Reply to Reviewer #2:

I must apologize upfront for having such a short review-I have been in the field for most of the month. I did give the paper a good read. All in all I concur with the first anonymous reviewers on recommendations. As for what I have to add, I temper my comments with the understanding that for this journal, the paper is really a report on SPRINTARS and their methods for improving computational cost relative to an inline model version for future reference. They then apply their system to the problem of inverse modeling. The paper is all in all well written and clear for which the authors should be commended. Theory is reasonably well laid out, although their intro has a distinct European and Asian focus. The only major concern I have is in section 5, regarding the use of fine and coarse mode aot from MODIS (see below). But all in all I recommend accepting with moderate revisions. I don't need to see the paper again.

We thank a referee for careful reading our manuscript and for giving useful comments. We have deliberately confirmed and considered your comments. We believe that we have made adequate answers to your comments. We also believe that we have made adequate corrections to comments by the referee #1. In revised manuscript, the changes are highlighted by yellow makers.

Comments:

1) Section 2 & 3: I do wonder about the practicality of having a 4D var system applied to a nearly 3 degree horizontal grid spacing-at that spatial resolution and that the correlation length for aerosol particles is about 150 km, I bet 2D var would work just as well with 98% cost savings (indeed, they should just try it as a control). But, as I noted above, this is largely a descriptive journal so I have no problem with the underlying premise.

At first, we should note that the SPRITNARS/4D-Var data assimilation (DA) system is capable of adapting to finer resolutions (e.g., T106 and T213). In this study, we focus on the inverse modeling of aerosol emissions with satellite measurements and the variational method. The inversion of aerosol emissions with the satellite measurements over the downwind regions (i.e. over the ocean) requires 3-7 days of assimilation windows. This is one of the main reasons why the 4D-Var data assimilation (DA) method, which uses an adjoint model of CTM, was applied to the aerosol inverse modeling. In addition, the 4D-Var DA system can assimilate all observations measured in the period of the assimilation window simultaneously. It is difficult for sequential DA method (e.g., 2D-Var and 3D-Var) to do that. 2D and 3D-Var methods would show comfortable performance for modification of initial condition in an operational aerosol forecasting using observations measured at a given time with much lower computational cost, as the reviewer mentioned.

2) For section 4, I suggest that not so much time be spent on the angstrom exponent, rather they compare fine and coarse mode AOT to aeronet a much more straightforward and robust metric. For Table 1, the use of a regression as a metric is ambiguous. I would recommend plot of rmse as a function of AOT. Also in the table text define the acronyms in the table (that goes for all tables actually). Table 2 and 3 can also be combined to save space.

One of the main aims of section 4 is verification of the off-line version of SPRINTARS (OFS). All statistics were calculated between OFS and the original on-line version of SPRINTARS (ONS). The performance and reproducibility of SPRINTARS are beyond of the scope of this paper, and no observation data (e.g. aeronet data) was used in the verification. The reproducibility of SPRINTARS are discussed in other papers (Takemura et al., 2000, 2005, and 2012), and SPRINTARS is involved in Aerosol Comparisons between Observations and Models (AEROCOM, http://aerocom.met.no/Welcome.html). Also see the reply to the comment #3.

In addition to Angstrom exponent, which is commonly used as an indicator of the aerosol size distribution, we also calculated statistics for AOTs of each individual aerosol. We think that statistics for fine- and coarse-mode AOTs are duplicative and have no need to be shown.

We showed regressions to investigate trends (over- or under-estimate) and bias of OFS against ONS, and left those values in the revised Table 1. RMSEs were calculated to each aerosol, and shown with mean AOT values in Table 1. Moreover, we added definitions of the acronyms in the revised Tables 1 and 6.

Main sources and the regions used to calculate the emission amount were different between dust and sea salt. In order to prevent readers' misunderstanding, we decided that Tables 2 and 3 remind separated.

3) Another interesting thing to consider that while indeed there are small differences between the inline and offline model (say 6-8%), it would help if it was framed into the context of some aeronet data. How does a 6% change compare to the rmse for a few key stations instead of looking at bulk? In fact, a map of the aeronet sites used would be very helpful. The earlier sections are a bit terse on this issue. For section 5, what is not clear to me is if they are using fine and coarse mode AOT from MODIS for over land in their analysis. If so, it needs to be retracted in the next draft. The fine/coarse aot over land has no skill as even admitted by the algorithms. Section 5 is an area where the authors need to be more careful. Again, as the point of the paper is to demonstrate the model, not invert source functions I am not too concerned. But, satellite error characterization is 3/4 of the problem, and there is precious little information on this point. So, they authors should spend more time discussion how their simulated observation relates to real data and associated error. Also, clearly (from figure 6) the da system is doing its job. But the background difference between CR and NR is really quite large in Asia and in places where the AOTs are quite large. Thus, while on average spatially and temporally the two models are very similar, as depicted in figure 6 b the difference for over a week could be as much as a factor of 2. This needs a bit more work and description than is in there than what is simply presented in Figure 8. Are there certain meteorological conditions that lead to the skewness of figure 8? Is one 10 day run for the inversion experiment sufficient? It might be one way for one period, another later or earlier.

Thank you for your suggestions. The accuracy of the off-line version of SPRINTARS (OFS) depends on variability of the input meteorological data, which is linearly interpolated to the model time step. Dust and sea salt emissions that are strongly affected by variable data (i.e., wind speed at 10 m) showed relatively larger differences.

As the reply to the comment #2, in section 4, in order to validate the performance of OFS (i.e., consistency with ONS), we calculated statistics using values at the whole model grids with 3-hour interval. We also discussed horizontal distributions (Figure 3) and vertical distributions (Figure 4) of deviations, and comparisons of dust and sea salt emissions in each major source region (Tables 2 and 3) to investigate the regional dependency of the performance. Because reproducibility of SPRINTARS is beyond this paper, we think that a comparison with specific aeronet data is not necessary. In the future study in which aeronet data is assimilated, we need to investigate how difference between ONS and OFS at aeronet sites affects assimilation results.

We do know the problems of AOT over land, and mentioned the notice of the land product in the manuscript (Lines 21-25 on Page 3441) referring to Remer et al. (2008). In addition to the larger uncertainty in the land product, the MODIS product cannot provide AOT over bright surface regions (such as desert) except for the Deep Blue product. This leads the much less coverage of the land product than the ocean product. In this study, as mentioned in the manuscript (Line 25 on Page 3441), Experiments 1-3 (E1-3), which assimilated fine- and coarse-mode AOTs over only the ocean, were performed to assess the real capability of the system. Moreover, we also conducted two additional “sensitivity” experiments (E4 and E5) as mentioned at Lines 25-28 on Page 3441. In the sensitivity experiments (E4 and E5), we assumed the case if the MODIS product could provide fine- and coarse-mode AOTs over the land with the same coverage and accuracy as over the ocean, and evaluated how much that land product has impacts on the inversion (because the real land product would have much less impact comparing with the ocean product due to the not separated (total) AOT, the lager uncertainty and the less coverage). This is also because we use the same error over land than over ocean. We added more description about the sensitivity experiments in the revised manuscript. Also see the replies to the specific comments #8 and 9 by reviewer #1.

"In the sensitivity experiments (E4 and E5), we assume the case if we could obtain fine- and coarse-mode AOTs over the land with the same frequency and accuracy as those over the ocean."

The simulated observations are derived from the nature run (NR), which is a proxy of the "true" state, and used in the same manner as the real observations. To bring the OSSEs settings closer to the real case, we assigned the error of the MODIS product estimated by Kaufman et al. (2005) to the observation error. Information about the deviation between NR and CR is included in the background error. We added a new paragraph in section 5 to describe the setting of the inverse modeling more carefully. Also see the reply to the general comment #2 by reviewer #1.

"We assigned scaling factors of aerosol emissions to control parameters. The scaling factor allowed increases or decreases of the existing emissions, and could not detect missing sources. Because the CR was generated by emissions perturbed by the scaling factor (as shown by Equation 15), detection of missing aerosol sources is beyond the scope of this experiment. Emissions of dust and sea salt aerosols, which have several particle bins, were adjusted as total emissions (not each emission of their bins). The initial aerosol conditions were not included in the control parameters, and the CR, NR and ARs were initialized with identical aerosol fields. The background errors were based on the setting of the CR. We assigned 200% for SO2 and carbon emissions. For the natural aerosol emissions (dust and sea salt), 300% of uncertainty was assigned. Temporal and special correlations were not considered in this study. The experimental period was 10 days (21–31 May 2007) based on average lifetime of aerosols in atmosphere. The assimilation window was the full 10 days period. "

The control run (CR; before assimilation) was generated by the randomly perturbed emission. Consequently, the deviation between NR and CR has regional and temporal dependencies. For example, over East Asia where various and large aerosol emissions are situated, there are larger differences. Another example is that we find larger differences in the period when dust storms occur. That mainly contributed to the large difference between CR and NR and the skew in Figure 8. Note that the difference and skew are fine with the OSSEs, because the main aim of the OSSEs is to estimate the performance of the DA system and the simulated observations by assimilating the simulated observations to CR. In Figure 8, there also are skew in inversion results (E2 and E4) (less than CR). Regional dependencies of assimilated data due to cloud and land/ocean leaded these skew. On the other hand, the perfect experiment (PE), which assimilated the simulated observations over the globe (all sky, ocean and land), shows almost symmetric distributions. Meteorological conditions have little contributions to the skews, since CR, NR, and ARs used the identical meteorological field. We added more description to Figure 8 in the revised manuscript.

"The CR shows skewed distributions due to the emissions randomly perturbed by the scaling factors which have the maximum limitation. The inversion results by E2 and E4 also show skewness. Regional dependencies of assimilated observation data due to cloud covers and land/ocean lead this skewness. On the other hand, the PE, in which the simulated observations over the globe are assimilated, achieves symmetric distributions and significant …"

We decided the assimilation window based on average lifetime of aerosols in atmosphere (about a week). In the future studies, we plan assimilation experiments with the longer assimilation windows and the real observations. We added more detailed description about the assimilation window in section 5 as shown above.

[References]

Kaufman, Y. J., Boucher, O., Tanré, D., Chin, M., Remer, L. A., and Takemura, T.: Aerosol anthropogenic component estimated from satellite data, Geophysical Research Letters, 32, L17804, 10.1029/2005GL023125, 2005.

Takemura, T., Okamoto, H., and Maruyama, Y.: Global three-dimensional simulation of aerosol optical thickness distribution of various origins, Journal of geophysical, 105, 2000.

Takemura, T., Nozawa, T., Emori, S., Nakajima, T. Y., and Nakajima, T.: Simulation of climate response to aerosol direct and indirect effects with aerosol transport-radiation model, Journal of Geophysical Research - Atmospheres, 110, 10.1029/2004jd005029, 2005.

Takemura, T.: Distributions and climate effects of atmospheric aerosols from the preindustrial era to 2100 along Representative Concentration Pathways (RCPs) simulated using the global aerosol model SPRINTARS, Atmospheric Chemistry and Physics, 12, 11555-11572, 10.5194/acp-12-11555-2012, 2012.