Reply to referee #2

I support the publishing of the paper after taking the following aspects into account. I am a little bit surprised to see which papers are used to support the goal of this study, since the authors belong to the official product team, I believe they should be more careful of this issue. For instance, the authors mainly used Yan et al., 2016, Bilal et al., 2014, 2015, Wei et al., 2019, as the most important starting point of this work, the work of Yan et al., 2016 somehow fits the goal of this study, Wei et al., 2019 mainly raised that Dark-Target cannot provide aerosol product over the bright surfaces, which, to my understanding, does not fit to support why we need a new mask. The work of Bilal et al., 2014, 2015 are using a single scattering assumption for the aerosol model to support the retrieval of haze conditions, if the authors cite these papers, does it mean the authors support the idea of using a single scattering assumption for aerosol retrieval? I will strongly suggest the authors re-check the published papers carefully.

We thank the reviewer to encourage us to perform a more thorough background study. We cite these paper as they all mention the problem of DT missing retrievals over eastern China and tried different methods to solve the issue. It doesn't mean that we support these solutions because in this study we developed a new solution that can be applied to operational production in near real time. We now state this explicitly, and we added following citations:

Li S, Chen L, Xiong X, Tao J, Su L, Han D, Liu Y. Retrieval of the haze optical thickness in North China Plain using MODIS data. IEEE transactions on geoscience and remote sensing. 2012 Oct 10;51(5):2528-40. DOI: 10.1109/TGRS.2012.2214038

Zhang X, Wang H, Che HZ, Tan SC, Shi GY, Yao XP, Zhao HJ. Improvement of snow/haze confusion data gaps in MODIS Dark Target aerosol retrievals in East China. Atmospheric Research. 2020 May 30:105063. DOI: 10.1016/j.atmosres.2020.105063.

Chen W, Fan A, Yan L. Performance of MODIS C6 aerosol product during frequent haze-fog events: A case study of Beijing. Remote Sensing. 2017 May;9(5):496. DOI: 10.3390/rs9050496

Eq (1) and (4) should be re-format

We added parentheses and reformat the equations.

Section 4.1, I believe that the paper from Yang et al., (2020) is some early test of this mask issue, however, the paper is in Chinese and I can only read the abstract part, how the heritage of new mask from MERSI has been adapted to MODIS, taking potential problems of instrument differences.

MODIS and MERSI have very similar design and the two used channels are very similar: 0.645, 0.856, and 2.11 μm in MODIS and 0.654, 0.869, and 2.13 μm in MERSI. Thus, after testing, we adapted the threshold used in MERSI (ρ2.11< 0.08).

Section 4.2, I think it makes not much sense to compare the operational aerosol models, which to my understanding, were derived from other sites compared to the new regional model, which is specifically derived for China, the issues are how they define the region in which the “regional type” can be used? Or they define the whole of China, or the Eastern part, as regions with a "regional type"? A geographic figure to show regions assigned to “regional type” is helpful. Is the “regional type” a single-model or also a mixture of fine and coarse modes, if so, why there is only one mode presented in Table 1?

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.

We applied a single aerosol model to the entire study area that is shown in Figure 13 and 14, calling it our “regional” model, as opposed to the more general global models used by the operational algorithm. From Figure 4, we see that subareas 1, 2, and 3 are relatively similar especially for AOD bins < 2, and AOD > 3 in Area 3 is represented by only one data. Thus, we aggregated aerosol models from all three subareas into one and applied to the entire study region. The regional model is bi-model distribution that contains both fine and coarse mode as what Figure 5 shows. However, although including a coarse mode portion, this aerosol model is still fine mode dominate and is listed as fine mode model shown in Table 1. During the study period, the dominate aerosol type is pollution. Thus, we didn’t change the coarse mode aerosol model, which is made from dust retrievals by AERONET.

Is there any problem with Fig 9 (b), the correlation coefficient is so low? There is quite a large reduction of values fall into EE, can the author explain a bit more about this? We can also see that in Fig 9 (d), the performance of the additional points seems to be worse than the operational products, why?

Thanks for pointing this out, we found that there are two outliers, which cause this issue. We removed these points from our analyses due to these two outlier values are way above 5. These two points values reduced largely after adapting the new LUT and values are within 5.

We will reiterate here what we wrote to Reviewer #1 who raised similar concerns.

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.

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. 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.

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.

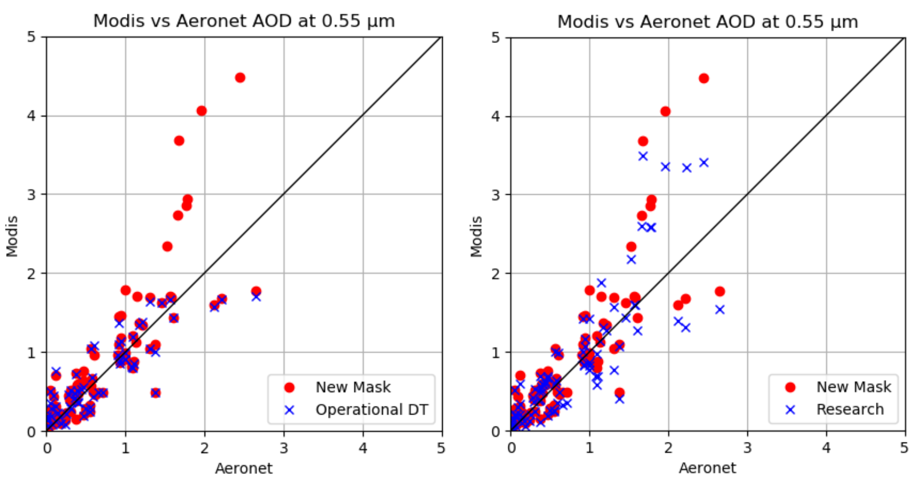


Figure 9 Comparisons of the MODIS DT AOD at 0.55 µm against collocated AERONET observations during January, February, and March 2013 over China. Three datasets are used operational DT AOD (Operational DT), an intermediate AOD retrieved using the same LUT as the operational DT but with modified masking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

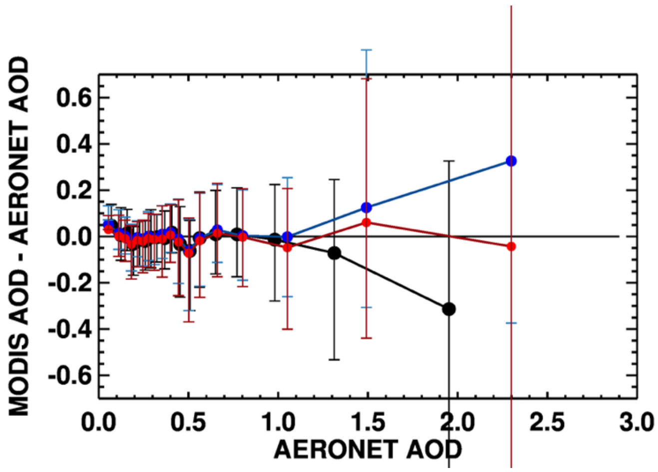


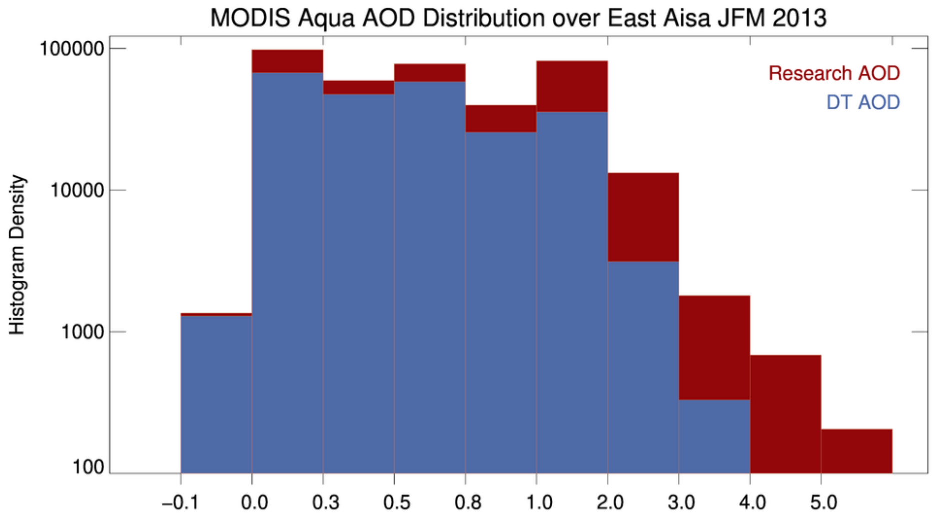
Figure 10 Bias between MODIS and AERONET over land AOD at 0.55 µm as function of AERONET AOD at 0.55 µm. Black represents the operational DT AOD, blue represents the AOD using the operational LUT but with new masks (New Mask), and red represents the research AOD. The dots are the mean bias within each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

Fig 12 should be updated, it is better to have some transparency for the overlap regions.

The reason for adding transparency is to see both bars when one is higher than another. However, after we updated Figure 12 regarding the smaller domain over China (25° to 40° N and 105° to 120° E). The updated figure shows that in all bins, red bars are higher than blue bars. Thus, we do not use transparency plot due to it is unnecessary in this figure.



MODIS Aqua AOD Distribution over East Aisa JFM 2013
Histogram Density
100000
10000
1000
100
Research AOD
DT AOD
–0.1
0.0
0.3
0.5
0.8
1.0
2.0
3.0
4.0
5.0