Review of RAMS-MLEF Atmosphere-Aerosol Coupled Data Assimilation: A Case Study of a Dust Event over the Arabian Peninsula on 4 August 2016

The focus of this paper is to use coupled atmosphere-aerosol data assimilation to improve forecast initial conditions using ensemble data assimilation (Maximum Likelihood Ensemble Filter). The focus is on a dust case study to demonstrate the use of coupled data assimilation in the difficult to predict littoral zone. Coupled aerosol-atmosphere data assimilation is very new, so I think this work is a useful contribution. However, I believe for it to be published additional information and substantially more analysis needs to be provided. There are several comments that I think should be addressed:

1) In the introduction (page 2, line 31-32), it is mentioned that the goal of this study is to improve representation and forecast of aerosol in the littoral zone. This is the only place in the whole paper that this is mentioned. If this is a main goal of this work and part of what makes this analysis unique from previous aerosol-atmosphere coupled data assimilation studies, it would be beneficial to expand on this in the introduction and mention what the results mean for this goal in the conclusions. It would be useful to address why coupled data assimilation is the right strategy for this problem and why you chose this case study with the goal in mind.
2) One of your main conclusions is that the assimilation doesn't improve the representation of the Saudi Arabian plume because of lack of observations in the interior of the plume. In your introduction, you discuss various MODIS retrieval algorithms for AOD, including the Deep Blue retrieval over land. The Deep Blue MODIS retrieval would provide the over land observations for the Saudi Arabian plume. Why did you not assimilate this dataset as well? It might be worth rerunning the experiment with these observations included.
3) Can you describe in more detail how the ensembles are formed? Do the ensembles account for differences in boundary conditions for both the atmosphere and aerosol components? Along this same line, I think you should include some analysis of what the ensemble spread looks like for your experiment since this is important in determining the kind of impact the observations will have. Do you need some spin-up time for your experiments to generate sufficient spread? Did you do any analysis to determine if the spread is enough? This is a very short experiment, so more details to see what is produced in 6 DA cycles would be useful.
4) You cite an observation error from Liu et al. 2011, which just refers to the observation error from Remer et al. 2005. I would recommend citing Remer et al. 2005 as well since this is the source of the observation error that you are using. Also, the observation error that they use is actually 0.03 + 0.05 *AOT over water and 0.05+0.15*AOT over land. There are more recent error evaluations for collection 6 MODIS products (Levy et al 2013; Sayer et al 2013) There are also MODIS error evaluations specifically focused on aerosol assimilation that you should consider for your work (Zhang and Reid, 2006; Hyer et al 2011; Shi et al 2011).
5) In Figure 6, it would be helpful to include the analysis of both experiments (ATMONLY and ATMAOD) since you are trying to compare the two experiment. If you don't show it because the difference in AOD was minimal between first guess and analysis for ATMONLY experiment, then it would be worth mentioning this. How do the analyses from ATMONLY and ATMAOD compare throughout the experiment?

6) The conclusion that including AOD in the assimilation brings valuable information is not surprising since it's directly related to aerosol and has been shown to be of value in many other data assimilation papers, including some papers that were cited. It would be useful to get a better understanding from your work of the impact of the meteorological observation on aerosol. It doesn't seem like they are doing much for this case? Would you then just recommend weakly coupled data assimilation? Does the aerosol have any feedback on the meteorology? I think more discussion/analysis along these lines are needed since this is the important part of this work.
7) For the analysis increment, you only show results from level 11. Why did you choose this level to look at? Is most of the aerosol contained in this model level? Are the increments similar throughout the model profile? I think you should expand on this analysis. I think it would also be useful to show the first guess, in addition to the increments, to get a better sense of how large of an impact the observations have.
8) You mention in the analysis increments section of the results that it is difficult to identify relationships between different variables. If you only allowed one additional variable at a time in your data assimilation (ie. aerosol and wind, aerosol and water vapor etc), perhaps you could get a better understanding of what the different cross-correlations are doing.
9) You verify your results against the MERRA renanalysis product. You might want to include other products for comparison. The Navy also has an aerosol reanalysis product (NAAPS) and the ICAP multi-model ensemble is also a very useful product for comparison.
10) It would be useful to mention in the paper that you are using the ensembles to provide you with the forecast error covariance, including the cross-component elements that are needed for the strongly coupled data assimilation.
11) Lat/lon information on some of your figures would be helpful (Figure 3,4) and it is difficult to read the values on the colorbars in Figure 8.