

General comments:

The paper uses a symmetric rain rate to define the radar reflectivity error in the assimilation algorithm based on the symmetric rain rate referring to the symmetric error model in satellite all-sky assimilation. The paper is well-structured but still, there are many ambiguous sentences in the paper which need to be rewritten/clarified.

Response: we appreciate your kind acknowledgment of this study. We hopefully addressed the reasons behind use analyses of NCEP GFS, without reflectivity assimilation. We also clarified the definition of representation error. Moreover, some ambiguous sentences will be rewritten in our revision.

major revisions:

One crucial aspect absent in the paper is that the reflectivity error in a DA system is a representative error of this system. It is important to know that the reflectivity error is indicative of the system's overall accuracy. Therefore, with any change made to the DA system such as adjustments in the NWP model settings or modifications to the forward model, the reflectivity error need to be recalculated or recalibrated. However, in this paper, all calculations are founded on a free forecast. The equivalent reflectivity is derived from the 6-hour model forecast.

Response: the 6-hour forecasts were not free forecasts, at least not exactly free, because the WRF model were initialized by the analyses. Although radar reflectivity was not assimilated, numerous observations, such as several satellite radiances and station measurements, were assimilated in order to improve the moist, wind and temperature. At convective scale, the environmental model variables were of more importance than the cloud and precipitation particles (Fabry and Sun 2010). The updated hydrometeors cannot survive in numerical weather model if the environmental model variables are unsuitable. Thus, the model results in this study can represent the 6-hour growth of model errors in statistics.

We can attempt to estimate the climatological representation error on basis of 6-hour forecasts initialized by the analyses, because most current data assimilation method pursue the statistical optimization in climatology. We will address clearly this issue in our revision. Moreover, using similar model runs, without data assimilation, to investigate the error structure of reflectivity has been acceptable by scientific community (Zeng et al., 2021).

Reference:

Fabry, F. and Sun, J.: For how long should what data be assimilated for the mesoscale forecasting of convection and why? Part I: on the propagation of initial condition errors and their implications for data assimilation. *Monthly Weather Review*, 138, 242–255. <https://doi.org/10.1175/2009MWR2883.1>, 2010.

Zeng, Y., Janjic, T., Feng, Y., Blahak, U., de Lozar, A., Bauernschubert, E., Stephan, K., and Min, J.: Interpreting estimated observation error statistics of weather radar measurements using the ICON-LAM-KENDA system, *Atmos. Meas. Tech.*, 14, 5735–5756, <https://doi.org/10.5194/amt-14-5735-2021>, 2021.

Besides, defining a more accurate reflectivity error is expected to enhance the assimilation results. However, the paper did not present any plots related to the implementation of the newly defined reflectivity error model in a DA system and its comparison with the constant error (which, as stated in the paper, is deemed unsuitable for radar assimilation).

Response: attacking the non-Gaussian error distribution of reflectivity is the main goal of this paper. Unveiling the non-Gaussian distribution (Figs. 3 and 4), how to build the symmetric error

model (Figs. 7 and 8) and exhibiting the qualitative (PDF distributions in Fig. 9) and quantitative (JSDs in Table 2) improvements on Gaussianity of PDF already constituted a complete logic, which emphasizes the effects of symmetric error model on OmBs of reflectivity and falls better in the scope of *Atmospheric Measurement Techniques*. Especially, we concluded that the symmetric error model can improve the strong non-Gaussian PDF of reflectivity OmBs. Due to these considerations, we also changed the title of this resubmission to “Improving the Gaussianity of Radar Reflectivity Departures between Observations and Simulations by Using the Symmetric Rain Rate”.

Recently, several studies related to the structures of observation error have been published, without application of error structures on data assimilation (Waller et al., 2016; Waller et al., 2019; Zeng et al., 2021, and so on...). They demonstrate the importance of error statistics.

In our own conceit, this resubmission could be a timely and useful work in error structures of radar reflectivity, which may encourage readers to build a more effective symmetric error model based on their own assimilation and prediction systems.

We are well aware of the fact that building the symmetric error model is to improve the analysis in reflectivity assimilation. Based on our very preliminary results, we found that the symmetric error model is very sensitive to some parameters (positive impact obtained when the tuning parameter is 0.25 and the lower boundary is 3 dBZ in Eq. 6), due likely to the inconsistency between the symmetric error model and the forecast and assimilation systems. Considering the length and completeness of this study, we maintain that this resubmission focuses on the improvement on Gaussianity of radar reflectivity by using the symmetric error model. The numerical experiments about application of this symmetric error model on reflectivity assimilation could be another topic, which is our ongoing research.

References:

- Waller JA, Simonin D, Dance SL, Nichols NK, Ballard SP.: Diagnosing observation-error correlations for Doppler radar radial winds in the Met Office UKV model using observation-minus-background and observation-minus-analysis statistics. *Mon. Weather Rev.* 144: 3533–3551.
<https://doi.org/10.1175/MWR-D-15-0340.1>, 2016.
- Waller, J. A., E. Bauernschubert, S. L. Dance, N. K. Nichols, R. Potthast, and D. Simonin: Observation Error Statistics for Doppler Radar Radial Wind Superobservations Assimilated into the DWD COSMO-KENDA System. *Mon. Wea. Rev.*, 147, 3351–3364,
<https://doi.org/10.1175/MWR-D-19-0104.1>, 2019.
- Zeng, Y., Janjic, T., Feng, Y., Blahak, U., de Lozar, A., Bauernschubert, E., Stephan, K., and Min, J.: Interpreting estimated observation error statistics of weather radar measurements using the ICON-LAM-KENDA system, *Atmos. Meas. Tech.*, 14, 5735–5756,
<https://doi.org/10.5194/amt-14-5735-2021>, 2021.

Line 40: "The error statistics associated with radar reflectivity, consisting of both the instrument error and representation error": Could you please clarify the meaning of "representation error" in this context?

Response: for an observation network, ‘a small-scale phenomenon may be misrepresented as a much larger scale phenomenon. Thus, the error of representativeness is a measure of the error caused by the misrepresentation of all scales smaller than the grid spacing of numerical model’, cited from Daley (1991), which seems to be the original definition of representation error. Similar definition and classification of observation error is also used by Waller et al. (2016). The observational operator error was not included in representation error. This manuscript distinguished between the representation error and the observational operator error because we followed the original definition of representation error (Daley 1991).

A recent study (Janjić et al., 2018) reported that the representation error consists of three components, mismatch between scales in observations and model results, errors of observation

operator and quality control or preprocessing. Unfortunately, the definition and terminology of representation error are still not unified. In our revision, we can update the definition of representation error as reported by Janjić et al. (2018).

References:

Daley, R.: Atmospheric Data Analysis. Cambridge University Press, 457 pp, 1991.

Janjić, T., McLaughlin, D., Cohn, S.E. and Verlaan, M.: Conservation of mass and preservation of positivity with ensemble-type Kalman filter algorithms. *Mon. Wea. Rev.*, 142, 755–773, <https://doi.org/10.1175/MWR-D-13-00056.1>, 2014.

Waller JA, Simonin D, Dance SL, Nichols NK, Ballard SP.: Diagnosing observation-error correlations for Doppler radar radial winds in the Met Office UKV model using observation-minus-background and observation-minus-analysis statistics. *Mon. Wea. Rev.* 144: 3533–3551.

<https://doi.org/10.1175/MWR-D-15-0340.1>, 2016.

Line 58: "It is clear in reflectivity assimilation, where errors including representation errors and operator errors increase with the precipitation amount." Firstly, in a scientific text, clarity is essential. It would be beneficial to provide a reference to support this claim. Secondly, could you please clarify what is meant by "operator error"? Is it synonymous with forward model error, which refers to the model converting the NWP model output to radar reflectivity? When discussing representative error in a DA system, it typically encompasses NWP model error, forward model error, and other factors. Why is operator error excluded from the representative error of a DA system?

Response: according to the latest definition of representation error, this sentence will be rewritten as "It is clear in reflectivity assimilation, where representation error including mismatch between scales and observational operator error increases with the precipitation amount." For the mismatch between scales, the heavier precipitation (stronger convection) usually shows lower predictability, leading to larger OmBs. Another evidence is that the heavy precipitation usually has lower ETS than little precipitation. For the observational operator of reflectivity, the cold process in strong convection, including ice-phased and mix-phased hydrometeors, complicates the projection from model variables to reflectivity. The OmBs of reflectivity in cold process are usually larger than those in warm process. The references associated with this claim, such as Sun and Zhang (2020) and Jung (2008), were already cited in this manuscript. We will explain this issue in our revision.

The operator error is synonymous with forward model error, as well as the observational operator error. They mean the error introduced in the transformation from model variable to observation variable. To avoid misunderstanding, we will change 'operator error' to 'observational operator error' or 'error of observation operator' in our revision.

As also seeing the above response, the original definition of representation error in Daley (1991) is misrepresentation between scales, which is different to the error of observation operator. To unify the terminology and definition, we will follow the definition in Janjić et al. (2018) which is the latest publication.

References:

Jung, Y., Xue, M., Zhang, G. F., and Straka, J. M.: Assimilation of simulated polarimetric radar data for a convective storm using the ensemble Kalman filter. Part II: Impact of polarimetric data on storm analysis. *Mon. Wea. Rev.*, 136, 2246–2260, <https://doi.org/10.1175/2007MWR2288.1>, 2008.

Sun, Y. Q. and Zhang, F. Q.: A New Theoretical Framework for Understanding Multiscale Atmospheric Predictability. *Journal of the Atmospheric Sciences*, 77, 2297–2309, <https://doi.org/10.1175/JAS-D-19-0271.1>, 2020.

Line 119: "The WRF model has been nested in one-way with a coarse resolution of 9 km and a

fine resolution of 3 km”: The nested domain should be inside the parent domain.

Response: the inner domain of WRF model was inside the outer domain in this study. We will rewrite the descriptions related to model settings in revision.

line 125: “The GFS analyses at 0000 UTC and 1200 UTC in the 6 months are used to drive the WRF model.”: what does this sentence mean? 6 months analysis?

Response: it means that all analyses at 0000 UTC and 1200 UTC from April to September were used. There were 366 analyses from NCEP GFS in total. We will rewrite this sentence in our revision in order to avoid misunderstanding.

Fig3: The plots 3a and 3b, as well as 3c and 3d, appear to be identical. This should not be the case. Please review the plots.

Response: except the thousands of samples along the abscissa, Figs. 3a and 3b are identical. Removing the missed simulations and their corresponding observations is the difference between ‘any-reflectivity’ (Fig. 3a) and ‘both-reflectivity’ (Fig. 3b). Other samples were not changed at all. The missed simulations and their corresponding observations located along the abscissa in Fig. 3a, as well as Fig. 3c. Thus, the Figs. 3a and 3b should be identical except the thousands of samples along the abscissa in Figs. 3a and 3c.

What is the purpose of excluding the false and missed events? and defining the ‘both-reflectivity’? Ultimately, all data points need to be accounted for in defining the standard deviation.

Response: we illustrated what give rise to the strong non-Gaussian error distribution of reflectivity in comparison of ‘both-reflectivity’ and ‘any-reflectivity’ scenarios. The thousands of samples along the abscissa in Fig. 3a mean numerous missed simulations (observed, but not simulated), leading to the skewness and high peak of PDF in Fig. 4. This feature also relates to the “zero gradient” effect as stated in Introduction.

Moreover, defining the ‘both-reflectivity’ also can explain that the radar reflectivity has a stronger non-Gaussian distribution than the satellite radiance in cloudy sky. As discussed in lines 157-161, the ‘both-reflectivity’ scenario, which is a simplified scenario, exhibits similar distribution to the nonprecipitating cloud affected satellite radiance. Thus, the ‘any-reflectivity’ scenario has a stronger non-Gaussian feature than the satellite radiance in cloudy sky. This conclusion is supported by the binned standard deviations (Fig. 8), where the difference between the first two bins is much greater than the other bins.

In addition, we built the symmetric error model and discussed the properties of predictor in ‘any-reflectivity’ scenario, as shown in Figs. 7-9. We did not investigate the symmetric error model in ‘both-reflectivity’ scenario.

Fig 8C: The black dashed line depicts the logarithm of sample numbers that fall below 2 after 12.5 mm h^{-1} , indicating that the number of samples is less than 100. If this is indeed the case, it implies that the number of samples in these bins is insufficient for calculating the standard deviation. As demonstrated, the number of samples in certain bins can reach up to 10^6 , revealing a significant inconsistency in standard deviation definition across bins. Therefore, it is advisable to establish a sample number limit for standard deviation calculation. I would recommend a limit of 10^3 or 10^4 samples.

Response: we only fitted the linear regression function from 0.5 to 9 in Fig 8c, as well as listed in Table 1, where the sample number is larger than 10^3 . We also described that reason in line 270-273.

We will write more sentences to explain the straight line used in the symmetric error model when the sample number is less than 10^3 .

minor correction:

Line 34: “in an idea model” → “in an ideal model”

Response: it is a typo. We will correct it in our revision.

line 60: simulations, usually called Observations → simulation, defined by Observation

Response: We will correct it in our revision.