Reviewer #3

General comments:

This study focuses on investigating the impacts of DA update frequency and observation number on the non-Gaussianity of model simulation error, in a case of strong convection. Results shown in the manuscript show that the non-Gaussianity of error can be reduced by increasing the DA frequency and number of observations, which could possibly improve the performance of EnKF. While the results are impressive, there are several problems the authors may need to address before the manuscript is published. I hereby recommend a major revision.

We would like to thank Reviewer #3 for raising interesting points that helped expand the discussion and to add interesting results. Also the comments helped to clarify several aspects that were not clear in the original version of the manuscript.

Specific comments:

1. Model configuration: The authors did not provide enough information about the model's configurations. In line 59, the model used in this study has a horizontal resolution of 1 km, 50 vertical sigma levels, and a size of 180 km by 180 km (Fig. 1a). I wonder what the range and resolution of the vertical sigma levels are defined. According to my knowledge, models with higher horizontal resolutions should also have higher vertical resolutions. The number of vertical levels of the model introduced in this study is probably too coarse for 1km-scale simulations. (also see specific comment 2)

There were missing important configuration aspects in the original version of the manuscript. To address this point we add more details about the vertical discretization. 50 vertical levels extend up to 18 km elevation with a variable grid spacing from 140 m to 790 m in a hybrid sigma-z terrain-following coordinate.

2. According to the paper, better results of radar data assimilation were obtained with vertical localization scale of 2 km and horizontal localization of 4 km (Line 70 – 73). In this sense, a 1:2 ratio of the horizontal to vertical resolutions of the model could give more reliable simulation results. i.e., If, in this study, the model's horizontal resolution if set as 1 km, then the vertical resolution could be set as around 500 m.

We agree with the reviewer. The vertical resolution is variable with higher resolution close to the surface. On average the vertical resolution is approximately 360 m which satisfies this condition. The vertical grid spacing is less than 500 m up to 8-km height, then it becomes more than 500 m beyond 8-km height and reaches 790 m at the model top.

3. The authors mentioned that the non-Gaussianity reduced by 40% when assimilation window length shortened from 5 minutes to 30 seconds. What are the main benefits from the reduction? The authors claimed that this could improve the performance of the EnKF, without showing any evidence. It might be better by simply showing the error of precipitation output simulated in different experiments.

We would like to thank the reviewer for raising this important point. Following this comment we include a new figure (Figure 8) showing the RMSE and bias of the analysis and first guess with respect to the observed maximum reflectivity. We also include a discussion of the results.

To evaluate the impact of assimilation frequency on the distance between the analysis and first guess to the observations in a more systematic way, we compute the root mean squared error (RMSE) and bias for reflectivity observations (Fig. 8). The computation of the RMSE and bias between the model and the observations is done by comparing the column maximum of the reflectivity for each horizontal grid location and time. The RMSE and bias are computed only over grid points at which the observed maximum reflectivity is over 5 dBZ. The time series of RMSE shows a better fit to the observed reflectivity for shorter assimilation windows. The impact of 4D DA is not so clear, 1MIN-4D slightly outperforms the 1MIN but 5MIN-4D and 5MIN perform similarly (Fig. 7a). This is partially because in 4D data assimilation the analysis results from the assimilation of all the observations within the assimilation window, while to construct this figure, only the observations at the analysis time were considered. The bias, computed as the mean difference between the model and the observations does not seem to be consistently affected by the assimilation frequency (Fig. 8b). These results are in agreement with those observed in the time series of KLD for different variables. However, we should be cautious with the interpretation of these results since increasing the observation number can lead to both a reduced KLD and a better fit to the observed quantities, not necessarily implying a causal link between these two effects.

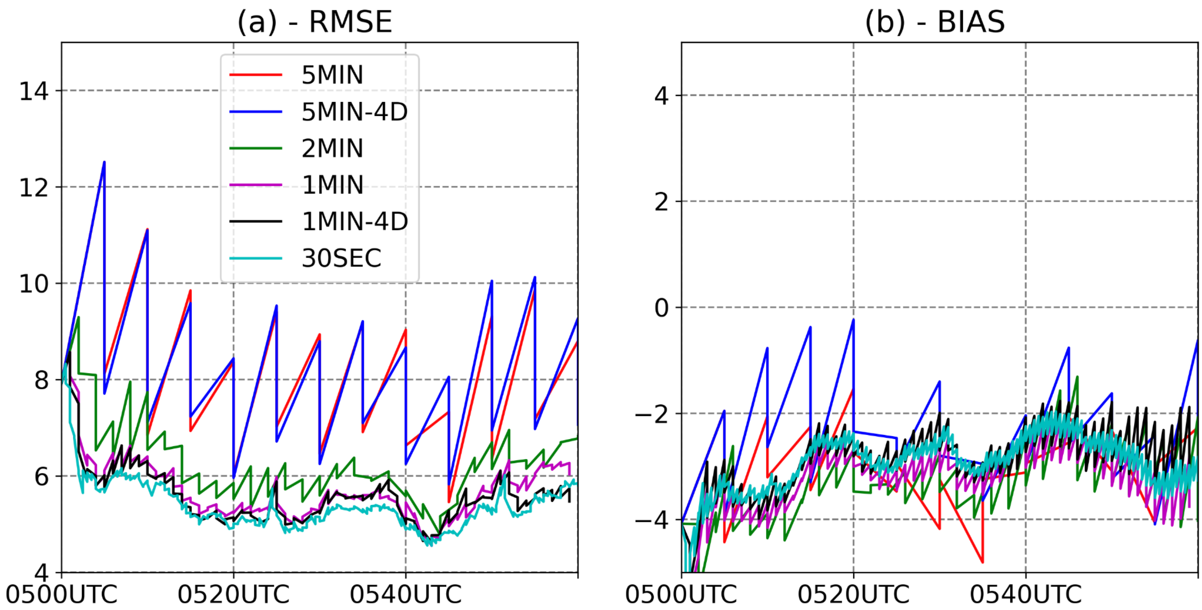

Figure 8: Sawtooth time-series of the root mean squared error (dBZ, a) and bias (dBZ, b) of the maximum reflectivity of the analysis and first guess for the 5MIN (red), 5MIN-4D (blue), 2MIN (green), 1MIN (magenta), 1MIN-4D (black) and 30SEC (cyan) experiments.

4. 2e–h: The authors marked the location of the maximum KLD for vertical velocity at the lower troposphere in Fig. 2f, but middle troposphere in Figs. 2e, 2g and 2h even though a maximum KLD center is not obvious in Figs. 2g and 2h. If the authors intended to emphasize the improvement of KLD in the middle troposphere, they should consider the KLD in the middle troposphere in all cases.

We would like to thank the reviewer for pointing this out. The explanation in the previous version of the manuscript was not clear. Our intention is not to focus on the middle troposphere but to show examples of distributions associated with KLD maxima within convective clouds.

We restrict the search of the maximum KLD to the grid points at which the ensemble mean reflectivity is over 30 dBZ to investigate departures from the Gaussian within convective clouds where radar data impact is larger.

5. 5b and 5h: The main highlight of this figure (which is also that of the manuscript) is improvement in the KLD with higher DA frequency. However, Fig. 5 also shows obvious increase in the KLD of specific humidity in the 30SEC experiment, in both the raining and non-raining cases. And, it seems that the authors did not make any discussion on these results. While there are great improvements in KLD of most grid points, especially that of vertical velocity when the authors focus on convective-scale simulation, why is the same improvement not obtained for specific humidity? Are errors generated from the more frequent DA update?

We thank the reviewer for bringing this interesting point. We add a remark to the discussion in Figure 6 (previously Figure 5), about the different impacts of frequent updates on specific humidity and other variables.

There are some exceptions to the general reduction in non-Gaussianity with increased update frequency. Specific humidity in non-raining grid points shows larger KLD in the 30SEC than in the 5MIN experiments. This is also the case for the raining grid points at upper levels in the second half of the experiment. Also the KLD in W in the non-raining grid points at middle and upper levels is slightly larger in the 30SEC experiment.

We also analyse this in more detail in Figure 7 which is a new figure showing the impact of the forecast and the assimilation step on the KLD.

To investigate the effect of the analysis update on non-Gaussianity we present the time series of the KLD of the analysis and first guess vertically and horizontally averaged over the “raining” and “non-raining” grid points (Fig. 7). At most times and variables over the “raining” and “non-raining” grid points KLD is reduced during the assimilation step. Experiments with longer windows experience more KLD growth during the forecast as expected, but also a larger reduction at the analysis step, which is not as effective as the more frequent updates in reducing the analysis KLD. As noted before, the specific humidity over the “non-raining” grid points behaves differently, in this case, KLD increases during the assimilation step for almost all times and experiments leading to larger KLD at shorter assimilation windows (Figs. 7b,f). In this area mostly “non-raining” observations are assimilated to suppress spurious clouds. Interestingly in the “non-raining” grid points 5MIN-4D is the experiment providing the lowest KLD for all variables (Figs. 7b, d and f).

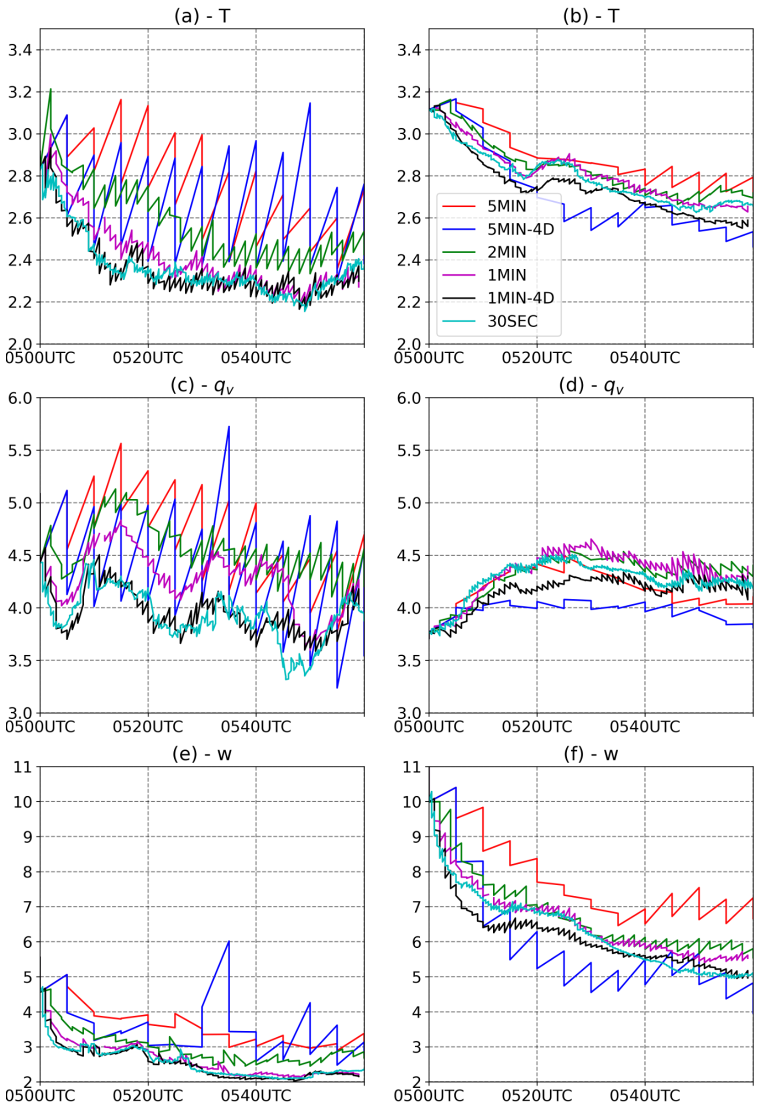

Figure 7: Sawtooth time-series of the KLD (10^{-2}) of the analysis and first guess over the rainy (<0dBZ, a,c,e) and non-rainy (>30dBZ, b,d,f) grid points for temperature (a,b), specific humidity (c,d) and vertical velocity (e,f) and for the 5MIN (red), 5MIN-4D (blue), 2MIN (green), 1MIN (magenta), 1MIN-4D (black) and 30SEC (cyan) experiments.

6. Line 67: "… Climate Forecast System Reanalysis Saha et al. (2010)" à "… Climate Forecast System Reanalysis (Saha et al., 2010)"

We change this following the reviewer's comment.

7. Line 99 and Eq.(1): "where P(x) and Q(x) are two . . . " à should be "where p(x) and q(x) are two . . ."?

We agree with the reviewer and change the sentence accordingly.

8. Lines 140–141: Wrong use of "so that".

Following the reviewer's suggestion, we change the sentences in the following way: "However, KLD in 5MIN-4D is larger than that in 30-SEC or 1MIN-4D, **indicating that** DA frequency is equally important."

9. Lines 144: "raining" and "non-raining" grid points sound better than "rain" and "non-rain" grid points to me.

We change the definition to "raining" and "non-raining" as suggested by the reviewer.