Interactive comment on "An over-land aerosol optical depth data set for data assimilation by filtering, correction, and aggregation of MODIS Collection 5 optical depth retrievals" by E. J. Hyer et al.

Anonymous Referee #1

Received and published: 11 November 2010

1. This is a very ambitious paper that develops corrections to the MODIS AOD product relative to AERONET over land, with the aim of providing better correlation with the sun photometer observations. This approach offers a good, practical way to produce an "assimilation quality" global satellite AOD product. This study is very carefully done, and demonstrates good understanding of the products used. It represents an important contribution to the literature.

Thank you for the kind words.

2. You might comment on how this approach compares with "neural net" methods (e.g., Radosavljevic et al., IEEE Geosci. Remt. Sens. Lett., 2010), which match the MODIS radiances directly to AERONET, completely avoiding any physical basis for the retrieval, but achieving very high correlation with the "ground truth" data.

We have added a discussion in Section 5.3 comparing our approach to developing AOD corrections with neural net approaches.

3. The title might be: ". . . assimilation by filtering, correcting, and aggregating MODIS Collection 5. . . "

We appreciate the suggestion, but do not find it useful or necessary to change the title of the paper at this stage.

4. Introduction, paragraph 5, lines 2-3. I think you could make a case that MODIS represents the state-of-the-art for global aerosol optical depth, but I don't think that would be true for aerosol properties (which usually mean intensive properties). Same comment for Section 2.1, line 2. Over land, MODIS assumes aerosol properties, based on an AERONET-derived seasonal/regional climatology.

We have clarified the paragraph to specifically refer to MODIS AOD. In the time since this project was begun, a tremendous amount of research has been done on aerosol retrievals from many different sensors, but in terms of over-land retrieval of AOD, the MODIS c5 product remains the most robust and well-characterized algorithm for global applications.

5. Figure A1. I think I know what it is, but you might mention what the blue lines represent. Also, what actual quantity is plotted on the Compliance scale. For Figure A1(c), you might mention that when AOD_A is very high, there is likely a

plume over the AERONET site, in which case the larger MODIS sampling region would typically yield a lower spatial-average value.

Caption to Figure A1 has been amended as requested.

6. Section 2.4, equation 1. Actually, Abdou et al. use an envelope of *the larger of* 0.05 or 0.2 * AOD. The MODIS team uses the sum of an absolute and a relative criterion for their envelopes.

Section 2.4 was amended to correct our misapprehension of Abdou's method.

7. Figure 1b. I'm not sure if you allow for the negative AOD values the MODIS Collection 5 data contain over land. I can understand why you might truncate the data at zero, but you might mention explicitly how you handle this. (Same question for Table 1c.) Also, if the black dashed lines represent the 25th and 75th percentile values of the actual data within each AERONET AOD bin, I'm not clear why they go to zero near the origin.

Our treatment of negative AODs from MODIS is: we use them in our L2 analysis. In our L3 product, if the mean AOD after aggregation is negative, it is truncated to 0. In Figure 1b, both X and Y axes actually extend to -0.05. You can see the gap on the left-hand side of the contours, because AERONET data do not have negative values. Section 2.1 and the description of Figure 1(b) have been modified to include this information. As for the percentiles, they do not actually go to zero, but they decline uniformly down to AERONET AOD about 0.05, at which point the 25th percentile of MODIS AOD is negative, and the curve changes shape. See attached a blow-up of the bottom left corner of Figure 1(b) that illustrates this behavior.
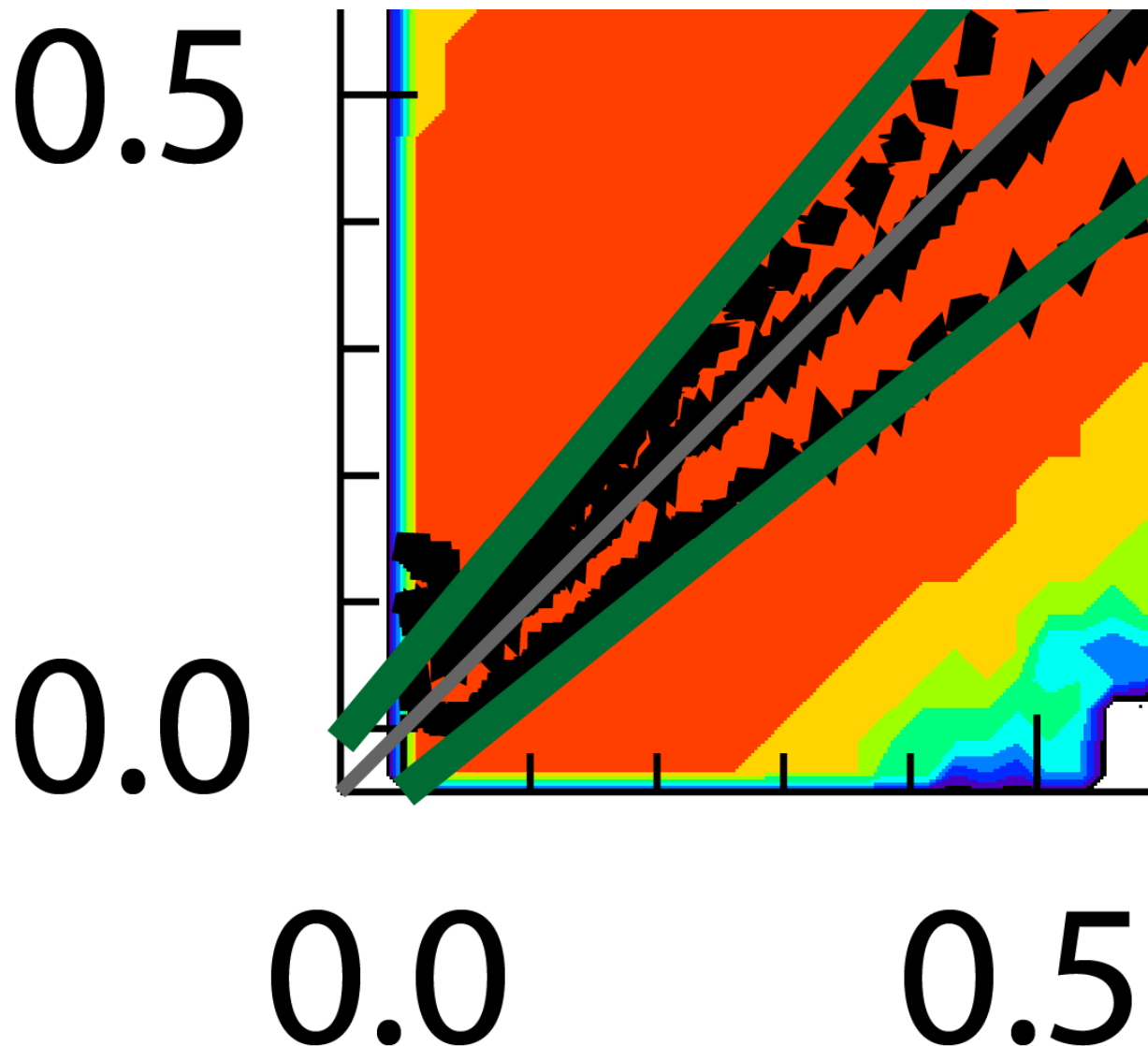
8. Section 3.0, last paragraph. Levy et al., ACP 2010 updates some of the references given here, specifically for the over-land AOD retrievals. Also, next-to-last sentence, might add: ". . . should be used for scientific analysis, and the results here corroborate that these data provide systematically better agreement with AERONET."

Added clause as requested.

9. Section 3.1, paragraph 3, line 1. "through" not "though" Very nice work here.

Typo corrected; again, thank you for the compliment.

10. Section 3.1, last paragraph, sentence 3. There are also correlations between

scattering angle, latitude, and view angle in the MODIS observations that might be of relevance here.

Non-stationarity of scattering angle with latitude is now noted in the discussion.

11. Section 4.1, first paragraph end. Just for interest, can you tell if the higher-slope values for Aqua at high-AOD sites are due to more intense wildfires in the afternoon vs. the morning?

More intense afternoon burning, different distributions of forest vs non-forest fuels, and varying optical properties depending on near-source aging of smoke: each or all of these might contribute to the differences between Terra and Aqua in high-AOD cases in South America. A more fine-grained analysis like that of (Eck et al. 2009) might have a better chance at separating these effects, which is beyond the capability of the methodology used in our study. This is an interesting question, and we hope that with application of more detailed data, a satisfactory explanation can be developed.

12. Section 5.1, equation 7. This is clearly a good approach. I'm a little surprised you use the same three fitting parameters for the global range of surface types. Perhaps this is be more detail than is needed for good assimilations, and I do realize that your thresholds remove the extremes. FYI, I found it difficult to keep track of all the dimensions of information encoded in Figure 5a.

Early on when we began examining the MODIS product, we examined the variation in error characteristics between land cover types. While clear differences were found, the correspondence between vegetation type from the MODIS land cover map and the AOD errors was limited in its quantitative skill. When we altered our approach to use albedo data instead, it was clear that we were getting much closer to the actual source of error in the MODIS AOD. While there is almost undoubtedly an additional term of error in MODIS AOD associated with vegetation structure (BRDF, but more importantly shadowing), the term is of lesser importance compared to the primary error in surface reflectance characterization, which we capture effectively with the 3-parameter model.

13. Section 5.1, equation 7. Also, I guess there is no need to include view-angle dependence – I can see that Figure 4 suggests minimal view-angle bias in tau_M, but I'm wondering if the application of equation 7 for all cases could introduce one?

No. The view-angle and scattering-angle biases, as shown in Figure 4, are very likely real, but too variable to usefully correct for. The application of the albedo correction does not increase these biases.

14. Section 5.2, first paragraph. You might mention here that over land, the MODIS algorithm essentially assumes particle properties based on an AERONET climatology.

As such, there might also be some interesting aspects of the way this climatology is derived that produce the larger differences in the Sahel and S. America.

We added a mention of the MODIS treatment of aerosol properties and a citation to (Levy et al. 2007) in this section.

15. Figures 9, 10, and 12. The scale bars are difficult to discern. For Figure 10, I'm wondering whether it might be more useful to show in the right column the number of days, on the same scale as is used for the left column, rather than showing differences. In the current version, I don't see any structure in the rightcolumn plots, and it is difficult to make comparisons between the left-column and corresponding right-column plots. For Figure 11, again it is difficult to see more in the multiple plots than just the seasonal variation in MODIS coverage, e.g., at high northern latitudes. Perhaps showing just a few plots, for selected regions and/or seasons, but at higher resolution, would convey more of the key differences.

These figures have been modified. The figures in the paper will now show only JJA 2008, and the original figures with all seasons will be moved to supplementary material. I compared and contrasted the two ways of showing data coverage for Figure 10, as you suggested, but found the ratio method more effective.

16. Section 6.2, "Quality assurance filtering" paragraph, last sentence. For clarity: ". . . and positive errors dominate for the excluded data."

Corrected.

17. Section 6.2, "Exclusion of partially cloud retrievals" subsection, last sentence. This is interesting. I'm not surprised that it preferentially removes high-biased cases, but if the filter does not change the overall compliant fraction, does this mean that it removes proportionately equal fractions of initially compliant and non-compliant cases? If so, and if the identification of partly cloudy cases is working well, is the AERONET data cloud-contaminated to roughly the same degree as the MODIS data for these events?

Yes, the excluded partially cloudy retrievals have similar compliance to the remaining data set. But these retrievals are still worth removing, even though the overall compliance does not improve, because when left in, they conceal compensating negative errors. Basically, the MODIS algorithm is tuned in such a way that the positive biases associated with incomplete cloud clearing are balanced by negative biases over dark surfaces (e.g. negative AODs in South America). When the partially cloudy retrievals are removed, the negatively-biased retrievals become easier to diagnose and correct.

18. Section 6.2 overall. You might mention the rationale for performing the filters in

the particular sequence chosen. I can see that the order might not matter for some tests, but it might make a difference for others, such as the "buddy check'.

A brief description of the order of filtering has been added to Section 6.2. The only meaningful choice in the ordering of filters is whether to apply textural filtering before or after QA filters. In keeping with our general conservative approach, we applied the textural filters last, after all of the QA, cloud, albedo, and other filters. This choice could arguably impose a cost in data volume without a corresponding benefit in product quality, but it is a simply expressed decision. The textural filters serve as both a QA check because of the known relationship between texture of reflectances and residual cloud bias, and they also serve, along with the data volume constraints we impose, to assure the representativeness of our L3 AOD values.

19. Section 7(a), line 4. Perhaps: ". . . a prognostic RMS error model with a noise floor is more appropriate."

Added.

20. Section 7(g). See Point 14 above. Perhaps there is a bit more worth saying here.

All signs point to the regional biases we identified having their origins in the tables of aerosol optical properties used by the c5 retrieval. But the evidence from this study neither proves that point nor suggests a clear course of remediation, beyond "get better tables of optical properties." So we will not extend our conclusions beyond the evidence from our study. In response to the comments of Reviewer #2, we did add a reference in Section 7(e) to a paper in preparation that uses the MODIS and MISR retrieved AOD to identify regional/seasonal differences that need detailed *in situ* observation to resolve.

21. Section 7(h). If Basic QA reduces the data volume by 50% to achieve a 50% improvement in compliance, whereas the additional 10% filtering produces an additional 30% improvement, I'm wondering what would happen if you imposed the filtering step producing the greatest proportional compliance improvement first.

The MODIS Mandatory QA flags are a blunt instrument; however, they help. We debated at great length whether to use "Good" (QA=2) retrievals in our analysis, since they are statistically not much worse than "Very Good" (QA=3) retrievals. In the end, however, we decided to maintain consistency with our general philosophy of "a few good measurements," and adopt the recommendation of the MODIS Science Team. It is not far-fetched to imagine a different approach to the QA Flags that preserves more data, but there will always be some cost in data quality.

22. Section 7(i). Would it be appropriate to reference here Zhang and Reid (ACPD

2010) about MODIS calibration?

Modified this point to include citation of that recent paper, which is quite relevant to that discussion.

23. Section 7, very last paragraph. Whether the QA procedures reduce biases and random error in themselves seems less important than that the final product meet specific compliance criteria for assimilation, i.e., for this application, it seems that the question is not whether the data are "better," but whether they are "good enough."

Good point. Changed text to reflect desirable properties of a product for data assimilation: "low bias and low random error."

Anonymous Referee #2
Received and published: 21 December 2010
The manuscript entitled "An over-land aerosol optical depth data set for data assimilation
by filtering, correction, and aggregation of MODIS Collection 5 optical depth retrievals"
by E.J. Hyer et al. presents a comprehensive quantitative evaluation of MODIS Collection 5 over-land aerosol product using AERONET observations considering the influence of MODIS detected clouds, snow cover, MODIS viewing geometry, as well as surface boundary conditions and assumed microphysical models. The paper extensively analyzes MODIS regional biases with thorough explanations and quantifications of regional bias variability. This work establishes important criteria for determining satellite aerosol data quality that is needed for many transport model applications and climate studies such as model data assimilation, model regional aerosol behavior constraints as well as global aerosol climatology analysis. The paper is of very high quality and certainty appropriate for publication in Atmos. Meas. Tech.
General comment:
The paper expands Zhang and Raid, 2006 MODIS data assessment over the ocean to more complex land retrievals. As the authors noted, land surfaces in general have strong spatial variability, and observations near aerosol sources have both temporal and spatial variability that may not be fully captured by point observations at AERONET stations.
As a minor revision, the authors should comment on the importance of field experiments and other surface observations to constrain the physics of aerosol processes near sources that are also needed to evaluate and constrain satellite retrieval algorithms.

Thank you for this very generous review.
We agree that more detailed measurements are necessary to resolve some of the behaviors of the MODIS AOD retrieval identified in this study. In the revised manuscript, we make this point clearly in Section 7(e), and refer to work in preparation by a colleague that uses the MODIS and MISR retrieved AODs to identify locations and seasons where detailed observations might produce the greatest gains in understanding.

S. Kinne (Editor)

stefan.kinne@zmaw.de

Here is my assessment:

the paper is in excellent shape and both reviews (as well as mine - see attachment -) are highly favorable. There are a few minor corrections (especially those mentioned by reviewer 1 and by me that should be addressed) Assuming that the authors take care of these comment ... this paper is acceptable for publication.

Please also note the supplement to this comment: http://www.atmos-meas-tech-discuss.net/3/C2401/2011/amtd-3-C2401-2011-supplement.pdf

Interactive comment on Atmos. Meas. Tech. Discuss., 3, 4091, 2010.

General comments

The paper is long, which is a bit discouraging to start it off. However, the major points are nicely summarized.

We apologize for the length of the paper. We worked hard to make it concise; perhaps not hard enough.

The error understanding refers to location and times of AERONET samples, which do only represent a sub-sample of MODIS events, as demonstrated for the cloud and possibly also the snow contamination. The regional analysis provides a lot of useful insides in retrieval assumptions and deficiencies.

It is known that MODIS retrievals have problems over South America also in terms of seasonality. The biomass model, as I recall, is tied to African (Savannah) biomass observations and therefore much too absorbing in type. This leads to AOD overestimate early in the biomass season (overestimates). However, large AOD events in the later part of the biomass season over South America (possibly as high AOD are co-located with clouds) are missed so that the then retrieved AOD values are too low. This 'artificially' advances the biomass season forward in time and should be responsible for a large RMS.

This is a good example of the kind of detailed examination that is needed to truly root out the sources of errors in retrieved AOD, that lie beyond the scope of the current study. We hope that the work we have done will motivate and in some ways facilitate these detailed investigations.

For the central Africa discussion I miss comparison at Ilorin, which has a strong biomass signal at year's end.

The auxiliary materials include detailed graphical and statistical comparison at Ilorin. See /l2_site_pages/aqua/Africa_above_equator/ps108_ilorin.pdf

I like also in that context that the AERONET sites are introduced with lat lon … but for some sites (e.g. Taipei, Xianghe) this info is missing and should be added.

Fixed.

The surface albedo bias correction approach (using only AERONET_MODIS deviations at low AOD values) is interesting but at the same time also a bit confusing. Are we not throwing everything in one basket with a single global parameterization, while there are strong seasonal and regional differences in surface albedo strengths and ratios? Maybe corrections could be stratified by vegetation type and season?

The albedo fields used were based on 16 days of data only, and were actually produced on an 8-day schedule. So seasonal variation is very much included in this correction. As for additional differences (such as by vegetation type): early on when we began examining the MODIS product, we examined the variation in error characteristics between land cover types. While clear differences were found, the correspondence between vegetation type from the MODIS land cover map and the AOD errors was limited in its quantitative skill. When we altered our approach to use albedo data instead, it was clear that we were getting much closer to the actual source of error in the MODIS AOD. While there is almost undoubtedly an additional term of error in MODIS AOD associated with vegetation structure (BRDF, but more importantly shadowing), the term is of lesser importance compared to the primary error in surface reflectance characterization, which we capture effectively with the 3-parameter model.

While it is nice to see that overall compliance improved with all these (innovative) efforts it is a bit disappointing that the improvement seems rather marginal. On the other hand the real measure of improvement may be somewhat hidden by the limitation to compare to AERONET samples.

Given the tiny changes in regionally averaged AOD that can yield climatically significant changes in forcing, we believe these improvements are meaningful. For our data assimilation application, the elimination of outliers is vitally important, even though the damaging effects of outliers do not show as clearly in bulk statistics. Just as importantly, the sources of error we identified can be used in the ongoing effort to build the best possible retrieval of aerosol properties from satellite.

With respect to the figures, I like the pdf-quartile lines in Figure 1a, I would prefer white fonts for the site names on the dark blue background in Figure 5a, the colors ('green'?) in 5c are difficult to see, for curiosity I also would look at slope/correlations sub-samples with тm (not тa) >1.4. I also wonder if we could increase data-volume/coverage with MODIS deep blue data?

1) **Figure 5: changed.**
2) **In the auxiliary material, you will find MODIS-AERONET scatter plots for each AERONET site. For single AERONET sites, the data volumes are small enough that the behavior at high MODIS AOD can be readily examined.**
3) **Deep Blue would indeed provide a great boost in the coverage available from MODIS. We are awaiting the results of a study on these data being undertaken by Yingxi Shi at the University of North Dakota.**

Overall this is a great paper. Congratulations!

Thank you again.