**Review for Assimilation of MODIS Dark Target and Deep Blue Observations in the dust aerosol component of NMMB/BSC-CTM version 1.0**

Overall, I think this is a nice paper that outlines the implementation of an LETKF for constraint of dust aerosol optical depth in the NMMB/BSC-CTM. This work is a first step towards implementing an operational data assimilation system for dust aerosol and, as noted by the authors, multiple species in the future. The paper outlines the dust emission scheme used in the model, which is the basis for generating ensemble members, the LETKF implementation, the observations used in the system for assimilation (MODIS Dark Target and Deep Blue) as well as evaluation of the analysis and forecast. There was clearly a lot of work that went into implementing and testing the new system. I think this paper is worthy of publication, but there a few things that need to be addressed that will add to the clarity of the paper. Overall, I think the figures need to be reworked. They are quite small and the font on the labels is too small to read on most of them, making it harder to evaluate the results. I also have some specific comments and questions below that would help to clarify the results that are presented.

**Specific Comments:**

1. Page 4, Line 98-100, what about the UK Met Office? They assimilate dust AOD in their unified model.
2. Page 4, Lines 107-109, limited spatial correlations have been shown in some studies, depending on what they are sampling, but do you suspect this would be much longer for big dust transport events, especially coming off of the Sahara over the Atlantic ocean? This looks like it would be the case based on MODIS observations.
3. Page 5, I think it would be helpful to include units on the variables in the vertical dust mass flux equation. Also, what value do you use for C? This is constant globally? Is the source mode coefficient how you distribute the mass among the size bins? What threshold do you typically use for the friction velocity (when not perturbed for the ensembles)?
4. Page 5, Line 165, what were the main sources of uncertainty identified in the evaluation efforts?
5. Page 6, to clarify your data assimilation approach, you might want to mention some specifics about the 4-dimensional extension of the LETKF and why you chose to use the extension since you are assimilating observations regularly over 6 hour intervals with the NRL MODIS product. Do you expect to incorporate observations in the future that are asynchronous? Did you test at all the performance of the LETKF versus LETKF with the extension? This would be interesting.
6. Page 7, Line 210-214, what units do you use to define the distance in the localization and what localization factor do you use? It's hard to tell from this how much localization is used.
7. Page 7, Ensemble perturbations in the vertical flux, You are perturbing the distribution of dust emissions among the size bins, but the total mass flux is held fixed? Are the perturbations that you show in Figure 1 the same for all locations or does this vary by grid or region? It might be good to change the solid red line with the error bars in Figure 1 to make it easier to distinguish from the ensemble perturbed lines. Maybe to a dotted or dashed line? Also, it might be useful to show somewhere what sizes the bins correspond to.
8. Page 7, Ensemble perturbations in the threshold friction velocity perturbation. Again, do the random perturbations vary with location or are the same perturbations applied everywhere?

This matters as it will determine your covariances and how an observation spatially impacts your model state.

9. Page 8, Lines 237-239, if the structure of your source perturbations is temporally and spatially constant, you are essentially specifying your background covariances, much in the way a variational approach operates. As you mention, this is the first stage of development, so I think that's a reasonable first means for generating the ensemble and will probably help you do well near source regions, but you may have problems for transport events.
10. Page 8-9. MODIS Dark Target, I would increase the size of Figure 2, it's too small to see. It would probably also be useful to see some sort of summary of the observations over the experimental time period, perhaps a data count to see where your simulations are being constrained or a mean of your observations. Also, I'm concerned about using over-land AE as a filter for dust. It's been shown that this product is pretty binary (see Levy et al. 2010) and more problematic for coarse mode aerosol than fine. Have you checked to make sure you aren't getting other aerosol in there, like biomass burning aerosol? Perhaps this could be contributing to some of the bias that you are seeing.
11. Page 10, Numerical Setup. The Control is the exact same model as the ensemble free run, the only difference in the ensemble free run is you have perturbed dust emissions (either in the distribution in the bins or threshold friction velocity) and the control is a single run?
12. Page 11, Lines 358-360. I suspect your insensitivity to ensemble size is a result of how you are generating the ensembles themselves (you are sampling from a specified distribution) and also maybe you are heavily localizing (can't tell without units though). This will likely change as you add other perturbations to your system and you may find that you need a much larger ensemble as 12-24 members is quite small.
13. Page 11, Lines 365-367. Did including vertical localization make much of a difference for AOD assimilation? I'm not sure if you tested it without, but I would think that this wouldn't have much impact for a column-integrated observation.
14. Page 12, your use of error as observation minus model is a bit confusing to me. The bias for example would have a negative value when the model is biased high. Typically, you would use your estimator (model) minus the expected value (observation). I would suggest flipping this so that your bias maps in Figure 11 and stat tables/bar graphs don't confuse the reader into thinking the model is biased low when the opposite is true.
15. Page 13, Section 7.1 I think it would be beneficial in Figure 5 to also show the difference between the DA experiments and your ensemble free run (or control). The difference between the DT+DB simulation and the free run is pretty clear, but harder to see with the DT run. Also, I assume this is dust AOD only? If so, you should probably put that in the Figure caption and mention that in the text as well (Page 13, line 444). Are these differences persistent over the entire simulation since you only show one month?
16. Figure 6, Does the DT simulation's coefficient of variation look similar to the DT+DB? If so, you might want to mention that in the text. If they are different, you should probably show both. Also, does the mean AOD change much with the different perturbation schemes (Figure 6 and 7)?

17. In Figure 6, I'm surprised that you have considerable spread in places that I wouldn't expect, like near the poles in the Southern hemisphere. Are the ensemble members being inflated as part of the data assimilation?
18. Page 14, Lines 460-462. This sentence implies the more spread the better since you'll just push towards the observations. However, your goal is to really have sufficient spread that represents the uncertainty in the system. Have you tried to determine whether or not the spread that you are generating is representative of the uncertainty?
19. Figure 8, I would remove the color bars here for each subplot to save space and increase the individual plot size and font size (same for all the figures). I also wonder if you increase the number of bins in your scatterplot, whether the asymmetry that you talk about would be more apparent.
20. Figure 9, The analysis increments that you are showing are in dust AOD? If so, you should add that to the figure caption or labels.
21. Page 15, Section 7.2 For AERONET sites in transport regions, such as La Parguera, it looks like the dust AOD has decreased with data assimilation compared to the control. However, the analysis increments shown in Figure 9 show an increase in AOD. Perhaps the prior state has decreased so much with the near-source corrections that the increase observed over the oceans still produces an AOD at sites impacted by transport that is still less than the control? I'm curious what you found with that.
22. Figure 11, I would put one colorbar at the bottom of each column of figures then maybe add one label at the top of each column (Control, DA-NRL, DA-NRL-DB) and add one label on the y-axis for each row (Bias, RMSE, Corr, FRGE). That way you can increase the size of each map and make the labels larger. Also, it's so small that it is impossible to see any difference in the circle sizes and there is no reference to use to determine what number of samples the circle size corresponds to.
23. Figure 14 and 15 need to be fixed, the labels are way too small to be able to read. It makes it hard to evaluate your forecast results.
24. I wonder if you might want to show in your statistics bar graphs the average dust AOD as well to give some context to how large the errors really are and maybe considering adding error bars (maybe through bootstrapping) to your statistics to test if the differences are statistically significant.

**Technical Corrections:**

1. Page 3, Line 67, change "to different model inter-comparison" to "in different model inter-comparison"
2. Page 3, Line 73, saying the community resorted to data assimilation makes it sound kind of negative. Maybe you could say something like…because of these large uncertainties, the atmospheric composition community has begun to make use of data assimilation for better characterizing and predicting….

3. Page 3, Line 79, you might want to cite the Sessions et al. 2015 paper after the sentence where you mention that assimilation of aerosol observations is now operational at many forecasting centers.
4. Page 10, Line 325. You should probably cite the AERONET uncertainty
5. As a suggestion on your equations, you may want to go through and make sure the variables are consistent across equations. For example, in equation 4 the size bins 1 through 8 are indicated with a *b* while in equation 1 they are indicated with a *k*. Later *k* refers to ensemble members. This might confuse the reader. Also, it would be useful to include units with your variables.
6. Page 13, regions for validation (Lines 430-440). I think in Figure 4 it would be good to list the regions associated with each box. You can probably just put this in the figure caption and say which color box goes with which region, to tie the map to Table 2.
7. For Figures 5,6,7, the colorbars are the same on the different subplots within each figure, so I would only show the colorbar once to save space and make the maps larger. They are too small to see clearly.
8. In the caption for Figure 10, you should mention that this is the analysis AOD and not the prior.
9. Page 17, Line 567-568. This sentence isn't very clear. You are referring to the Sahara? Better temporal evolution, reflected by the increase in correlation with AERONET over time?