[Response to referee report by Anonymous Referee #3, 21 Sep 2025]

(the original referee comments are denoted in blue and italic texts)
This study conducted observation system simulation experiments (OSSEs) idealized for high-frequency radar reflectivity data assimilation (DA) with local ensemble transform Kalman filter (LETKF) and compared 3 experiments: (i) 3D-LETKF every 5 min, (ii) 4D-LETKF every 5 min with observations every 30 sec, and (iii) 3D-LETKF every 30 sec. As a result, (iii) had the smallest non-Gaussianity of first guess ensemble and the best analysis accuracy. Since it is interesting to clarify the advantages of high-frequency DA with the idealized OSSEs, this study is valuable to be published. However, the causal relationship between the non-Gaussianity of first guess ensemble and the analysis accuracy is not clear even in the OSSEs conducted in this study. Therefore, it should be clarified that this study does not investigate pure impact of non-Gaussian distribution but imply the advantages of high-frequency DA partially in the viewpoint of non-Gaussianity. To prevent misunderstanding about it, I recommend major revision. The order of the following comments is not related to importance.

We appreciate the referee for the valuable comments. We have revised the manuscript accordingly. Below are the responses to each comment.

Specific comments:

1. L4-5, L50-51, L75-77, and L286-287: The OSSEs conducted in this study do not completely exclude the impact other than non-Gaussianity because the DA method, the number of assimilated observations, the ensemble spread, etc. are different between the 3 experiments compared in this study. This study does not clarify the pure impact of non-Gaussian distribution but imply the advantages of high-frequency DA partially in the viewpoint of non-Gaussianity. The sentences should be revised not to be misunderstood.
We appreciate the important comment. We agree that the experimental design did not exclude the impact of the difference in ensemble spread among the three cases, which is significant in some parts of the domain. Also we agree that the number of assimilated observations is different between 4D-5MIN and 30SEC because of the threshold of gross error and the number of first guess ensemble members for the assimilation. We have rewritten the sentences to make the scope of this study clearer (L4-5, L53-54, L77-79, and L314-317 in the revised manuscript).

2. L20-23: This description is not necessarily correct in 4D-EnKF. Since the experiments with 4D-EnKF are conducted in this study as well as 3D-EnKF, the development of 4D-EnKF and its advantage also should be explained.
I have added the description of 4D-EnKF (L23-26 in the revised manuscript).

3. L53-54: Why is this study useful for the non-Gaussian DA? Could you cite any previous studies?
I have rewritten the sentence to make the implication clearer (L55-56 in the revised manuscript).

4. L73-77 and L288-290: This experimental design does not completely exclude modification to the ensemble perturbations because spatial localization is applied to reduce the effect of sampling error. This limitation should be stated.
In this sentence we meant to focus on the ensemble perturbation as an approximation of the background probability distribution. We think it is not the ensemble perturbation but the spatial structure of background error correlation that is modified by spatial localization.

5. Figure 1: To add axis of height (km) is helpful to understand the height of convections shown in other figures.

We have revised Figure 1 adding the secondary vertical axis showing height.

6. L90-91: The forcing by the warm bubble should be stated more concretely and quantitatively here.

It was stated in the sentences which follow L90-91. We have rephrased the description to make it clear (L92-95 in the revised manuscript).

7. L137-138: Does it mean that 5-dBZ reflectivity is assimilated even where the first guess < 5 dBZ? In this case, the precipitation becomes stronger in the analysis. Is it no problem?

It does not occur as the adjustment to 5 dBZ is also applied to the first guess. We have added the explanation about it (L139-141 in the revised manuscript).

8. L143-146: This description is redundant and difficult to be understood. Does it mean that reflectivity is assimilated only where at least one ensemble member > 10 dBZ in the first guess?

We have rephrased the description and moved it to the earlier part just after the description of the treatment of reflectivity (L141-142 in the revised manuscript).

9. L160-161: Why was the potential temperature perturbation over the entire domain in addition to the perturbation in the warm bubble?

It was applied to add nonzero spread of the dynamical variables in the area outside the convective cell. We have added the explanation (L162-164 in the revised manuscript).

10. L161-162: Were the Gaussian perturbations added in the warm bubble and the whole domain at the first assimilation cycle? If so, why is the first guess ensemble expected to be non-Gaussian? How to add the perturbations should be explained more clearly.

In the previous version of the manuscript, the expression "the first guess ensemble *at the first data assimilation cycle* is expected to have ..." was not clear and confusing.

In fact, the first time step when the radar reflectivity is assimilated is 00:10:00, as it takes about 10 minutes from the initial time for the convective cell to develop enough to be observed. The expected non-Gaussianity comes from the development of perturbation for that initial 10 minutes interval.

We have revised the description (L164-165 in the revised manuscript). Also, as the response to the comment 12 below, we have added Fig.3 to show the time series of the number of assimilated observations.

11. L183: What determines this kernel bandwidth? Could you cite any previous studies?

The equation (9) is derived in the Silverman (1986) textbook and commonly used for kernel density estimation. I have added the reference.

12. L191-192: The ensemble spread in the 5MIN-4D case should also be stated. In addition, it is better to show the time series of the ensemble spread until 00:50:00 in all cases to confirm that the filter divergence has not occurred.

We have added the maximum value of the ensemble spread in the 5MIN-4D case in the text (L211-212 in the revised manuscript). We have also added Fig.3, which shows the time series of the total number of assimilated observations in 5-minutes window and ensemble spread in reflectivity averaged over the grid points where the true reflectivity value is over 10dBZ.

13. Figure 5: To make the discussion in Section 3.2 deeper, the ensemble spread should also be shown in addition to ensemble mean and the difference from the nature run. I think main difference between the 3 experiments is the ensemble spread.
We have added the contours of ensemble spread in reflectivity and surface accumulated precipitation in Fig. 5 and 11 (Fig.6 and 12 in the revised manuscript).

14. L213-214, L276-277, L279-280, L283-284, and L309-314: If the ensemble spread or the KLD is largely different between the 3 experiments, these descriptions are not precise for ensemble forecasts. The difference should be shown also for the ensemble forecasts.
The ensemble spreads of the three experiments shown in Fig.5 and 11 (Fig.6 and 12 in the revised manuscript) are mostly not different from each other, though they are different in analysis reflectivity and vertical velocity. It further supports the conclusion in those sections that the different data assimilation frequency has a limited impact on precipitation forecast. We have added the additional sentences (L235-237 and L304-306 in the revised manuscript).

15. Figure 6: The time of the 100 samples should be shown as well as the position of the grid point.
We have added the time to the caption of Fig. 6 (Fig.7 in the revised manuscript).

16. L221-223 and L304-306: The nonlinear cross-variable relationship between graupel mixing ration and vertical velocity looks caused by difference of non-Gaussianity of these variables (vertical velocity is closer to Gaussian). If this interpretation is correct, it should be stated.
Our interpretation is rather that nonlinearity appears as non-Gaussianity. When there is a nonlinear relationship between two variables caused by the nature of the system dynamics, they have different degrees of non-Gaussianity in the probability distribution, even though one of them may be nearly Gaussian.

17. L227-228: Could you show the mathematical definition of "the mutual information between the ensemble members of graupel and vertical velocity at every grid point, after removing the linear dependency"?
We have added the description and mathematical formula to calculate it in Section 2.5.

18. Figure 7: The caption should include "in the 5MIN-3D case".
We have revised the caption as suggested.

19. L230-231 and L306: The nonlinear relationship between the two variables may be one cause of the error of the analysis. However, it may simply be caused by smaller impact of assimilation in the part of smaller ensemble spread. Or, the smaller number of observations assimilated with low-frequency may also make the large error of the analysis. It is better to show various possible causes.
We agree that the impact of the difference of first guess ensemble spread in vertical velocity may be significant and the impact of nonlinearity can't be easily evaluated separately from it. We have revised the sentences to mention both possibilities (L255-256 and L333-334 in the revised manuscript).

20. L237: "pertrbed" -> "perturbed"
We have corrected the typo.

21. L253-255, L266-269, L323-324, and L325-326: These disadvantage of 5MIN-4D and 30SEC are general and should be found also in the experiments in Section 3. Why are they *found only in the additional experiments in Section 4?*

We have a speculation that the side effect appears in those experiments because of the large bias in the first guess caused by a biased background profile, which is the main difference from the experiment in Section 3.