Reply to referee #1

The authors changed the inland water and snow mask of original MODIS DT algorithm to increase data coverage in China. The aerosol model and the aerosol layer scale height are also changed to increase the accuracy of AOD inversion. The article has a complete structure and clear logic. However, I still have some major comments I hope the authors will explain before publishing as follows:

We thank reviewer for his/her suggestions. Below is our point to point response.

1. The inland water and snow mask threshold setting of MODIS is strict. Is it a strict mask to reduce the inaccuracy of inversion, or just a misjudgment? From the comparisons in Fig.9b, although the original DT algorithm can retrieve AOD from the additional sample size from new mask, the accuracy has decreased. After using the new aerosol model and aerosol layer scale height (Fig.9c), there is not a significant improvement in accuracy, and the percentage of "within EE" also dropped. If the mask causes a decrease in the inversion accuracy, please comment on the impact of changing the mask conditions on the availability of AOD.

   The reviewer has identified the esoteric problem with aerosol remote sensing with a sensor like MODIS or VIIRS. The multi-spectral measurements contain sufficient information to accurately retrieve some aerosol characteristics, such as AOD, if the other properties involved in the retrieval, such as particle size, absorption and surface reflectance, can be sufficiently constrained. The algorithms must resort to a pre-computed Look Up Table (LUT) and empirical constraints on surface reflectance. The standard global Dark Target algorithm has fine-tuned its assumptions to constrain its retrieval accuracy to well-defined error bars, in a global sense. However, part of the algorithm's success is based on a careful masking of situations that will not match assumptions. In some situations, such as the China-in-winter example that we focus on here, this masking results in the unfortunate loss of a significant number of retrievals with high aerosol loading. This biases the overall statistics of the resulting AOD and reduces product availability for applications that require day-to-day AOD monitoring.

   Thus, the main purpose of this exercise reported on in this paper is to increase the number of AOD retrievals over China. We know a priori that to bring back the once-masked high AOD will immediately introduce a degradation of overall accuracy, because that is the reason these opportunities were rejected in the first place. We also know that adjusting the LUT and other assumptions to improve the quality of the new retrievals will likely shift old retrievals to poorer accuracy. Thus a priori goals are (1) to bring back high AOD, (2) make adjustments to reduce **new biases** introduced by the new retrievals and (3) to minimize new error and scatter introduced across the range of AOD.

   There are trade-offs in trying to meet all three goals. We proceeded with this trade-off by making increased number of retrievals (goal 1) and managing new biases (goal 2).

The new algorithm shows significant increase of number of retrievals along with reduced bias when aerosol loading is high. The result is a success in what we set out to do, but unfortunately at the expense of increasing the scatter across the range of AOD.

To make the success criteria clearer, we have modified the text (lines 68-71) and Figures 9 and 10 in the revision and have also added Table 2.

2. The author divided geographical areas to obtain three aerosol types but does not rule out that aerosol type during the observation period find significant changes, such as dust weather. It is suggested that the author use cluster analysis or other methods to classify aerosol types or discuss the frequency and contribution of special weather. Compared with the original MODIS model, how does the aerosol model proposed in this study contribute to AOD inversion?

   The focus of this study is that DT loses data coverage over winter. Thus, we focus on January to March when spring dust storms are rare. Nevertheless, we generated the aerosol model using AERONET's size distribution which include both fine and coarse particles and represent the averaged aerosol properties, as measured by AERONET, over the region we selected during the study months. The average AERONET models for each of the geographical areas turned out to be sufficiently similar so that we can use the overall mean and not differentiate from region to region. Cluster analysis would tell us the same thing.

   We realize now that the way we presented this work leads to the wrong expectations for a reader. In the revision we approach the description of finding a regional model differently, explaining right up front the conclusion that there is insufficient evidence to use more than one aerosol model for all of China. Then the plots of AERONET-derived size distribution from the different parts of China are put into better context.

   The regional model has a stronger AOD dependency in terms of absorption when compared with non-absorbing and moderate models from the standard retrieval. It is slightly more absorbing in low AOD compared to the standard non-absorbing model (and similar to the moderate model) and its absorption decreases with increase of AOD. The regional model becomes less absorbing than both standard models when AOD > 2. Thus, the new aerosol model results in a slight increase of AOD when AOD < 2 and lower AOD when AOD > 2.

3. In this research algorithm, the authors changed the aerosol layer scale height in the vertical profile in order to obtain better inversion results. However, the scale height is not always 0.5 km in all weather conditions. How did the authors choose the scale height under different weather conditions? If the scale height is always set to 0.5km, how much biases can be caused when retrieving the AOD from January to March 2013?

We agree that the best approach is to change aerosol layer height for each retrieval scene. However, there is no information for us to get the prior knowledge of aerosol layer height nor does the current DT algorithm structure support a selection of aerosol layer height for each retrieval. Our analyses show that when the scale height is set to 0.5 km, the low bias when AOD < 0.5 (as seen in Fig. 9 and Figure 10.) is small. Analyses regarding this issue is also discussed in the reply of next question.

4. Figure 9 shows the validations of the research algorithm, but the advantages of the new algorithm cannot be clearly seen. It is recommended to show the advantages of the new algorithm point by point based on the results of Fig.9b and c like Fig.11. Similarly, the ordinate of Fig10 is too large to see the advantage of the new algorithm. Please adjust the ordinate to a reasonable range.

   Thanks for the suggestion, we modified Figure 9 and Figure 10 and added a Table to show statistics.

   Figure 9 shows how three AOD data sets, namely operational DT AOD, intermediate AOD retrieved using the same LUT as the operational DT but with modified masking (New Mask), and AOD from the research algorithm, compare to each other. All statistics are shown in Table 2. Figure 9a overlays the operational DT AOD onto the New Mask AOD. We can easily identify paired data from the two datasets. The slight differences between two paired data points are expected because these data points represent spatially averaged MODIS AOD and temporally averaged AERONET AOD. When within the averaging criteria new MODIS AOD become available in the New Mask AOD, the averaged value will be slightly different.

   Figure 9a shows there is a large (50%) increment in **the number of retrievals** in the New Mask AOD when AERONET AOD > 1 (19 points from operational DT and 30 from New Mask). **This is the primary goal, as explained in 1. above**. These additional points show that our algorithm successfully retrieved AOD from many high aerosol scenes that are not retrieved in the operational algorithm. However, these extra AOD are highly overestimated with a mean bias of 0.26 when AOD > 1 while the Operational DT shows a negative mean bias of -0.196. Figure 9a also shows multiple New Mask AOD exist without corresponding Operational AOD when AOD is around 0.8 (about 20% increment in number of AOD < 1), which are most likely due to change of snow mask. Figure 9b overlays the research DT AOD on top of the New Mask AOD. We notice that there are large reductions of AOD values when New Mask AOD are above 1.5. The mean bias of the entire data set reduced from 0.16 in New Mask AOD (0.26 from AOD > 1) to 0.076 in the Research AOD (0.097 from AOD > 1). **Reducing bias is the second goal from 1. above**.
   The RMSE also reduced from 0.517 to 0.45, although it is still larger than in the operational DT algorithm. This is the tradeoff we are forced to live with.

These are significant reductions in bias after applying our new aerosol model with a much stronger AOD dependent absorption and using a reduced aerosol layer height. We can also see from Figure 9b that when AOD is lower than 0.5, there are no obvious low points from Research AOD when compared with New Mask AOD, meaning that the change in aerosol model and aerosol layer height has minimum effects when AOD is low. Thus, although we are forced to use one aerosol layer height in the retrieval process that is representative of heavy aerosol loading conditions, the impacts of this choice are small on AOD retrievals when aerosol loading is low. A similar conclusion is also shown in Figure 10. We changed the y-axis data range in Figure 10 to better illustrate data when AERONET AOD is small. We can see that when AERONET AOD is less than 0.5, the mean error pattern and standard deviation of the bins from three data sets are closely following each other. But they diverge at AOD > 1.

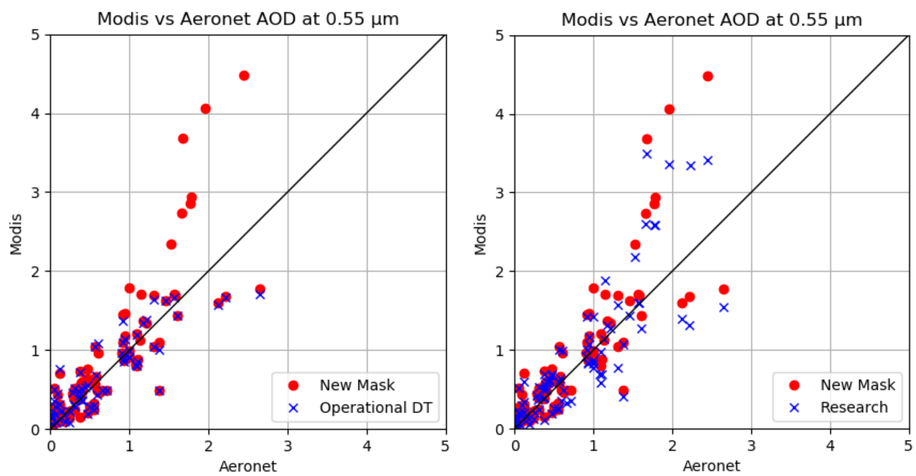

Figure 9 Comparisons of the MODIS DT AOD at 0.55 µm against collocated AERONET observations during January, February, and March 2013 over China. Three datasets are used, operational DT AOD (Operational DT), an intermediate AOD retrieved using the same LUT as the operational DT but with modified masking (New Mask), and AOD retrieved with the full regional research algorithm (Research). a) Operational DT AOD overlay on New Mask AOD, b) Research AOD overlay on New Mask AOD.

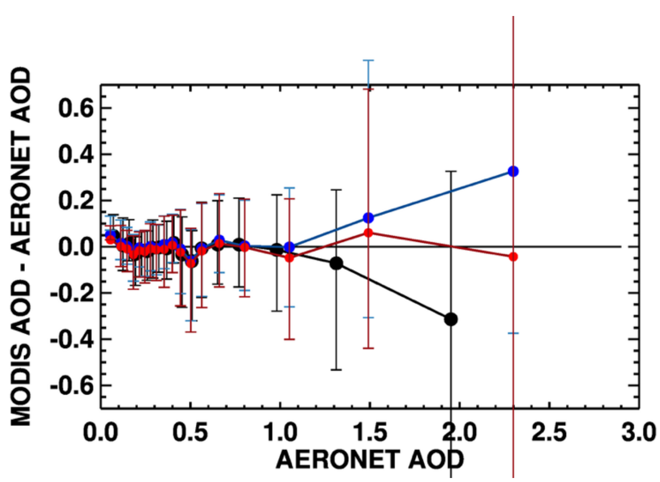

Figure 10 Bias between MODIS and AERONET over land AOD at 0.55 µm as function of AERONET AOD at 0.55 µm. Black represents the operational DT AOD, blue represents the AOD using the operational LUT but with new masks (New Mask), and red represents the research AOD. The dots are the mean bias within each AERONET AOD bin, and the bars represent the standard deviation of the bias.

Table 2 Statistics of validation between Operational DT AOD, AOD using the operational LUT but with new masks (New Mask), and Research AOD against AERONET during January, February, and March 2013 over China. Numbers in parentheses are the statistics for AERONET AOD > 1.

	% within EE	N	R^2	Mean Bias	RMSE	Slope	Offset
Operational DT	40.91	66(19)	0.754	0.003 (-0.196)	0.286	0.75	0.151
New Mask DT	30.34	88(28)	0.700	0.161 (0.260)	0.517	1.01	0.098
Research DT	33.71	89(30)	0.701	0.076 (0.097)	0.450	0.96	0.081

5. MODIS products have a resolution of at least 10km, and the research algorithm in this paper seems to be unlimited on the spatial resolution. So, why use 0.5 resolution for comparison in the 2013 winter characteristics analysis? Please explain.

The gridded product represents the mean states of the aerosol loading over a region and within a time window, while pixel level data show variations of the aerosols loading at certain spot over a period of time. In this section, we want to investigate the change in aerosol spatial distribution due to increasing high AOD retrievals over winter. It shows the bulk impact of research products. Thus, we want to use gridded data and 0.5 by 0.5 degree grid box to insure enough data points in each grid. The impact on pixel level data is shown in the AOD histogram (Fig. 12).