Response to reviewers

Comment 1:

- l. 36: "lunched" should be "launched"

Our response to comment 1:

Corrected. (**L36, L46**)

Comment 2:

- Fig. 6: Which images are model equivalents and which ones are observations? This information is missing.

Our response to comment 2:

The figure caption was polished in the revised manuscript. Explanation to each subplot was added. The revised figure caption is "***Figure 6: Synthetic (the first column) and observed (the second column) visible images and the corresponding probability density distribution functions (the third column) for two selected cases. The first panel (a1-a3) is the results for the case at 06:00 UTC on 1 September 2020. The second panel (b1-b3) is the results for the case at 06:00 UTC on 15 September 2020.***". The revised figure caption should avoid the ambiguity. (**L256-260**)

Comment 3:

- l. 158 / Fig. 8: You say that you did not need the ice particle effective radius, as it is not required for the Baran scheme. However, you compared the Baran scheme to the Baum scheme, which does require an effective radius. So which effective ice particle radius did you use to generate Fig. 8?

Our response to comment 3:

Thank you for pointing this out. The Baran scheme does not have a dependence on the effective radius of ice clouds ($\mathrm{Re_{ice}}$). However, the Baum scheme does depend on $\mathrm{Re_{ice}}$. Since the state variables of the CMA-MESO model does not include $\mathrm{Re_{ice}}$, $\mathrm{Re_{ice}}$ was explicitly calculated following Hong et al. (2004) and Yao et al. (2018) to facilitate the radiative transfer simulations based on the Baum scheme. The method was summarized in the following. (**L368-375**)

$$Re_{ice} = \min(11.9 \times 0.75 \times 0.163 \times M_i^{1/2}, 500 \times 10^{-6}) \qquad \text{(R1)}$$

where M_i denotes the ice crystal mass, which is calculated by Equation (R2),

$$M_i = \frac{\rho_a q_i}{N_i} \qquad \text{(R2)}$$

where ρ_a denotes the density of air. q_i denotes the mixing ratio of ice crystals. N_i denotes the concentration of ice crystals which was approximated by Equation (R3),

$$N_i = \min(\max(5.38\times10^7(\rho_a \times max(q_i, 10^{-15}))^{0.75}, 10^3), 10^6) \quad \text{(R3)}$$

Comment 4:

- l. 271: Typical snow albedo values are higher in the visible spectrum (see e.g. Gardner, A. S., and M. J. Sharp (2010), A review of snow and ice albedo and the development of a new physically based broadband albedo parameterization, J. Geophys. Res., 115, F01009). However, taking into account the the snow may not fill the full pixel, you could claim that 0.2 is used as a lower limit for the surface albedo of snow-covered surfaces.

Our response to comment 4:

Thank you for this comment and for recommending this paper, which is really helpful. The surface albedo of snow in the visible spectral band varies with the physical properties of snow. With the increase of the average radius of ice grains, the surface albedo is decreased (Gardner and Sharp, 2010). In addition, the surface albedo of dirty snow which includes absorbing particles and old snow which includes some melting water is less than that of pure snow (Xu and Tian, 2000; Gardner and Sharp, 2010). In general, the lower limit for the surface albedo of snow-covered surfaces in China is suggested to be 0.2 (e.g., Fig. 3 of Xu and Tian, 2000, shown below).

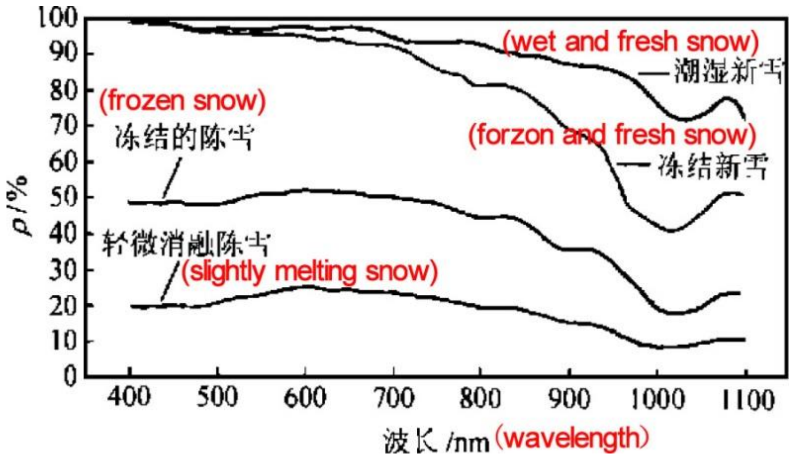

Fig. 3 in Xu and Tian. The spectral distribution of the albedo for different snow state.

Therefore, the threshold test was performed by the following formula,

$$\omega \le 0.2/3.14 \quad \text{(R4)}$$

where $0.2/3.14$ denotes the BRDF for a Lambertian radiator. (**L267-274**)

Comment 5:

- Section 3.3: It seems to me that to detect instrument problems actually the cloud-free pixels (blue lines in Fig. 7) would be enough -- they clearly show the change in calibration. Only clear sky reflectances are used for these blue lines, so no model state is involved -- right? Is there an advantage of using also the model state (looking at the red/black line in Fig. 7)?

Our response to comment 5:

It is true that the time series of the O-B biases for cloud-free pixels could reflect the abrupt change in calibration. This is reflected by the blue lines in Fig. 7. To simulate the TOA reflectance for cloud-free pixels, some of the model state variables derived from the CMA-MESO model were provided to the RTTOV-DOM radiative transfer model. These model state variables include the temperature profile, the humidity profile, the surface pressure, the 2-m height temperature and humidity, the 10-m height U- and V-wind components, and the surface temperature.

The abrupt change was also revealed by the red and black lines in Fig. 7. Since the O-B biases were positively related to the observed reflectance which is proportional to the calibration coefficient, the mutant signal was amplified when cloudy pixels were involved (remember that the clouy reflectance is generally larger than cloud-free reflectance). Therefore, involving cloudy pixels for monitoring the performance of the instrument should have some advantages. When cloudy pixels were involved, the state variables provided to RTTOV-DOM not only include the clear sky variables mentioned above, but also include cloud variables, i.e., the mixing ratio of cloud droplets, the mixing ratio of rain, the mixing ratio of ice, the mixing ratio of snow, the mixing ratio of graupel, the grid-scale cloud fraction, and the effective radius of liquid water clouds explicitly calculated based on some parametrizations. (**L324-325**)

Comment 6:

- Table 2, caption: "Land+Scean" should be "Land+Sea"

Our response to comment 6:

Corrected. (**Table2 in Page 20**)

Comment 7:

- Fig. 10: A background grid or zero lines would be very helpful here to better see the asymmetry.

Our response to comment 7:

We added background grid to Fig. 10. (**L528-531**)

Comment 8:

- l. 435ff: So if I understand correctly, there are four bias correction coefficients, gamma_land,cloudy, gamm_land,clear, gamma_ocean,cloudy and gamma_ocean,clear. It would be useful to state that explicitely here.

Our response to comment 8:

Yes you are right. There are four bias correction coefficients corresponding to two underlying surfaces (land and sea) and two cloud masks (cloud and cloud-free). In addition, some of the cloud masks derived from the FY-4A CLM product could be also "uncertain". Therefore, two extra bias correction coefficients were involved to deal with the "uncertain" scenarios for land and sea separately. Therefore, the bias-corrected reflectance O' is calculated by Equation (R5),

$$O' = O(1 + \gamma_{clm}^{sfc}) \quad \text{(R5)}$$

where γ_{clm}^{sfc} denotes the bias correction coefficient. The subscript clm denotes cloud mask, which is either cloud-free (clr), or cloudy (cld), or uncertain (uct). The superscript sfc denotes the surface type that is either land or sea. To be specific, γ_{clm}^{sfc} represents one of the six bias correction coefficients including γ_{clr}^{land}, γ_{cld}^{land}, γ_{uct}^{land}, γ_{clr}^{sea}, γ_{cld}^{sea}, and γ_{uct}^{sea}. (**L448-462**)

Comment 9:

- l. 470ff: The bias correction coefficients are now called mu instead of gamma. It would be good to stick to one name.

Our response to comment 9:

Thank you for pointing this out. In the revised manuscript, μ was replaced by γ here and elsewhere. (**L448-462; L509-527**).

In addition, the simulated reflectance, originally denoted by *r*, was denoted by "*B*" in the revised manuscript. The modification was to maintain consistency with the so-called "*O-B*" analyses in section 5.

Comment 10:

- l. 443: "For the ensemble forecast, the synthetic image was generated by averaging the seven visible images simulated from seven ensemble members.". I do not understand. I thought you would just use the additional enemble members to get more matches (pixels that are cloudy in both O and B or cloud-free in both O and B) that can be used for the computation of mu_clr, mu_cld (l. 472/473). However, I cannot see how you would use the ensemble average for this purpose. As an example, suppose the model is actually unbiased and produces clouds with the same reflectance distribution as the observation, and the clouds are just at the wrong locations. Then in the ensemble average the reflectances will be lower (because if there is a cloud, it will in general not be present in all members) and if you use these low, ensemble averaged reflectances in the computation of mu_cld you will get a positive, potentially even large value (and not the value 0 that would be correct for no bias). Are you really using the ensemble average in this way and if yes, why is it working? Or did I misunderstand how you use the ensemble average? Then please provide an equation with the ensemble average reflectances that clearly shows what you are doing.

Our response to comment 10:

For the Ensemble Kalmam Filter (EnKF) method which has been widely used for the DA of satellite radiance data, the observation increments were calculated using the ensemble mean in the observation space (e.g. Equation (5) of Zhou et al., 2022).

$$\Delta y_n = \left(y_n^{\mathrm{p}} - \bar{y}_{\mathrm{p}}\right)\left(\sigma_{\mathrm{u}}/\sigma_{\mathrm{p}}\right) + \bar{y}_{\mathrm{u}} - y_n^{\mathrm{p}},\ n = 1, \ldots, N \quad \text{(R6)}$$

where Δy_n denotes the observation increment for the nth ensemble member, y_n^{p} the first guess of the observed variable for the nth ensemble, $\bar{y}_{\mathrm{p}}$ the ensemble mean of the first guess of the observed variable, σ_{p} the first-guess sample error standard deviation of the observed variable, σ_{u} the updated standard deviation of σ_{p}, $\bar{y}_{\mathrm{u}}$ the ensemble mean of the posterior estimate (i.e., the analysis) of the observed variable.

Therefore, to maintain consistency with the ensemble-based DA methods, the bias correction method should be performed based on the ensemble mean of the first-guess reflectance, denoted by $\overline{B_{clm}^{sfc}}$, which was generated by Equation (R7),

$$\overline{B_{clm}^{sfc}} = \frac{1}{N_{ens}} \sum_{l=1}^{N_{ens}} B_{clm}^{sfc}(l) \quad \text{(R7)}$$

where N_{ens} denotes the number of ensemble members, i.e., seven in this study. l is the index for an arbitrary ensemble member. (**L463-478**)

We agree with your comments that the ensemble averaging could decrease the reflectance for a pixel classified to be cloudy for a deterministic forecast in cases where the cloud did not occur for all ensemble members due to the displacement errors. As a result, the bias correction coefficient estimated by the ensemble forecasts is larger than that estimated by a deterministic forecast due to the underestimated first-guess reflectance (Table 1 and Table 2). In other cases where clouds occur for all the ensemble members, the uncertainty of the ensemble mean should be smaller than the uncertainty of a single ensemble member. In general, the number of the matched cloudy pixels was larger for the ensemble forecast than the deterministic forecast. Since the O-B departure was dominated by the cloudy error characteristics, the PDF of the bias-corrected O-B departure satisfied the unbiased Gaussian function better for the ensemble forecast than the deterministic forecast. (**L479-486, L500-596, Fig. 10, L528**)

Comment 11:

- Section 5: To demonstrate that you need to distinguish between land/ocean and cloudy/cloud-free, it would be great to provide values for the bias correction coefficients for the two dates. And to show how helpful the ensemble is, it would be good to provide information on the the number of matching pixels with / without ensemble.

Our response to comment 11:

We added the bias correction coefficient γ, the number of matching pixels N_{match}, and the number of pixels in the observed images N_{obs} for cloudy and cloud-free scenarios over land and sea surfaces. (**Table 1 in Page 19 and Table 2 in Page 20**)

Other changes

In addition to the changes mentioned above, we corrected some typo errors and polished the writings throughout the manuscript to avoid ambiguities and misunderstandings. All these revisions were marked in red in the track-changes file.