Response to reviewer

- **General comment:**

This paper compares the FY-4A/AGRI 0.65-um visible reflectance (O) with the model simulations generated from CMA-MESO forecasts using the RTTOV (B). The potential sources contributing to the differences between O and B, such as the unresolved aerosol processes, the ice scattering models, are analyzed.

The paper is relevant to the cloud remote sensing field, as the growing international fleet of next-generation geostationary imagers can be expected to aid in our understanding of the diurnal cycles of clouds and aerosols. Well understood and characterized the biases of their observations will therefore be well received by the community. However, the authors make what I think are several unsubstantiated assertions (see my detailed comments). I recommend major revisions before reconsidering for publication. My general and specific comments are below.

- **Our reply to general comment:**

The authors thank the anonymous reviewer for the constructive comments. We made major revisions to the manuscript, including the evaluation of the forecasts of CMA-MESO, discussions on the spatial and seasonal variations of the O-B biases, corrections on some typo errors (e.g., abbreviations, inappropriate usage of “evaluation”, etc.). As a result, the outline of the manuscript was changed compared with the initial version. Some of the revisions were made according to the reviewer’s comments. Some were made by the authors spontaneously to improve the readability. The point-by-point response to the reviewer’s comments were provided below.

- **General Comment 1:**

A comparison with the model simulations cannot be called an “evaluation”, especially when the model simulations are not as accurate as expected. Currently, the RTTOV forward-operator for clouds and/or within the visible and shortwave infrared spectral ranges is still questionable, and the forecasts from CMA-MESO also lack adequate evaluations.

- **Our reply to general comment 1:**

Thank you for pointing this out. It is true that the reflectance simulated from the forecasts of CMA-MESO model using RTTOV cannot represent the true reflectance due to the deficiencies of the CMA-MESO and RTTOV models. Therefore, the title of this manuscript was changed to “Exploring the characteristics of FY-4A/AGRI visible reflectance using the forecasts of the CMA-MESO model and its implications to data assimilation”. (L1-3 in the revised manuscript)

- **General Comment 2:**

As (1), if the authors persist in characterizing the biases of AGRI reflectance observations by comparing with the model simulations, the performances of RTTOV forward-operator and the forecasts from CMA-MESO should be evaluated first.

- **Our reply to general comment 2:**

We agree with the reviewer that the evaluation of the CMA-MESO model and the RTTOV-DOM forward operator is necessary when addressing the O (FY-4A/AGRI 0.65-um visible reflectance) - B (model simulations generated from CMA-MESO forecasts using the RTTOV) differences. Accordingly, two main revisions were made to address this problem.

(1) Evaluation of the forecasts of CMA-MESO

The forecasts of CMA-MESO model were evaluated by comprehensive observations.

Firstly, the cloud mask diagnosed from the forecasts of the CMA-MESO model was compared with the spatiotemporally collocated cloud mask products derived from the Himawari-8 geostationary satellite. Quantitative analysis indicated a Fraction Skill Score (FSS) larger than 0.7 in most cases (Fig. 5 in the revised manuscript, L232-234), implying that the CMA-MESO model performs well in forecasting cloud locations.

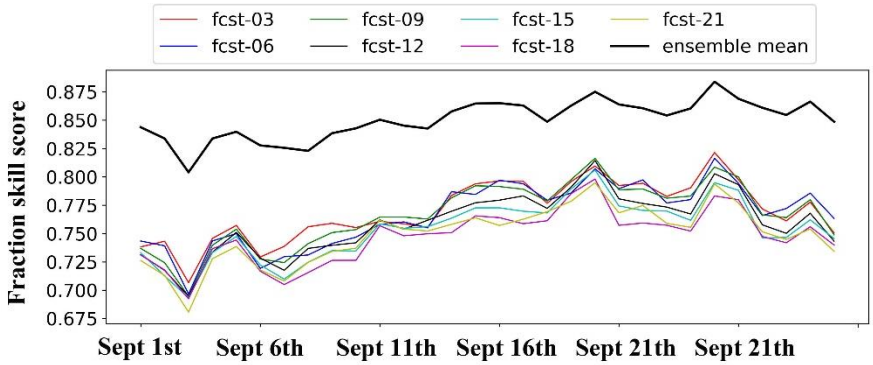

Figure 5: Fraction Skill Scores (FSSs) of cloud cover for the short-term forecasts of CMA-MESO with different forecasting lead times. The FSSs were calculated with a square of length of 2.

Secondly, the one-hour accumulated precipitation was compared with the observations provided by the multi-source observed precipitation products in Chinese mainland. The results indicated that the CMA-MESO model could reproduce the precipitating areas with adequate accuracy (Fig. 6 in the revised manuscript, L243-247). Quantitative analysis of the Enhanced Threatening Score (ETS), a commonly used metrics for evaluating the accuracy of precipitation forecasting, revealed better performance of CMA-MESO in forecasting the light to moderate rain (Fig. 7 in the revised manuscript, L248-251). Despite that some spatial discrepancies of the core of heavy rain were revealed between the forecasts and observations, the domain-averaged precipitation agrees well with the observation (Fig. 8 in the revised manuscript, L263-265). The results imply that the forecasts of CMA-MESO should have certain reliability, especially when it comes to the domain-averaged quantities (remember that the O-B biases were calculated on the domain average).

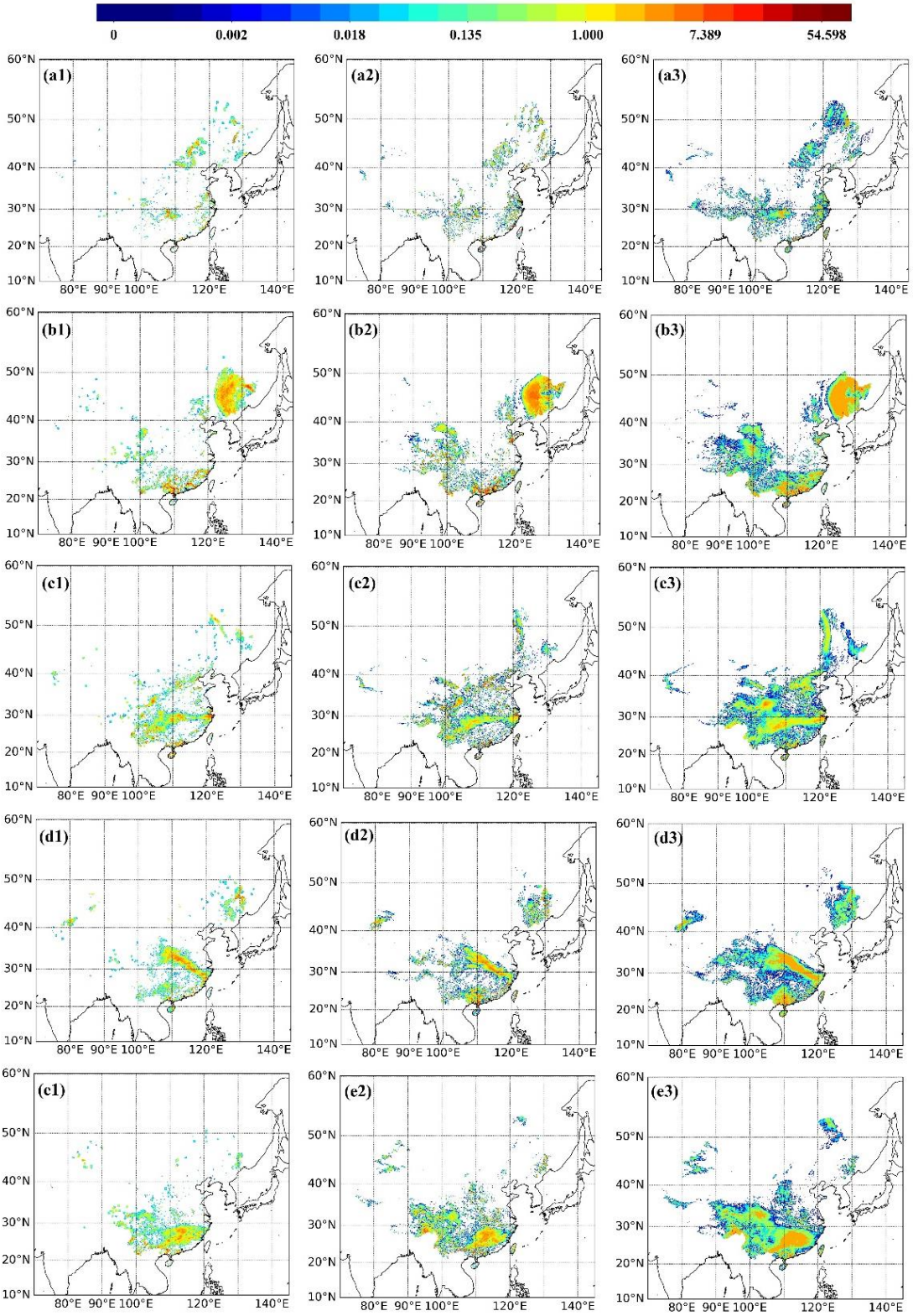

Figure 6: Accumulated one-hour precipitation at 06:00 UTC derived from the multi-source observed precipitation products (the first column), the 6-h forecasts of CMA-MESO (the second column), and the ensemble forecasts including seven ensemble members which have different forecasting lead times. From top to bottom, each panel corresponds to the results on 2 September 2020 (a1-a3), 8 September 2020 (b1-b3), 14 September 2020 (c1-c3), 20 September 2020 (d1-d3), and 26 September 2020 (e1-e3).

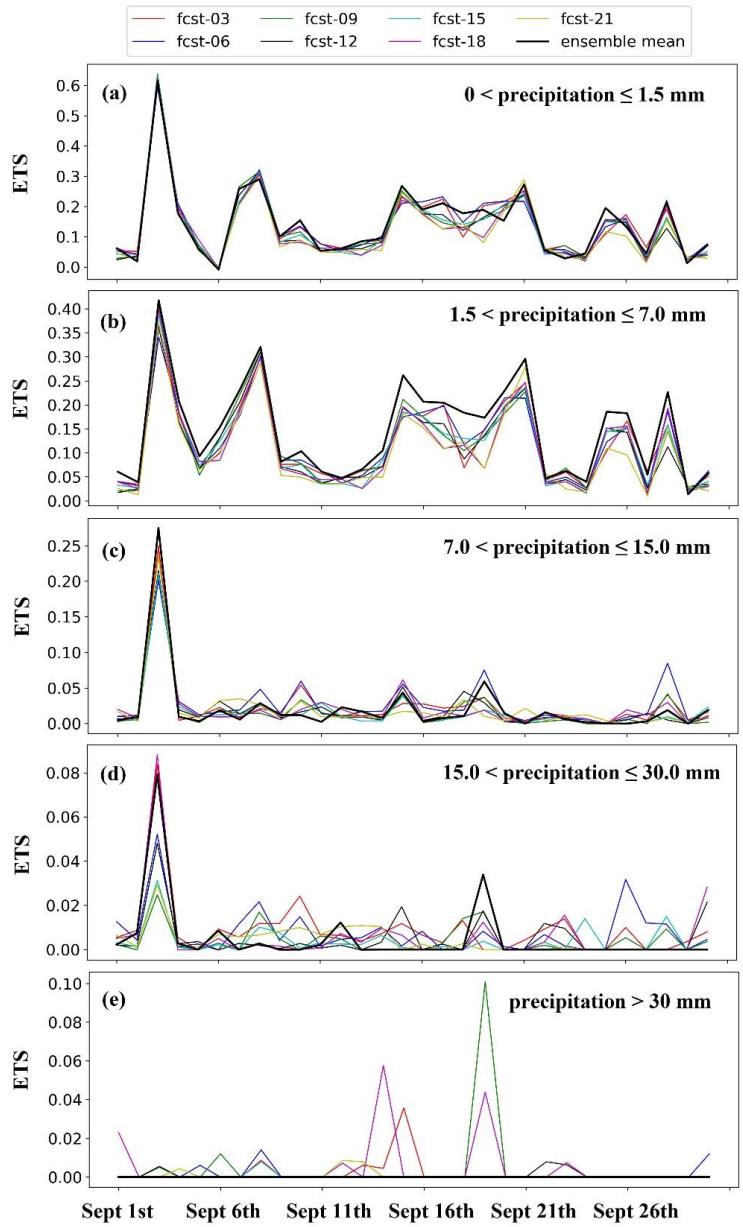

Figure 7: Enhanced Threatening Score (ETS) for the accumulated one-hour precipitation at 06:00 UTC derived from the forecasts of CMA-MESO with different forecasting lead times. From top to bottom, each precipitation range represents the (a) light rain, (b) moderate rain, (c) heavy rain, (c) rainstorm, and (e) heavy downpour. Fcst-03 means the 3-h forecasts at 06:00 UTC, and so forth.

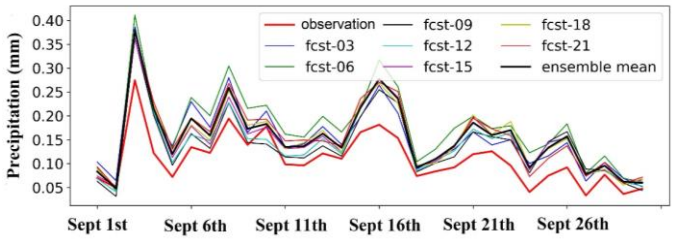

Figure 8: Domain-averaged one-hour accumulated precipitation at 06:00 UTC derived from the short-term forecasts of CMA-MESO with different forecasting lead times. Fcst-03 means the 3-h forecasts at 06:00 UTC, and so forth.

The evaluation using two observations both revealed the advantages of an ensemble forecast over a deterministic forecast in reducing the errors in cloud simulations. The implication to the bias correction of data assimilation is that the B derived from an ensemble forecast should be a better choice than from a deterministic forecast.

(2) Discussions on the deficiencies of RTTOV (L266-294)

We feel helpless to evaluate the performance of the RTTOV model. The largest challenge comes from a lack of accurate observations corresponding to the real atmosphere state variables. Therefore, some of the major deficiencies were discussed in the revised manuscript.

Although great strides have been made to RTTOV, we have to admit that the simulations of RTTOV in cloudy regions are still questionable. Potential errors include the ignored 3D radiative effects, the uncertainties of the cloud optical properties in the build-in cloud schemes, etc. For example, the pre-assumed cloud particle size distribution (PSD) inherent in the cloud schemes in RTTOV is inconsistent with that of NWP models, not to mention the representativeness of the pre-assumed PSD in real cases. To illustrate this problem, a sensitivity study was performed by RTTOV configured with two different ice schemes, i.e., the Baum and Baran schemes. Distinct differences were revealed for the simulated reflectance (Fig. in the revised manuscript, L287-289), which confirms the uncertainties in the cloud optical properties of the RTTOV model.

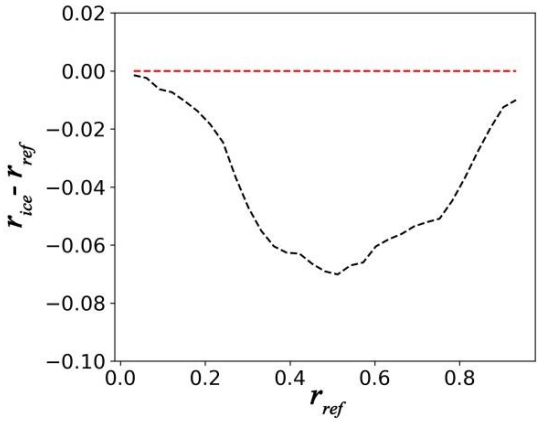

Figure 9: The impact of ice cloud schemes on the TOA reflectance. r_{ref} denotes the reflectance simulated by RTTOV configured with the ice scheme of Baran et al. (2015). r_{ice} denotes the simulations based on the SSEC/Baum ice scheme.

To the best of our knowledge, we do not know the actual representativeness of the optical properties derived from the built-in cloud schemes. Sophisticated evaluation will be needed to address the performance these build-in cloud schemes in real cases.

Currently, B derived from any of the NWP models and forward operators will enviably suffer from errors. As long as B is systematically biased, correcting the biases in O based on B and other predictors, which is the routine operation for the data assimilation of infrared and microwave satellite observations, will be questionable. Therefore, we have to admit that the O-B method is a measure of last resort due to a lack of sufficient reference observations for comparing with the satellite observations. Whether the bias correction brings benefits to the numerical weather prediction should be tested by data assimilation in real-world whether systems and should be evaluated by comprehensive observations.

- **General Comment 3:**

The bias characteristics are not well analyzed. (1) How about the spatial distributions or seasonal variations of AGRI biases? (2) Do they have differences before and after the FY-4A satellite's U–turn at the vernal and autumnal equinoxes?

- **Our reply to general comment 3:**

(1) Spatiotemporal variation of the O-B biases (L178-211)

To better characterizing the spatiotemporal variation of the O-B biases, extra simulations were performed for March, June, and December. We did not perform the simulations for the whole 2020 year mainly because the radiative transfer simulation is rather computationally comprehensive. On our linux cluster which is equipped with 2.20 GHz Xeon Silver 4214 CPU, it will take approximately 30 min ~ 1 hour (32 CPUs for parallel computation) for the RTTOV-DOM (V12.3) to generate a synthetic visible imagery which includes 2501×1671 pixels (the CMA-MESO grids). We think the results for March, June, September, and December could reveal some seasonal variation characteristics of the O-B biases.

Based on the four-month simulations over the CMA-MESO domain, the temporal (Fig. 3 in the revised manuscript, L193-196) and spatial (Fig. 4 in the revised manuscript, L209-211) variation characteristics of the O-B biases were explored. The results indicate different spatiotemporal variations of the O-B biases for the four months, which is closely related to the spatiotemporal of the performance of the CMA-MEOS model, the variation of aerosol properties, etc.

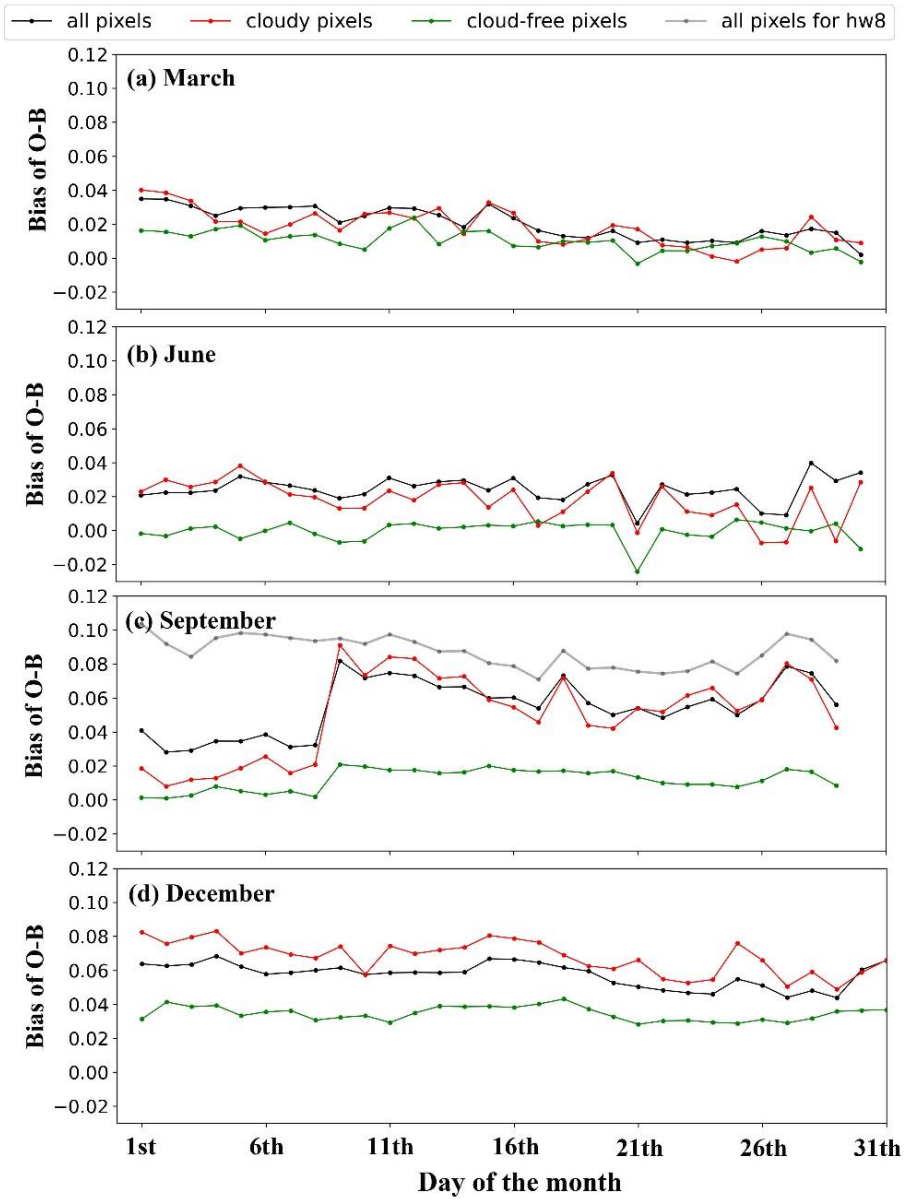

Figure 3: Time series of the O-B biases for the cloud-free, cloudy, and all pixels for FY-4A visible observations in (a) March, (b) June, (c) September, and (d) December. The time series of the O-B biases for Himawari-8 visible observations for all pixels was also provided in (c) for comparison with that of FY-4A.

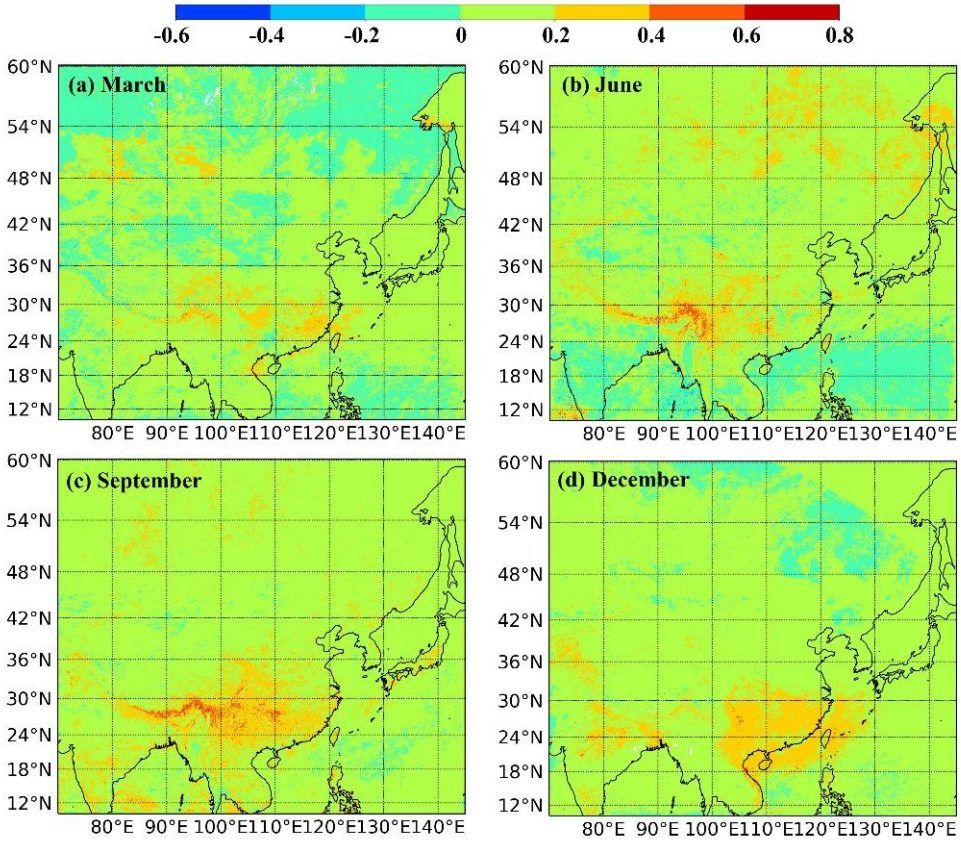

Figure 4: Spatial distribution of the O-B biases for FY-4A visible observations in (a) March, (b) June, (c) September, and (d) December.

(2) Characteristics during U–turn at the vernal and autumnal equinoxes (Fig. 3, L193-196)

In the Northern Hemisphere, the vernal and autumnal equinoxes falls about 20 March and September 22 or 23, respectively. During this period of time, the Sun crosses the celestial equator, leading to changes in the sun-satellite geometries. Checking through the time series of the O-B biases in March and September, we see no differences during these two days or around when compared with other dates (Fig. 3(a) and Fig. 3(c) in the revised manuscript, L193-196). Therefore, a tentative conclusion could be drawn that the temporal variations of the O-B biases do not have differences before and after the FY-4A satellite's U–turn at the vernal and autumnal equinoxes.

- **Specific Comment 1:**

Lines 16, 22, 33 and 72: The abbreviations (FY, TOVS, and so on) should be given full name when first appeared in the abstract and text.

- **Our response to specific comment 1:**

The abbreviations were checked throughout the revised manuscript, and the full names were given the first time they appeared in the abstract and text. E.g., FY is the abbreviation of “Fengyun”, TOVS is the abbreviation of Television infrared observation satellite Operational Vertical Sounder, and RTTOV is the abbreviation of Radiative Transfer for the Television infrared observation satellite Operational Vertical Sounder, etc. (L17, L23-24, L35, L76-77, etc. in the revised manuscript)

- **Specific Comment 2:**

Line 85: Himawari-8 satellite should be introduced because not all readers know it is the first one of the Japanese next-generation geostationary satellite.

- **Our response to specific comment 2:**

In the revised manuscript, we added some introductions to the Himiwari-8 satellite. (L106-107 and L152-164 in the revised manuscript)

“…… Himawari-8, the first one of the Japanese next-generation geostationary satellite ……” (L106-107)

Himawari-8 was launched by the Japan Meteorological Agency on October 7, 2014. The geostationary satellite carries the Advanced Himawari Imager (AHI) which provides radiance observations covering visible to infrared spectra and completes a full-disk scan every 10 minutes. The Himawari-8/AHI band 3 (0.55 μm - 0.72 μm) is close to the FY-4A/AGRI channel 2, which contains critical information on clouds, aerosols, and underlying surfaces (Bessho et al., 2016). In this study, the Himawari-8 cloud mask products gridded at 5 km resolution in September 2020

were used to evaluate the performance of the CMA-MESO in predicting cloud locations. To ensure the spatial collocation between the observations and simulations, the Himawari-8 cloud mask products were interpolated to the CMA-MESO grids by a bi-linear interpolation. In addition, the Himawari-8 reflectance data gridded at 5 km × 5 km resolution in September 2020 were used to explore the stability of Himiwari-8 visible observations. The observed reflectance data were interpolated to the CMA-MESO grids to ensure spatial collocation with the simulated equivalents. Since the synthetic imageries for the FY-4A and Himaeari-8 were derived from the same forecasts from CMA-MESO, the differences in the time series of O-B biases for the two satellites should reveal some different characteristics of their corresponding visible instruments. (L152-164)

- **Specific Comment 3:**

Line 82: How about the spatial coverage of CMA-MESO, or the region of interest in this study?

- **Our response to specific comment 3:**

In the revised manuscript, the spatial coverage of CMA-MESO was shown by Fig. 1, which is also the region of interest in this study. (L95-97 in the revised manuscript)

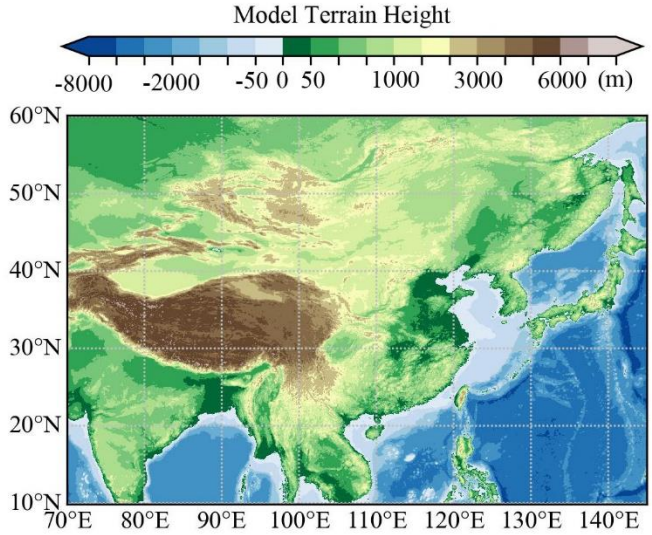

Figure 1: The domain coverage of the CMA-MESO model, which includes 2501×1671 horizontal grids with a horizontal grid spacing of 0.03 °.

- **Specific Comment 4:**

Lines 96 and 117? Here, the authors give two cloud mask definitions. Which one will be used for Tables 1 and 2?

- **Our response to specific comment 4:**

We are sorry for not making it clear here. The two cloud masks are defined for the synthetic imageries (observations, O) and observed imageries (simulations, B) separately. For the observed imageries, cloud masks were directly derived from the cloud mask products. For the synthetic imageries, cloud masks were dragonized from the CWP with a threshold value of 0.01 kg m^{-2}.

For spatiotemporally collocated observations and simulations, the O-B biases were calculated for the pixels which are designated to be cloudy and cloud-free for both O and B. The O-B biases for the cloudy and cloud-free pixels were further used to correct the systematic biases of the corresponding scenarios separately. (L136-139 in the revised manuscript)

- **Specific Comment 5:**

Lines 201-203: I can't understand this sentence. Aren't the "microphysical properties therein" "cloud variables"?

- **Our response to specific comment 5:**

We were intended to say that "Compared with the infrared and microwave radiance observations, the visible reflectance is much more sensitive to cloud variables, regardless of the type of cloud hydrometeors or the vertical location of clouds. In contrast, the infrared radiance data are only sensitive to cloud-top properties due to strong absorption effects of clouds in infrared spectra".

This part was revised in the revised manuscript. (L334-336 in the revised manuscript)

- **Specific Comment 6:**

Figure 6: Readers can hardly identify the differences between observed and model simulated reflectance. The authors are suggested using a different colormap or adding figures to show their differences.

- **Our response to specific comment 6:**

This figure was modified in the revised manuscript. The first and second columns shows the observed and simulated imageries. They share the same colorbar to make a fair comparison between them. The third column shows the O-B departure of the imageries. We think the revised figure could depict the differences between O and B more clearly. (L367-370 in the revised manuscript)

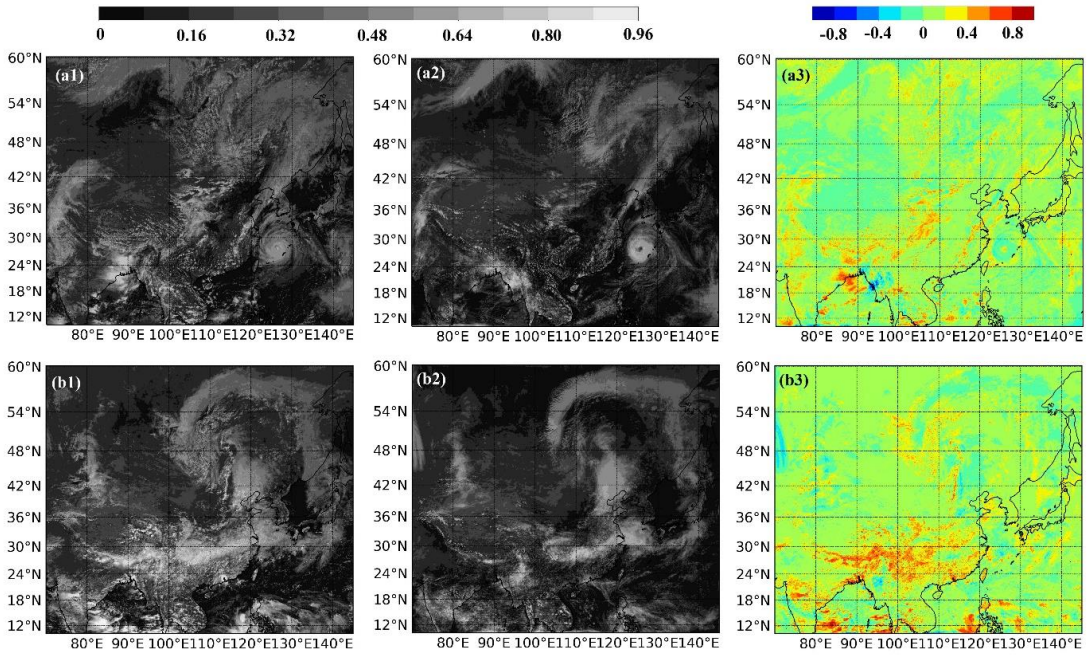

Figure 11: The observed and bias-corrected reflectance at 06:00 UTC on 1 September 2020 (a1-a3) and 15 September 2020 (b1-b3). From left to right, the three columns correspond to the observed imageries (a1-b1), the ensemble mean synthetic imageries (a2-b2), and the observation-minus-simulation imageries (a3-b3).