Atmos. Chem. Phys. Discuss., 15, C8891–C8898, 2015
www.atmos-chem-phys-discuss.net/15/C8891/2015/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "The importance of temporal collocation for the evaluation of aerosol models with observations" *by* N. A. J. Schutgens et al.

A. M. Sayer (Referee)

andrew.sayer@nasa.gov

Received and published: 3 November 2015

1 General comments and recommendation

This paper uses global model runs and observational aerosol data to illustrate that sampling effects on climatologies of aerosol optical properties (AOT, AE, SSA) often do not tend to zero even when fairly long averaging periods are considered. This is an important step in the recent discussions within the community towards understanding sources of error in climatological aerosol data and model/observation discrepancies, particularly since the currently-available remote sensing observations have some sig-

nificant limitations in their temporal sampling. So this study is relevant and important. The use of three separate model data sets helps to generalise the results somewhat, and give an idea about model diversity (i.e. this is an issue for all the models, not one specific model). I have no major problems with the work presented, although have a number of comments, below. I therefore recommend that the paper be accepted for publication in ACP following revision to account for these comments. The main comment of note is the use of 6-hour vs. 24-hour comparisons, and how it would be very useful to be able to state to what extent 6-hour data are better than 24-hour data for these types of comparison (it could help set a new standard approach). I am happy to review the revised version if the Editor feels that this is necessary, and the authors are welcome to contact me (Andrew Sayer, andrew.sayer@nasa.gov) if they have any queries about these comments.

2 Specific comments

General: I think that the Copernicus style guide requires acronyms defined in the Abstract (e.g. AOT, AE, SSA) to be defined in the main text at first use as well. Similarly MODIS, AERONET, AEROCOM and others are presented without definition.

P26193, L20 onwards: Another study which might be worth mentioning here is Smirnov et al. (2002), which looked at the diurnal variability of AOT from AERONET sites. So this was similar in principle to the Kaufman study cited, although shows the breakdown of results from individual sites directly.

P26194, L7-12: You might add that it is more common for satellite-satellite comparisons to use common spatial sampling on a daily basis already (e.g. the various comparisons between SeaWiFS and other data presented in Sayer et al., 2012), since

in that case both data sets have temporal sampling limitations. So these issues are known about in the satellite community, it's more that the modelling community haven't started accounting for these sampling issues as often yet. (And hopefully the present study will contribute to this changing.)

P26197, L1: I would suggest Levy et al. (2007, 2010) as better references for the MODIS Collection 5 aerosol products, since Remer et al. (2006) (mistakenly given as 2005 in the paper) was an analysis of Collection 4, not Collection 5, data. The Levy papers provide some algorithm/validation details for Collection 5 (although focused on land; there isn't an ocean C5-focused paper to my knowledge).

P26197, L24-27: I agree that it's a good idea to be using these bias-corrected data sets for probing these sampling effects on the models, and in that sense 6-hourly output would seem to make sense, since that's the temporal output of the NRL product. However, a large number of model comparisons are done just with the standard MODIS products, which are provided on a daily (or longer) basis, and so these comparisons are made on a daily/monthly/yearly basis. So, my question: is it feasible to add a discussion of how much the results change if the temporal compositing of the remote sensing data is not 6-hourly but 24-hourly? I think that this would be helpful. If it turns out that things don't change much, then this suggests that people who are currently doing monthly/yearly comparisons may be ok if they just go to daily remote sensing products instead (which are readily available). On the other hand if 6-hour and 24-hour results are different, then this is useful information since it means that us satellite data providers should really consider providing 6-hourly level 3 output (as opposed to the current daily/8-day/monthly now) as a standard, because these 6-hourly products are not as highly used (i.e. although the NRL and AORI bias-corrected data sets exist, a lot of people still go to the un-bias-corrected NASA standard products). As is noted later in the manuscript, standard practice

for AEROCOM output is also daily rather than 6-hourly output, so this would be a change which the modelling teams would have to adopt as well. So the answer to the 6-hour vs. 24-hour question could be in my view an important finding and action item coming out of the manuscript, if the analysis could be extended to answer it.

P26202, L7-11: Clouds as well as snow/ice are important in winter.

P26202, L18 onwards: This paragraph indicates an artificial contrast between land and ocean can be created in AE data, as a result of a minimum AOT threshold in the AORI bias-correction and filtering algorithm. I think that it should be made clear that this is a specific feature of the AORI data set (from the current text it is implicit if one is familiar with the literature and MODIS products, but a less familiar reader may infer incorrectly that this is a feature in the standard MODIS products, which is not the case). Therefore this result may not be transferable if one were using the standard MODIS products, or an alternative bias-correction technique which used a different (or no) AOT threshold.

Also of relevance to this section is that the latest MODIS Collection 6 Dark Target products do not provide AE over land, only over ocean (Levy et al., 2013). This was a result of analyses (e.g. Levy et al., 2010) which indicated that there was little skill in this parameter. So this result about the artificial contrast is not transferable to the most current version of the MODIS data used. (NRL and AORI bias corrections have not yet been made available for Collection 6, so it's reasonable to use the older Collection 5 data, although the caveat could be made in the paper that the results will change a bit once these bias-correction data sets are updated). Note also that MODIS Deep Blue (Hsu et al., 2013), not used in the present study, DOES provide AE over land and appears to have some skill (Sayer et al., 2013). So this information is still available from standard MODIS Deep Blue products over land (albeit not in the Dark Target

products which are used in the current versions of the NRL/AORI products).

P26203, Section 5.4: The plots accompanying this (Figures 10 and 11) are interesting in that the errors, as one would expect, are smallest when the observational data coverage is high. I wonder if the authors could add a map or some discussion showing what the observational data coverage is for the data sets used. This way if one wanted to put a filter of, say, 50 % coverage when considering where to compare, we could get an idea for how much of the world this would throw away, and where this would happen. I did something similar to this in the paper Sayer et al. (2010), which the authors cite earlier in their manuscript, although this was fraction of grid box filled rather than fraction of time series filled, and the observational data I used then had poorer temporal sampling. I realise that the point of the present study is to encourage people to use coincident temporal sampling, but the point of my suggestion here is that seeing where these low-coverage areas are most likely to occur can also provide guidance for understanding results from older studies where such temporal collocation was not performed.

P26203, L18: The minimum AERONET AOT threshold for a successful Level 2.0 inversion with SSA is 0.4 (at 440 nm), not 0.2 as stated here. I'd suggest adding a reference to Dubovik et al. (2000) here, together with a bit more discussion about this topic.

P26204, L22: This discussion and associated figure could also be used to argue for model and satellite output being provided at 6-hourly, rather than daily, time scales (see my prior main comment). This is an important figure since it basically shows a longitudinal dependence of error from the definition of a 'day' relative to UTC.

P26205, L6-8: I'd like the authors to expand a bit on this comment in the Conclusions. If sampling errors on daily data are larger than observational errors, then that suggests that the first step in improving the utility of observational data sets is not to reduce observational errors but to improve observational coverage. Although cloud/surface issues will preclude complete coverage, this could be used to make the argument for a multi-sensor bias-corrected AOT data set a priority. I believe that NRL either have or are creating a combined MODIS Terra/Aqua data set (not sure if MISR is being included as well or not, but this would fill some gaps in MODIS Terra Sun glint holes over the ocean). It (the whole paper in fact) is also an argument for the development of AOT data sets from geostationary sensors (particularly the second-generation such as GOES-R, Himawari-8, MTG) as well as DSCOVR/EPIC at the L1 point to get at diurnal variability. It'd be good to have some more discussion of these possibilities, to help point to the need for them as a complement to our existing polar orbit sensor data records.

Table 2: From the referring text (P26201) and caption, I'm not certain what the different rows in this table refer to. Why isn't there just one number per model? Can the text and caption be clarified?

Figures, general: Can we have a map of the NRL and AORI global AOT fields, to see how similar they are? (Or at least a brief discussion in the text about this.)

Figures 1,2, 4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15: Can font size on labels and legends be increased?

Figure 2 (and other AOT maps later e.g. Figure 7): For the panels showing AOT, green is not necessarily an intuitive color scale for showing AOT variations: something

based on red/brown might be better (e.g. the reverse of the scale used in Figure 13). Also, a discrete colour scale (e.g. 10 levels from 0-0.5) might be easier to pick out the exact AOT at different locations from the different models. At present for example it's hard to resolve some of the shades. Also, it might make sense to put the AOT and AOT standard deviation on the same range (currently AOT is 0-0.5 while standard deviation is 0-1); again, with different scales, it's hard to compare, and the overall impression from just looking at the hues is that the AOT standard deviation is small where from the bottom row we can tell this is actually often not the case.

References

Hsu, N. C., M.-J. Jeong, C. Bettenhausen, A. M. Sayer, R. Hansell, C. S. Seftor, J. Huang, and S.-C. Tsay (2013), Enhanced Deep Blue aerosol retrieval algorithm: The second generation, J. Geophys. Res. Atmos., 118, 9296–9315, doi:10.1002/jgrd.50712

Levy, R. C., L. A. Remer, S. Mattoo, E. F. Vermote, and Y. J. Kaufman (2007), Second-generation operational algorithm: Retrieval of aerosol properties over land from inversion of Moderate Resolution Imaging Spectroradiometer spectral reflectance, J. Geophys. Res., 112, D13211, doi:10.1029/2006JD007811

Levy, R. C., Remer, L. A., Kleidman, R. G., Mattoo, S., Ichoku, C., Kahn, R., and Eck, T. F. (2010), Global evaluation of the Collection 5 MODIS dark-target aerosol products over land, Atmos. Chem. Phys., 10, 10399-10420, doi:10.5194/acp-10-10399-2010

Levy, R. C., Mattoo, S., Munchak, L. A., Remer, L. A., Sayer, A. M., Patadia, F., and

Hsu, N. C. (2013), The Collection 6 MODIS aerosol products over land and ocean, Atmos. Meas. Tech., 6, 2989-3034, doi:10.5194/amt-6-2989-2013

Sayer, A. M., N. C. Hsu, C. Bettenhausen, Z. Ahmad, B. N. Holben, A. Smirnov, G. E. Thomas, and J. Zhang (2012), SeaWiFS Ocean Aerosol Retrieval (SOAR): Algorithm, validation, and comparison with other data sets, J. Geophys. Res., 117, D03206, doi:10.1029/2011JD016599

Sayer, A. M., N. C. Hsu, C. Bettenhausen, and M.-J. Jeong (2013), Validation and uncertainty estimates for MODIS Collection 6 "Deep Blue" aerosol data, J. Geophys. Res. Atmos., 118, 7864–7872, doi:10.1002/jgrd.50600

Smirnov, A., B. N. Holben, T. F. Eck, I. Slutsker, B. Chatenet, and R. T. Pinker (2002), Diurnal variability of aerosol optical depth observed at AERONET (Aerosol Robotic Network) sites, Geophys. Res. Lett., 29(23), 2115, doi:10.1029/2002GL016305

Interactive comment on Atmos. Chem. Phys. Discuss., 15, 26191, 2015.