Vienna, Austria
[3]CMA Earth System Modelling and Prediction Centre, China Meteorological Administration, Beijing, China
[4]School of Environmental Science and Engineering, Nanjing University of Information Science and Technology, Nanjing, China

Correspondence to: Yong Wang (yong.wang@nuist.edu.cn)

Abstract. Weather forecasts are inherently uncertain due to the chaotic nature of the atmosphere and unavoidable errors. Ensemble forecasting is the established approach for quantifying the uncertainty. However, it is both computationally expensive and inherently prone to under-dispersion, as it simulates multiple atmospheric trajectories with a finite number of members. In this study, we propose a novel paradigm that achieves uncertainty quantification by directly predicting forecast errors, thereby bypassing the need to simulate multiple trajectories. We employ a denoising diffusion probabilistic model for this task, as its generative capabilities are well-suited for learning high-dimensional distributions. By stochastically sampling from the learned distribution and adding the generated errors to the physics-based nowcast, an ensemble nowcast can be constructed efficiently without the need for perturbation generation or parallel model running. The proposed approach is applied to 10-meter wind speed nowcast, which is important but has received relatively limited attention in diffusion-based weather forecasting studies. Results show that the diffusion model effectively captures the spatial structure and probabilistic characteristics of forecast errors, leading to improved deterministic accuracy and a well-calibrated ensemble. In addition, different noise schedules for the diffusion process are systematically evaluated. The results indicate that the Cosine schedule provides the most reliable performance for uncertainty prediction, offering practical guidance for configuring diffusion models in weather forecasting applications.

1 Introduction

Weather forecasts are inherently uncertain because of the chaotic nature of the atmosphere and unavoidable errors in observations, initial conditions and numerical models (Lorenz, 1965; Palmer, 2002; Slingo and Palmer, 2011). Accurately characterizing this uncertainty is critical for many operational applications, including severe weather warning, wind power management and aviation operations (Zhu et al., 2002; Toth et al., 2007). Over recent decades, quantifying forecast uncertainty has gained growing attention (Chen et al., 2025). Ensemble forecasting has become the primary approach to this problem and is now an integral part of modern weather prediction systems (Du, 2007; Alley et al., 2019). It generates

https://doi.org/10.5194/egusphere-2026-1438
Preprint. Discussion started: 21 April 2026

multiple forecast members by perturbing initial conditions or model configurations, with each member representing a possible evolutionary trajectory of the atmosphere (Wang et al., 2011). The spread among these members provides an estimate of forecast uncertainty (Van Schaeybroeck and Vannitsem, 2016). In recent years, ensemble methods have been increasingly extended to the nowcasting range (0–6 hours), where rapid forecast updates are essential for high-impact weather monitoring and warning (Wilson et al., 2010; Wastl et al., 2018; Bojinski et al., 2023).

Despite the progress made by both dynamical and AI-based ensemble forecasts, they still face a common limitation: under-dispersion (Schultz et al., 2021). The ensemble often fails to fully characterize the true probability distribution (Buizza et al., 2005; Lang et al., 2024). This limitation stems from the basic idea behind ensemble generation: the ensemble methods simulate multiple trajectories precisely because forecast errors are unknown in advance (Ehrendorfer, 1997). By generating a set of possible atmospheric evolutions, they aim to approximate the probability distribution of forecast errors and thus quantify uncertainty (Zhuang et al., 2021). However, a finite number of members cannot fully represent the true distribution, which inevitably leads to under-dispersion. This trajectory simulation paradigm becomes particularly challenging in the weather nowcasting, where atmospheric processes evolve quite rapidly at high resolution and small-scale variability can lead to large deviations between ensemble members (Berenguer et al., 2011; Chkeir et al., 2023; Schmid et al., 2023). As a result, adequately sampling the range of possible evolutions through a finite set of trajectories remains difficult (Kann, 2012; Sun et al., 2014; Zhang et al., 2023).

An alternative perspective comes from the definition of a meaningful ensemble forecast: for an ensemble to provide a useful estimate of uncertainty, the spread among its members should reflect the magnitude of errors in the predicted atmospheric state (Fortin et al., 2014; Feng et al., 2019). This points directly to forecast errors as the quantity of interest (Feng et al., 2024). Recent works such as DiffCast and StormCast have taken a step in this direction (Pathak et al., 2024; Yu et al., 2024). They first use a backbone network to produce deterministic forecasts, then apply a diffusion model to predict the residual errors. However, these approaches still operate within the trajectory simulation paradigm. They rely on a backbone model to simulate atmospheric evolution, with the diffusion model acting largely as a post-processing correction. Nevertheless, their success provides empirical support for the idea of predicting forecast errors directly.

Diffusion models (DMs) have shown remarkable capability in learning complex high-dimensional stochastic distributions (Price et al., 2024; Zhong et al., 2024a). This property makes them well suited for representing the uncertainty of forecast errors and generating realistic error samples (Li et al., 2024; Shu and Farimani, 2024; Mardani et al., 2025). However, most existing studies on DM-based nowcasting have focused primarily on precipitation and convective weather (Gao et al., 2023; Leinonen et al., 2023; Zhong et al., 2024b; Asperti et al., 2025). Other important variables such as near-surface wind have received relatively limited attention (Xiao et al., 2023; Zanetta et al., 2024).

In this study, we propose a new framework for wind speed ensemble nowcasting based on direct modelling of forecast error distributions. Instead of generating members through perturbations or predicting the residuals of an AI forecast model, we directly simulate the error probability distribution of a physically based deterministic forecast produced by a dynamical system. A denoising diffusion probabilistic model (DDPM) is used to learn the conditional distribution of forecast errors (Ho

https://doi.org/10.5194/egusphere-2026-1438
Preprint. Discussion started: 21 April 2026

et al., 2020; Turner et al., 2024). This formulation represents uncertainty as stochastic realizations of forecast errors conditioned on a physically consistent forecast, avoiding the need to repeatedly simulate atmospheric evolution while still providing a probabilistic description of forecast uncertainty. Unlike conventional statistical post-processing methods that mainly correct systematic biases, the proposed diffusion model explicitly learns the full conditional distribution of forecast errors and generates stochastic samples that represent the uncertainty structure of the forecast. By sampling from the learned error distribution and adding the generated errors to deterministic nowcast, an ensemble nowcast can be constructed efficiently.

The results demonstrate that the proposed approach can effectively represent the uncertainty of wind forecasts. This study offers a new perspective for probabilistic nowcasting and highlights the potential of DDPMs for uncertainty prediction in weather forecasting. By adding stochastically generated errors to the deterministic nowcast, the method improves forecast accuracy and constructs a well-calibrated ensemble nowcast.

This article is organized as follows. Section 2 describes the dataset used in this study. Section 3 details the configuration of the diffusion model and the methodology employed. Section 4 presents the verification results of the stochastically generated errors and the resulting ensemble nowcasts. A summary and conclusions are given in section 5.

2 Dataset

This study trains a DDPM leverages nowcasts of 10-meter wind speed from the Seamless Integrated Weather Prediction and Applications system (SIVA, detailed descriptions are in Zhu et al., 2025). The learning objective of DDPM is the error of SIVA nowcast, defined as the difference between forecast and the corresponding analysis field. This analysis constitutes a robust gridded observation product, generated by applying topographic adjustments and dense-station-based correction to 3D NWPs. The dataset contains hourly-updated wind speed nowcasts with a 1-hour temporal resolution (up to a lead time of 6 hours) and the corresponding errors on a 652 × 632 mesh grid with a 1 km resolution (Fig. S1). It covers the period from 1 October 2021 to 30 June 2023 over the domain (115.35°E–122.33°E, 29.88°N–35.81°N). Owing to computational constraints, the original high-resolution data was subdivided into smaller 256 × 256 patches to facilitate model training. The data was spilt into nonoverlapping parts for training (1 October 2021–30 April 2023), validation (1 May 2023–31 May 2023), and testing (remainder).

Since wind speed is log-normally distributed, we applied a logarithmic transformation in data preprocessing. This operation yields forecast errors that are approximately Gaussian distributed, which is conducive to DDPM training. An additional advantage is that it eliminates the need to explicitly enforce a positivity constraint when reconstructing the final wind speed forecast from the predicted errors.

https://doi.org/10.5194/egusphere-2026-1438
Preprint. Discussion started: 21 April 2026

3 Method

3.1 Denoising Diffusion Probabilistic Model

Many recent studies have leveraged the probabilistic generative framework of DDPMs for estimating uncertainty in weather forecasting (Li et al., 2024; Wang et al., 2024; Andrae et al., 2025; Cachay et al., 2025). A DDPM consists of a forward diffusion process and a reverse denoising process. The forward process is a Markov chain defined by gradually adding Gaussian noise to the data, which ultimately diffuses the original data into a Gaussian distribution. Formally, given an original image x_0, the noisy image x_t at diffusion step t is obtained as:

$$q(x_t|x_0) = \mathcal{N}(x_t; \sqrt{\bar{\alpha}_t}x_0, (1-\bar{\alpha}_t)\mathrm{I}) \quad (1)$$

where $\bar{\alpha}_t = \prod_{i=1}^{t}(1-\beta_i)$, and $\{\beta_t \in (0,1)\}_{t=1}^{T}$ is a predefined variance schedule controlling the amount of noise added at each step. The reverse process is to reconstruct the original x_0 from a pure noise through:

$$p_\theta(x_{t-1}|x_t) = \mathcal{N}(x_{t-1}; \mu_\theta(x_t, t), \sigma^2\mathrm{I}) \quad (2)$$

where $\mu_\theta(x_t, t)$ and σ is the mean and variance of the distribution at the previous step, respectively. Following Ho et al. (2020), the μ_θ is derived from a noise prediction neural network:

$$\mu_\theta(x_t, t) = (x_t - \frac{1-\alpha_t}{\sqrt{1-\bar{\alpha}_t}}\epsilon_\theta(x_t, t))/\sqrt{\alpha_t} \quad (3)$$

where $\epsilon_\theta(x_t, t)$ is a function approximator to predict the noise ϵ that conditioned on the current step t. Consequently, the training objective simplifies to minimizing the difference between the true and the predicted noise:

$$\mathcal{L}(\theta) = \mathbb{E}_{x_0,y,t,\epsilon}(\|\epsilon - \epsilon_\theta(x_t, y, t)\|^2) \quad (4)$$

In this study, DDPM is applied to predicting the forecast error (x_0) given a deterministic wind speed nowcast (y). Here, x_0 represents the predicting target and the conditional input y provides the necessary meteorological context (the forecast itself). Unlike AI-based forecasts, which may themselves deviate from physical constraints, the dynamical nowcast provides a physically consistent baseline whose errors carry interpretable physical meaning. This distinction is important: approaches such as DiffCast learn the residual of an AI-based forecast, which can be influenced by the model's own biases and may not directly reflect the intrinsic uncertainty of the atmospheric state. By contrast, conditioning on a physically based nowcast ensures that the error distribution we learn is rooted in the actual physics of the atmosphere, offering a more direct and interpretable path to uncertainty quantification.

To enable efficient inference, we employ the Denoising Diffusion Implicit Model (DDIM) sampling scheme (Song et al., 2022). It achieves comparable generation quality to standard DDPM with only 200 sampling steps, a substantial reduction from the typical 1000 steps process:

$$x_{t-1} = \sqrt{\alpha_{t-1}}\left(\frac{x_t - \sqrt{1-\alpha_t}\times\epsilon_\theta(x_t,t)}{\sqrt{\alpha_t}}\right) + \sqrt{1-\alpha_{t-1}-\sigma_t^2}\times\epsilon_\theta(x_t,t) + \sigma_t\epsilon_t \quad (5)$$

https://doi.org/10.5194/egusphere-2026-1438
Preprint. Discussion started: 21 April 2026

Although extensively studied for image generation, the Linear, Cosine, and Sigmoid noise schedules have not been adequately evaluated for weather forecasting (Nichol and Dhariwal, 2021; Jabri et al., 2023; Turner et al., 2024). Therefore, a systematic comparison is performed to assess their relative efficacy in this domain.

3.2 Model Training

The model was trained on a cluster with two NVIDIA 5090 GPUs, with a batch size of 18 per GPU, for 3,637,000 iterations, which took approximately 67 hours. The AdamW optimizer was used with β_1=0.9, β_2=0.99, and a weight decay of 0.01 (Kingma and Ba, 2017; Loshchilov and Hutter, 2019). The learning rate follows a cosine decay schedule, starting from an initial value of 0.0001 after a linear warmup of 1,000 steps. Generating a 16-member ensemble from the trained model takes approximately 4 minutes.

As described in Section 2, the prediction target of DDPM is the error of a given wind nowcasting. This error is progressively corrupted into Gaussian noise through the forward diffusion process. The corresponding nowcasting time series serves as the essential contextual condition and is fed into the denoising U-Net network together with the noised error image. The U-Net is a denoiser trained to reconstruct the forecast error for a given nowcast by iteratively denoising a random Gaussian noise input. At each scale transition in the U-Net, the feature maps are processed through a sequence of operations: two residual blocks, a temporal attention layer, and a convolution that performs either down-sampling or up-sampling. To incorporate physical constraints, the model input is formed by concatenating the wind nowcast with its noised error counterpart. As a result, the learning process for denoising is explicitly guided by the prior knowledge embedded in the physical forecast.

The error predicted by the denoiser is a random sample from the learned empirical distribution. By repeating this sampling process, an ensemble of errors is generated to represent its uncertainty. Combining this stochastically predicted error ensemble with the deterministic wind nowcast results in an ensemble nowcast of wind speed. The workflow of the DDPM in this study is shown in Figure 1.

https://doi.org/10.5194/egusphere-2026-1438
Preprint. Discussion started: 21 April 2026

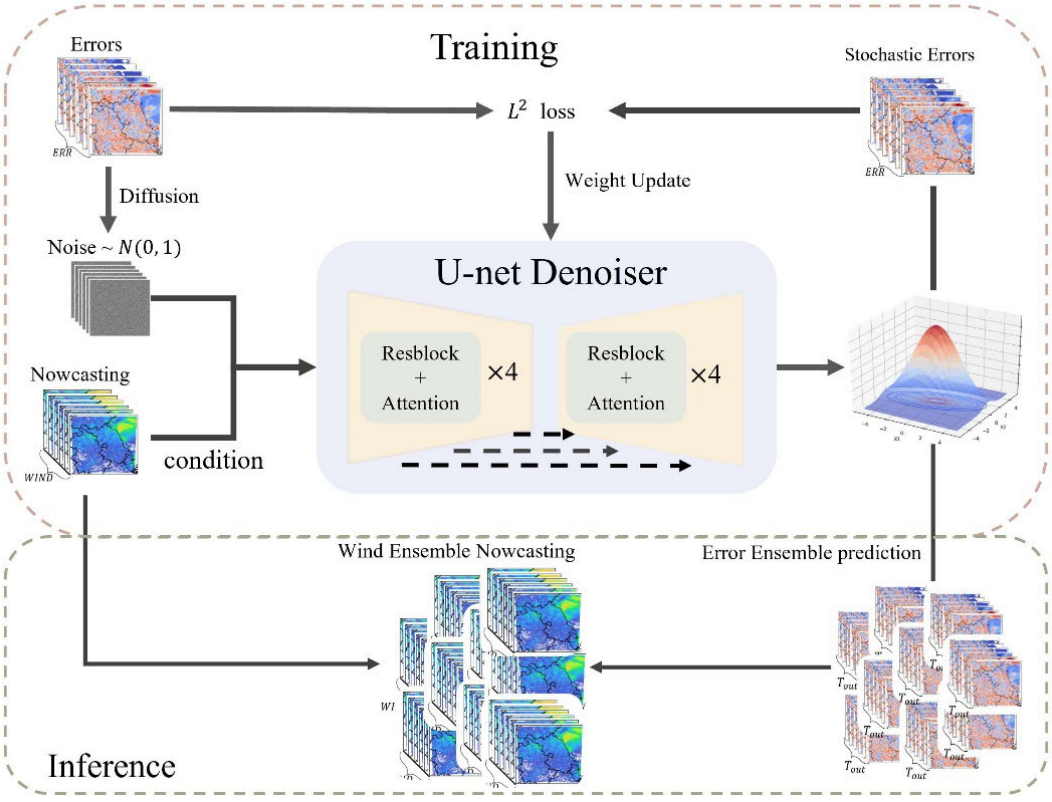


Fig. 1. Framework of DDPM.

3.3 Evaluation Metrics

For verification, the SIVA forecast errors are designated as the reference for evaluating the DDPM, while the analysis fields serve as benchmark for the ensemble nowcasts. This study employs standard probabilistic evaluation metrics to evaluate both the DDPM-generated errors and the ensemble nowcasts derived therefrom.

These metrics comprise the Talagrand diagram (also known as the rank histogram), reliability diagram, Continuous Ranked Probability Score (CRPS), Brier Score, Brier Skill Score, Root-mean-square-error (RMSE) and ensemble spread (Hamill, 2001; Jolliffe, 2004; Ebert, 2009). To assess the physical consistency of the predicted errors, the joint probability distribution between errors and the corresponding wind nowcasts is evaluated. Furthermore, the power spectrum is employed to verify whether both the generated errors and the constructed wind speed nowcast retain their expected physical characteristics.

Given the negligible sensitivity of key verification metrics to ensemble size, an ensemble with 16 members was adopted to ensure statistical robustness while maintaining computational tractability. It should be noted that no corresponding reference ensemble is available for the nowcasts, as the operational SIVA system is purely deterministic.

https://doi.org/10.5194/egusphere-2026-1438
Preprint. Discussion started: 21 April 2026

4 Results

4.1 Evaluation of Stochastic Generating Errors

As outlined in the bottom part of Figure 2, the DDPM generates an ensemble of 16 independent error estimates for a given wind nowcast through independently sampling. Based on robust physical principles, wind nowcast establishes an inherent linkage between forecast error and wind speed. The DDPM explicitly uses the nowcast as a conditioning input, thereby incorporating this physical linkage as a constraint. Consequently, its predicted errors are expected to adhere to the same physical relationship. This section presents an analysis based on the statistical evaluation metrics of the test set (June 2023).

The agreement of both the joint and marginal probability distributions with the benchmark (Fig. 2) demonstrates that the errors predicted by DDPM are physically consistent and statistically robust. This confirms DDPM effectively estimates the conditional distribution of forecast errors given the wind nowcast (Fig. 2). Notably, the three noise schedules produce joint distributions with marked differences. Furthermore, the errors predicted by Linear schedule deviate from the benchmark, introducing bias into the constructed wind nowcast. Although Cosine schedule yields distributions (both joint and marginal) closest to the benchmark, it still struggles to capture errors associated with high wind speeds ($\log_{10}\text{Wind} > 2$) in the 6-hour forecast. This limitation reflects a broader tendency of stochastically generated samples to be overly smooth, potentially missing localized or extreme features. Both Cosine and Sigmoid schedules accurately reconstruct the marginal distributions of the joint probability. This implies that for time-series prediction with DDPMs, employing an appropriate noise schedule is crucial. Specifically, avoiding an excessively rapid decay of the data signal during diffusion leads to a more accurate reconstruction of the target probability distribution.

https://doi.org/10.5194/egusphere-2026-1438
Preprint. Discussion started: 21 April 2026

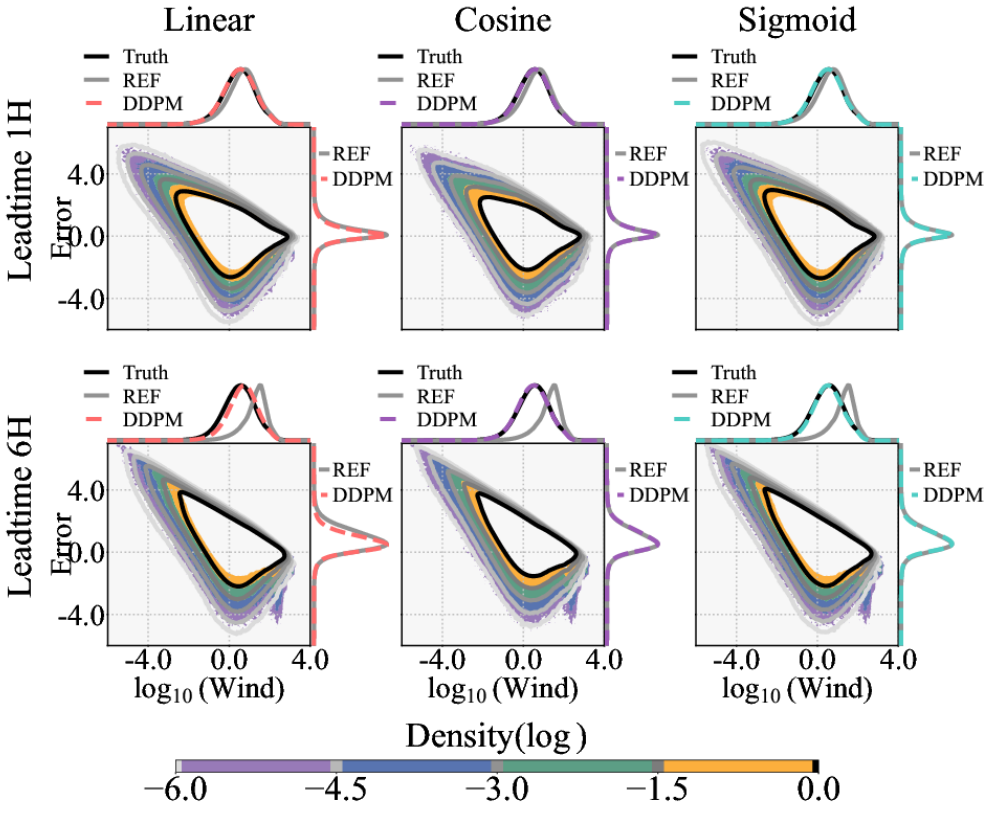


Fig. 2. Joint distribution of error and wind speed for the 1-hour (top row) and 6-hour (bottom row) forecast. The shaded contours and contour lines represent the deterministic reference (REF) and ensemble predictions of DDPM, respectively. Marginal distributions of logarithmic wind speed (top x axis) and error (right y axis) are shown for observations (Truth, black solid), the deterministic forecast (REF, grey solid), and the DDPM ensemble (dashed).

Figure 3 shows the energy spectra of the forecast errors at different lead times. The energy spectra show that the errors predicted by DDPM closely match the benchmarks across all scales, indicating that the model accurately reproduces the spatial structures. This confirms that the stochastically predictions constitute conditional samples from a distribution that adheres to physical constrains. The Cosine schedule yields the best match to the benchmark energy at small scales for all lead times, followed by the Sigmoid schedule. The Linear schedule exhibits the most pronounced energy deficit at fine scales, which highlights its known limitation in high-resolution generation tasks.

Among the three noise schedules evaluated, the Linear schedule yields the least accurate error predictions, exhibiting systematic bias in the distribution and a poor representation of temporal dynamics of errors. The Sigmoid schedule provides better control of overall bias, but its depiction of temporal evolution is less effective than that of the Cosine schedule. The Cosine schedule behaves the most stable performance and produces the most physically reasonable temporal evolution. These results suggest that a moderate amount of additional noise in the later stages of the forward process can enhance the prediction of distribution tails. It is important to note, however, that the effect of noise intensity is not physically constrained

https://doi.org/10.5194/egusphere-2026-1438
Preprint. Discussion started: 21 April 2026

but rather represents an inherent uncertainty in the noising process itself. Therefore, the selection of an appropriate noising schedule requires careful consideration in practice applications.

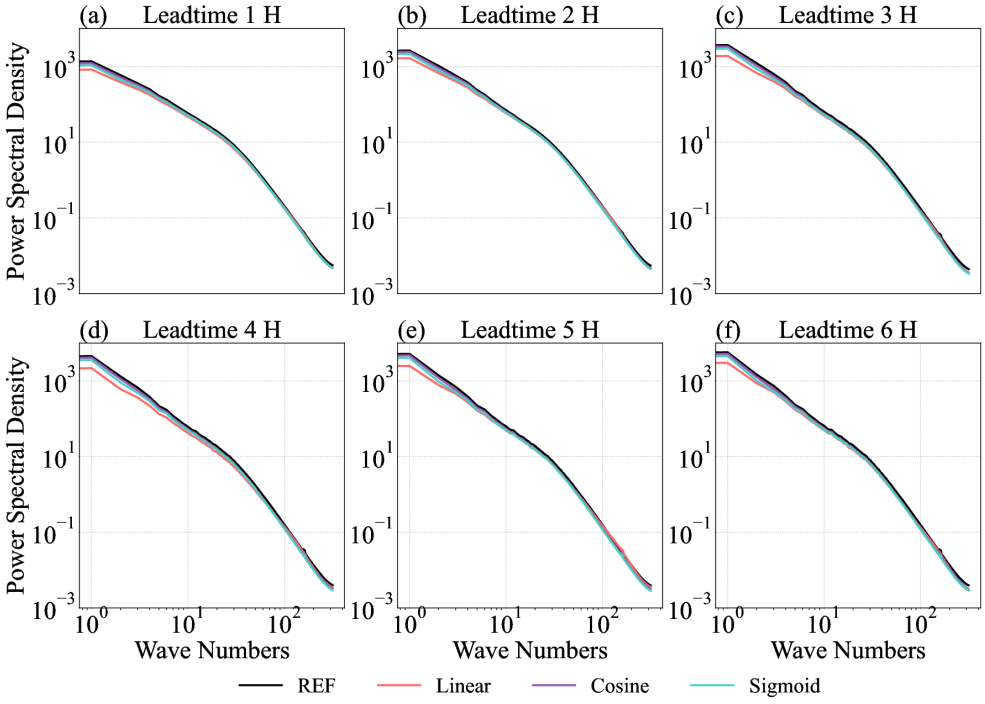


Fig. 3. Comparison of power spectra density for Errors in different forecast lead time. The black solid lines denote the errors of deterministic nowcast (REF). While the orange, purple and teal solid lines denote DDPM predictions with different noising schedule (Linear, Cosine and Sigmoid).

The results confirm that DDPM can effectively learn the conditional distribution of forecast errors, with noise schedule selection playing a critical role in reconstruction quality. A key insight is that moderate noise in later diffusion stages enhances tail prediction, yet this remains an unconstrained design choice requiring careful empirical tuning.

4.2 Verification of Ensemble Nowcasts

Evaluations of the generated errors reveal that DDPM captures the physical characteristics of forecast errors, thereby learning more than just their statistics. This capability allows it to estimate the forecast error distribution and thereby quantify uncertainty independently of any perturbation scheme. Adding these generated errors to the deterministic nowcast effectively calibrates the forecast, as these errors closely approximate the actual forecast errors. More importantly, this approach successfully constructs an ensemble nowcast without requiring perturbation or parallel integration. It thus achieves dual objectives: calibrating the forecast while simultaneously providing an estimate of its uncertainty.

The energy spectra of wind speed provide an intuitive metric for assessing how accurately the ensemble nowcasts represent physical characteristics of the ground truth across spatial scales. The DDPM-based ensemble nowcasts exhibit

https://doi.org/10.5194/egusphere-2026-1438
Preprint. Discussion started: 21 April 2026

energy spectra that are closer to the benchmark than the deterministic nowcast (REF) across all scales. This improvement is particularly pronounced at the largest scales. This result demonstrates that the calibration process preserves the physical characteristics of the deterministic nowcast and enhances its accuracy at large scales (Fig. 4). The spectral analysis thus suggests that the DDPM, by leveraging the convolutional layers of its U-Net architecture, effectively captures the spatiotemporal relationships between forecast errors and the underlying wind speed field.

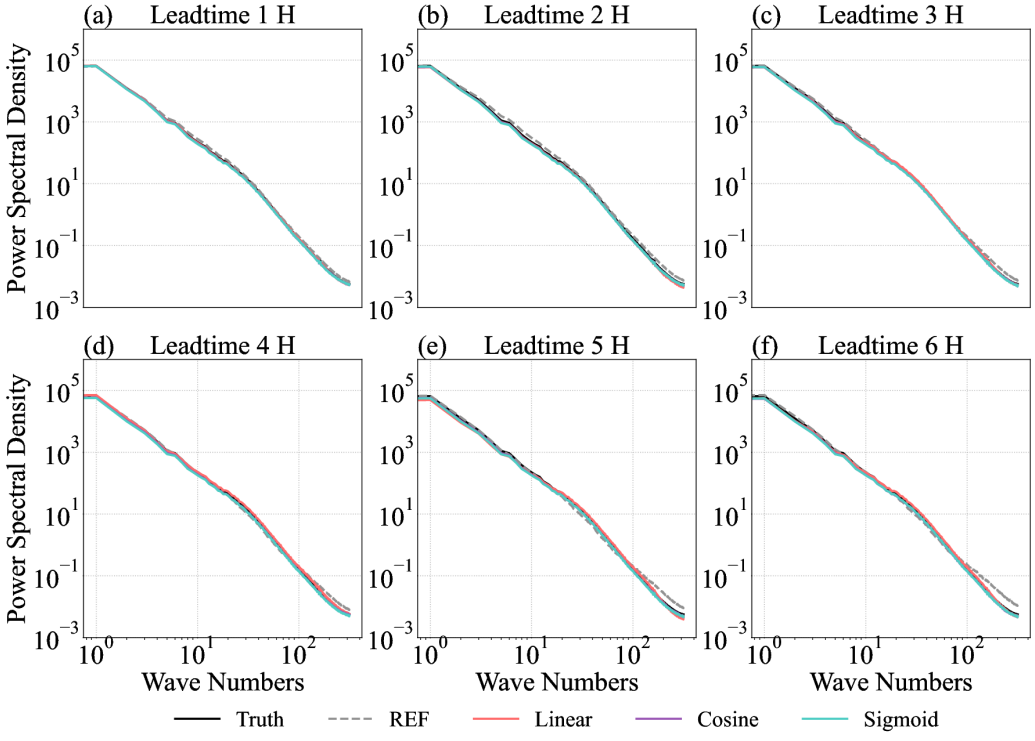


Fig. 4. Comparison of power spectra density for wind speed nowcasts in different forecast lead time. The black solid and grey dashed lines denote analysis field (Truth) and deterministic nowcast (REF), respectively. While the orange, purple and teal solid lines denote DDPM ensemble nowcasts with different noising schedule (Linear, Cosine and Sigmoid).

The accuracy and reliability of the wind speed ensemble nowcasts are then evaluated using standard verification metrics. These metrics clearly delineate performance differences among the three noise schedules. The Linear schedule exhibits a significant positive bias (Fig. 5a), which directly leads to its elevated RMSE and CRPS compared to the other two schedules. This bias stems from an overly rapid decay of the original signal during forward noising process, which excessively corrupts data and makes it more difficult for DDPM to reconstruct an accurate generation during denoising. Consequently, the verification scores for this strategy show less smooth progression with the lead time increasing (Fig. 5b). The Cosine schedule achieves the best performance across all verification metrics (BIAS, CRPS, RMSE), with the Sigmoid schedule ranking second. These results point to a key design principle: an optimal noise schedule must transition smoothly to guide the denoising process while preserving sufficient signal information in the later stages to prevent degradation.

https://doi.org/10.5194/egusphere-2026-1438
Preprint. Discussion started: 21 April 2026

Well-calibrated ensemble forecasts exhibit uncertainty (as measured by ensemble spread) that is, on average, commensurate with the magnitude of their errors. This property can be quantitatively characterized by the ratio of RMSE/spread. The ratio above 1 indicate under-dispersion, while it below 1 indicate over-dispersion. This ideal is best approximated by the Cosine schedule, as evidenced by its RMSE/spread ratio being closest to 1 (Fig. 5c, d). The observed under-dispersion across all three schedules is consistent with the conclusions from the previous section regarding the joint distribution of forecast errors and wind speed. Nevertheless, the consistent increase in ensemble spread with forecast lead time confirms that the DDPM successfully captures the temporal growth trend of forecast errors.

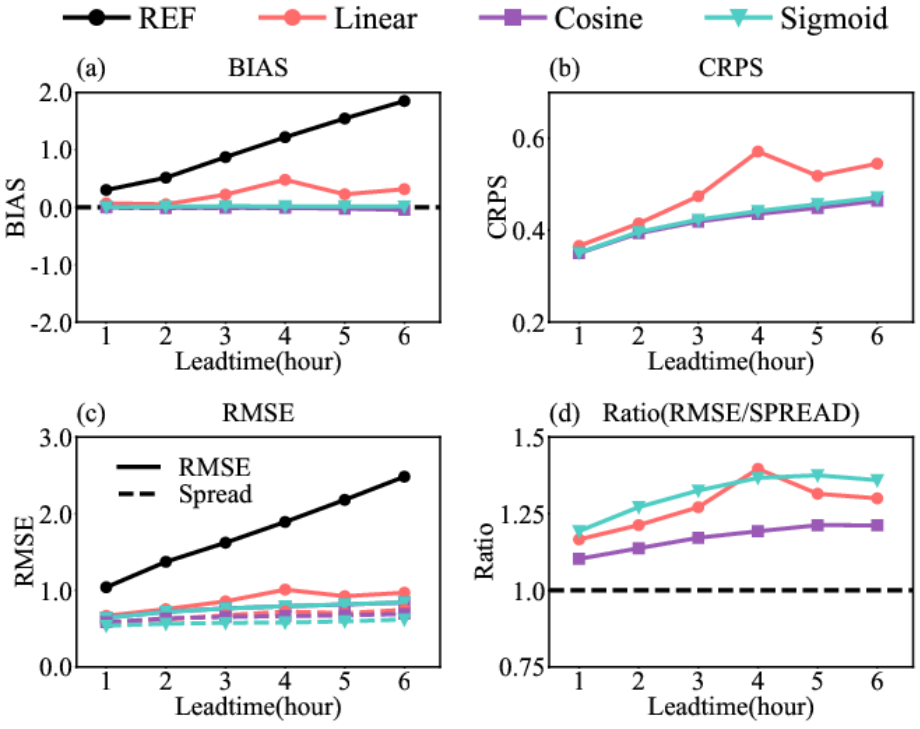


Fig. 5. Verification metrics for deterministic and ensemble mean nowcasts over the test period. (a) BIAS, (b) CRPS, (c) RMSE (solid lines) and ensemble spread (dashed lines). (d) Ratio of RMSE to SPREAD. In all panels, the black solid line denotes the deterministic nowcast (REF), while the orange, purple, and teal lines correspond to the ensemble nowcasts generated using the Linear, Cosine, and Sigmoid noising schedules, respectively.

Figure 6 presents the average RMSE of the REF, the RMSE of the ensemble mean for different noising schedules, and their corresponding ensemble spread. A key observation is that the RMSE for all DDPM-based forecasts is substantially lower than that of the REF. Among the three schedules, the Linear schedule produces the largest spread, which also increases most noticeably with lead time. However, this strategy also results in the highest RMSE among the three. As the preceding verification scores (Fig. 5) indicate, this disproportionately large spread is attributable to the significant positive bias introduced by the Linear noising process. In contrast, the Cosine schedule generates a more refined and physically coherent spatial structure in its spread field. This is particularly evident over land regions with low wind speeds and small forecast

https://doi.org/10.5194/egusphere-2026-1438
Preprint. Discussion started: 21 April 2026

errors, where it produces a clearer, more organized dispersion pattern that better reflects the expected spatial characteristics of a reliable probabilistic forecast.

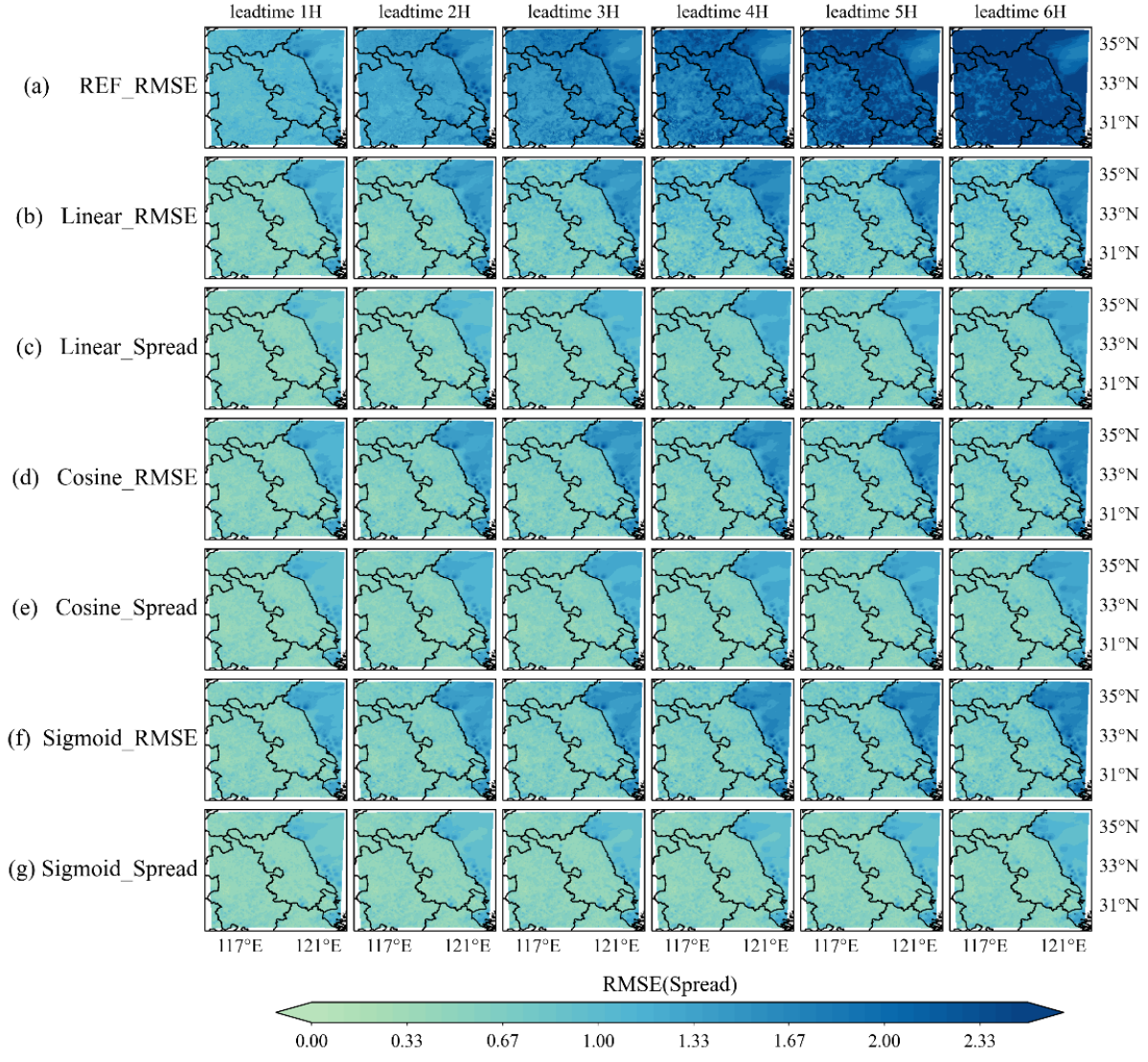


Fig. 6. Spatial distribution of RMSE and ensemble spread at each forecast lead time. (a) Deterministic nowcast (REF) RMSE. For each DDPM noise schedule (Linear, Sigmoid, Cosine), the RMSE of ensemble mean (b, d, f) and the corresponding ensemble spread (c, e, g) are shown.

To comprehensively assess ensemble reliability, the Brier Score and Brier Skill Score were calculated for different wind speed thresholds. The Brier score measures the mean squared error between the forecast probability of a given event (e.g., wind speed exceeding a threshold) and the corresponding binary observations. It ranges from 0 to 1, with values closer to 0 indicating more accurate probabilistic forecasts, and a value of 1 representing no forecast skill. The Brier skill score (BSS) quantifies the relative skill of a probabilistic forecast compared to a reference forecast (e.g., climatology, persistence). A

https://doi.org/10.5194/egusphere-2026-1438
Preprint. Discussion started: 21 April 2026

BSS greater than 0 indicates that the forecast outperforms the reference; a BSS equal to 0 suggests performance comparable to the reference; and a BSS less than 0 denotes inferior skill relative to the reference. A perfect forecast achieves a BSS of 1.

Figure 7 shows that the DDPM ensembles attain high Brier Scores across all wind speed thresholds, indicating their successful discrimination of occurrence probabilities by wind intensity. An exception is the Linear schedule, which shows a clear performance deficit specifically at lower wind speeds (wind > 1 m s^{-1}). The results of Brier scores and BSS confirm that the DDPM effectively discriminates between wind regimes of differing intensities and faithfully reproduces the conditional probability distributions associated with high-impact weather events. Furthermore, comparison of the three schedules reveals that the Linear schedule exhibits notable bias in predicting events near the core of the probability density, while the Sigmoid schedule shows bias in low-probability events. This further supports the findings presented in Fig. 2.

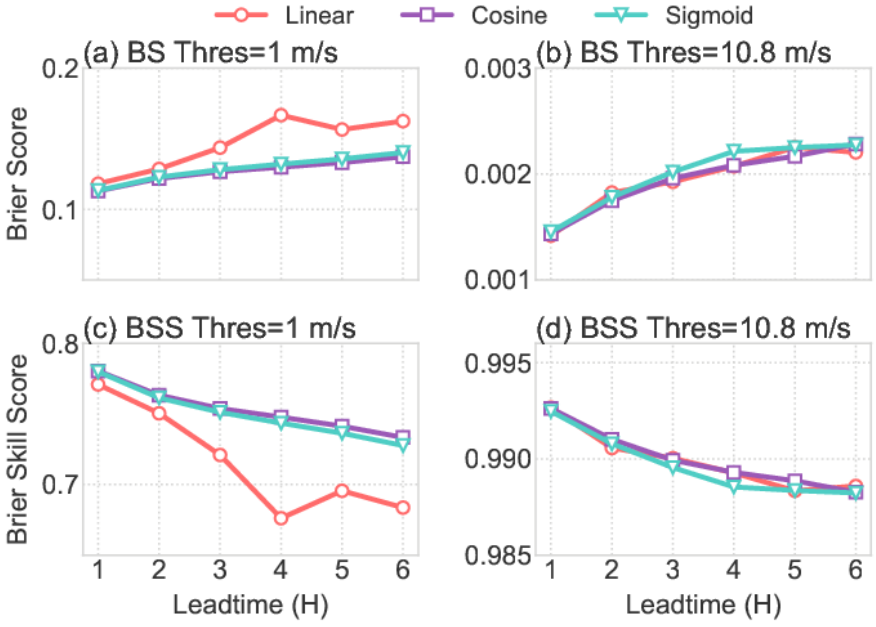


Fig. 7. Brier Score (BS) and Brier Skill Score (BSS) of the wind speed ensemble nowcasts as a function of forecast lead time for two wind speed thresholds. (a, c) Threshold: 1 m s^{-1}. (b, d) Threshold: 10.8 m s^{-1}. In all panels, results are shown for the three DDPM noising schedules: Linear (orange), Cosine (purple), and Sigmoid (teal). The reference value in calculating BSS is the climatology probability of the given threshold in the training set.

The Talagrand diagrams (rank histograms) illustrate the effectiveness of ensembles in quantifying uncertainty. A flat distribution indicates that the probability of the ground truth falling into any rank interval is equal, representing the calibration degree of an ensemble forecast (Fig. 8). The ensemble of the Cosine schedule maintains the most stable evolution of reliability with lead time. In contrast, the ensemble of the Linear schedule exhibits a pronounced positive bias specifically beginning at the 3-hour forecast and persisting beyond, which does not evolve smoothly, highlighting its weakness in temporal error modelling. The Sigmoid schedule, while more temporally stable than the Linear, suffers from significant under-dispersion (Fig. 5d) due to excessive noise in the later noising stages diluting signal fidelity.

https://doi.org/10.5194/egusphere-2026-1438
Preprint. Discussion started: 21 April 2026

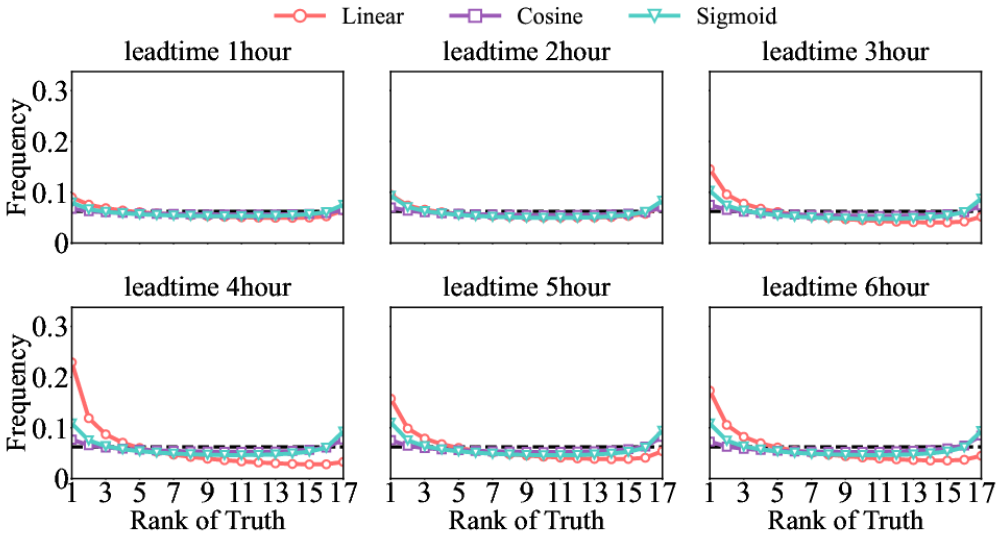


Fig. 8. Talagrand Diagrams of wind ensemble nowcasts for different schedules.

The reliability diagram visually represents the accuracy of probability estimates from an ensemble forecast for a specific event. This study compares the forecast reliability of the generated ensemble nowcasts at different wind speed thresholds. Figure S2 shows the reliability diagrams for the 1 m s^{-1} wind speed, which lies at the centre of the observed probability distribution. The forecast probability of a 1 m s^{-1} wind speed event closely matches the observed frequency, with the ensemble from Cosine schedule exhibiting the most stable reliability across lead times. The Linear strategy shows clear signs of overconfidence starting from 3-hour forecast, systematically overestimating the probability of winds exceeding 1 m s^{-1}. For forecasts beyond 4 hours, both Cosine and Sigmoid demonstrate higher and more stable probability estimating skill than Linear.

The ensembles maintain high reliability at high wind events with threshold 10.8 m s^{-1}, exhibiting temporal evolution characteristics consistent with those observed for the 1 m s^{-1} threshold (Fig. S2). This consistency demonstrates the robustness of the ensemble nowcasts across wind speeds. A noteworthy finding is that for forecasts beyond 3 hours, the probability estimates from both the Sigmoid and Linear schedules become unstable, revealing their weaker capability in estimating the tails of the probability distribution.

Although the DDPM ensemble nowcasts achieve high reliability for high wind speeds, their performance is less stable compared to that for lower wind speeds. This can be attributed to both the inherently larger forecast errors associated with high winds and the limitations of neural networks in predicting rare events. Strengthening the physical consistency between wind speed and error fields—for instance, through additional physics-based constraints—offers a viable pathway for improving this performance in future work.

https://doi.org/10.5194/egusphere-2026-1438
Preprint. Discussion started: 21 April 2026

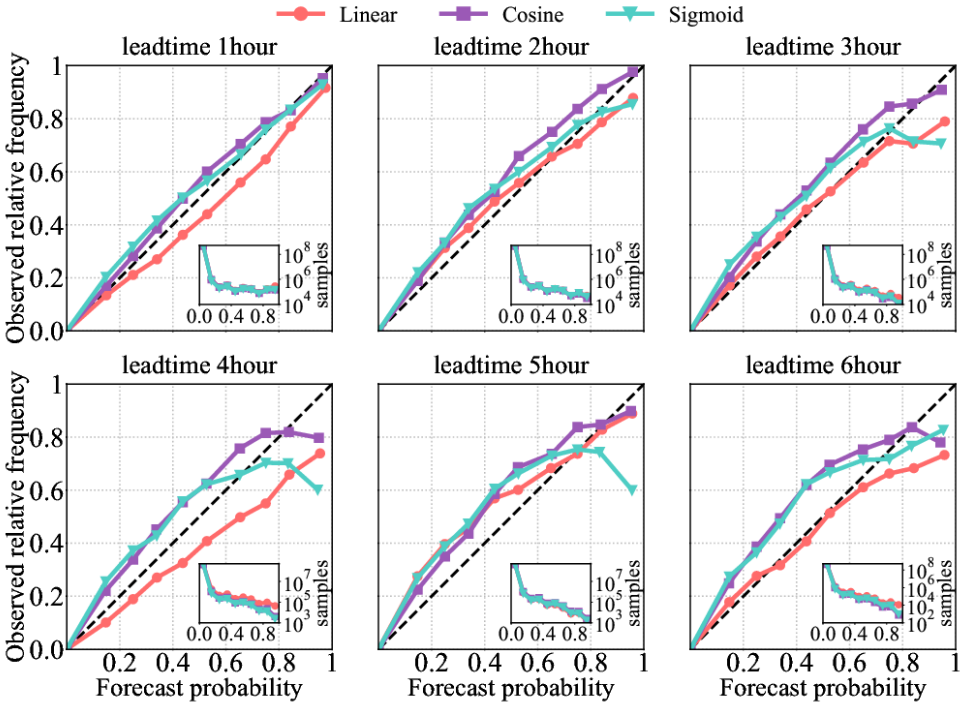


Fig. 9. Probability diagram of wind ensemble nowcasting in different lead time for three schedules with threshold 10.8 m s^{-1}. The inset in each subplot denotes the corresponding statistics samples.

This section evaluates the DDPM-based ensemble nowcasting across physical consistency, forecasting skill, and probabilistic reliability. The Cosine noising schedule outperforms Linear and Sigmoid across all metrics. Its reliability remains consistently high across all lead times, producing ensembles that are physically coherent and statistically reliable. The Linear schedule suffers from rapid signal decay and positive bias, while Sigmoid exhibits under-dispersion from excessive late-stage noise. Despite some under-dispersion and reduced stability at high wind speeds, the DDPM successfully captures the temporal evolution of forecast errors, validating its effectiveness for probabilistic nowcasting across wind speeds.

4.3 Case study

The above analysis is based on the statistical evaluation metrics of the test set. The performance of DDPM on individual cases is presented below. Given its superior performance in the preceding analysis, only the ensemble mean and probabilistic forecasts from the Cosine schedule are displayed here to compare with the truth and the deterministic nowcast.

Figure 10 shows the observed and forecasted surface wind speed changes from 08:00 to 13:00 UTC on 10 June 2023 (Case1). For this high-impact event over the central land region, deterministic nowcast predicted a strong wind band that resulted in a significant false alarm ratio. The forecast field also became progressively smoother as lead time increased (Fig. 10a, b). The ensemble mean, by contrast, reduced the bias and captured the spatial structure more accurately (Fig. 10c).

https://doi.org/10.5194/egusphere-2026-1438
Preprint. Discussion started: 21 April 2026

However, the DDPM is conditioned on this positively biased deterministic nowcast. When its generated errors are added back to correct the forecast, the resulting ensemble mean exhibits a negative bias. Despite this, the spatial structure of the probabilistic forecast aligned well with the ground truth (Fig. 10d). This indicates that the ensemble members effectively represented the forecast uncertainty of such wind events (Fig. S3).

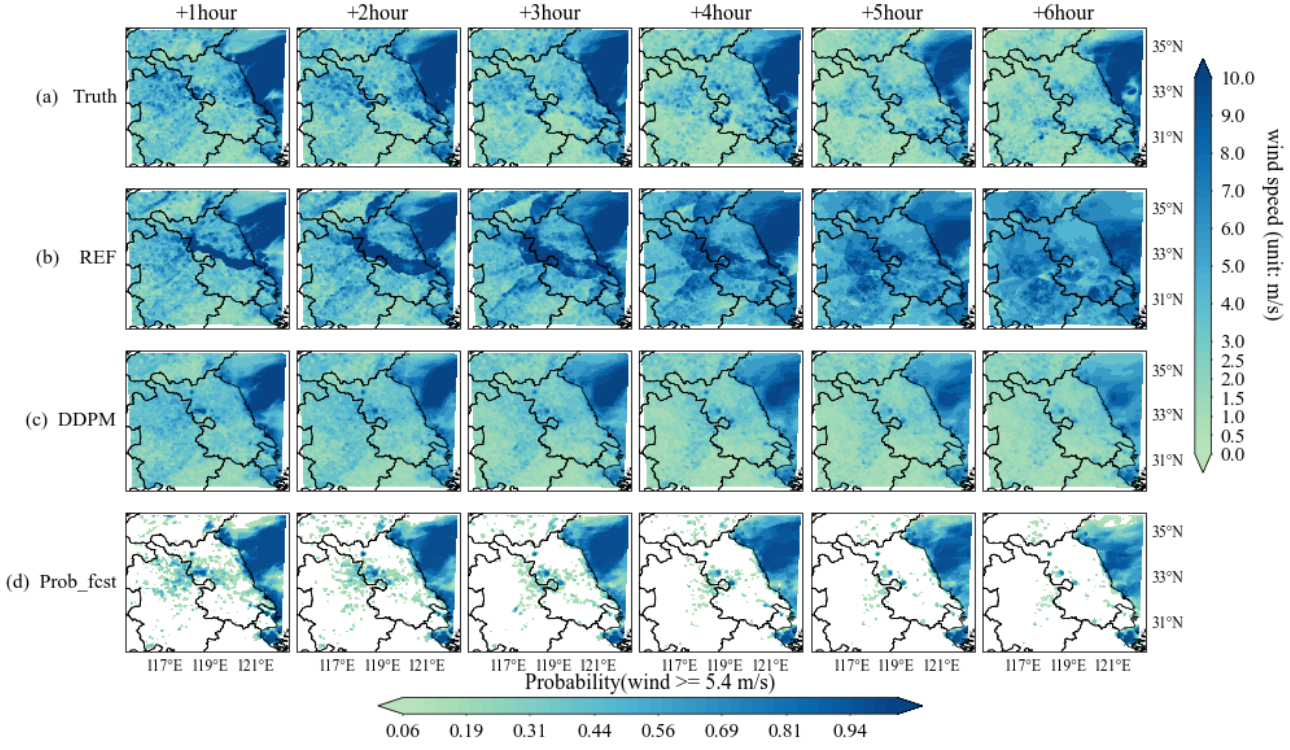
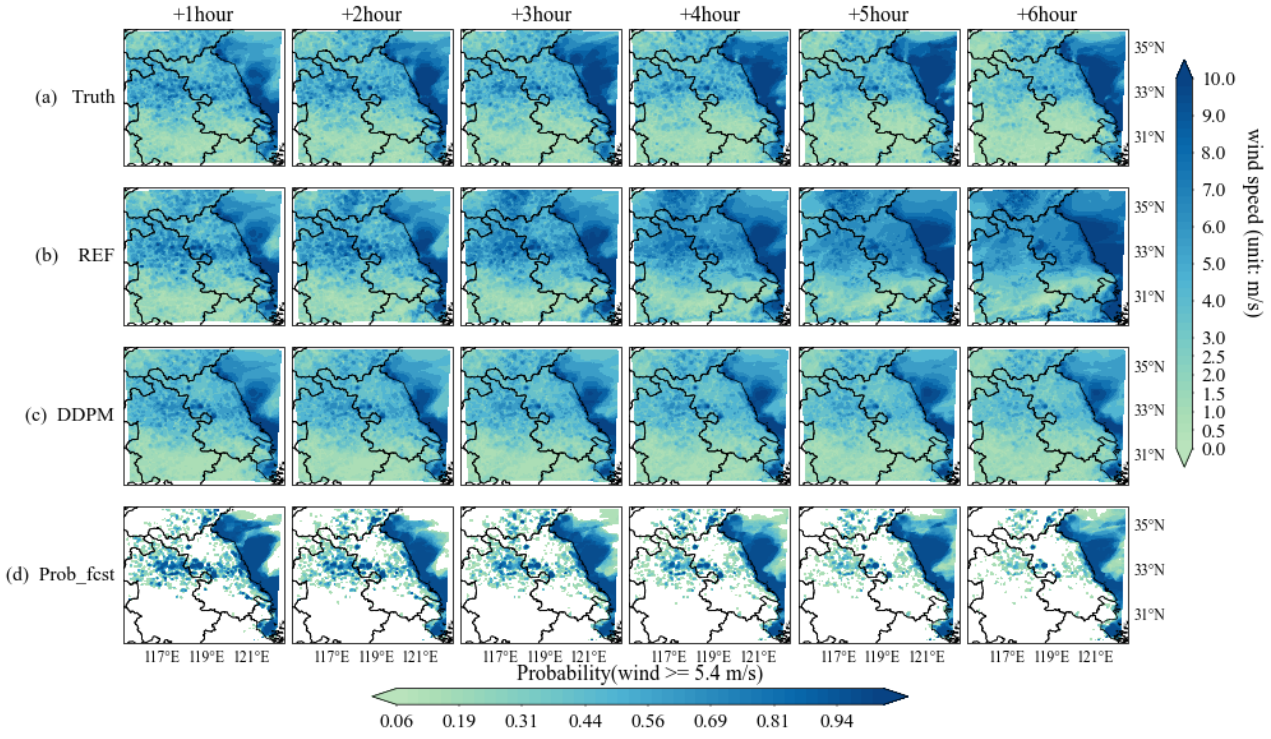


Fig. 10. Wind speed from 08:00 to 13:00 UTC on 10 June 2023. (a) Ground Turth, (b) Deterministic Nowcast (REF), (c) Ensemble Mean forecast of DDPM with cosine schedule and (d) probability forecast with threshold 5.4 m s^{-1}.

Figure 11 shows the other case (from 07:00 to 12:00 UTC on 24 June 2023, Case2). The deterministic forecast for this case exhibited higher accuracy than for Case1. The smoothing effect on the ensemble mean was also reduced, and the probabilistic forecast shows finer and more accurate details. This case benefits from smaller errors in the deterministic forecast itself, leading to improved ensemble performance. The spatial uncertainty represented by the ensemble members is also superior to that in Case 1 (Fig. S4). The comparison of these two cases suggests that while the model provides well-calibrated ensemble nowcasts, the representation of uncertainty for extreme values remains under-dispersive (Fig S5).

Case 1 is the sample with the highest wind speed in the test set, which also highlights the limited capability of DDPM in estimating uncertainty for extreme events. In addition to the inherent limitations of the model in predicting tail distributions, the scarcity of extreme samples further contributes to this issue. Therefore, future work should explore the incorporation of additional constraints to enhance the performance in this regard, particularly when training samples are limited.

https://doi.org/10.5194/egusphere-2026-1438
Preprint. Discussion started: 21 April 2026



340 **Fig. 11. The same as Fig. 10 but for case of 07:00 to 12:00 UTC on 24 June 2023 (UTC).**

5 Conclusion and Discussion

This study presents a novel approach to weather forecast uncertainty quantification by employing a diffusion model to predict forecast errors for 10-meter wind speed nowcasting. Leveraging the stochastic generation process of diffusion model to estimate error uncertainty, this work establishes a new paradigm for the application of diffusion models in this domain. By 345 adding the stochastically predicted errors to the deterministic nowcast, a well-calibrated ensemble nowcast can be produced. The performance of different noise schedules is also compared, providing insights into the configuration of diffusion models for forecast uncertainty prediction tasks.

Conditioned on a physically-based wind speed nowcast, the model produces forecast errors that preserve good physical consistency with wind speed. The spatial structure and probabilistic distribution characteristics of the errors are also 350 effectively retained. By accurately reconstructing the probability distribution of forecast errors, the model provides a reliable quantification of weather forecast uncertainty. However, for extreme events, the model still suffers from the issue of excessive smoothness. This limitation may stem from the use of a basic diffusion model architecture. In future work, introducing cross-attention layers could enhance the model's ability to learn the physical coupling between the forecast and the errors.

https://doi.org/10.5194/egusphere-2026-1438
Preprint. Discussion started: 21 April 2026

The ensemble wind speed nowcasts are constructed by adding the predicted errors to a physically-based deterministic forecast, which ensures that the resulting ensemble remains physically consistent. The spatial energy spectra of these ensembles also align more closely with the ground truth than those of the deterministic forecast. More importantly, because the errors are accurately predicted, the ensemble mean effectively corrects the deterministic forecast, leading to substantially lower RMSE and BIAS. The ensembles are shown to be well-calibrated through probabilistic verification, demonstrating that the diffusion model accurately estimates error uncertainty and, as a result, produces reliable uncertainty estimates for the wind nowcast. Nevertheless, the case study reveals the model's limitations in predicting extreme events. A promising direction for future work is to adopt a latent diffusion model, which operates in a compressed latent variable space. This would allow the model to more effectively uncover the latent structures linking forecasts and errors for extreme events, ultimately enhancing the representation of error uncertainty in the tails of the distribution.

Another contribution of this study is the comparative evaluation of commonly used noise schedules for diffusion models in weather forecasting applications. The results indicate that the Linear schedule yields the weakest predictive performance, while both Sigmoid and Cosine schedules achieve relatively high forecast accuracy. This suggests that the noise weights in the diffusion process should follow a smooth decay. Because the Sigmoid schedule requires more extensive experimentation to fine-tune the noise weights in the later stages of the diffusion process, the Cosine schedule proves to be more suitable for uncertainty prediction tasks in weather forecasting.

In summary, this study introduces a paradigm for probabilistic weather forecasting that shifts the focus from simulating atmospheric trajectories to directly predicting forecast error distributions. By coupling a dynamical model with a diffusion-based error generator, the proposed framework produces well-calibrated ensemble nowcasts while maintaining computational efficiency and physical consistency. The comparative analysis of noise schedules further provides reference for deploying diffusion models in practice. Future work will focus on enhancing the representation of extreme events through latent diffusion models, with the goal of extending this error-learning paradigm to additional variables and forecast horizons.

Code and data availability. To support reproducibility, the source code for the exact version of the model used in this study is permanently archived on Zenodo with the link https://doi.org/10.5281/zenodo.19673042 (Zhu, 2026a). The latest development version is additionally available in a GitHub repository: https://github.com/zywnuist/Diffwind/tree/master. The corresponding dataset has been deposited on Zenodo and can be accessed via: https://doi.org/10.5281/zenodo.19029543 (Zhu, 2026b).

https://doi.org/10.5194/egusphere-2026-1438
Preprint. Discussion started: 21 April 2026

Author contributions. Yanwei Zhu and Yong Wang equally contributed to this work. Yong Wang and Aitor Atencia proposed the method; Markus Dabernig designed the evaluations; Yanwei Zhu and Shuyan Zhou applied the method and perform the experiments; Yanwei Zhu wrote the manuscript draft and all authors reviewed and edited the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We are grateful to Nanjing University of Information Science & Technology for their support and technical assistance in this research. The authors gratefully acknowledge the financial support from the Postgraduate Research and Practice Innovation Program of Jiangsu Province (KYCX25_ 1595).

References

Alley, R. B., Emanuel, K. A., and Zhang, F.: Advances in weather prediction, Science, 363, 342–344, https://doi.org/10.1126/science.aav7274, 2019.

Andrae, M., Landelius, T., Oskarsson, J., and Lindsten, F.: Continuous Ensemble Weather Forecasting with Diffusion models, https://doi.org/10.48550/arXiv.2410.05431, 12 April 2025.

Asperti, A., Merizzi, F., Paparella, A., Pedrazzi, G., Angelinelli, M., and Colamonaco, S.: Precipitation nowcasting with generative diffusion models, Appl. Intell., 55, 187, https://doi.org/10.1007/s10489-024-06048-y, 2025.

Berenguer, M., Sempere-Torres, D., and Pegram, G. G. S.: SBMcast – An ensemble nowcasting technique to assess the uncertainty in rainfall forecasts by Lagrangian extrapolation, J. Hydrol., 404, 226–240, https://doi.org/10.1016/j.jhydrol.2011.04.033, 2011.

Bojinski, S., Blaauboer, D., Calbet, X., De Coning, E., Debie, F., Montmerle, T., Nietosvaara, V., Norman, K., Bañón Peregrín, L., Schmid, F., Strelec Mahović, N., and Wapler, K.: Towards nowcasting in Europe in 2030, Meteorol. Appl., 30, e2124, https://doi.org/10.1002/met.2124, 2023.

Buizza, R., Houtekamer, P. L., Pellerin, G., Toth, Z., Zhu, Y., and Wei, M.: A Comparison of the ECMWF, MSC, and NCEP Global Ensemble Prediction Systems, Mon. Weather Rev., 133, 1076–1097, https://doi.org/10.1175/MWR2905.1, 2005.

Cachay, S. R., Aittala, M., Kreis, K., Brenowitz, N., Vahdat, A., Mardani, M., and Yu, R.: Elucidated Rolling Diffusion Models for Probabilistic Weather Forecasting, https://doi.org/10.48550/arXiv.2506.20024, 24 June 2025.

Chen, J., Zhu, Y., Duan, W., Zhi, X., Min, J., Li, X., Deng, G., Yuan, H., Feng, J., Du, J., Li, Q., Gong, J., Shen, X., and Mu, M.: A Review on Development, Challenges, and Future Perspectives of Ensemble Forecast, J. Meteorol. Res., 39, 534–558, https://doi.org/10.1007/s13351-025-4909-4, 2025.

https://doi.org/10.5194/egusphere-2026-1438
Preprint. Discussion started: 21 April 2026

Chkeir, S., Anesiadou, A., Mascitelli, A., and Biondi, R.: Nowcasting extreme rain and extreme wind speed with machine learning techniques applied to different input datasets, Atmos. Res., 282, 106548, https://doi.org/10.1016/j.atmosres.2022.106548, 2023.

Du, J.: Uncertainty and Ensemble Forecast, Science & Technology Infusion Lecture Series, 42 pp, https://doi.org/10.25923/VPJE-W924, 2007.

Ebert, B.: Feature-specific verification of ensemble forecasts, The Centre for Australian Weather and Climate Research, Melbourne, Australia, available online at https://space.fmi.fi/Verification2009/presentations/tuesday/TUES_Session-6/O6.7_Ebert.pdf. (last access: 7 December 2025), 2009.

Ehrendorfer, M.: Predicting the uncertainty of numerical weather forecasts: a review, Meteorolo. Z., 6, 147–183, https://doi.org/10.1127/metz/6/1997/147, 1997.

Feng, J., Li, J., Zhang, J., Liu, D., and Ding, R.: The Relationship between Deterministic and Ensemble Mean Forecast Errors Revealed by Global and Local Attractor Radii, Adv. Atmos. Sci., 36, 271–278, https://doi.org/10.1007/s00376-018-8123-5, 2019.

Feng, J., Toth, Z., Zhang, J., and Peña, M.: Ensemble forecasting: A foray of dynamics into the realm of statistics, Q. J. Roy. Meteor. Soc., 1–24, https://doi.org/10.1002/qj.4745, 2024.

Fortin, V., Abaza, M., Anctil, F., and Turcotte, R.: Why Should Ensemble Spread Match the RMSE of the Ensemble Mean? J. Hydrometeorol., 15, 1708–1713, https://doi.org/10.1175/JHM-D-14-0008.1, 2014.

Gao, Z., Shi, X., Han, B., Wang, H., Jin, X., Maddix, D., Zhu, Y., Li, M., and Wang, Y.: PreDiff: Precipitation Nowcasting with Latent Diffusion Models, http://arxiv.org/abs/2307.10422, 28 December 2023.

Hamill, T. M.: Interpretation of Rank Histograms for Verifying Ensemble Forecasts, Mon. Weather Rev., 129, 550–560, https://doi.org/10.1175/1520-0493(2001)129%253C0550:IORHFV%253E2.0.CO;2, 2001.

Ho, J., Jain, A., and Abbeel, P.: Denoising Diffusion Probabilistic Models, http://arxiv.org/abs/2006.11239, 16 December 2020.

Jabri, A., Fleet, D., and Chen, T.: Scalable Adaptive Computation for Iterative Generation, https://doi.org/10.48550/arXiv.2212.11972, 14 June 2023.

Jolliffe, I. T. (Ed.): Forecast verification: a practitioner's guide in atmospheric science., Wiley, Chichester, 240 pp., 2004.

Kann, A.: On the skill of various ensemble spread estimators for probabilistic short range wind forecasting, Adv. Sci. Res., 8, 115–120, https://doi.org/10.5194/asr-8-115-2012, 2012.

Kingma, D. P. and Ba, J.: Adam: A Method for Stochastic Optimization, https://doi.org/10.48550/arXiv.1412.6980, 30 January 2017.

Lang, S., Alexe, M., Clare, M. C. A., Roberts, C., Adewoyin, R., Bouallègue, Z. B., Chantry, M., Dramsch, J., Dueben, P. D., Hahner, S., Maciel, P., Prieto-Nemesio, A., O'Brien, C., Pinault, F., Polster, J., Raoult, B., Tietsche, S., and Leutbecher, M.: AIFS-CRPS: Ensemble forecasting using a model trained with a loss function based on the Continuous Ranked Probability Score, https://doi.org/10.48550/arXiv.2412.15832, 20 December 2024.

https://doi.org/10.5194/egusphere-2026-1438
Preprint. Discussion started: 21 April 2026

Leinonen, J., Hamann, U., Nerini, D., Germann, U., and Franch, G.: Latent diffusion models for generative precipitation nowcasting with accurate uncertainty quantification, http://arxiv.org/abs/2304.12891, 25 April 2023.

Li, L., Carver, R., Lopez-Gomez, I., Sha, F., and Anderson, J.: Generative emulation of weather forecast ensembles with diffusion models, Sci. Adv., 10, eadk4489, https://doi.org/10.1126/sciadv.adk4489, 2024.

Lorenz, E. N.: A study of the predictability of a 28-variable atmospheric model, Tellus, 17, 321–333, https://doi.org/10.1111/j.2153-3490.1965.tb01424.x, 1965.

Loshchilov, I. and Hutter, F.: Decoupled Weight Decay Regularization, https://doi.org/10.48550/arXiv.1711.05101, 4 January 2019.

Mardani, M., Brenowitz, N., Cohen, Y., Pathak, J., Chen, C.-Y., Liu, C.-C., Vahdat, A., Nabian, M. A., Ge, T., Subramaniam, A., Kashinath, K., Kautz, J., and Pritchard, M.: Residual corrective diffusion modeling for km-scale atmospheric downscaling, Commun. Earth Environ., 6, 124, https://doi.org/10.1038/s43247-025-02042-5, 2025.

Nichol, A. and Dhariwal, P.: Improved Denoising Diffusion Probabilistic Models, https://doi.org/10.48550/arXiv.2102.09672, 18 February 2021.

Palmer, T. N.: Predicting uncertainty in numerical weather forecasts, Int. Geophys. Ser., 83, 3–13, https://doi.org/10.1016/S0074-6142(02)80152-8, 2002.

Pathak, J., Cohen, Y., Garg, P., Harrington, P., Brenowitz, N., Durran, D., Mardani, M., Vahdat, A., Xu, S., Kashinath, K., and Pritchard, M.: Kilometer-Scale Convection Allowing Model Emulation using Generative Diffusion Modeling, https://doi.org/10.48550/arXiv.2408.10958, 20 August 2024.

Price, I., Sanchez-Gonzalez, A., Alet, F., Andersson, T. R., El-Kadi, A., Masters, D., Ewalds, T., Stott, J., Mohamed, S., Battaglia, P., Lam, R., and Willson, M.: Probabilistic weather forecasting with machine learning, Nature, https://doi.org/10.1038/s41586-024-08252-9, 2024.

Schmid, F., Agersten, S., Bañon, L., Buzzi, M., Atencia, A., De Coning, E., Kann, A., Moseley, S., Reyniers, M., Wang, Y., and Wapler, K.: Conference Report: Fourth European Nowcasting Conference, Meteorolo. Z., 32, 83–87, https://doi.org/10.1127/metz/2022/1156, 2023.

Schultz, M. G., Betancourt, C., Gong, B., Kleinert, F., Langguth, M., Leufen, L. H., Mozaffari, A., and Stadtler, S.: Can deep learning beat numerical weather prediction? Phil. Trans. R. Soc. A., 379, 20200097, https://doi.org/10.1098/rsta.2020.0097, 2021.

Shu, D. and Farimani, A. B.: Zero-Shot Uncertainty Quantification using Diffusion Probabilistic Models, http://arxiv.org/abs/2408.04718, 8 August 2024.

Slingo, J. and Palmer, T.: Uncertainty in weather and climate prediction, Phil. Trans. R. Soc. A, 369, 4751–4767, https://doi.org/10.1098/rsta.2011.0161, 2011.

Song, J., Meng, C., and Ermon, S.: Denoising Diffusion Implicit Models, https://doi.org/10.48550/arXiv.2010.02502, 5 October 2022.

https://doi.org/10.5194/egusphere-2026-1438
Preprint. Discussion started: 21 April 2026

Sun, J., Xue, M., Wilson, J. W., Zawadzki, I., Ballard, S. P., Onvlee-Hooimeyer, J., Joe, P., Barker, D. M., Li, P.-W., Golding, B., Xu, M., and Pinto, J.: Use of NWP for Nowcasting Convective Precipitation: Recent Progress and Challenges, Bull. Am. Meteorol. Soc., 95, 409–426, https://doi.org/10.1175/BAMS-D-11-00263.1, 2014.

Toth, Z., Schultz, P., Mullen, S., Demargne, J., and Zhu, Y.: Completing the forecast: assessing and communicating forecast uncertainty, ECMWF Workshop on Ensemble Prediction, 7–9 November 2007, Reading, United Kingdom, available online at: https://www.ecmwf.int/sites/default/files/elibrary/2007/12792-completing-forecast-assessing-and-communicating-forecast-uncertainty.pdf, (last access: 17 February 2026), 2007.

Turner, R. E., Diaconu, C.-D., Markou, S., Shysheya, A., Foong, A. Y. K., and Mlodozeniec, B.: Denoising Diffusion Probabilistic Models in Six Simple Steps, http://arxiv.org/abs/2402.04384, 10 February 2024.

Van Schaeybroeck, B. and Vannitsem, S.: A Probabilistic Approach to Forecast the Uncertainty with Ensemble Spread, Mon. Weather Rev., 144, 451–468, https://doi.org/10.1175/MWR-D-14-00312.1, 2016.

Wang, R., Fung, J. C. H., and Lau, A. K. H.: Skillful Precipitation Nowcasting Using Physical-Driven Diffusion Networks, Geophys. Res. Lett., 51, e2024GL110832, https://doi.org/10.1029/2024GL110832, 2024.

Wang, Y., Bellus, M., Wittmann, C., Steinheimer, M., Weidle, F., Kann, A., Ivatek-Šahdan, S., Tian, W., Ma, X., Tascu, S., and Bazile, E.: The Central European limited-area ensemble forecasting system: ALADIN-LAEF, Q. J. Roy. Meteor. Soc., 137, 483–502, https://doi.org/10.1002/qj.751, 2011.

Wastl, C., Simon, A., Wang, Y., Kulmer, M., Baár, P., Bölöni, G., Dantinger, J., Ehrlich, A., Fischer, A., Frank, A., Heizler, Z., Kann, A., Stadlbacher, K., Szintai, B., Szűcs, M., and Wittmann, C.: A seamless probabilistic forecasting system for decision making in Civil Protection, Meteorolo. Z., 27, 417–430, https://doi.org/10.1127/metz/2018/902, 2018.

Wilson, J. W., Feng, Y., Chen, M., and Roberts, R. D.: Nowcasting Challenges during the Beijing Olympics: Successes, Failures, and Implications for Future Nowcasting Systems, Weather Forecast., 25, 1691–1714, https://doi.org/10.1175/2010WAF2222417.1, 2010.

Xiao, H., Wang, Y., Zheng, Y., Zheng, Y., Zhuang, X., Wang, H., and Gao, M.: Convective-gust nowcasting based on radar reflectivity and a deep learning algorithm, Geosci. Model Dev., 16, 3611–3628, https://doi.org/10.5194/gmd-16-3611-2023, 2023.

Yu, D., Li, X., Ye, Y., Zhang, B., Luo, C., Dai, K., Wang, R., and Chen, X.: DiffCast: A Unified Framework via Residual Diffusion for Precipitation Nowcasting, in: 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 27758–27767, https://doi.org/10.1109/CVPR52733.2024.02622, 2024.

Zanetta, F., Nerini, D., Buzzi, M., and Moss, H.: Efficient modeling of sub-kilometer surface wind with Gaussian processes and neural networks, https://doi.org/10.48550/arXiv.2405.12614, 13 November 2024.

Zhang, Y., Long, M., Chen, K., Xing, L., Jin, R., Jordan, M. I., and Wang, J.: Skilful nowcasting of extreme precipitation with NowcastNet, Nature, 619, 526–532, https://doi.org/10.1038/s41586-023-06184-4, 2023.

Zhong, X., Chen, L., Li, H., Liu, J., Fan, X., Feng, J., Dai, K., Luo, J.-J., Wu, J., and Lu, B.: FuXi-ENS: A machine learning model for medium-range ensemble weather forecasting, https://doi.org/10.48550/arXiv.2405.05925, 9 August 2024a.

https://doi.org/10.5194/egusphere-2026-1438
Preprint. Discussion started: 21 April 2026

Zhong, X., Chen, L., Liu, J., Lin, C., Qi, Y., and Li, H.: FuXi-Extreme: Improving extreme rainfall and wind forecasts with diffusion model, Sci. China Earth Sci., 67, 3696–3708, https://doi.org/10.1007/s11430-023-1427-x, 2024b.

Zhu, Y. W., Atencia, A., Dabernig, M., and Wang, Y.: Quantifying the analysis uncertainty for nowcasting application, Geosci. Model Dev., 18, 1545–1559, https://doi.org/10.5194/gmd-18-1545-2025, 2025.

Zhu, Y. W.: Code for denoising diffusion probabilistic model to generate 10-m wind speed ensemble nowcast, Zenodo [Model], https://doi.org/10.5281/zenodo.19673042, 2026a.

Zhu, Y. W.: Dataset for denoising diffusion probabilistic model to generate 10-m wind speed ensemble nowcast, Zenodo [Dataset], https://doi.org/10.5281/zenodo.19029543, 2026b.

Zhu, Y., Toth, Z., Wobus, R., Richardson, D., and Mylne, K.: The economic value of ensemble-based weather forecasts, Bull. Am. Meteorol. Soc., 83, 73–83, https://doi.org/10.1175/1520-0477(2002)083%253C0073:TEVOEB%253E2.3.CO;2, 2002.

Zhuang, X., Xue, M., Min, J., Kang, Z., Wu, N., and Kong, F.: Error Growth Dynamics within Convection-Allowing Ensemble Forecasts over Central U.S. Regions for Days of Active Convection, Mon. Weather Rev., 149, 959–977, https://doi.org/10.1175/MWR-D-20-0329.1, 2021.