

We sincerely thank the reviewers for their thoughtful and constructive feedback on our manuscript. Their insightful comments have significantly strengthened the rigor and clarity of this work. Below, we provide point-by-point responses to each comment. Should any response require further clarification or if additional concerns arise, we welcome the reviewer's continued guidance. We remain committed to refining the manuscript in accordance with the reviewer's recommendations.

Reviewer 1:

(1) As described in Abstract, “Moreover, the model sustains consistent performance across varying missing data lengths and continuity patterns, indicating robust generalization capabilities.”, the statement of “consistent performance” is based on the experimental results in section 4.4. However, the numerical values are not the same. How should this be understood?

Reply: The emphasis here on maintaining consistency is on the robustness of the model. Specifically, relative stability is demonstrated by the gradual and moderate decline in performance as the length of missing data increases and the continuity of missing segments deepens—yet all metrics remain consistently above a high baseline, with no abrupt or catastrophic degradation. Furthermore, no matter what kind of missing situation is encountered (such as an increase in the length of missing data or a deeper degree of continuity), the model can effectively reconstruct the missing echoes, rather than completely failing in certain cases.

(2) In Section 2.2, why is it necessary to separately establish an independent validation set to evaluate the model's ability in reconstructing the missing echoes?

Reply: Since this paper adopts an unmasked learning framework, the model will simultaneously reconstruct both the missing areas and the non-missing areas during the training process. Therefore, in the conventional validation process, the overall evaluation metrics are prone to be influenced by a large number of non-missing areas. The non-missing

areas usually have high quality, and their reconstruction errors are relatively small. Thus, the overall metrics cannot fully reflect the model's ability to restore the truly missing areas. Based on this, this paper sets up an independent validation set and focuses on evaluating the missing areas separately to more accurately measure the model's performance in recovering missing echoes. At the same time, the results of this independent validation can also be used to assist in parameter selection and model optimization during the model training process, thereby improving the robustness of model.

(3) Why wasn't the optimization of various hyperparameters carried out in the validation set during the period of March to May 2022?

Reply: Since the overall evaluation metrics were adopted in the validation set from March to May in 2022, and the non-missing areas accounted for a large proportion of the whole sample, the overall metrics were prone to be dominated by the non-missing areas, and thus could not fully reflect the model's actual recovery ability for the missing areas. Therefore, this paper additionally established an independent validation mechanism for missing areas to further evaluate the model's recovery performance in the missing areas and assist in the optimization of relevant hyperparameters.

(4) In the method of constructing the loss dataset, if the radar echoes within the selected area are set to zero using rectangles, will this cause the model to be more inclined to repair the rectangular missing areas?

Reply: This study does not rely on explicit missing mask during training and inference. This means that when the model is learning, it does not know whether the zero values in the input are real no radar echoes or missing rectangular areas. It treats all zero values uniformly, that is, by inferring their most likely values by capturing the intrinsic spatiotemporal evolution patterns and structural features of the radar echo sequence. If the model develops a preference for repairing rectangles, it first needs to be able to distinguish and identify the rectangular boundaries, which is precisely the information that this study deliberately avoids allowing the model to learn through mask-free training.

(5) Why was PixelShuffle used in the model design?

Reply: PixelShuffle can not only reduce the artifacts caused by upsampling, but also more effectively restore the spatial boundaries and high-frequency structures of radar echoes, while balancing the reconstruction quality and computational efficiency. Specifically, 1) It reduces the problem of checkerboard artifacts caused by the uneven coverage of the convolution kernel during deconvolution. 2) It can better preserve spatial boundaries and fine-scale structures by enabling feature learning in the low-resolution feature space and then restoring the spatial resolution through a channel-to-space rearrangement operation. This is more conducive to maintaining the continuity of echo boundaries and strong convective structures. 3) Compared with interpolation-based methods, PixelShuffle enables learnable feature reconstruction while maintaining higher computational efficiency.

(6) The weighted L1 loss in this paper adopts an artificial discrete weight design based on the reflectance threshold. Could you please explain the basis for setting these weights? Additionally, is this discontinuous segmented weight likely to cause unstable gradients during the optimization process and subsequently lead to a higher FAR in the experiments?

Reply: The design of the weighted L1 loss is mainly aimed at addressing the severe imbalance issue of strong and weak radar echoes in the samples. Since although the strong echo areas account for a relatively small proportion in the samples, they have higher importance in meteorological services, this paper increases the loss weight of the strong echo areas to guide the model to pay more attention to the recovery ability of the strong convective structure and the high reflectivity core area. The current weight design is indeed an empirical solution; however, the ablation experiment results show that this design can effectively improve the CSI and POD under strong echo conditions, indicating its effectiveness in enhancing the recovery ability of strong radar echoes. Moreover, we performed an ablation study comparing the standard L1 loss against the weighted L1 loss. The results consistently demonstrated that the standard L1 loss performs worse than the weighted L1 loss in terms of metrics such as PSNR, SSIM, and CSI. **The relevant descriptions can be found in Section 4.1 (lines 300–301, 328–332).**

(7) Why not adopt the model learning with missing mask instead of using mask-free learning?

Reply: In the actual application of radar mosaic data, the positions, shapes and temporal continuity of the missing radar echoes usually exhibit strong randomness and irregularity. Moreover, high-quality and accurate missing mask is often difficult to obtain stably in many cases. If the model overly relies on explicit mask information, it is prone to developing a dependency on specific missing patterns, thereby reducing its generalization ability for unknown missing patterns. At the same time, the traditional mask-based restoration methods essentially belong to local completion under the condition of known missing positions. The model needs to know in advance which areas have missing data. This study aims to explore a more general spatiotemporal sequence reconstruction framework, enabling the model to rely solely on the spatiotemporal evolution laws of the radar echoes, autonomously identify and restore abnormal areas, thereby reducing its reliance on prior missing information. Of course, we also recognize that mask-free learning still has certain limitations. For example, in large-scale weak echo or long-duration weak precipitation scenarios, there may be some confusion between the real low-value areas and the missing areas. Therefore, in future work, we will further explore related research such as uncertainty-aware modeling, attention-guided spatiotemporal restoration, etc.

(8) Will post-processing result in discontinuity at the boundaries between the missing and non-missing regions?

Reply: Since the model adopts an overall reconstruction strategy, the missing areas and the non-missing areas establish temporal and spatial consistency during the restoration process. Therefore, in the post-processing step, only replacing the original values of the non-missing areas will not cause obvious boundary discontinuities. Moreover, the visualization results in the experiment show that the model can maintain good spatial continuity and structural coordination near the missing boundaries.

(9) In the non-missing area impact analysis experiment of section 4.3, the statement “Collectively, these results demonstrate that the post-processing strategy effectively preserves structural fidelity in non-missing regions while maintaining robust reconstruction capability

in missing regions.” is not very appropriate. From the results, it only relates to the non-missing areas and does not address the missing areas.

Reply: Thank you for your reminder. The description here is indeed not precise enough. The sentence should be revised as follows: “Collectively, these results demonstrate that the post-processing strategy effectively preserves structural fidelity in non-missing regions while not affecting the reconstruction results in the missing areas, thereby improving the quality of the final restoration outcome.”

(10) This paper mainly verified the model’s adaptability to different lengths of missing data and changes in temporal continuity. But do these missing patterns overlap with the missing distribution during the training stage? Could the model’s generalization ability be derived from the memory of the statistical characteristics of the missing patterns rather than truly learning the spatiotemporal evolution laws?

Reply: The generalization experiments in this paper mainly focus on the model’s adaptability to different missing lengths and changes in temporal continuity. Although the training stage employed random missing simulation, the testing stage further designed different lengths and different temporal continuities of missing echoes to minimize the direct correspondence between the training and testing missing patterns. Moreover, due to the use of the mask-free learning framework, the model cannot directly rely on fixed missing positions or shapes to complete the restoration. Therefore, its restoration ability mainly stems from the learning of the spatiotemporal evolution laws of radar echoes rather than simple memorization of specific missing patterns. The experimental results show that even in cases of longer continuous missing, the CSI and POD of the model still show a certain decline, but overall remain relatively stable, indicating that the model possesses certain spatiotemporal dynamic modeling capabilities and robustness in missing pattern recovery.

Reviewer 2:

(1) The data only cover March to September. If the method is intended only for the rainy season, this should be stated clearly. Otherwise, the authors should test whether the model generalizes to other months, including winter or dry-season cases, to rule out overfitting to seasonal echo patterns.

Reply: Thank you for raising the important question regarding the seasonal applicability of the model to different seasons. To address this, we evaluated the model's generalizability using its pre-trained weights on precipitation data from non-rainy-season months in 2023 (i.e., January–February and October–December). Quantitative results are as follows: PSNR = 25.838; SSIM = 0.749. At 20 dBZ, CSI = 0.516, POD = 0.610, and FAR = 0.239; at 30 dBZ, CSI = 0.320, POD = 0.471, FAR = 0.309; and at 40 dBZ, CSI = 0.132, POD = 0.289, FAR = 0.401. These results indicate that the proposed method maintains reasonable generalizability beyond the rainy season (March–September). Nevertheless, quantitative metrics—including CSI, POD, and FAR—exhibit noticeable degradation at higher reflectivity thresholds (30 dBZ and 40 dBZ), primarily attributable to the low frequency and spatial sparsity of strong-echo events in non-rainy-season data. This scarcity of informative strong-echo samples limits the model's capacity to learn robust representations for such phenomena, thereby reducing reconstruction accuracy. **A detailed discussion of this limitation—including its implications for model calibration and potential mitigation strategies—is provided in Section 5 (lines 471–484).**

(2) The paper should explain how complete radar samples were selected before artificial gaps were added. Were samples with real missing sectors removed? How was completeness checked?

Reply: A very good question. Indeed, special clarification is needed on this matter. In practice, radar mosaic data gaps predominantly occur near the periphery of the mosaic domain due to data acquisition delay, limited radar coverage and beam blockage. To ensure the integrity and physical fidelity of the training samples—and to avoid conflating model-inferred

reconstruction with genuine data loss—we deliberately selected a high-coverage subregion centered on Wuhan, the capital of Hubei Province, where overlapping radar coverage is maximal and echo absence is almost impossible to happen. This spatially constrained sampling strategy minimizes the risk of embedding systematic missingness into the dataset.

The relevant explanations for data selection can be found in Section 2.1 (lines 91–94).

(3) Values below 10 dBZ are set to 0, and missing regions are also encoded as 0. This makes true no-echo pixels indistinguishable from missing pixels. The authors should test whether the model introduces false echoes in valid zero-valued regions, especially because FAR is relatively high.

Reply: Thanks for this insightful question. This study focuses on learning the spatiotemporal evolution patterns of radar echoes via supervised model training, and—critically—does not employ any missing-data masking mechanism. Consequently, during inference, the model interprets all zero-valued pixels uniformly as instances requiring reconstruction, leveraging its learned dynamical patterns to estimate the most probable reflectivity value at each such location. As a result, the dataset construction does not require explicit differentiation between meteorologically genuine “no-echo” pixels (i.e., physically valid zeros) and instrumentally missing pixels (e.g., due to radar blockage or data transmission failure). That said, validation experiments conducted on fully observed (non-missing) regions confirm that the method may occasionally generate spurious echoes in areas where zero values are physically meaningful. To mitigate this, the post-processing module was explicitly designed to minimize this risk to the lowest level.

(4) The post-processing step naturally improves RMSE and MAE in non-missing regions because observed values are copied back. This does not prove that the final fields are physically consistent. The authors should evaluate possible discontinuities or non-physical structures near the boundaries of reconstructed regions.

Reply: We sincerely thank the reviewer for identifying the potential issue of boundary discontinuity arising from post-processing—a critical concern that had not been adequately

addressed in our initial analysis. To quantitatively assess boundary continuity, we propose the Boundary Continuity Difference (BCD) index, a metric designed to measure local intensity discontinuity across the boundary region. Specifically, the BCD is computed over a symmetric 5-pixel-wide band centered on the boundary—both inside and outside the missing region—with lower values indicating superior continuity. Experimental evaluation yields a BCD of 6.41 for the raw model prediction and 6.32 for the post-processed output, confirming that post-processing does not exacerbate boundary discontinuity. This finding aligns with qualitative observations from localized visual inspection, where boundary transitions remain smooth and free of visible stitching artifacts. It is important to emphasize that the BCD captures only local intensity consistency at the boundary and should therefore be interpreted in conjunction with global error metrics (e.g., PSNR) and holistic visual assessment. Collectively, the quantitative and qualitative evidence indicates that the proposed post-processing step preserves structural coherence and introduces no perceptible discontinuities. **The supplementary explanation regarding this issue has been added in Section 4.3 (lines 386–395, 399–406).**

(5) The evaluation mainly uses synthetic missing regions. The authors should test the method on real missing radar patterns, at least qualitatively. Otherwise, it remains unclear how well the model generalizes to operational cases.

Reply: We fully concur with your viewpoint. We conducted tests on other real missing data (individual cases not from the dataset in this study), and the visualization results are presented in the attached file. For data with real radar echo missing, it is extremely challenging to obtain corresponding complete data for comparison. Nevertheless, we can still indirectly evaluate the quality of the model reconstruction based on the radar echoes of the preceding and succeeding frames. From the visualization results, it is evident that the overall reconstruction effect of the model is satisfactory. It basically restores the distribution characteristics of the radar echoes of the precipitation system, not only effectively reconstructing the major strong echo areas but also accurately depicting the weak and small echoes. Although there are issues such as the

weakening of the extreme echo intensity and the consistency of edge transition, the current restoration effect still demonstrates high value for practical applications. These existing deficiencies precisely serve as the objectives for future research and improvement. **The test results of the real missing data have been added as Figure 9 in the appendix (lines 503–504).**

(6) The novelty should be demonstrated more directly. Since a missing-value mask provides useful information, removing it may reduce performance. A comparison with an otherwise similar mask-conditioned model would clarify whether the mask-free approach is beneficial or only operationally convenient.

Reply: Thank you for giving this constructive suggestion. We fully acknowledge that incorporating missing-value masking as prior information is a well-established and effective strategy, and comparative analysis against masking-based approaches would indeed offer valuable insights into the distinct behavioral characteristics of the two techniques. However, the primary objective of this work is not to assert the superiority of mask-free inpainting over masking-conditioned methods, but rather to investigate a fundamental question: whether high-fidelity reconstruction of missing spatiotemporal regions can be achieved solely from contextual information—without explicit encoding of the missing mask as input. We expect to provide a technical reference for operational applications where the explicit missing mask cannot be obtained. Consequently, our study centers on establishing the conceptual feasibility, architectural viability, and practical utility of mask-free modeling, rather than conducting a head-to-head performance benchmark across masking mechanisms. Given that a rigorous comparative evaluation would necessitate substantial modifications—including redesigning baseline architectures, retraining all models under consistent protocols, and ensuring fair ablation controls—it falls outside the scope and resource constraints of the present investigation. **Nevertheless, this limitation is explicitly acknowledged in Section 5, and a follow-up study is planned to conduct a systematic comparative evaluation of current state-of-the-art masking-based approaches (lines 488–493).**

(7) Pixel-wise losses can smooth sharp convective structures. The authors should analyze

whether the model underestimates peak reflectivity, smooths intense cells or loses small-scale structures, especially at higher reflectivity thresholds.

Reply: Thank you for your suggestions. We agree that the loss function based on pixel level might smooth out the sharp convection structures. Moreover, the differences in overall indicators between the experiments with and without pixel-level loss in the ablation study in Section 4.1 also indicate that the pixel-level loss has more benefits than drawbacks for the results. Of course, regarding this issue, we will conduct optimization research later from aspects such as model structure, loss design, and post-processing. **At the same time, we have added relevant descriptions in Section 5 (lines 466–470).**