Atmos. Meas. Tech. Discuss., 7, C1107–C1118, 2014
www.atmos-meas-tech-discuss.net/7/C1107/2014/

Atmospheric Measurement Techniques Discussions
Open Access

AMTD
7, C1107–C1118, 2014

Interactive Comment

Interactive comment on "Application of spectral analysis techniques to the intercomparison of aerosol data – Part 4: Combined maximum covariance analysis to bridge the gap between multi-sensor satellite retrievals and ground-based measurements" *by* J. Li et al.

J. Li et al.

jl2862@columbia.edu

Received and published: 29 May 2014

We thank Dr. Andrew Sayer for his helpful and insightful comments and suggestions. Below we respond to the comments point-by-point. The manuscript was also revised accordingly.

This study builds upon previous work by the authors using principal component analy-

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

sis (PCA) and related techniques to analyse spatiotemporal variations in aerosol optical depth (AOD) from a variety of satellite sensors (here, MODIS, MISR, SeaWiFS, and OMI). Ground-based AERONET data are also incorporated. The main thrust of these papers is that by decomposing datasets (singly or jointly) in terms of principle components/singular vectors the spatial variability, inter-sensor differences, and trends can be analysed in a new way. This work is topically relevant to AMT and, along with their prior studies, is a novel (to my knowledge) way to look at this satellite AOD data. The main development from their prior studies seems to be the presentation of the Combined Maximum Covariance Analysis (CMCA) technique, to incorporate the AERONET AOD into the analysis. After the CMCA technique is applied, the decomposed results are analysed, and some comments made about the representivity of monthly mean AOD fields from the various satellite datasets.

I have some specific comments (below), but following these (mostly minor) revisions and clarifications which the authors should address, I favour publication of this paper in AMT. Several of my comments encourage the authors to emphasise more strongly the different mechanisms which may lead to apparent disagreements between sensor monthly AOD maps (retrieval biases, and observability issues related to swath width/cloud screening). The overall clarity, quality of language, and quality of presentation are fine (aside from a few figures which I feel should be redrawn). The authors are welcome to contact me in case of any queries about my comments. Andrew Sayer andrew.sayer@nasa.gov

Thank you for the general comments and the positive feedbacks. We answer the specific comments below.

Page 3505, line 21: I suggest rephrasing to avoid starting a sentence with ‘And’.

This sentence has been rephrased, by combining it with the previous rephrases into one sentence.

Page 3508, line 14: rather than ‘Ångström Relationship’ I think it would be better to say

AMTD
7, C1107–C1118, 2014

Ångström power law or something similar.

"Ångström Relationship" is changed to "Ångström power law".

Page 3508, line 16: I don't know that I would say MODIS was 'designed' to retrieve aerosol properties, given that's but one of the applications the sensor has been put to, and it is certainly far from an ideal instrument to retrieve many of the quantities we want to know about aerosols to the level of uncertainty we wish to know them at. Still, I appreciate what the authors are trying to say here, and am perhaps being a bit picky. I would suggest rewording to 'has been used to retrieve aerosol properties over land and ocean' or similar.

We changed "designed" to "has the capability of".

Page 3508, lines 20-21: Just FYI, in Collection 6, calibration improvements made by the MODIS Characterization Support Team (MCST) mean that Terra and Aqua should both be temporally complete for Deep Blue. But for Collection 5 data, the authors are correct that Aqua gives the more complete record.

Thank you for the information. It would be great to see the Terra DB product when Collection 6 becomes available.

Page 3509, lines 3-4: Again FYI, by the time that peer-review for this paper has been completed, the MODIS Aqua Collection 6 level 3 products should be available and so the authors should be able to use the 'merged' (and/or standalone DB/DT) dataset(s) as they may choose. However as this paper is in part a proof of concept of this new analysis type, I would not say that it is necessary for publication that the authors repeat the analysis using C6 (although I would encourage it, should the authors have the capability, as C6 will hopefully become one of the 'standard' datasets for the coming years).

Thanks again for the information. It appears that Level 3 C6 data are not available yet. But we are very interested in performing a similar analysis with C6 data in the near

AMTD
7, C1107–C1118, 2014

future.

Page 3509, lines 22-23: If I have understood this sentence correctly, the authors are talking about the narrow MISR swath meaning the revisit time may be too long to resolve small-scale temporal variability? If so, I would explicitly say 'lower spatial sampling' or similar rather than just 'lower sampling'.

We are sorry for the confusion. We did mean that MISR swath is narrower and it thus has a longer re-visiting time. This sentence has been revised to "MISR has a narrower swath width leading to a longer re-visiting time".

Page 3510, lines 2-3: As the SeaWiFS mission ended in 2010, I think it would be better to use past tense (i.e. 'covered' rather than 'covers') here.

Corrected.

Page 3510, line 19: I suggest 'agreement' rather than 'agreements'.

Corrected.

Page 3510, line 22: I suggest rephrasing to avoid starting a sentence with 'And'. 'Additionally,' may be suitable here.

"And" has been changed to "Additionally".

Page 3511, line 14: The Smirnov et al (2002) reference cited here is missing from the bibliography, and I think it should be Smirnov et al, RSE (2000) instead for cloud screening. The Level 2 quality assurance document is a different one, which is not cited here. Both should probably be cited: Holben, B.N., T.F. Eck, I. Slutsker, A. Smirnov, A. Sinyuk, J.S. Schafer, D.M. Giles, and O. Dubovik (2006). Aeronet's Version 2.0 quality assurance criteria Proc. SPIE, 6408(64080Q), DOI:10.1117/12.70652 Smirnov, A., Holben, B. N., Eck, T. F., Dubovik, O., and Slutsker, I.: Cloud-screening and quality control algorithms for the AERONET database, Remote Sens. Environ., 73, 337–349, 2000

AMTD

7, C1107–C1118, 2014

The citation of Smirnov et al. (2000) has been corrected and added to the reference list. Hoblen et al. (2006) has also been cited in the text and added to the reference list.

Page 3511, lines 26-27: It would be cleaner to write 'Gandhi College' instead of 'Gandhi_College'; AERONET site names sometimes have these extra characters so the filenames don't contain spaces, but that isn't necessary when writing text. Also, I think saying 'Singapore in Singapore City' is redundant. I'd change the end of this sentence to 'Mukdahan in Thailand, and Singapore.'.

Thank you for pointing out these mistakes. Most "_" characters in AERONET site names have been removed, although a few are retained because they do not seem to be the name of a place (e.g., IER_Cinzana, IMS-METU-ERDEMLI). "Singapore City" has been removed.

Page 3511, line 27: One of the sites listed in Table 1 is Izana, which is at around 2.4 km altitude in the Canary Islands. I think that this site is probably not useful for the analysis presented in this paper because of this high altitude, meaning that AERONET won't be sampling a lot of the total column AOD (as it will be physically located below the site). From the AERONET webpage (http://aeronet.gsfc.nasa.gov/new_web/photo_db/Izana.html): 'The observatory is situated on the top of a mountain plateau located in a pre-national park area protected by the sky law. IZO is normally above a temperature inversion layer, and so, free from local anthropogenic source influences.' So using this as a point of comparison against satellite data is not particularly representative; certainly a 2.4 km mountain can't be considered a good representation of this region (island chain) on a 1x1 degree spatial scale! The site of Santa Cruz Tenerife on the same island may be suitable, as it is low-lying and seems to have a good data volume (http://aeronet.gsfc.nasa.gov/new_web/photo_db/Santa_Cruz_Tenerife.html). If you compare typical AOD from Tenerife and Izana you can notice quite large differences.

Thank you very much for this helpful information. We replaced Izana with

Santa_Cruz_Tenerife in the analysis and updated the global analysis results. This change only has minimal effect on the results.

Page 3512, lines 18-19: I am not sure it is correct to say ‘the interpolation performs well without introducing much uncertainty’, as this isn’t something which is really quantified. The gap-filled time series at Minsk looks plausible but that could just be the case because what’s being used to fill in the gaps is basically a perturbation to the average seasonal cycle for the missing months. Again, I understand what the authors are trying to say here, but I don’t think this assertion is supported. I think they can say that the interpolation seems reasonable, for sure. Maybe they could do a test by manually deleting and interpolating some months from the AERONET time series, and quantifying the error resulting from the interpolation there. Of course it won’t be possible to know the uncertainty when there is a real data gap, but we’d be able to get some idea of typical magnitudes. Since the CMCA seems to be driven a lot by the month-to-month variability, which is in itself dominated at many locations by a seasonal cycle which is fairly repeatable between years, perhaps the reinstatement of the seasonal cycle in this interpolated data means that any errors introduced by the interpolation don’t have much effect on the outcome. Either way, I think that this aspect is something which the authors should quantify and discuss a bit more.

Thanks for this comment and suggestion. We agree that the example in the text does not necessarily indicate how well interpolation performs. We therefore replaced it with an experiment for the IER_Cinzana station, which has full data record and with relatively large variability. For stations with smaller variability, the results are even better. In addition, we performed a cross validation by taking out each data point once and interpolated it, for gaps with one, two and three consecutive missing data, respectively. The results indicate that or one or two missing data cases, the interpolation error (defined as the std between interpolated and original time series) is usually below 3% of the standard deviation of the original time series. However, for three data gap case, the error grows faster and can reach 7% for a few stations. We therefore restrict the gap to

AMTD
7, C1107–C1118, 2014
Interactive Comment

be two points maximum except for the manually added stations to cover representative aerosol source regions.

Page 3512, line 23: I am surprised about the hole in central Australia for MODIS. Deep Blue should be providing a fairly complete coverage there, and the NDVI should be low enough that the merging process the authors are performing fills in the holes in the DT data with DB there. This has been our experience with the Collection 6 data (see e.g. Figure 19 of Levy et al., AMT, 2013), and based on browse images of monthly Collection 5 data (available online at http://modisatmos.gsfc.nasa.gov/MYD08_M3/browse_c51.html), fairly complete Deep Blue coverage here seems to be the case for Collection 5 as well. The authors should check into this.

We checked this problem for Central Australia and found that this is due to the negative values in the MODIS DT retrievals. Specially, during the June to August months, the NDVI is between 0.2 and 0.3 for Central Australia and we thus average the DT and DB AOD. However, the DT AOD for this region is consistently negative while the DB AODs are quite small. As a result, the monthly mean average for some cases will become negative. Although negative AODs allowed in MODIS DT retrieval to account for statistical bias, negative values do not exist in any of the other datasets and converting negative AODs using Angstrom power law will lead to imaginary numbers, as a result we removed these grids with negative data.

The explanation of the problem for Australia has been added to the 2nd paragraph of Section 3.1. Page 3513, line 18: 'averaged' not 'avearged'.

Corrected.

Page 3514, line 8: The cross-covariance matrix is the matrix C? This is what I infer but it isn't stated explicitly. Also, shouldn't C be in bold and regular typeface rather than italic, as it is a matric?

Yes. We explicitly mentioned C as the cross covariance matrix in this sentence. C is also changed to bold regular font.

Page 3414, line 11: X_AERONETËĘT should be in bold regular typeface, as it is elsewhere, rather than italic, as it is a matrix.

X_AERONET was changed to bold regular font, so were the other matrix symbols.

Page 3514, line 17: I think it would be better to say 'in descending order of magnitude', to be explicit.

We changed it according to the above suggestion.

Page 3515, line 8: 'we divide V' rather than 'we divide the V matrix'

Corrected.

Page 3516, line 12: I think it would be clearer to say that the first 3 modes 'each' explain >10% of the variance, rather than 'all' explain it. Based on eyeballing Figure 5, it looks like together modes 1-3 explain about 70% of the total variance. If I understand things correctly, this is suggesting that 70% of the global variability of the aerosol field (on monthly time scales and spatial scales of 1 degree) can be explained in terms of three basic components. The authors mention the links to their previous work and say they won't discuss it here, but I think that it is worth mentioning this point again here briefly, as it is in my mind quite an interesting result (3 modes being so dominant).

"all" has been changed to "each" and we explicitly added the discussion that "these three modes together explain $\sim$70% of the total variance.

Section 4, figure 7 general discussion: This is a figure making quite an important point about the agreement between various satellite data products. Now as the authors are looking at monthly mean fields, there are two main things which could cause these disagreements. One question is retrieval biases at high and low AODs, related mainly to aerosol model and surface reflectance assumptions respectively, which will affect the peaks and troughs of seasonal cycles. Another is related to observability, which is in part a function of swath width and aerosol temporal variability but also things like cloud cover and each algorithm's success at being able to perform a retrieval at all. See for example Figure 1 and discussion of Reid et al (Atmos. Res., 2013, doi:10.1016/j.atmosres.2012.06.005) for about this in the context of south-eastern Asia (where availability of retrievals can be a challenge). And then the extent to which observability will be a problem will depend on things like the spatial and temporal scales the data are aggregated at. The authors hint at these topics a little on pages 3517-3518, but I think it would be good to mention these aspects more directly. It is my impression that the fact our validation papers against AERONET tend to dwell on level 2 comparisons (which focus mainly on retrieval-level errors rather than sampling-related representivity) means that observability issues have been given comparatively less attention when they can be important for studies such as this, which are using level 3 data. And the issue of observability is not limited only to the satellite datasets-the AERONET monthly means are (to my best knowledge) the means of daily mean data. So there are questions about how many observations does AERONET need to be representative of the aerosol burden on a given day, and how many days with observations are needed to be representative of a given month, the answers to which are probably dependent on the temporal variability of AOD at each individual site. In other words, AERONET can have this 'sampling uncertainty' on a monthly mean too.

Thank you for this helpful comment. We agree that these two types of uncertainties should be explicitly mentioned in the discussion. And actually the observability issue may be more dominant here, since every dataset has been validated using collocated AERONET data while significant differences in seasonality and events are still found in Level 3. We added a general discussion at the end of Section 4.1 emphasizing these two factors and the important role of observability in Level 3 comparisons.

Page 3518, line 7: I think the authors can just say 'Borneo' rather than 'Borneo Island'.

OK.

Page 3518, lines 9-15: I think it could be worth expanding on this statement a little. Specifically, as SeaWiFS lacks thermal IR channels (which are useful for cloud screening), one of the main methods of cloud detection over land is spatial variability (as clouds tend to be bright and spatially variable). Unfortunately, some aerosol plumes have a similar appearance. So there can be a tradeoff in cloud screening between errors of omission (i.e. missing a real cloud) and errors of commission (i.e. saying something is a cloud when it isn't). Our goal was to err on the side of caution and minimise the likelihood of clouds incorrectly being labelled as cloud-free. Thus, in some areas (including Borneo but also e.g. the Amazon which also often has small clouds and smoke aerosol plumes), coverage is reduced because sometimes these smoke plumes are incorrectly labelled as clouds and so the pixels discarded. And as well as low-level clouds, there can often be a thin cirrus layer above the area which can result in positive biases in the retrievals if not identified. So this is also related to the observability issue I mentioned above. The cirrus can affect AERONET observations too, see e.g. Huang et al. (JGR, 2012, doi:10.1029/2012JD017757), Chew et al. (Atmos. Res., 2013, doi:10.1016/j.atmosenv.2013.06.026), and others. On the topic of future AERONET deployment the authors bring up here, in the last few years 3 AERONET sites have been acquiring data in Borneo (Kuching, Pontianak, Palangkaraya) although unfortunately yes, there was nothing available during the analysis period of this study. See Salinas et al. (Atmos Env, 2013, doi:10.1016/j.atmosenv.2013.02.016) for some first results at Kuching.

Thank you for the information. The discussion on conservative cloud screening has been expanded according to the above comments. We also revised the AERONET deployment part by acknowledging the establishment of several stations in recent years.

Page 3518, line 12: 'biomass' not 'bioass'.

Corrected.

Page 3518, line 16: Here there is a jump between looking at the full dataset to looking at the (deseasonalised) anomaly dataset). I think it would be clearer if this part became a new subsection 4.1. Then section 4.1 could be renamed ‘Global analysis – full dataset’ and the new section 4.2 ‘Global analysis – anomaly dataset’ or similar.

OK. We changed the subtitles according to the above suggestion.

Page 3518, line 16: ‘anomaly’ not ‘anomly’

Corrected.

Page 3520, line 23: ‘Gangetic’ not ‘Gengetic’.

Corrected.

Page 3523, line 20: ‘did not’ rather than ‘do not’. Also, I think it is unfair to say that the event wasn’t captured by SeaWiFS or OMI. The widespread smoke from the fires is evident in the SeaWiFS and OMI records, it just isn’t as prominent as in the others. I think that part of the reason is that the SeaWiFS monthly mean has some gaps in this region (I had a quick look through the source data) as the heaviest smoke was flagged as cloud (see aforementioned cloud-screening difficulties due to lack of IR bands on the sensor) meaning that some of the heaviest smoke events were not present in the SeaWiFS data. This in turn propagates into lower level 3 fields. It is possible that the same was true for OMI, although I’m not familiar enough with the algorithm/dataset to say for sure.

“do not” has been changed to “did not”. We apologize for the inappropriate phrasing. It is true that some smoke signals also show up in Level 3 data for SeaWiFS and OMI, such as Figure 11. This sentence was thus changed to ” SeaWiFS and OMI did not fully represent the intensive Russian wildfire in August 2010. Their signals are weaker compared to MODIS and MISR for this event.”

References: Levy et al. (2013) is now published in final form in AMT (the present citation is to AMTD). So this should be updated to: Levy, R. C., Mattoo, S., Munchak, L. A., Remer, L. A., Sayer, A. M., Patadia, F., and Hsu, N. C.: The Collection 6 MODIS aerosol

products over land and ocean, Atmos. Meas. Tech., 6, 2989-3034, doi:10.5194/amt-6 2989-2013, 2013.

Thank you for the information. The reference is updated.

Table 1: This table should be checked, as some sites are not where the authors are saying they are. For example, IER Cinzana is in Mali, which is in Africa (not South America). Missoula is in the USA (not Europe). IMS METU-Erdemli is in Turkey, which could be classed as Europe or Asia (but not Africa, where the authors currently place it). There may be more errors, that was from a brief look. Similarly I am not sure why an 'others' category is needed – couldn't 'North America' be renamed 'North/Central America' and have these two sites listed there, or else make a new 'Central America' category for Mexico City and La Parguera? See also prior comment about formatting of site names (e.g. can't FORTH_CRETE be Forth Crete?).

We apologize for the mistakes. We checked the location of each station to make sure they are listed in the correct region. Many stations names are also replaced by the name of the place. The "others" category is changed to "Central America".

Figures 11, 13, 14: The presentation of these figures should be improved. The coast-lines are very coarse and basic, and there is a lot of white space around the edges of the regions mapped (it looks like the clipping for the data is smaller than the clipping for the map borders).

The resolution of the coastlines has been improved from 1 degree to 0.25 degree. However, some white space between the data map and coastline still exists and is caused by the coarser resolution of the data than the coastline mask.

Interactive comment on Atmos. Meas. Tech. Discuss., 7, 3503, 2014.