Atmos. Meas. Tech. Discuss., 6, C1571–C1576, 2013
www.atmos-meas-tech-discuss.net/6/C1571/2013/

Atmospheric Measurement Techniques Discussions

AMTD

6, C1571–C1576, 2013

Interactive Comment

Interactive comment on "Validation and empirical correction of MODIS AOT and AE over ocean" *by* N. A. J. Schutgens et al.

N. A. J. Schutgens et al.

schutgens@physics.ox.ac.uk

Received and published: 11 July 2013

Response to Jeff Reid's review.

Based on Jeff's comments and those of the other reviewers, we have now included in the introduction explicit definition of our statistic metrics (e.g. median as a bias) and why we did so. We have also substantially altered the text of those sections that pertain to our methodology.

After reading Jeff's comments and discussing some issues with him offline, we (the authors of the paper) feel that several of the issues Jeff raises are the result of him interpreting our text differently from what we intended. We have therefor undertaken to rewrite those parts with the aim of removing any ambiguity. We would like to thank Jeff for his review that gave us the opportunity to clarify certain important points. Below are our responses in detail.

Q: 1) I do take exception to Nicks statement's in the paper that compared to our papers, this present analysis is "more complete" and "better performed" (I do agree it is an extension). Ultimately, what we have is a difference in point of view as to how our error models should be constructed. What they have done is more complex, but after reading this paper I would not change the way we operate-namely we look at multiple retrievals against a single AERONET measurement. And I certainly would not agree that they performed a more robust analysis-although it does address a few interesting issues. We have personally explained our point of view to Nick that there is merit in this method in that we want to understand the regional variability around a single site. Has is shown in the paper, the correlation length of most aerosol features is well in excess of the 50 km rang ring that they and we use. Thus, variability within that range ring has meaning from a retrieval noise point of view. Error is error, and we will build up or stats any way we can. Now, the authors have a point that we need to be careful that a single site or good clear days do not bias the sample-we agree. But the way we have dealt with this is by adding dimensions to the error model. Over water, the dimensions are fine/coarse mode partition (which this paper does not address and hence cannot be called more compete see comment 4), plus wind, and cloud cover. Over land (contrary to what is stated in this paper), Hyer does in fact perform bias correction based on albedo, view angle and region.

A: We could not find these quotes in the paper. Of course, we have tried to improve upon the excellent work by Zhang & Reid and Shi et al., but an actual comparison of the corrected datasets (theirs and ours) is beyond the scope of the current paper. Note that our paper extends on their results by presenting corroborative evidence of the Maritime Aerosol Network and a corrected AE observation (and AE random error estimate).

We also note the following (and we have changed the paper to stress this better): 1) as

Fig 2 (revised paper) shows, using the full dataset, instead of a smaller independent subsample, yields larger AOT biases at high AOT. Also, we find lower biases at high windspeed (by $\sim$ 0.02). 2) Like Zhang & Reid and Shi et al, we find that the correction algorithm should be different for low and high AOT. The 'breakpoint' in our algorithms is, however, different. They use 0.2 (), we use $\sim$0.05. Our value has been objectively estimated (i.e. through minimising a specific error metric, as detailed in the paper). We found poor correction algorithms for our dataset if we put the breakpoint at 0.2. 3) We seem to find similar random errors (fig 20, revised paper), but apparently (or so we understand from Edward Hyer's comments), our approach much reduces any skew that others have seen in these data (Fig. 17, revised paper). This allows us to interpret the error in a standard (Gaussian) way.

Instead of using fine mode AOT (or fraction) as a predictor of random error, we have used AE (Angstrom exponent). The reasons for our choice of AE are explained in the paper.

Finally, we differ with Jeff in our interpretation of the spatial correlation of (MODIS observed) aerosol. Rather than allowing the study of retrieval noise, this noise is itself correlated over long distances (Fig 1, revised paper). Including all possible MODIS observations co-located with a single AERONET observation adds no extra information. Sure, one increases the sample size substantially, but it will not be an independent sample.

Q: 2) The statistics presented (and in particular the plotting axis) are at times poorly defined and unclear. There are 23 plots and little synthesis. The presentation as a whole in fact is unclear, with statements like "we can make a correction" and then we are referenced to the appendix. Our corrections are simple and on par with the real uncertainty of the system. I am curious how they came to their conclusions think these require more explanation. In their appendices they should define all variables. In many of the plots, "MODIS error" is listed. But really what I think they mean is mean bias. But is unclear what it is they are really presenting. The authors may want to

AMTD

6, C1571–C1576, 2013

Interactive Comment

go through these. Second, how they construct error estimates appears to be mean absolute difference, whereas most folks utilize root mean square error. Also in this paper, they look at the error distribution and the calculation of a median, which is fine if you want to map the distribution of error. But for our application, data assimilation, we need to use MAE and or the RMSE (or if the bias is known, the RMSD). Median is the most likely value, but the mean is the most representative. What is really mucked up here is that the standard deviation of their histograms is the RMSD, so why not just provide it? Perhaps the mean bias and an RMSD can be provided in a table across several dimiesions?

A: Jeff is right that we did not properly define our error metrics (bias & random error), this is now corrected in the introduction. In particular: - MODIS error refers to the distribution of MODIS-AERONET AOT values, or one particular value - bias refers to the median of this distribution - random error refers to half the interquantile range 15.8 to 84.2%. of the standard deviation of this distribution We also explain why we use quantiles (to reduce the impact of outliers on what is almost but not really Gaussian error statistics).

In the axes labels of two Fig. (10 & 14, revised paper), the term 'error' is used while bias was meant (the captions were correct) . This has been corrected. (In the majority of the plots we do show error distributions, in agreement with our definition of the term 'error').

Our random error definition (based on the interquantile range) is not the mean absolute difference but very similar to RMSE or RMSD (as far as we know, RMSE and RMSD are identical), provided RMSE (or RMSD) is calculated for an unbiased distribution. In other words, our random error is independent of any bias (due to the use of the interquantile distance). Note that our random error agrees with the standard deviation for a Gaussian distribution (whether its mean is zero or not).

The median is, by the way, not the most likely value of a distribution (that would be

AMTD

6, C1571–C1576, 2013

the mode). The median divides the distribution in two equally sized samples, one with values below the median, the other with values above the median. In the case of near-Gaussian errors with outliers we use it as a robust estimate of the bias.

Q: While diagnostic error is useful, what we really need is prognostic RMSE. This is an important point and one that should not be overlooked- everything we have done has been in the context of data assimilation. There is no way that we can take the error models presented in this paper and apply them to our data assimilation problem. What would be wonderful is if the authors could calculate RMSE as a function of MODIS AOT and compare that to the numbers in Yingxi Shi's paper. Here we have to use RMSE (over RMSD) because after bias correction, we don't know what the bias is! If we did, we would correct for it-hence in the face of the unknown it has to be RMSE. I really want to know if by their sampling if they get a different number. We can also provide the authors with the mean AOT data from our DA grade product, and they can demonstrate a difference.

A: We agree that, in principle, the prognostic RMSE (or a similar metric) is what matters most. However, we found it was easier to derive a fitted function for diagnostic error (i.e. as function of AERONET AOT) than prognostic error (as function of MODIS AOT). The reason is presumably that in the latter case, the independent variable AOT has a large random variation (due to the MODIS AOT error of course). On the other hand, once a diagnostic error has been determined, the corrected MODIS AOT can be used instead of AERONET AOT to yield prognostic error. Note that our random error estimates are very similar (Fig 20).

Q: 4) Going through the paper, it strikes me there are some subtleties on remote sensing retrievals that might be missed by the authors. First and foremost, much of the bias we found was microphysical, and that could be corrected by application of a correction term based on the retrievals own fine mode fraction. Thus, when you show global statistics, and have a mean bias near zero, much of this is a result of offsetting penalties from fine and coarse mode dominated aerosol airmasses (again, this analysis is not more complete than what we did). Second, choosing as an example AOTs for above 0.5 is not representative of the global marine atmosphere. The authors may think they are doing the retrieval a favor by using higher AOTs and thus better signal. But for AOTs>0.5 one enters a multiple scattering regime, hence errors multiply. For comparison against AERONET and why some sites “don’t work out” it is both error on the MODIS side plus non-representativeness on the AERONET side. For example, sites like Coconut Island are impacted by upslope winds near the island of Oahu, and hence higher boundary layers and higher AOTs than just a few kilometers out to sea. It is for this reason in part that the site was later pulled. You mayy want to talk to the AERONET guys abotu yoru findings. Soem of yoru flaggged sites looked ok to us too as long as we dont use coastal retreivals.

A: We agree with Jeff that intensive aerosol properties (size or species) are partly to blame for the observed biases. We pointed this out in the paper and it is of course the reason for using AE in the correction of AOT bias. We have presented our reasons for using AE instead of fine mode fraction in the paper.

Secondly, our example of how sub-sampling affects biases for AOT>0.5 is only for instructional purposes. Nowhere in the actual validation/correction have we limited ourselves to AOT>0.5

We agree with the reviewer that correlation of MODIS and AERONET may be bad due to several reasons. The purpose of selecting only cases with relatively high correlation (> 0.5) was to retain only those co-located observations pairs that made ’sense’. If for instance the MODIS - AERONET AOT difference is entirely due to noise, there is no purpose in using that site to estimate biases. We have studied the impact of our threshold (0.5) by increasing it to 0.75. This removes another 20% of all MODIS-AERONET data pairs but has negligible impact on the resulting correction.

Interactive comment on Atmos. Meas. Tech. Discuss., 6, 3765, 2013.