**Atmospheric
Measurement
Techniques**
Discussions

Interactive comment on "Aerosol retrieval experiments in the ESA Aerosol_cci project" *by* T. Holzer-Popp et al.

J. Reid (Referee)

reidj@nrlmry.navy.mil

Included here are selections from my "pre-review" which was submitted in the first stage of the review process. Being a pre-review, the authors were not bound to respond, and I see from the latest round by and large they hadn't responded to the major comments. Much of what I said then holds still today, so by and large I submit it again. Bottom line is that this is really just a report for the project, but without any solid substantial scientific results-one month of data is not enough to say concretely anything about products-especially with the coverage issues associated with AATSR . Alternating 6 months is the minimum before other than that from operational NASA products point of view, the ESA product lines have a long way to go. This is by no means a snarky comment, but

a simple fact. I encourage the team members to apply lessons learned from the NASA development side, and in due time I think the ESA time series too will be valuable climate data sets. This paper is a report on the first round of intercomparisons and sensitivity tests of the ESA Aerosol_cci project. In summary, they examined September 2008 globally for a number of sensors and algorithms. For this one month, global plots are given for a) the native algorithms b) Algorithms run with identical aerosol microphysical models. c) A repeat of b) with a climatological aerosol prior to select the optical model, and d) a repeat of c with identical cloud masks (it is not entirely clear to be that it was not b)). From this the do brief comparisons and concluded that in general the retrievals improved by the use of the updated optical models, and it was a mixed bag for the use of the climatological prior. In the most general sense, this paper represents a progress report for the project. A three month followed by a 1 year examination is in the works. They do a nice job explaining the different algorithms and what they did. In so much that this is really a report, in itself I have no objections to it moving on to fuller review. This said, the science of this report, and in particular in regard to methodologies for verification and evaluation, is really quite poorly described. First and foremost, while I understand how much work is required for a number of small shops to do a global analysis (the amount of data to be moved around is enormous), one month is not enough to say anything substantial with global conclusions. This is especially true with such narrow swath instruments such as used here. It is like trying to verify MISR with 1 month of data, something the MISR team would not do. What is presented is the grossest of "sniff tests" for the products. Since using improved optical models appears to help, this is not entirely wasted time, but if algorithms are to be 'judged' to determine which one will go into production, they are doing a disservice to the developers and sponsors alike. Specific issues for consideration include

1) Based on Figure 5, it seems that there is very little data making it into the composite plots. Contextual bias is likely severe. A map of how many samples actually go into the map is certainly required. I imagine it is on the order of 3 or 4 over the non-arid parts of the world. Is this really enough to say anything concrete?

2) For AERONET, the scientists here have fallen into the same trap as the MODIS folks. Namely, MODIS optimized their retrieval to the global data set. But, the overwhelming majority of sites is in the US and Europe. Thus, when regional studies were conducted, the algorithm failed everywhere except... eastern US and Europe. You can show an improvement in bulk scores by using their optical models, but in reality make things worse in certain locations. Certainly, they need to break things down to areas of a dominant species (smoke, biomass burning, etc). Also for AERONET, they are using log scales. Their RMSE is large enough that they can plot this linearly. As we do (look at Yingxi shi's papers) I would do a linear plot, with a regression line and RMSD bounds.

3) Over both land and water, it is clear from the plots that the lower boundary condition algorithm components are deeply flawed for all algorithms. Over land this is understandable. Over water, this is about as slow of a pitch possible. Even more discussion of the lower boundary condition needs to be included, as it is what is driving verification bus.

4) In general I found the figures small and very hard to read. Some effort needs to be made in enlarging and improving figure quality. I see they took the comment that they should use one color bar, but now the figures are a bit awkward. My suggestion (To one and all in our community) is the money spent on Adobe Illustrator is money well spent.

_____

Interactive comment on Atmos. Meas. Tech. Discuss., 6, 2353, 2013.