General statement

We would like to thank the editor for coordinating the review of our work and the peer-reviewers for their valuable comments on our study. In the following, we will address the referees' comments and present our plans and ideas for revising the manuscript. For clarity, our responses are highlighted in blue.

Referee comment #2

The manuscript introduces a paradigm for uncertainty quantification in short-term weather forecasting, specifically focusing on high-resolution, 10-meter wind speed nowcasting. Traditional uncertainty quantification relies heavily on dynamical ensemble prediction systems (EPS) that simulate multiple atmospheric trajectories by perturbing initial conditions or physical parameterizations. While effective, EPS is computationally intensive for real-time applications and frequently suffers from under-dispersion. To overcome these constraints, the authors propose bypassing the generation of multiple physical trajectories altogether. Instead, they leverage a generative artificial intelligence framework—specifically a Denoising Diffusion Probabilistic Model (DDPM)—to directly model and predict the full conditional distribution of forecast errors, using a deterministic physical forecast as the conditioning background.

Data preprocessing includes a logarithmic transformation of wind speed data, which successfully converts skewed wind speed errors into a near-Gaussian distribution and mathematically restricts the reconstructed wind speed fields to positive values. A core contribution of the paper is the empirical evaluation of three distinct noise scheduling frameworks (Linear, Cosine, and Sigmoid) within the diffusion process. The experimental results demonstrate that the choice of the noise scheduler is critical for structural and statistical fidelity. The Cosine schedule outperforms the alternatives across multiple verification metrics, achieving optimal error scores and well-calibrated histograms. Furthermore, spatial power spectrum analysis confirms that the samples generated via the Cosine schedule preserve the correct kinetic energy cascade and spatial structures of the wind fields.

The manuscript is well-structured, the physical treatment of data via logarithmic transformation is sound, and the comparison of noise schedules (Linear, Cosine, and Sigmoid) provides valuable empirical insights for the atmospheric modeling community. However, there are critical limitations regarding the temporal coverage of the dataset, the severe seasonal bias in the verification phase, the lack of traditional statistical baselines, mathematical inconsistencies in the description of the diffusion framework, and several data visualization/structural clarity issues that must be addressed before publication.

Reply:

We thank the reviewer for the thorough and constructive assessment of our work. We will carefully address all the concerns raised and provide clarifications where needed. The revisions will be traceable in the manuscript.

Major Concerns:

#Comments 1

(Lines 85–87): The total historical record used spans from 1 October 2021 to 30 June 2023. This represents less than two full years of atmospheric data. Since deep generative models like DMs typically benefit from a substantial and diverse set of samples to properly map high-dimensional chaotic fields, a dataset shorter than two years may increase the risk of overfitting to the specific intra-annual anomalies of those two cycles, potentially limiting the model's capacity to generalize to long-term climate variability. It would be highly beneficial if the authors could provide a brief justification as to why a dataset shorter than 24 months is sufficient for high-resolution (1 km) spatial error generation, or alternatively, consider expanding the training set by incorporating additional years of archival data from the SIVA system.

Reply 1:

Thanks for the comments. We agree that a longer training period would generally be desirable, and we fully acknowledge the concern about the limited temporal coverage.

We would like to clarify an important distinction between our task and typical time-series forecasting with diffusion models. In many DM-based forecasting applications, the model is asked to predict a future state x_{T+t} given a past state x_T, meaning that the temporal evolution of the field must be learned from the data. In our case, the DDPM does not predict the future state directly; rather, it learns a conditional mapping from the deterministic forecast y (which is already valid at the target time) to the corresponding forecast error $e = x_{true} - y$. The input and output are contemporaneous—both refer to the same valid time. The temporal dimension enters only through the training samples, not through the prediction task itself. This conditional formulation alleviates the need to learn temporal dynamics, and consequently reduces the demand for a long continuous time series.

To provide some empirical evidence that the model can generate spatial errors sufficiently, we compared the daily-mean wind speed from the DDPM ensemble mean, the deterministic SIVA forecast, and the analysis field for the test month (June 2023). As shown in Figure 1, the ensemble mean consistently lies closer to the analysis than the deterministic forecast does, indicating that the model learns useful error structures even with the current dataset.

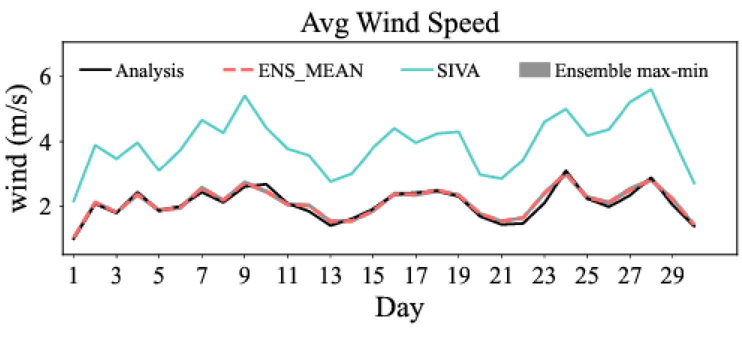

Fig.1. Daily mean wind speed over the test period (June 2023) at lead time 6h, comparing the analysis field (black), DDPM ensemble mean with cosine noising schedule (dashed orange), SIVA nowcast (light blue), and the ensemble maximum and minimum (shadow).

As noted in our response to Comment #2, we will extend this analysis to additional months (e.g., January) to further assess seasonal consistency once the retrained model becomes available. And we acknowledge that a longer training period would be beneficial, especially for rare extreme events. This limitation will be noted in the Conclusions and the revision can be traceable in the manuscript.

#Comments 2

(Lines 88–90 and Line 167): The authors state that the dataset was split chronologically, leaving the "remainder" for testing (which corresponds strictly to June 2023, as confirmed in Line 167). As defined in Lines 85–86, the spatial domain covers East China (115.35∘E−122.33∘E, 29.88∘N−35.81∘N). In this region, June is climatologically dominated by a meteorological regime characterized by stationary front dynamics. Evaluating the DDPM's performance exclusively during this single month could introduce a seasonal bias. Unless a specific seasonal behavior is being targeted (which should then be explicitly mentioned in the manuscript), it is recommended to evaluate the model across a broader range of seasonal conditions. Given the relatively small size of the dataset, a comprehensive evaluation might be challenging; however, adopting a validation strategy that ensures cross-seasonal representation would greatly help to demonstrate the model's generalizability across different atmospheric regimes.

Reply 2:

Thanks for the comment. We will retrain the model using a training period that excludes the test months: October 2021 – December 2022 and February 2023 – April 2023. We then will evaluate the Cosine schedule (the best-performing scheme) on both January 2023 (winter) and June 2023 (summer). The results will be compared to check for seasonal consistency and discussed in Section 4.

#Comments 3

(Lines 68–70 and Lines 159–161): In Lines 68–70, the authors state that, unlike conventional statistical post-processing methods, their proposed model explicitly learns the full conditional distribution of errors. Later, in Lines 159–161, they explain the lack of a reference ensemble by noting that the operational SIVA system is purely deterministic. While this clarifies the absence of a dynamical ensemble comparison, it would be highly valuable to include a reference baseline for probabilistic verification. To more clearly demonstrate the added value and skill of the conditional estimation provided by the DDPM, the model could be benchmarked against standard reference baselines. Even a simple climatological distribution or a persistence-based error model would serve as an insightful benchmark to quantify the actual statistical improvement and contextualize the computational investment required by the diffusion architecture.

Reply 3:

Thanks for the suggestion. We agree that a reference baseline would help to demonstrate the added value of the DDPM. We will implement a climatological baseline, where ensemble members are generated by randomly sampling from the historical distribution of SIVA forecast errors (pooled from the training set). The performance of this climatological baseline will be compared to the DDPM using the same verification metrics (e.g., CRPS, reliability diagrams). The results will be presented and discussed in Section 4. A brief description of the baseline method will also be added in Section 3.

#Comments 4

(Lines 105–106): In Lines 105–106, the text states that μθ(xt,t) and σ2 are the mean and variance of the distribution at the previous step. While the explicit formulation to obtain μθ is provided in Equation 3, it would be helpful if the authors could explicitly define how σ2 is handled or computed.

Reply 4:

Thanks for the comment. In our implementation, we follow Ho et al. (2020) and set $\sigma^2 = \beta_t$, where β_t is the predefined noise variance at step t (the same as in the forward process). We will add a brief clarification after Equation (2). The revision will be traceable in the revised manuscript.

#Comments 5

(Line 109 and Equation 4): In Line 109, the network is introduced as ϵθ(xt,t), yet in Equation 4, it appears as ϵθ(xt,y,t). To avoid ambiguity, the authors should clarify whether the conditional variable y should be consistently included in the notation throughout the text.

Reply 5:

Thanks for pointing it out. The conditional variable y (the deterministic nowcast) is used throughout the diffusion process. To avoid ambiguity, we will consistently write $\epsilon_\theta(x_t, y, t)$ in both the text and Equation 4. The revision will be traceable in the manuscript.

#Comments 6

(Figure 1 / Section 2.3): The flowchart in Figure 1, which outlines the overall framework of the model, can be somewhat difficult to follow, even when cross-referenced with the text explanation in Section 2.3. To further enhance its clarity and make it more intuitive for the reader, it is suggested to refine the representation of the data flow. Specifically, explicitly linking the visual blocks to the mathematical variables used in the text—such as the physical background (y), the forward/reverse processes (xt), and the final output—would significantly improve the diagram's transparency. Explicitly referencing the corresponding equation numbers (Equations 1–4) within the flowchart components would also greatly assist the community in understanding and reproducing the exact methodology.

Reply 6:

Thanks for the suggestion. We will redesign the flowchart to make it more self-explanatory. Specifically, we will:

- Add numbered steps (①, ②, ⋯) to indicate the start of training and the main stages.
- Explicitly link the visual blocks to mathematical variables.
- Reference the corresponding equation numbers within the flowchart.

The revised figure will be included in the manuscript.

Minor Concerns:

#Comments 7

(Figure 6 / Section 4.1): In Section 4.1, specific conclusions are drawn based on the spatial patterns shown in Figure 6. However, the current color palette utilizes very soft gradients and smooth transitions, which can make it challenging for the reader to visually discern the differences highlighted in the text. It is suggested that the authors consider changing the color scale of Figure 6 to a high-contrast or perceptually uniform divergent palette to visually enhance and further substantiate the claims made in the manuscript.

Reply 7:

Thanks for the suggestion. We will change the color palette of Figure 6 to a perceptually uniform diverging scheme to better highlight the spatial differences discussed in the text. The revised figure will be included in the manuscript.

#Comments 8

(Supplementary Material / Text Structure): The manuscript contains a prolonged text explanation discussing a specific figure that has been placed entirely in the Supplementary Material. When a figure requires such an extensive textual breakdown and is central to supporting the core arguments, forcing the reader to search for it in a separate document can disrupt the flow of reading. To make the paper more self-contained and streamline the reading process, it is recommended to move this specific figure from the supplementary material into the main manuscript, embedding it close to its corresponding text discussion.

Reply 8:

Thanks for the suggestion. We agree that a figure requiring detailed discussion should not be placed in the supplementary material. We will move this figure to the main text and update the cross-references accordingly. The revision will be traceable in the manuscript.