Review Comments

This study provides a data set that constrains the relative contribution of each of nine major source regions to size resolved dust emission, atmospheric loading, optical depth, concentration, and deposition flux by integrating an ensemble of global model simulations with observational constraints on the properties and abundance of atmospheric dust. The authors show that current models might on average overestimate the contribution of North Africa sources and underestimate the contribution of Asian dust. The manuscript is well written, and results are clearly presented. This study is a valuable contribution to improving global dust cycle in models and constrain dust impacts on the Earth System. I only have minor comments and recommend publication after they have been answered.

General Comments:
(1) The analysis for global dust cycle in AeroCom Phase I simulations by Huneeus et al. (2011) is a benchmark for research on dust modeling and has been widely used to show the large diversities in simulating global dust cycle. Since new features and parameterization schemes have been developed and added to global climate models for the past ten years, how the fast-developing climate models affect the results? Would you expect better agreement with inverse model results for CMIP6 models?

(2) In Figure 2, the model ensemble is better than AeroCom Phase I model results compared with the inverse model results. It is interesting to see that the estimation for Southern Sahara and Sahel from model ensemble is quite close to the inverse model. I wonder if the authors could explain a little more. The AeroCom simulations are for dust cycle in the year 2000. In the companion paper, the models in the ensemble are all nudged to reanalysis during 2004-2008. For each model, dust size range is extended to 20 μm. Many models use the dust emission scheme of Kok et al. (2014) and have dust size distributions consistent with Kok et al. (2017). How much would the models selected in the model ensemble affect the results? I wonder how the authors interpret the differences between the inverse model and model ensemble for dust loading and DAOD. Do representations of dust transport and deposition in the model play a role here?

(3) The authors talked about using dust extinction profiles from CALIOP and CATS to further constrain dust vertical profiles. How about dust concentration measurements from aircraft campaigns, such as ATom? Is the inverse model able to take measurements from aircraft campaigns? What is the limitation arising from the lack of constraints on dust vertical profile?

(4) In Section 4, the authors give a short discussion on the limitation arising from biases in dust transport. It would be nice if the authors could add more discussion on the limitation arising from representations of dust transport and deposition in models.

Specific comments:
(1) It seems that Eqs 5-8 are quite similar to Eqs 1-4, just for particle size bins. I would suggest the authors remove Eqs 1-4 or Eqs 5-8 for simplicity. These equations all contain θ, ϕ, and P only for showing coordinate. I would suggest the authors show something simpler, such as $\check{f}_{\tau_{r,s,k}} = \check{\tau}_{r,s,k} / \sum_{r=1}^{N_{sreg}} \check{\tau}_{r,s,k}$, which is more friendly to readers.

(2) Line 338 and 340,  Figs. 6d-h instead of Fig. 6d-h.

(3) Line 497, I think the authors are talking about Table 3 instead of Table 2.