This manuscript proposes an approach to handle the non-Gaussian error distribution of reflectivity OmBs (dBZ), which adopts the idea of the symmetric error model in all-sky radiance data assimilation. This work demonstrates that the symmetric error model built by the rainrate predictor, can improve the Gaussianity of OmB distribution, by using six-month composite reflectivity data and simulated products. Moreover, the reflectivity OmBs present a more complicated error model that can be fitted by a three-piecewise function, compared to the satellite radiances, since the radar reflectivity is often discontinues. This manuscript is well structured and could be a valuable contribution to the radar and data assimilation communities. I have several comments as below. I'd like to recommend minor revision to this manuscript.

We appreciate the constructive comments from referee #1 and reply all of them in the following blue words.

1 This work compares two OmB data, the maximum composite and the reflectivity at 1 km. Results show that two OmB data have similar features, such as horizontal distributions and PDFs. The rainrate derived from the reflectivity at 3 km is then used to build the symmetric error model of the maximum composite. However, the maximum composite and reflectivity at 1 km and 3 km, respectively, could be different. The correlations between the derived rainrate and the maximum composite (or/and the reflectivity at 1 km) are needed to clarify the potential inconsistent usages of data.

Response: The maximum composites and rainrates derived from 3 km reflectivity are not exactly identical, but both of them are highly associated with the strength of convections. The large maximum composite and heavy rainrate can indicate a strong convective system, and vice versa. This study used derived rainrates to describe the heteroscedasticity of maximum composites in terms of the convective strength and demonstrated the symmetric error model can improve the Gaussianity of OmBs of the maximum composites.

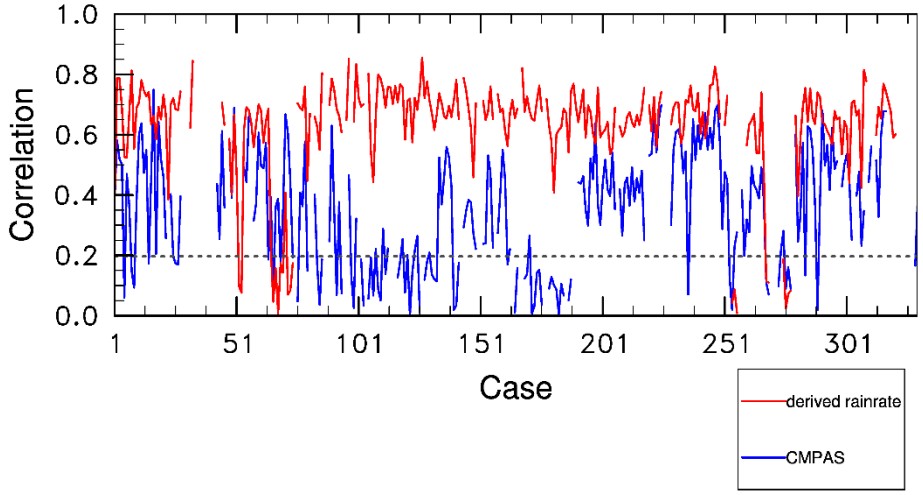

Figure R1. the absolute correlations between the maximum composites and two rainrate data in six months. The blue and red lines represent the rainrates derived from reflectivities at 3 km altitude and the CMPAS rainrates. The dash line shows the 95% confidence.

The red line in Figure R1 shows that the absolute correlations between the derived rainrates and the maximum composites are evidently high (>0.75) in most precipitating cases, despite some cases present low correlations. Thus, using derived rainrates to describe the heteroscedasticity of the maximum composite is rationale, similar to the cloud liquid water or liquid water path for satellite radiances. We do not give the correlations between the derived rainrates and reflectivities at 1 km. Because the sample amount of reflectivities at 1 km is much less than other data and this study did not build a symmetric error model of reflectivities at 1 km.

We also gave the absolute correlations between the CMPAS rainrates and the maximum composites in Figure R1. The absolute correlations of CMPAS rianrates decrease obviously because the independent errors, including the sampling and representative errors from the third-party data, increase rapidly. However, the CMPAS rainrates can build a similar symmetric error model to the derived rainrates and can further improve the Gaussianity of OmBs of the maximum composites in comparison with the derived rainrates. Thus, the differences between two rainrate data allow us to investigate how the accuracy of predictor affects the symmetric error model.

2 There various types of data used in this work, especially with different horizontal resolutions (e.g., the 5-km CMPAS and 3-km WRF products). Thus how the interpolation performed to deal with the inconsistent resolutions needs clarification. It is possible that the interpolation increases the OmB variances. The authors emphasize that all data are collected in mountainous areas. Does the interpolation consider the effects of terrain? Moreover, how about the resolution of CVMR and CAPPI used in this study?

Response: The horizontal resolutions of CVMR and CAPPI are 1 km. The coarsest resolution among various data is 5 km. Therefore, we used Euclidean distances as weights to interpolate data from fine resolutions to coarse resolution.

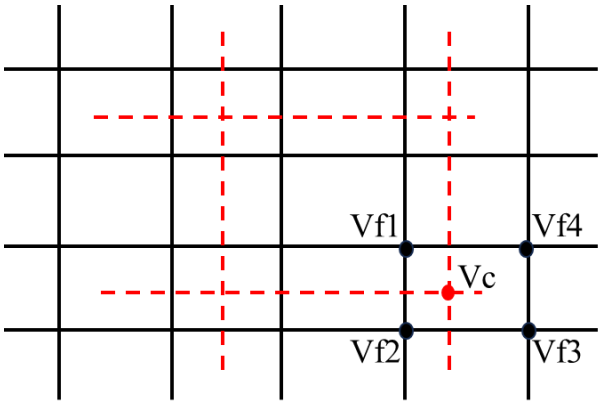

Figure R2. the schematic plot of interpolation.

As shown in Figure R2, the black grids represent the fine resolution data, such as 1 km radar observations and 3 km WRF products. The 5 km resolution radar observations and WRF products at red grids (V_c) are weighted average of the nearest four black grids (V_{f1}, V_{f2}, V_{f3}, V_{f4}):

$$V_c = a_1 V_{f1} + a_2 V_{f2} + a_3 V_{f3} + a_4 V_{f4}$$

where a_1, a_2, a_3 and a_4 are weights computed by the distances between a red grid and

the nearest four black grids.

As aforementioned, we do not take into account the effects of terrain in this interpolation. The only effect of terrain is the blockage mountainous areas in this study, which significantly reduces the sample amount of CAPPI at 1 km altitude.

3 The reflectivity OmBs highly depend on the forward operator. More elucidations about the forward operator are needed, in order to clarify how the OmB is derived.

Response: The algorithm of diagnostic reflectivity (dBZ) included in UPP softward package is based on rain, snow, and graupel mixing ratios was designed by Stoelinga (2005):

$$Z = 10\log_{10}(Z_{er} + Z_{es} + Z_{eg}) \qquad (R0)$$

Following some assumptions, the reflectivity contributed by rain droplets (Z_{er}) is given by:

$$Z_{er} = \Gamma(7)N_{r0}\lambda_r^{-7} \qquad (R1)$$

$$\lambda_r = \left(\frac{\pi N_{r0}\rho_l}{\rho_a q_{ra}}\right)^{0.25} \qquad (R2)$$

where N_{r0} is 8×10^6, ρ_l and ρ_a are the liquid water density and dry air density respectively. The q_{ra} is the rainwater mixing ratio in background.

Assumed snow particles are spheres, the reflectivity contributed by snow is given by:

$$Z_{es} = \alpha\Gamma(7)N_{s0}\left(\frac{\rho_s}{\rho_l}\right)^2\lambda_s^{-7} \qquad (R3)$$

$$\lambda_s = \left(\frac{\pi N_{s0}\rho_s}{\rho_a q_{sn}}\right)^{0.25} \qquad (R4)$$

where α is 0.224, N_{s0} is 2×10^7, ρ_s is the density of snow 100 kg m^{-3}. The q_{sn} is the snow water mixing ratio in background.

Similarly, the contribution of graupel particles can be obtained:

$$Z_{eg} = \alpha\Gamma(7)N_{s0}\left(\frac{\rho_g}{\rho_l}\right)^2\lambda_g^{-7} \qquad (R5)$$

$$\lambda_g = \left(\frac{\pi N_{g0}\rho_g}{\rho_a q_{gn}}\right)^{0.25} \qquad (R6)$$

where α is also 0.224, N_{g0} is 2×10^7, ρ_g is the density of graupel 400 kg m^{-3}. The q_{gn} is the graupel water mixing ratio in background.

According to above formulas (R0-R6), the reflectivity predicted by model can be computed by the rainwater, snow water and graupel water mixing ratios. Although it is a single moment algorithm, it can serve as a forward operator, converting model variables to reflectivity. We will add a few sentences in revision to elucidate this algorithm.

4 Figure 8, it is interesting that the logarithmic rainrates (Figure 8c) has a different distribution than the rainrates (Figure 8a) and CMPAS rainrates (Figure 8b), for the magnitudes of rainrates larger than 10 mm h^{-1}. Why the three-piecewise fitting function for the logarithmic rainrates does not capture the decrease trend for those

magnitudes larger than 10 mm h^{-1}?

Response: We argue that the difference between the logarithmic rainrates and other rainrates for the large magnitudes of predictors mainly results from the rapid decline of sample sizes at the tail of the logarithmic rainrates. The logarithmic transformation not only smooths the small rainrates (near to zero), but also smooths the large rainrates (near to the maximum). The head and tail of standard deviation distribution in Figure 8c are smoother than the others in Figure 8a and 8b and more samples concentrate in the middle of the logarithmic rainrates. The logarithmic rainrates then obtained more samples in total than the derived rainrates and CMPAS rainrates. The tail of the logarithmic rainrates is then closer to the upper boundary of reflectivity. Therefore, the numbers decline very fast at the tail of logarithmic rainrates, as shown by the black dash line in Figure 8c. The standard deviations of reflectivities become zero when reflectivities are closer to the boundary (Bishop 2016; Bishop 2019).

In theory, the linear regression may be inappropriate for this symmetric logarithmic rainrates. This is the reason that the R^2 of logarithmic rainrates is smaller than those of derived rainrates and CMPAS rainrates, as listed in Table 1. We still used the linear regression to build the symmetric error model because: 1 a straight line at the tail of symmetric rainrates was used to prevent an irrational fitting function; 2 we noticed the JSD of logarithmic rainrates in Table 2 is the smallest. Thus, we keep using the linear regression in this manuscript. A more appropriate fitting function for logarithmic rainrates is an interest topic in future.

References

Bishop, C. H.: The GIGG-EnKF: ensemble Kalman filtering for highly skewed non-negative uncertainty distributions. Q J R Meteorol Soc, 142: 1395–1412. https://doi.org/10.1002/qj.2742, 2016.

Bishop, C. H.: Data assimilation strategies for state‐dependent observation error variances. Q J R Meteorol Soc, 145, 217 – 227, https://doi.org/10.1002/qj.3424, 2019.