

**Comments on the manuscript titled “Potential Influences of Neglecting Aerosol Effects on the NCEP GFS Precipitation Forecast” by Jiang et al.**

This study evaluated the National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS) forecast bias in different precipitation (light rain, moderate rain, heavy rain and very heavy rain) by comparing the ground-based observations in three countries. Then the correlations between GFS precipitation forecast errors and the aerosol loading are investigated extensively to examine the potential impact of neglecting aerosol-cloud-interaction (ACI) on the operational rainfall forecast. The main result is that the GFS overestimates light rain, and underestimates moderate rain, heavy rain, and very heavy rain, which is partly due to the neglecting ACI process in GFS. The study fits within the scope of the journal, and the information and arguments are generally clear enough to be followed. Although the current study does not fully established the causal relationship between the ACI and the bias of precipitation forecast of GFS due partially to a lack of sufficient information, it should still be commended for confronting a highly-challenging task to make this first attempt to evaluate the numerical weather prediction forecast errors in terms of the potential effects of aerosols. Therefore, I'd recommend accepting this manuscript if the following comments are properly addressed.

Thank you very much for your constructive comments and suggestions. Our point-by-point replies are given below and the corresponding revisions are shown in the revised manuscript.

**Major Comments:**

1. As shown in figure 3, the magnitude of underestimation in light rain and overestimation in heavy rain by GFS are all similar over three counties, but the aerosol loading in China is significantly higher than in other two countries. If the aerosol is one of the major factors causing the bias in the GFS precipitation simulation, why there is no obvious difference in the magnitudes of the bias among the three countries?

***Response:***

First, the intention of Figure 3 is to show that the GFS model overestimates light rain and underestimates heavier rain. Second, of course, these model biases are caused by many factors, including initial dynamic settings and weather regimes. But it is beyond the scope of this paper to explore all possible causes. Comparing the model performance globally according to aerosol loading only is not sufficient because the model performance may differ for different regions. Our focus is on identifying any potential contribution of neglecting aerosol effects to the biases. The relationship between model performance and AOD was thus further investigated. This is also why we compared results from three countries. In each country, the standard deviation of the daily precipitation difference as a function of aerosol optical depth is presented in Fig. 5. Each

point represents a grid box. The significant positive correlation between standard deviation and AOD illustrates that neglecting aerosol effects may contribute to the model forecast bias. Third, the non-linear impact of aerosols on precipitation may also differ according to meteorological conditions, aerosol components, and the interactions between thermal and dynamic conditions. This is why we then focused on one specific region, Fujian Province, and did a long-term statistical evaluation of rainfall forecasts to mitigate these fluctuations in the model forecast accuracy.

2. For the study of the aerosol invigoration effect on the warm and cold based mixed clouds, please clarify the cloud top temperature is for convective core area or for whole convective clouds (including anvil areas). As those studies by Rosenfeld et al. [2008] and Fan et al. [2013], only the decrease of cloud top temperature for convective core with increasing of aerosol loading can be attributed to the aerosol invigoration effect.

***Response:***

The cloud-top temperature (CTT) obtained from CloudSat data are used to study the impact of aerosols on the cloud development of different cloud types. Based on the definition of deep mixed-phase clouds with warm bases shown in Table 1 (cloud-base temperature  $> 15^{\circ}\text{C}$ ), the CTT analyzed is mainly associated with the convective core although the stratiform/anvil areas cannot be totally ignored. Both the aerosol thermodynamic effect (i.e., convective invigoration) illustrated by Rosenfeld et al. (2008) and the microphysical effect (mainly the role of more but smaller longer-lasting ice particles) emphasized by Fan et al. (2013) contribute to the decrease in CTT. The point of analyzing CTT as a function of AOD for different cloud types here is not to figure out which role is more dominant, but to find out whether the CTT decreased or increased and whether the cloud is more suitable for precipitation or not.

**Table 1.** Definitions of warm- and cold-base mixed-phase clouds and liquid clouds.

	Cloud-base temperature ( $^{\circ}\text{C}$ )	Cloud-top temperature ( $^{\circ}\text{C}$ )
Deep mixed-phase clouds with warm bases	$> 15$	$< -4$
Shallow mixed-phase clouds with cold bases	$0-15$	$< -4$
Liquid clouds	$> 0$	$> 0$

3. Some of descriptions are too detailed and may not be necessary.

***Response:***

We have modified the descriptions accordingly. A brief description of the model setting, which is relevant to this study, has been given. Also, detailed descriptions of the MERRA-2 analysis in section 2.2.1 have been shortened.

**Minor Comments:**

1. Line 95: The description of “ARI are only considered offline and are not coupled with the dynamic system” is confused.

**Response:**

A seasonal climatological tropospheric aerosol background with a large horizontal resolution is used for both longwave and shortwave radiation. There is a current effort underway to change this to a monthly background. The temperature change caused by aerosols is not coupled to each forecast interval. Therefore, it is not coupled with the dynamic system.

2. Part 2.1: Since this study only used the simulation results and the details of GFS has been widely described, thus I'd suggest cutting the description in section 2.1 and paying more attention to the potential error of GFS precipitation forecast.

**Response:**

Lines 121 to 144: We have modified the descriptions accordingly. A brief description of the model setting, which is relevant to this study, has been given.

3. Section 2.2.1: Such a detailed description on MERRA-2 aerosol reanalysis is not necessary. What is the spatial resolution? Same with the CPC data?

**Response:**

Lines 157 to 167: This part of the manuscript has been shortened. The spatial resolution of the MERRA-2 reanalysis is  $0.625^{\circ} \times 0.5^{\circ}$  and that of CPC data is  $0.5^{\circ} \times 0.5^{\circ}$ . The data matching strategy is described in the newly-added section 2.3.1.

4. Line 251-255: Please give the observed time of the sounding data.

**Response:**

It is twice a day (at 00 UTC and 12 UTC). This information has been added to line 207.

5. Section 3.1.1: From figure 2, the systematic bias is found in three counties, such as the overestimations are found in north, west of China, and underestimations are found in east China. Could you explain this?

***Response:***

The GFS model tends to overestimate light rain and underestimate heavier rain. In the northern and western parts of China, it seldom rains and when it rains, it is mainly light rain. So the GFS model tends to overestimate precipitation in these parts of China. In eastern China, it rains more and deep convective precipitation is common. So the GFS model tends to underestimate rain in this region.

6. Line 340: Clarify the meaning of Z.

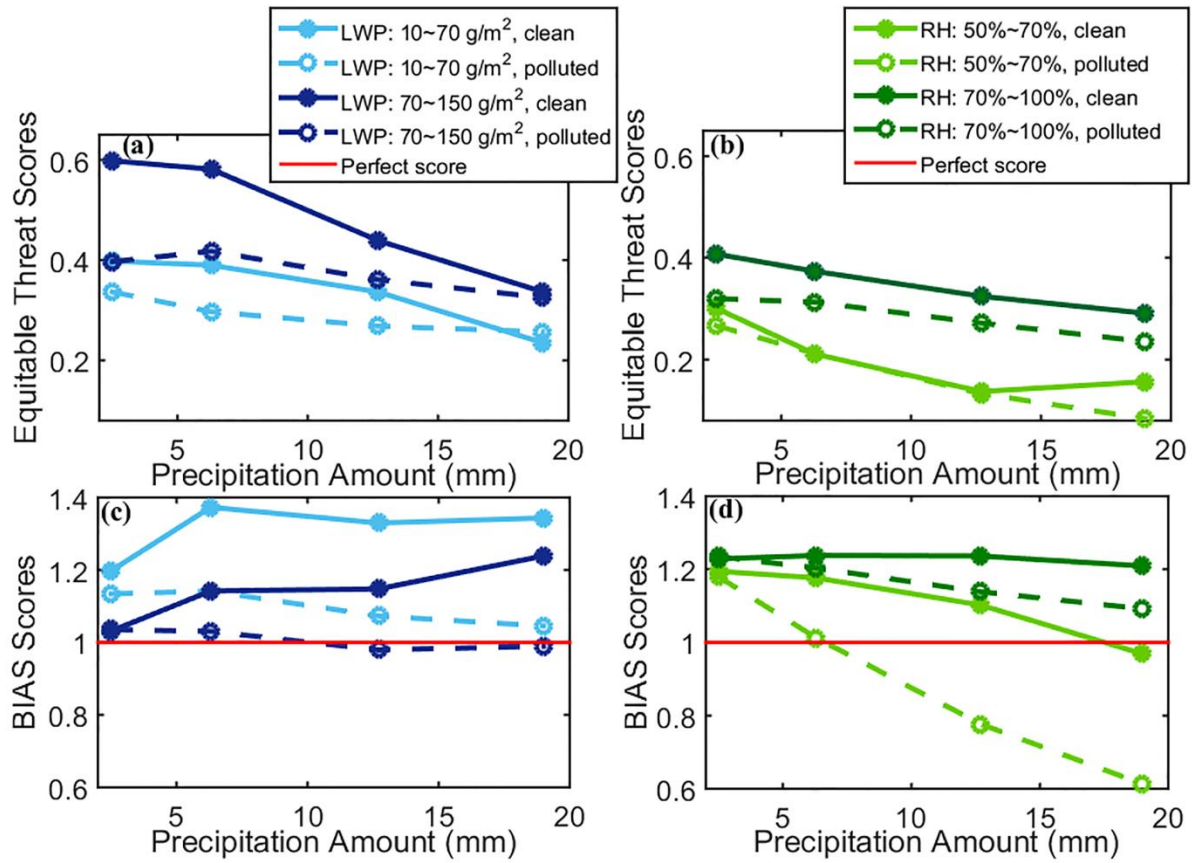
***Response:***

Line 325: The Z-score is the number of standard deviations from the mean value of the reference population. When 95% of the values fall within two standard deviations from the mean, a normal probability distribution is defined (according to the 68-95-99.7 rule). The p value is set as 0.05 in this study, therefore, the mean difference is not significant at a two-sigma level when  $Z < 2$ .

7. Line 385: in figure 6, please clarify the definition of the low, middle and high cloud mixing ratio, and the definition of the low, middle, high and very high AOD conditions. And why the thresholds of 5 and 20 are selected.

***Response:***

Figure 6 has been redrawn. We adopted the suggestion from another anonymous reviewer to replace the cloud mixing ratio at 850 hPa with LWP and RH to better show the different large-scale humidity conditions. The ETS and BIAS in the new Figure 6 are calculated for certain LWP or RH conditions and the top and bottom one-third of AOD values are defined as polluted and clean subsets of data. A threshold is used in the contingency table when calculating ETS and BIAS. The definition of hits or misses is based on the forecast rain amount above a certain threshold. In the new Figure 6, more commonly used precipitation amount thresholds have been used (i.e., 0.01, 0.25, 0.50, 0.75 inches).



**Fig. 6.** Equitable threat scores (a, b) and bias (BIAS) scores (c, d) as a function of precipitation amount for fixed ranges of liquid water path (LWP; a, c) and relative humidity (RH; b, d) under clean and polluted conditions. The LWP is divided into two categories: 10–70 g m<sup>-2</sup> (light blue) and 70–150 g m<sup>-2</sup> (dark blue). Data are from August 2015 in the U.S, China, and Australia. The RH is divided into two categories: 50–70% (light green) and 70–100% (dark green). Data are from year 2015. For a given LWP or RH condition, the top and bottom one-third of AOD values are defined as polluted and clean subsets of data, respectively. The solid lines represent the clean scenario and the dotted lines represent the polluted scenario. The horizontal red lines in (c) and (d) represent perfect scores.

8. Line 394-396: how to draw the conclusion of “the underestimation for heavy rainfall increases as AOD increases for low and middle cloud mixing ratio conditions” from figure 6d.

**Response:**

This sentence has been deleted.

9. Line 457: Although the long-term data are used, the seasonal variations in aerosol loading, cloud properties and meteorological parameters may result in the nominal relationship as shown in figure 12.

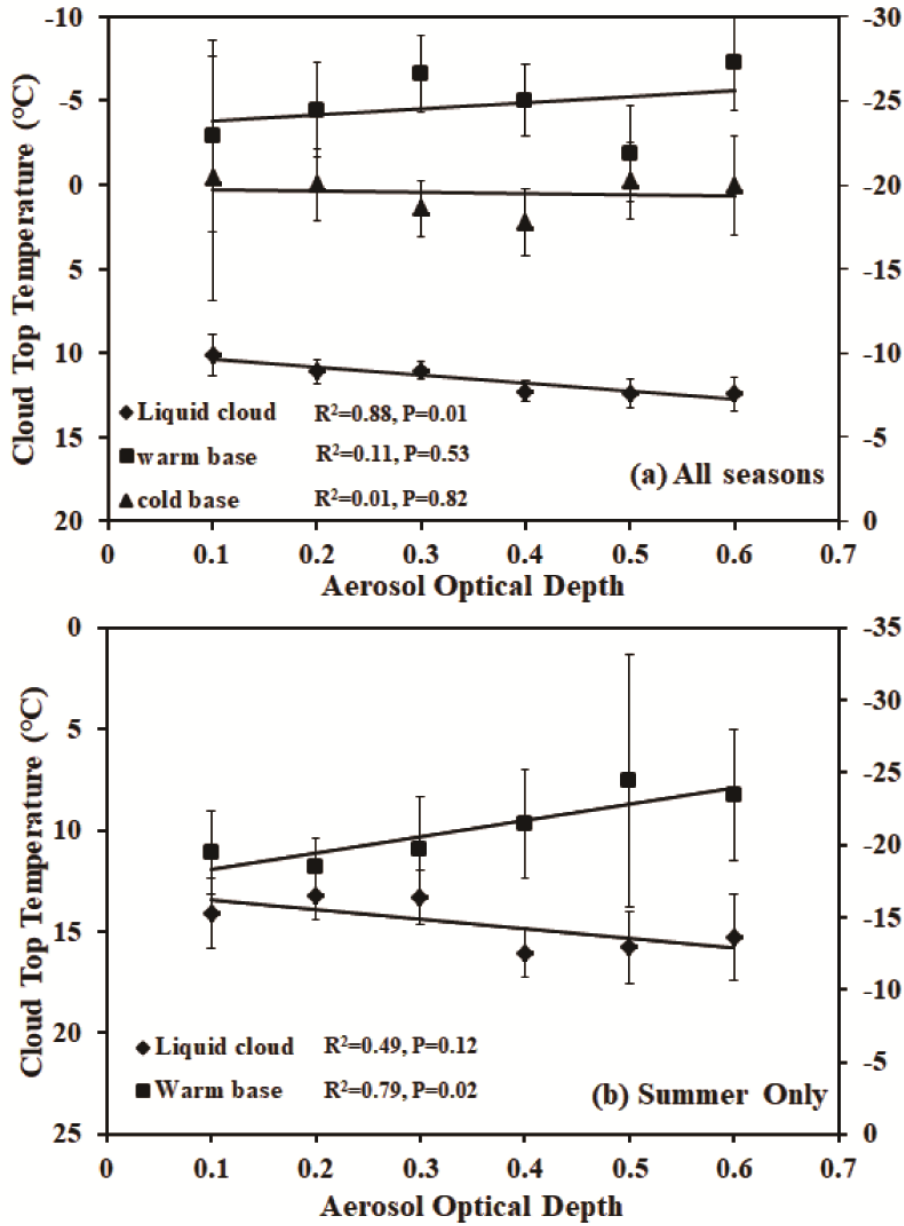
***Response:***

Line 434: Seasonal variations in aerosol loading, cloud properties, and meteorological parameters may influence aerosol-cloud-precipitation interactions. This is why we examine the impact of aerosols on clouds and precipitation for certain cloud types and ranges of LWP values. In Figure 12, the cloud effective radius as a function of AOD under different LWP conditions for liquid clouds is shown. The randomly-mixed samples are rearranged according to AOD. The figure shows some perturbations caused by changes in AOD.

10. Line 479-485 and figure 13: Is the relationship statistical significant? Please give P values in figure 13.

***Response:***

We have included P values in the new Figure 13.



**Fig. 13.** Cloud-top temperature as a function of aerosol optical depth for (a) liquid, warm-base mixed-phase, and cold-base mixed-phase clouds in all seasons, and (b) liquid and warm-base mixed-phase clouds in summer in Fujian Province, China. Diamonds represent liquid clouds, squares represent warm-base mixed-phase clouds, and triangles represent cold-base mixed-phase clouds. Right-hand ordinates are for warm-base and cold-base mixed-phase clouds. Data are from 2006–2010.

11. Line 485: It is either significant or not significant, based on the confidence level the authors choose. Therefore, I advise the authors to use stronger or weaker correlations, or higher or lower slopes, but not the more or less significant.

***Response:***

Lines 461 to 464: This sentence has been rewritten as “The negative slope of the linear relationship between CTT and AOD for warm-base mixed-phase clouds and the positive slope of the linear relationship between CTT and AOD for liquid clouds are both stronger in summer (Fig. 13b).”

12. Figure 8a: change the “Total” to “All”

***Response:***

Done.