

We would like to thank Alexei Lyapustin and the two anonymous referees for their comments on our manuscript. We are happy that the referees feel positive about our work, and have revised the manuscript to accommodate their suggestions, as listed below. Our comments are in italic type.

Review 1:

This paper describes a validation analysis of the recently released over-land SeaWiFS aerosol retrieval dataset. The manuscript is very well written, although could be a little shorter to accommodate an average reader's attention span, like mine. It is a very detailed and comprehensive study covering different validation aspects and providing comparison to both AERONET data and MODIS and MISR datasets. I recommend this paper for publication. I also have several recommendations to slightly improve this paper:

- 1) A brief mention of SeaWiFS original resolution and aggregated resolution during aerosol retrieval would be helpful.
- 2) A definition of QAs=1-3 with short description of differences among different QA levels would benefit the paper, as well as the user community.

To address these points (and those of Reviewer 2), we have rewritten the start of Section 2. We mention more explicitly the previous Deep Blue papers (Hsu et al., 2004, 2006), where the general Deep Blue methodology is described, and also outline some SeaWiFS-specific information and general points. In the interests of brevity and readability we did not wish to include an extended algorithm section in this paper, as we felt that for most users the previously-cited papers would provide sufficient information. We do include a brief summary of the differences. However, our group are presently working on a paper to describe this 'second generation' Deep Blue algorithm as applied to SeaWiFS and the forthcoming MODIS Collection 6 which will address this in more detail.

- 3) p. 2176: Definition, or explanation of the expected error (EE) as given is incomprehensible and should be simplified, e.g. as "1 sigma based on AERONET as truth".

We have done this in the revised manuscript.

Review 2:

The manuscript presents a comprehensive evaluation and validation of AOD derived from SeaWifs land retrievals. The paper is well written; the approach is clear and well described. The possible reasons for retrieval biases (e.g., surface, viewing geometry) are investigated. The authors also investigate effects of QA, spatial averaging, dependence on AOD, and temporal dependence to provide users important information on the range of applicability of SeaWifs data on both regional and global scales. Intercomparison with other datasets is also shown. Error estimates are well defined but probably should be moved into the description of the methodology. I highly recommend the paper for publication in AMT.

General comments on a scientific approach:

1. In my opinion paper needs to add some detail regarding the algorithm description. It was not particularly clear how "Deep Blue" was adapted to SeaWifs considering differences in MODIS and SeaWifs sampling and resolution.
2. How was the "Deep Blue" algorithm used to retrieve AOD at 550 and 670 nm?
3. What surface database was used?

To address these points (and those of Reviewer 1), we have rewritten the start of Section 2. We mention more explicitly the previous Deep Blue papers (Hsu et al., 2004, 2006), where the general Deep Blue methodology is described, and also outline some SeaWiFS-specific information and general points. In the interests of brevity and readability we did not wish to include an extended algorithm section in this paper, as we felt that for most users the previously-cited papers would provide sufficient information. We do include a brief summary of the differences. However, our group are presently working on a paper to describe this 'second generation' Deep Blue algorithm as applied to SeaWiFS and the forthcoming MODIS Collection 6 which will address this in more detail.

In specific reference to the reviewer's point 1: the algorithm actually required very little adaptation to work on SeaWiFS, as the relevant bands used for AOD retrieval are close. The level 1 spatial resolutions are different (1 km for MODIS, 4.5 km for SeaWiFS at the sub-satellite point), but the same principle is applied to both (retrieve at full resolution, aggregate to coarser cell resolution and use the cell statistics to define quality flags).

Review 3:

The paper is very well written. The paper provides massive amounts of data in large no of figures that are very clearly and logically presented. This is a great accomplishment. Yet, I am sorry to say that at the end all the tables, figures, and text manage to convey precious little information to the reader. The paper reads more like a comprehensive, well written project report rather than a scientific paper. Do we need all these figures to learn what the statistics given in Table 2 and 3 adequately capture? Wouldn't it be enough to provide just few illustrative examples. Rest of the figures and the data could be made available through a website to the few people who maybe interested in such detail.

As an aside to the above, we do have many additional figures which we can provide to interested parties. Related to the reviewer's comment, we have made textual changes to hopefully make the manuscript more easily informative to a data user, and removed some figures (see response to previous reviewers, and responses to comments below).

Is there a need for so much discussion of AE when it is clear that the quantity is not well measured? This is all the more so since AE is of rather marginal utility- useful just to separate small and large particle aerosols except even when it is well measured.

We agree with the reviewer's comments about AE utility, however, wanted to be sure to explicitly show these results as it is a quantity which users are frequently interested in. We have tried to emphasise the issues with AE, particularly for low-AOD conditions. We do not feel that the length of AE discussion in the manuscript is excessive.

Fig 8 is a good example of a plot that conveys so little information that it could be captured in a single sentence. The purpose of a figure should be do just the opposite- to provide information that will otherwise take lots of words to convey.

We agree with the reviewer on this point, and have removed both Figure 8 and Figure 6 (which we also feel could be dealt with as effectively in text).

From my perspective the two most interesting figures of the paper are figs 10 and 11. I would have liked to see more discussion of these figures with some attempt to sort out which results the reader should consider more reliable.

We have extended the discussion in Sections 4 and 5 slightly, to expand on this point.

What a reader would have liked to know if the inter-annual variability of the mean AOD in different geographical regions are similar to that derived from other techniques. But it appears that this information has been deliberately left out of this paper.

We decided not to talk about interannual variability much in this paper; we have also submitted a paper looking at trends in AOD from SeaWiFS, and comparing these with other satellites and AERONET, which touches on interannual variability a little. This paper (Hsu, N. C., Gautam, R., Sayer, A. M., Bettenhausen, C., Li, C., Jeong, M. J., Tsay, S.-C., and Holben, B. N.: Global and regional trends of aerosol optical depth over land and ocean using SeaWiFS measurements from 1997 to 2010, Atmos. Chem. Phys. Discuss., 12, 8465-8501, doi:10.5194/acpd-12-8465-2012, 2012.) is now cited in our revised manuscript to address this point. The aforementioned paper also discusses trends over ocean. We are planning further studies to investigate interannual variability in different regions from SeaWiFS and other datasets (one such study is presently in review), and because of these factors and to not increase the length of the manuscript too much, have not included time series in this paper. If the reviewer is interested, we would be happy to work with them on this topic.

Figs 12-20 are simply massive data dump. It is not clear how useful are correlation numbers when the data are so heavily clustered around one corner. It would have been lot better if the data would have binned and mean difference and std err of the mean difference in each bin would have been provided instead.

We think this is an excellent idea, and have replaced Figs 12-20 with one new figure based on this suggestion.

My basic recommendation is that the paper should be revised to increase the ratio of information to data. I will leave it upon the authors to decide whether they want to provide more information or less data to increase the ratio. If they want to do the former I will suggest focussing on the variability of the mean- in space and time- rather than on the variability of raw data around the mean.

We feel we have achieved this in the revision, by adding to the text in some areas, and simplifying/removing some figures which were less useful.