

Dear reviewers,

Thank you for the valuable comments, suggestions and commentary. We understand we have a product that many people use, and we have tried to produce an accessible and user-friendly document. We hope that we have succeeded in answering most of your concerns. Although, we want to stress that this was not supposed to be a comprehensive “validation” of the new product, we have provided more evaluation in the revised version of the paper. We are grateful to have such honest and detailed reviews.

We want to note that since this paper was originally submitted in November, 2012, there have been some updates and bug fixes implemented into the C6 aerosol retrieval code. Therefore, in the revised manuscript, some tables and figures will include updated information. In general, there are no major changes to the overall conclusions.

---

## Response to: Interactive comment by J. Reid

reidj@nrlmry.navy.mil

Received and published: 1 March 2013

Recommendation: Major Revisions

Dear Dr. Reid,

As always, your constructive criticism of both the paper and the algorithm it describes are very valuable. Yet, there is a difference between reviewing the merits of this paper and reviewing the algorithm itself. Because, whether we like it or not, the algorithm described here will become operational. The point now is to ask whether or not this paper is an adequate description of the algorithm and products for the community. Many of your comments concerned this paper itself. For example, we agreed that we should process and display other months besides January and July, and have expanded the test data set to include April and October 2008.

Other comments concerned the algorithm, not the description of the algorithm. At this point, we cannot make changes to the C6 code. Had there been an issue raised about a severe problem with C6 MODIS algorithm or products, of course we would have immediately “stopped the presses”. Plus, we are sure that you would have contacted us directly if there had been such a problem. It is unfortunate that there is no external review process in the MODIS operational protocol that can review potential Collection updates in time to make adjustments before choosing an irrevocable path forward. We welcome your criticisms of the algorithm itself and will keep in mind these criticisms as we move forward towards Collection 7, *under very limited budgets*. We are in a maintenance mode of operation now, with <1.5 people working hard just to keep up with calibration drift and other upstream issues. We understand that the ATMD open discussions allows for public expression of scientific ideas that may go beyond the immediate review of the paper itself, but we also have to keep focus on the question of the merits of this specific paper.

That said, since we had time since your original review in early 2013, where possible we made upgrades to the algorithm itself, and documented that in this paper.

Overview: This is an overview paper on the much anticipated MODIS data collection 6 (Col 6) aerosol retrieval. The paper is long, outlining numerous changes to the algorithm, with additional appendices in the back outlining specifics. Given the importance of the algorithm and its documentation, I heartily endorse the publication of a paper like this in the open access peer reviewed literature. Too often such documentation has been buried in hard to find reports. However, for publication, I would recommend the addition of further algorithm information as supplemental materials.

Thank you for supporting the idea of a long, descriptive, “documentation” of the algorithm and products. What kind of supplemental materials did you have in mind? Should any of the appendices be supplemental material?

Science wise, it appears that Col 6 has attempted to address a number of historical trouble areas for the algorithm. First they are working on cirrus contamination over land and water. I am grateful that the MODIS team finally acknowledges that the band of high AOT around the high-mid latitude oceans is largely artifact,

and that they have apparently done much to improve the algorithm. Lots of modelers erroneously tuned their sea salt emissions to match this artifact; I hope that more reasonable sea salt numbers will follow. A second major improvement updates the optical models so that, from broad global plots, it appears that some of the consistent biases in Central Africa and East Asia have been improved. Other important improvements include an a) update to the Rayleigh/gas absorption treatment; b) the extension of the latitude of retrieval; c) revised high AOT-cloud mask bias to allow retrievals over thick aerosol plumes. The recently developed 3 km MODIS product is also described, as is the joining of Col 6 dark target and deep blue algorithms for a “global” product.

An important consideration regarding this manuscript is that it does not address verification in any rigorous fashion. No doubt there will be several papers from within and outside the MODIS team that will appear in the next year after the product introduction. This paper merely describes the algorithms and demonstrates many large differences from previous versions. I don't have any direct objection to this course of action in particular, but that said, as in previous version updates, the algorithm has been updated based on the examination of relatively few data points. Within this paper, the authors demonstrate a monthly average for January and July, but omit the boreal spring and fall which tend to coincide with Asian dust and pollution outbreaks and monsoonal biomass burning seasons. So, based on this paper, we don't ultimately know how Col 6 is going to behave in key regions and time frames. Again, given the importance of the Col 6 algorithm to the earth science community, I would add additional plots for the spring and fall or, as we at NRL have done, break data down into the natural monsoonal aerosol periods of Dec-May, June-November.

Your suggestion to analyze spring and fall data is valid and important to do. As you might guess, the science development team (us) and the operational algorithm team (e.g. MODAPS) are two different groups. We have ability to process “operational algorithm – like” on our own machines (e.g., a few granules, a few days), but leave the heavy processing to MODAPS. This includes testing on multiple months. Originally we were limited to the bandwidth that was provided to us, which was up to six months of data. So to test the impact of new calibration drift, we chose the six months to be Jan/July for 2003, 2008, and 2010. Even as we made changes to the algorithm, we derived for the same set of months to compare with known “baselines”. The same set of months was used for other science testing (e.g. clouds, land processes, etc) as requested by different MODIS product development teams.

We had processed two additional months (April and October, 2008), which are now included and discussed within the paper.

One area of great interest is the changes to the calibration. It has been clear to us for some time that there have been instabilities in the MODIS radiance calibration. I know that part of the delay in getting Col 6 out has been to give the calibration team enough time to sort out trends in the calibration drift. But, the paper punts on this issue.

Yes, this was a “punt” on the issue. But you gave us back the ball. So now, we are including some discussion about the calibration and its effects. We state that it is not a calibration paper, and are preparing a separate detailed paper about the issue. However, in this paper, we added discussion to show how that Terra's calibration updates are a big deal for aerosol retrieval, and that given the eight months of data processed (across 7 years), we expect Terra and Aqua will trend together over the time series.

Specific comments: 1) First sentence of introduction. I was surprised that the MODIS aerosol team chose not to use the accepted definition of “aerosol”. Please reference the AMS glossary of meteorological terms and use it to correct the text in the introduction. I would also not say (next sentence) that “aerosols drive the hydrological cycle” as, personally, I think water vapor has something to do with it. But you can say that aerosol particles are an important component of the hydrological cycle.

Done

I make these points not to be snarky, but the paper does appear to incorporate a lot of grandstanding. Also keeping with past practices, the aerosol team shows an unfortunate propensity to cite their friends, to the exclusion of other more relevant work. Certainly, when outside work is cited it tends to be in a modest manner. The authors and developers for Col 6 should leverage this extensive body of knowledge. Col 6 is a flagship product for NASA. They need to be balanced. The body of the NRL/UND work, which has included the most rigorous global verification analyses, has been largely unreferenced except in token ways, and it is

unclear from this document whether the aerosol team has learned any lessons from us on bias correction.

We understand that we should leverage more of the NRL/UND work. We note, however, that we expect biases in our product, which may not be correctable within the along-orbit, swath framework. This is a result of the information available as a near-real time algorithm, on a pixel-by-pixel basis. Bias correction, while very important, does not mean the same thing to every user. Assimilation does not work well with data, which are biased, discontinuous, missing, or otherwise not behaving well. On the other hand, there are users who need spatial coverage, so it is not a disaster if data accuracy is compromised. Real aerosol plumes (in smoke plumes, within cloud fields, etc) need not be continuous, nor match well with ground-truth AERONET or other sunphotometer data. Our job is to produce the best product possible, using the resources available to us. Other groups, including the NRL/UND team, have to do what they need to do to use the product, whether it be throwing out "suspect" data or performing extensive bias correction.

Fortunately, we have, and expect to continue, such an arrangement where we learn from the work of others. We have a retrieval built on assumed physical information. There are assumptions about the sensor, aerosol optical properties, radiative transfer, surface characteristics, and retrieval mathematics. Each of these pieces has been studied separately and documented as well as possible. However, biases in one of these assumptions will lead to biases in the results, and the combination of biases is not necessarily linear. The work of the NRL/UND has been invaluable in characterizing these biases, and often contributes to finding possible causes. We appreciate it greatly.

2) Section 2: I very much appreciate the authors reiterating the MODIS history and "fun facts" on the nature of product levels, pixel size, retrievals etc. . . in a more conversational way. It seems that users often forget these issues, and it is good to repeatedly point them out. However, in this section there is reference to the previous ATBD. Is this work in lieu of a new ATBD or will a new ATBD be generated? I think it would be helpful if the authors could note what we should expect.

There will eventually be a C6 ATBD. The Deep Blue product did not have an ATBD for Collection 5. The C6 ATBD will be a single entity that covers both dark-target and Deep Blue. In the meantime, we used this venue to highlight the dark-target algorithm and product changes. Currently, we do not have a deadline for providing an ATBD.

3) Page172 line1-10. This is not exactly our point of view. Bottom line was: yes in fact we can assimilate (which don't get me wrong is a huge achievement), but we do that by largely throwing out 40% of the data points and making, at times, 30% bias corrections. In the next paragraph, it is pointed out that perhaps not all sites perform equally, which comes off as a waffle. Instead, at this point the authors should simply state the facts. That is, over 30-50% of the world, the MODIS algorithm performed quite well. It was certainly never "comparable to sunphotometers" as that would suggest parity, which no satellite product can achieve. But, the data was clearly within your pre-specified error bar. Another 30% of the data were ok if used with caution and with bias correction, but 20% of the data were utterly unusable. Recall, a bulk comparison against AERONET favors those sites where there are lots of sites and lots of data. From Shi's paper, we know those sites are where MODIS performs well. However, there are large tracks of land with few sites, and there MODIS tends to perform poorly.

You have a valid way of presenting facts. We adopted much of this paragraph in the revision. Yet, the dark-target algorithms are designed to work over clear-sky, vegetated, dark soil, and deep ocean scenes. Anything else and the algorithm assumptions may be violated. Bias correction is a tool for making the data more usable, and we appreciate that it provides some good information. Sometimes, the data may be useful even when they cannot be assimilated.

4) Section 3: As mentioned in 1) it seems as if the aerosol team largely did not learn from the NRL/UND over-land bias correction work. If we can do an outside empirical correction to help alleviate such issues as microphysical bias and lower boundary condition, (which even we admit is a gross hack), why can't they improve the algorithm? I understand that they want to have a single scene retrieval, but if using external surface reflectances (like Hyer uses) improves the retrieval, why not do it? The argument of a "physical retrieval" does not hold, as the authors admit that there is no value in the over-land size information. The authors cite 'data volume issues' preventing making adequate corrections, but Moore's law has been in our favor. If the PEATE can reprocess a year of MODIS data in a week, why can't GSFC, with all its computational resources, mine the data? We also noticed that snow bias is never mentioned, although we

have seen this to be a problem. With the increase in the latitude of retrieval, the authors must respond to the question “what is your assessment of snow bias working its way into the climatology?”

Ouch... Yes, there are a number of efforts (including your NRL/UND team) to mine the data. Remember, while data assimilation and weather forecasting is a major use of MODIS data, there are also many other applications that require different stream of data. We have chosen to be conservative in our retrieval but not extremely conservative. In other words, sometimes we will provide retrieval on the edges of scene acceptability (e.g. semi-arid scenes, and/or scenes near clouds), because we value aerosol “coverage”. There is no doubt a bias near clouds, but some of that is real aerosol enhancement and not simply contamination. Aerosol over snow is an interesting issue. During Jan 2013, there was a severe pollution episode over Beijing, China, which was not retrieved at all by MODIS. It turned out the snow mask was so conservative that even patchy snow was being thrown out by the aerosol retrieval. The presence of snow was confirmed by evaluating the MODIS snow product. As for the high latitude retrievals, we do not see evidence for increased snow contamination in the data. Thresholds to detect and filter out pixels having minimal snow signal (i.e. from melting or aged snow deposits) are stringent. We cannot prove in all cases, but we do the best we can.

Now to the bigger issue of including outside data sources (e.g. land surface reflectance/albedo maps, etc), There is no doubt that it can improve retrieval accuracy in many cases. However, there is a circular logic for using “atmospherically corrected” surface reflectance properties. In a climatological sense it is probably okay, but day by day? Greening, burning, plowing and urban construction are all examples of rapid changes that may not be captured by using climatology of atmospherically corrected data. At the same time, a systematic use of wrong aerosol model assumptions will lead to bad surface information. While we cannot escape the issue of wrong assumptions, the aerosol retrieval is not relying on heavily processed upstream data.

It is not anymore a problem of computational power to use input surface assumptions, but it is a problem of historical (and future aerosol processing). With the advent of newer sensors (e.g. Suomi-NPP/VIIRS) that have shifts in wavelength bands, not having a surface albedo or reflectance map allows for portability of algorithms across multi-sensors. Newer retrieval algorithms (e.g. MAIAC) may be very successful in dealing with more surface types, but have to compensate in other ways (cannot be run in near-real time, cannot be run if missing yesterday's data, etc).

So, we don't mean to be snarky here, but we had to make tough decisions. Prior to C5 implementation, we were a team of seven, nearly full time researchers. As commiserate with our “algorithm maintenance” state, we are a team that adds up to less than two fulltime positions. We have access to computational resources, but they are not ours. We have used the Atmosphere-PEATE (e.g. University of Wisconsin) to help us with our MODIS testing, but it is not their primary job. We chose specific items to study and improve. On the other hand, we do the best we can to offer a user-friendly and well-documented product. We know it is not perfect. Yet, we rely on users, such as the NRL/UND team, to “add value” to the product. There will hopefully be a collection 7, and we plan to leverage our users to help us optimize this effort.

5) Section 3.2.4 There has been much evidence that the QA flags are not really QA flags, but reflect “goodness of fit”, which does not actually correlate that well with data quality against AERONET. This should be discussed.

We believe that these QA flags, are in fact representing Quality Assurance. One of the criteria is “goodness of fit” (e.g. small fitting error), but that is not the only criteria. Essentially, the QA plan or QA “logic” follows the steps of the computer code. If at some point a test “fails”, then this is noted. The default QA Confidence (QAC) is 3. However, the QAC will be degraded based on the most severe failure. Thus the QAC represents the “satisfaction” of the retrieval, and it may or may not be related to the data quality compared to AERONET. Also, we showed in Levy et al., (2010) over land that QAC flags correlated very well with agreement with AERONET. For C6, we have assigned higher confidence to over-ocean retrievals, so that the very clean retrievals (low AOD) are not overlooked.

6) Section 3.2.6: Two months of data are insufficient to illustrate the differences in the old and new algorithms. Period. I fully acknowledge that the paper is not a validation paper, but if they want to demonstrate differences, presenting in this way is clearly non-representative and possibly misleading. As it

stands, I think there is considerable framing bias in the bulk statistics (although I very much appreciate the global difference plots). If the authors want to cut the number of plots and tables, then choose 1 year and verify on the more meaningful Dec-May and June-Nov time frames. If the biggest AOTs are in the spring and late summer/early fall, why are these periods not presented? January and July are pretty tame AOT wise (not many large events). In this light, Figure 7 has little meaning; while certainly the number of points within compliance is functionally the same, now the algorithm exchanges low bias for a high bias. I would prefer the authors present RMSE and RMSD as a function of AOT for the two algorithms, as this is much more robust and less subject to misinterpretation than presenting a correlation coefficient (I see they used  $r$  and  $r^2$  for obvious reasons). Also, the authors need to show data for major global regions, (most of the data currently presented appear to be from US and European sites where the algorithm works well), and plots normalizing RMSE by region would have great value in displaying the changes from Col 5 to Col 6.

We now include plots for the spring (April 2008) and Fall (Oct 2008) months (new Fig 1).

The co-location statistics are now Fig 11. As part of the new Fig 11, we also provide MODIS AOD error versus AERONET AOD as box-whisker plots (like as in Levy et al., 2010). While the global statistics (percent with EE; slope, intercept) have not changed much, the high bias at low AOD has been improved.

Yes, most of the data comes from U.S. and European sites. This is where the AERONET data are. When we provide a full validation paper of the entire Collection 6 dataset, we will break down by site, season, and other parameters. Or we may rely on our super-users (i.e. the NRL/UND team), who have the infrastructure to do the kind of validation that is envisioned here.

7) 3.3.1 Sentence 1: I am glad the wind speed correction was included. I appreciate the plot of 1 day in July, but please add the monthly average of this particular correction to give context with the other data presented. It would be clearer if there was one large figure showing the baselines, and then an associated cage for each correction that is made (surface, microphysics, cloud etc.) with a final total correction at the bottom. Minor comment, it is Zhang and Reid, 2006, not 2011, that noted the surface wind error. Similarly, the cloud bias issue listed in section 3.3.2 is also from the 2006 work. For collection 5, the work is dealt with in detail by Shi et al., 2011.

We like the idea of a figure showing each major “correction”, and then a final figure, which we have now included within the manuscript.

8) Section 3.3.3, I do not wish to be petty, but I want to point this out. When it comes to verification work you tend to cite your fields when in fact, collectively, our group has been reiterating to you for over 6 years that the QA flag does not correlate with verification. We have also been the ones pointing out (through our prognostic error modeling), the asymmetry in error, and why, if MODIS reports a clean AOT, it probably really is clean. What you fail to recognize is that these issues will lead to a representativeness bias when you construct a climatology. We have prognostic error models for “good” and “very good” data, and we use both categories of data. However, generally the MODIS team recommends that people only use the “very good data”. Through your formulation (and your weighting mechanism in section 4), you have probably biased your climatology.

But the QA flag DOES correlate with agreement with AERONET, at least over land, (Levy et al., 2010). For C5, it didn't over ocean, and we understand why. This is because the retrieval consists of two pieces of information: AOD and size parameter, and initially the product QA flag would degrade in situations where AOD might be very good, but size parameter not. It led to confusion. We are now recommending users use only “very good” over land, and this is based on validation with AERONET, as well as on decisions made during processing. But we are recommending over ocean use of all QA > 0, based on comparisons with ground sunphotometry including MAN, decisions during processing, and a decision to link QA flags to AOD and not size parameter.

We have known for a long time that when MODIS reports clean AOT, it is probably clean. However, we have not found an easy way to deal with this within the statistics. Yes, the asymmetry in error (and in reported AOT) will lead to error in climatology. This is the reason we allow negative AOT in the retrieval; it helps de-bias the statistics. As we are finding out that these negative AOTs are not random, and constrained in particular regions, then we recognize that some of the problems are systematic.

In this iteration of the MODIS algorithm (C6), we have taken special effort to recalibrate the QA flags. We have greatly increased the QA confidence for over-ocean data, and have slightly reduced the confidence over land. The full evaluation will determine whether the QA tuning has been successful. Yet, while we attempt to skew our data towards a climate/statistics-based interest, we must ensure data coverage to those willing to sacrifice accuracy.

9) Section 3.6. I am not sure doing an “average” of DB and DT is the best approach here. I would think based on internal machinery doing one or the other based on an error model would be the way to go.

You are absolutely correct here. The DT/DB merge was implemented because there has been strong community advocacy for a merged product. We spent some time analyzing the two products when both retrieved with high QA. This turns out to be only a tiny fraction of Earth’s land surface. The QA for DT decreases over brighter surfaces where DB’s QA is high, while the reverse is true over darker surfaces. When they both retrieved at high QA and near an AERONET station, there were no patterns suggesting that one was systematically better than the other. Taking the average became a first guess after recognizing that DT was biased high and DB biased low in these semi-arid regions. Feel free to come up with a different solution.

10) For the level 3 data, I suggest the algorithm team include the median and geometric standard deviation. In a 1 degree box on a single day this is unlikely to be useful, but in the monthly+ products, this is a much more robust way of describing the AOT. Also, while the authors correctly note that pixel weighting can bias towards a sensor, they should recognize they have the very same issue in verification against AERONET.

It turns out the official 1° Level 3 data does include median and standard deviation as a product. However, the point of this exercise was to show what happens to the mean with the removal of the pixel weighting. This information is now explained in the text. Yes, it is true that pixel weighting adds a bias to the results, and that, in essence, it is a pixel weighting exercise to compare two sensors. The verification exercise entails that both sensors are cloud free, which is a bias.

11) Section 5-not much to say here. The paper sort of punts the topic to (potential) later documents. But that, I think, is okay as the 3 km product in some ways is a different beast.

Different beast, and there are two AMT papers that discuss the 3 km product.

12) The lack of long term trend analyses and descriptive statistics in this paper on the calibration issues, which from a climate trend point of view is extremely important, is a significant shortfall of this paper. In order for Zhang and Reid (2010) to determine the size of the MODIS calibration bias and remove it, it took nearly the entire data record. I do not think that testing on a few months of data is going to characterize the calibration change over time. There may be a baseline change, but for trends, you will have do as we did, and engage in more rigorous long term analyses.

Absolutely true. As response to one of your earlier comments, this was a “punt” on the issue. This is not a calibration paper, and we are preparing a separate detailed paper about the issue. However, we added discussion to show how that Terra’s calibration updates are a big deal for aerosol retrieval. We cannot prove it at this time, but given what we see from the test data we processed across 7 years, we expect Terra and Aqua will trend together over the time series. The upstream changes to Terra’s calibration are pushing the two MODIS sensors into better agreement. We have revised the paper with three figures. A new Fig. 21 shows which Terra bands have changed the most in the new calibration. A new Fig. 22 shows the large impact on Terra’s AOD retrieval. Finally a new Fig. 23 shows a time series of monthly mean AOD from the eight test months, demonstrating that Terra and Aqua may march together.

Hope this review helps. I think I can speak for many and say that the MODIS AOT retrieval is one of the greatest achievements in the field. But, with a little more effort, its impact can be even greater. Jeffrey S.

Thank you. This paper describes part of a long term process, requiring many years of effort. Again, based on the constraints of satisfying multiple audiences, and was intended as a “modest” improvement to the MODIS dark-target products, we feel it is best to put out the product as soon as possible.

---

## Response to: Interactive comment by M. J. Garay (Referee)

michael.j.garay@jpl.nasa.gov Received and published: 6 March 2013

Thank you for the valuable comments, suggestions and commentary. We understand we have a product that many people use, and we have tried to produce an accessible and user-friendly document. We hope that we have succeeded in answering most of your concerns. Although, we want to stress that this was not supposed to be a comprehensive "validation" of the new product, we have provided more evaluation in the revised version of the paper. We are grateful to have such honest and detailed reviews.

We want to note that since this paper was originally submitted in November, 2012, there have been some updates and bug fixes implemented into the C6 aerosol retrieval code. Therefore, in the revised manuscript, some tables and figures will include updated information. In general, there are no major changes to the overall conclusions.

### General Comments:

This paper presents an overview of the MODIS Collection 6 (C6) Dark Target (DT) aerosol products. In particular, it describes a number of the changes that users can expect to see relative to the widely used MODIS Collection 5 (C5) aerosol products. The authors provide a readable introduction to the product content and the theory behind the retrieval approaches. In addition, some comparisons are made that show how the C6 and C5 products compare to one another, both globally and with regard to commonly used validation datasets. New merged datasets, a higher resolution (3 km) product, and the methodology used to construct the Level 3 (gridded) products are also discussed in some detail.

In a lot of ways the paper reads like a "conversational" Algorithm Theoretical Basis Document (ATBD), which is a good thing, in my opinion. ATBD's tend to be written and read only by the algorithm teams themselves, but they often contain information that is important for the larger scientific community. For this reason alone, I feel this is a significant paper, appropriate for publication in Atmospheric Measurement Techniques. My overall opinion of the paper is less harsh than that of the other referee (Jeff Reid), who had some frank criticisms of the paper in terms of the extent of the data analysis. However, as a member of an algorithm development team for a different satellite instrument, I am perhaps more sympathetic to the situation of the authors. Although the paper does not present any extensive validation of the new algorithm, it does present a number of examples of the changes observed by the algorithm team as modifications were made in the retrieval approach, and it discusses the background and motivation for these modifications. This provides valuable insight into the strengths and weakness of the MODIS DT aerosol retrieval algorithms from the perspective of the developers themselves. That said, I do agree, however, that critical analysis of the performance of the MODIS C6 DT aerosol retrieval appears to be lacking in a number of areas. In my view, the paper would be greatly strengthened by including a more careful consideration of some of the results already shown, as well as expansion of the analysis along the lines suggested by Jeff Reid.

Specific Comments: This is a rather long paper, so I will provide my comments broken down by section.

### Introduction

With regard to the "spatial resolution and repeatability," I was surprised not to see a reference to at least Levy et al. (2009), if not Shi et al. (2011), and the work of others who have looked at the spatial statistics of MODIS aerosol retrievals in great detail.

We have added these references.

### 2.2 Basic concepts of the MODIS aerosol retrieval algorithm

Perhaps I'm missing something, but it's not clear to me how, only knowing the total AOD and fine mode fraction, the AOD can be derived at other wavelengths. One would need to know something about the optical properties of both the fine and coarse mode aerosols, at least. In that case, it would be possible to derive the AOD at other wave- lengths and infer the Ångström Exponent (AE). This is an important point because certain research groups like to construct a MODIS-derived Aerosol Index (AI) by multi- plying the AOD by the

AE (e.g., Lebsack et al., 2008). Unlike the equivalently named AI from TOMS/OMI, the intention here is to derive a parameter that is related to the aerosol number concentration, which in turn provides an estimate of the cloud condensation nuclei (CCN), in the manner of Nakajima et al. (2001). However, what Nakajima et al. (2001) did was retrieve the AOD in two AVHRR wavelengths separately, allowing in some sense the “direct” derivation of the AE, independent of an aerosol model. The terminology is also somewhat confusing when the authors compare the DT land and water approaches. The authors state that an aerosol “model” is retrieved over land, while an aerosol “mode” is retrieved over ocean (p. 169). It’s not clear what distinction is being made. In both cases, a look up table (LUT) is employed and a specific aerosol “model” is retrieved. Even over water the MODIS retrieval provides information on the AOD at 0.55  $\mu\text{m}$  and the model (fine mode and coarse mode) that yields the minimum error relative to the observations – AE is not an independently derived parameter because using an aerosol model in a LUT prescribes the spectral slope of the AOD. If the intention is to highlight that the LUT itself is prescribed both regionally and seasonally over land, while a single LUT is used over the global oceans, then this point needs to be made more clearly.

This has been confusing forever. We have tried to explain a little differently in the paper. But to summarize:

**We know exactly** the optical properties of fine and coarse aerosol “types”. Over ocean, these fine and coarse aerosol types are represented as fine and coarse lognormal modes. The retrieval chooses a combination of fine and coarse modes to retrieve a bi-modal aerosol combination that matches TOA spectral reflectance. On the other hand, over land, both fine and coarse types are already assumed to be bi-modal (gleaned from AERONET climatology, Levy et al., 2007). The *fine-mode* is dominated by the *fine-mode*, whereas the *coarse model* is dominated by the *coarse mode*.

We know the properties (assumed for making LUTs) of fine and coarse mode types over ocean, as well as fine and coarse –dominated types over land. This means, that given the fine mode (or model) –fraction, we can determine everything to know about fine and coarse aerosol properties and spectral dependence of the total AOD. Therefore, we can derive AE.

We also know exactly what aerosol properties were assumed when creating the LUT, so if we want to know how many particles are larger than a particular threshold size, we can determine this information. In C5, we used the term “CCN” to refer to the number of particles of the fine mode, which were greater than 0.03 micron in radius. This was a bad idea, and for C6, we now denote this parameter with the cryptic name “PSML003” (Particles of the small mode having radius greater than 0.03 micron).

With regard to the retrieval of fine mode fraction (FMF) over water, I was surprised there is no discussion of (or even a reference to) the results reported by Anderson et al. (2005), who found that MODIS AE provides better agreement with suborbital measurements than MODIS FMF, at least in the ACE-Asia region for the Collection 004 algorithm. I do not believe these results were addressed in the study by Kleidman et al. (2005), which is referenced here. Again, this issue is important because members of the research community wish to use the MODIS aerosol products as a means for assessing the effect of aerosols on clouds (e.g., Lebsack et al., 2008), and the anthropogenic aerosol fraction (e.g., Yu et al., 2009). While the MODIS algorithm team cannot possibly be expected to anticipate all the possible uses of their product by the community, I believe that it is important for the algorithm team to help the community use their products properly, and a paper like this serves as an appropriate forum for such discussions. In light of this, I applaud the decision to change the SDS “Cloud\_Condensation\_Nuclei” to “PSML003\_Ocean,” as a step in the right direction (although from Table 3 it is difficult to determine whether the SDS “Cloud\_Condensation\_Nuclei\_Ocean” remains in the product or not). The removal the intensive properties from the Level 3 products will also provide a useful safeguard on some potential “inappropriate” uses of the MODIS data products in the future.

We agree. We now include reference to Anderson et al., (2005). The goal of the MODIS aerosol retrieval over ocean is to get the correct spectral dependence of the AOD. Because the retrieval is allowed to fiddle with the selection of aerosol modes to get to this result, we can see how retrieval of AE should be more robust than getting FMF. Plus, as Kleidman et al., (2005) discusses, FMF can have many definitions, so one should not expect uniform validation of FMF. AE, on the other hand, as long as wavelengths are consistent, should be less ambiguous.

Thank you for accepting the changes to PSML003. It was a mistake to still include Cloud\_Condensation\_Nuclei in the list. We will check to make sure that it is named correctly in all versions

of the documentation. As a side note, our development team has discussed renaming many of the SDSs to something more in-line with the physics of the parameters. However, a significant portion of our operational users would have had difficulties adjusting to renames of the most popular parameters.

### 2.3 Evaluation of the C5 MODIS aerosol products

The expected error (EE) for the FMF is given as  $\pm 0.20$  over ocean with a reference to Kleidman et al. (2005). Looking at that paper, it is unclear where this value comes from as comparisons are made against both the AERONET “Dubovik” and “O’Neill” inversions, with differing results. It is also unclear to me what the implication of a  $\pm 0.20$  FMF fraction sensitivity is. One interpretation is that the MODIS retrieval provides effectively three ranges of FMF over ocean (0.0–0.4, 0.3–0.7, and 0.6–1.0, for example, minimizing the overlap). It would be helpful to have a discussion of this issue from the algorithm team.

You are correct in that these things were not referenced or discussed clearly. More in the “spirit” of what Kleidman et al. did in their paper, and also like we did in validation paper over land (Levy et al., 2010), we derived an Expected Error (EE) envelope that would encompass 66% of the global collocations. This was done for FMF, and this is where the  $\pm 0.20$  came from. We never published a similar C5 validation paper over ocean to document this exercise. We have updated this information in the text.

I realize that I may be expressing a minority view, but I still take issue with the retrieval and reporting of negative AODs over land in the MODIS product. I am aware that other consumers of the product view the inclusion of negative AODs as a significant advance that provides more “statistically robust” results. I have two issues with this approach. First, this violates the physics of the retrieval itself, implying that less radiance is viewed by the instrument than for a perfectly clean (i.e., aerosol free, or Rayleigh- only) atmosphere. Algorithmically, it is surprising that this can ever occur given that some fraction of the pixels (20% at the low end over land) are already discarded as “unsuitable” for retrievals. The exclusion should handle the types of situations (e.g., cloud shadows) that could lead to extremely small radiances. Second, careful examination of the statistics shows a significant increase in the sampling exactly at the -0.05 AOD cutoff employed. This can be seen in Figure 5, but is somewhat clearer in Figure 5b in Kahn et al. (2009), with finer bin resolution. This is not consistent with the idea that there’s a tail that is tapering off quickly at low (negative) AOD, as would be the case for a Gaussian distribution centered somewhere around 0. It might be useful to plot the MODIS DT-land AOD data from -0.05 to +0.05 in equal increments of 0.01 to better see this behavior. The unequal bin widths used in Figure 5 in this paper make interpretation difficult.

While we understand your concerns, we feel there is no choice to retrieve negative values. The uncertainty is large enough that, yes, there are conditions which the scene appears “cleaner” than an aerosol-free atmosphere. Cloud shadows maybe an issue, but it is more often the assumptions related to surface reflectance. If the surface in a particular wavelength is brighter than is expected via the VIS-to-SWIR surface constraints, then the algorithm will compensate with less aerosol loading. In clean situations, this may require a negative optical depth to compensate. That is why we have a large uncertainty  $\pm(0.05)$  on conditions of low AOD. Admittedly, even with this argument, the histogram shows far too many retrievals of negative AOD, including “large” (exactly = -0.05) cases. This means that we still have not sufficiently accounted for systematic biases to our surface reflectance assumptions. Most of these extreme low biases are located in a few places (e.g. Australia, Brazil). If we didn’t retrieve negative, there would be no information over these regions. We could manipulate the QA confidence, but again, these clean areas represent very clean aerosol conditions, that might be otherwise ignored by the average user.

We have created revised histogram plots to better show this binning around zero.

### 3.1 Changes common to both land and ocean algorithms

The title of this section might be modified to include the term “radiative transfer,” since that is the primary focus of the changes discussed. The detail provided in this section is exemplary and this is precisely what makes this paper so valuable to the community.

Thank you, and we have changed the section title:

The authors note that the angular/stream resolution of both radiative transfer codes was increased, but it would be helpful to note what the final choice was for comparison and future reference.

Given the likelihood of slight differences in the fabrication of the detectors, I was surprised to see a single set of MODIS central wavelengths (CW) provided in Table A1, rather than a set of CW for Terra and a separate set for Aqua. It is encouraging to see the improved agreement between the numbers shown in Table A1 for C6 and those found in Table 2a in Kahn et al. (2007), which were calculated for Terra.

We have done testing, and determined that in all seven channels used for aerosol retrieval, the differences in CW between Terra and Aqua are  $\leq 0.0004 \mu\text{m}$ . The largest difference in Rayleigh optical depth is in band #3 (which is 0.0005). Therefore, we felt that splitting the difference was sufficiently accurate so that a single algorithm could be run with both sensors. We have now included the wavelength information for each sensor separately in the appendix A.

The “King factor” the authors provide (0.0279) is actually the Rayleigh “depolarization factor” (Young, 1980). As discussed by Young (1982), there is already enough confusion in the terminology, so it is better to follow standard practice here. In fact, this depolarization factor is weakly dependent on wavelength, so the value adopted by the MODIS aerosol team is really applicable only to a wavelength of 0.600  $\mu\text{m}$ . At 0.450  $\mu\text{m}$  it is 0.02883 and at 2.0  $\mu\text{m}$  it is 0.02702. These results are taken from Table 2 in Tomasi et al. (2005).

Yes that is correct. We have revised the text to mention this. The differences between 0.027 and 0.029 are insignificant for our calculations.

The authors report that the radiative transfer code changes, taken in sum, “resulted in an overall global mean AOD increase of approximately 0.01 over land, and a decrease by 0.005 over ocean” (p. 175). Are these results for both Terra and Aqua? Do these results reflect an annual mean value, or a comparison of a single month of data? Moreover, these numbers provide one example of what I believe is a lack of critical evaluation of the results. Taking the mean values of C5 AOD from Tables 2 and 4 in this paper, the change due to the radiative transfer alone results in a nearly 6% increase in the AOD over land and a 4% decrease in the AOD over ocean. The EE (for C5) is presented as  $\pm(0.05 + 15\%)$  over land and  $\pm(0.03 + 5\%)$  over ocean (pp. 171-172). How does one relate a change in the global mean AOD to the EE? In other words, is a  $\sim 5\%$  change in the global AOD “acceptable” given the reported uncertainty envelope?

These values represent a global mean for one month only. Although similar results have been found for all months tested, and for either Terra or Aqua.

Yes, these values are relatively large as compared to expected global means ( $\sim 5\text{-}10\%$ ). When it comes to EE, this 0.01 does not make much of a difference at all. It does increase the bias slightly over land (puts more retrievals above the 1-1 line), and decreases bias slightly over ocean (reduces number of retrievals above the 1-1 line). However, this 0.01 is a (mostly) systematic change across the board, and should not significantly impact those working with applications of monitoring or aerosol assimilation. Changes of  $\pm 0.01$  in AOD are common with just about any modification to the retrieval, which becomes VERY important if one cares about global means or radiative forcing. In other words, we believe that trying to get MODIS global AOD to within  $\pm 0.01$  is impossible, given the inherent uncertainties in a multi-wavelength, single angle instrument. On the other hand, standing by our RT calculations is important. When we have to repeat our RT exercises with other instruments, and with new algorithms, we have repeatable RT. So, a 5% change (or even a 10 or 20% change) in global AOD is acceptable within our EE envelope.

There is a point here that repeats itself several times in your review. The MODIS instrument is capable of returning only one piece of information over land, and two over ocean. In order to produce this information, we have to employ a great many assumptions. The net effect of these assumptions is an inherent uncertainty in the product that we categorize with EE and regression statistics based on validation against ground-based sunphotometry. As we constrain and improve our assumptions, and adapt to upstream changes, the MODIS results will fluctuate within this uncertainty. Changes on the order of 0.01 are fully expected. Changes as high as 0.02 would not be surprising. This is an inherent characteristic of this type of sensor.

And yes, this uncertainty is unacceptable to constrain direct radiative forcing, **on its own**. That’s why MODIS aerosol products have to be used in conjunction with the results of other instruments including other satellite sensors and suborbital information, and with models. Dr. Reid’s mention of the NRL/UND work on bias correction is just one example of a methodology that starts with the MODIS product, with all of its biases

and uncertainties, and makes the product more useful for a specific family of applications. We have made this explicit in the paper conclusions.

### 3.2.6 Comparison with C5 products and AERONET

Again, when discussing the results in this section, the authors neglect any critical evaluation. Expressing the change in the global mean land AOD (for Aqua) from C5 to C6 as a percentage, the values are +3% for January and +17% for July. As discussed above, this change to the global mean AOD seems large given the quoted EE. While it is not unexpected that the global AOD would increase given the removal of the additional CO<sub>2</sub> correction from C5, apart from some changes to the radiative transfer, the only other apparent significant changes from C5 to C6 are the cloud masking logic (including pixel call back) and a bug fix in the Quality Assurance (QA). So what are the possible reasons for a change in the global AOD over land of this magnitude? I was surprised not to see any regression plots of C6 vs. C5 that might shed some light on this. The histogram in Figure 5 suggests that C6 has a significantly greater fraction of higher AOD pixels compared to C5. How did the MODIS team verify this was not due simply to poorer cloud screening leading to a greater amount of cloud contamination, for example? Comparison with AERONET (e.g., Figure 7) would not show this because the AERONET data themselves are cloud screened. In the difference plots in Figure 6 it is notable that the majority of the land areas (with sampling) have a pinkish hue, indicating an overall increase in AOD that could very well be indicative of a cloud screening issue. However, the commentary on this section focuses instead on places with extremely high AOD, which the authors associate with heavy smoke. The results for Siberia and Canada are consistent with other smoke climatology databases (e.g., <http://www-misr.jpl.nasa.gov/getData/accessData/MisrMinxPlumes/>), but not every heavy smoke region (e.g., the Amazon) shows a large increase in AOD in July, and the region around Beijing is unlikely to be impacted very often by heavy smoke.

After additional testing, and some more updates, the over-land AOD actually decreases or increases by 0.015 or 10% in some months. We agree that this is huge. On the other hand, it really doesn't make much difference to percentage within EE, as minimum EE is 0.05.

We characterize the differences between C6 and C5 with histograms, global maps and AERONET collocations. I am not sure we needed to regress C6 vs C5, especially as the results are straightforward. We have modified the relative frequency histograms to regular frequency histograms, so we can better see which class of AOD increases or decreases. Over land, like the global maps, some months go up, some go down. Using a new Fig 2 (for July 2008, Aqua), we show that over land, AOD is increased due to gas correction, and increased because of the new cloud mask. Some of the increases are due to letting in more smoke, but others may be due to adding more clouds. The analysis that we presented in old Fig 3 (new Fig 8) suggests that when updating cloud masks we are recovering both high and low AODs. However, statistically, where there are a few new high AODs, that this gives a relative increase to the mean AOD.

We agree with the reviewer that we should have not have associated high AOD only with smoke. It can be smoke (where we know better cloud mask tuning yields better aerosol retrieval). However, high AOD can also be pollution (e.g. Beijing) and or dust. We cannot claim that the new cloud screening is better everywhere.

### 3.3.1 LUT and wind speed dependence

It might be worth noting that MISR also includes a wind speed dependence in the aerosol LUT over water. The reference would have to be Kahn et al. (2007) or the MISR aerosol ATBD.

Done

The description of the model used to provide the winds is incomplete. I believe the model is the National Centers for Environmental Prediction (NCEP) Global Data Assimilation System (GDAS), which is provided at 1° horizontal resolution as a MODIS ancillary product. However, I think that technically this is just an “analysis” result, as opposed to a “reanalysis,” but that may be an issue of semantics. Careful checking of the documentation shows that no “2-meter wind speed” parameter exists. I would guess that the MODIS team is using the “U-component of wind at 10 m AGL” and the “V-component of the wind at 10 m AGL” to derive a wind speed at 10 m AGL (not 2 m AGL as specified a number of times in the text). There are parameters provided at 2 m AGL (humidity and temperature), but not wind.

Yes, you are correct. It is the 10 m AGL winds we are using.

### 3.3.2 Cloud masking, sediment masking, and pixel selection

Although there are different standards, it might be good to stick with the band designations utilized by the MODIS calibration team (e.g., Wu et al., 2012). Here bands longward of 1.0  $\mu\text{m}$  are designated shortwave infrared (SWIR), rather than near infrared (NIR), as is done in this section.

Okay.

### 3.3.3 Quality assurance

In my experience, the criterion used for determining a “Rayleigh-like” atmosphere is too loose. At first glance, the approach seems reasonable given that the Rayleigh optical depth at 0.856  $\mu\text{m}$  is 0.01623 (Table A1). 10% of this value is just 0.001623. Since for low AOD the radiance is essentially linearly related to the AOD itself, the criterion appears to be putting a lower limit on the MODIS AOD retrieval of approximately 0.002. However, since the Rayleigh optical depth at 0.553  $\mu\text{m}$  is 0.09480, the 10% criterion really implies a lower limit to the MODIS AOD retrieval of about 0.01. Moreover, vicarious calibration by the MODIS team suggests that the calibration is good to within about 2 – 3% (Wu et al., 2012), so it seems that the algorithm may not be taking full advantage of the instrument sensitivity.

During the MODIS over-ocean retrieval, the first guess is that the observations and LUT must match at the 0.86  $\mu\text{m}$  channel. This is why we believe trying to retrieve when the TOA reflectance is less than 1.1 \* Rayleigh is pointless. The difference between 0.01 (at 0.55  $\mu\text{m}$ ) and 0.00 is meaningless. Nonetheless, it is “clean”, and should be counted. While it might be interesting in the future to try and increase precision, given all of the other uncertainties in the retrieval, going after this 0.01 is not going to be fruitful. On the other hand, we like that you stated the 1.1\*Rayleigh as being approximately 0.002 in AOD units, and have added pieces of your discussion in the paper.

### 3.3.5 Comparison with C5 products and AERONET

The change of -0.025 in the global mean AOD over oceans from C5 to C6 is huge! This is a 16.7% decrease in January and an 18.2% decrease in July. In fact, it is difficult to believe the magnitude of this change given the EE over ocean in C5 is reported as  $\pm(0.03 + 5\%)$ . The authors contend that this change is “a result of the new protocol of assigning QAC for low signal cases” (p. 193). Trying to piece together the causes for this change from Figure 11 alone is challenging. The sampling increased by  $5.76 \times 106$ , which is a single tickmark in the upper (linear) portion of the histogram plot. In fact, this is about the magnitude of the difference in sampling for the modal (peak) value in the 0.050 AOD bin. Although the increase in the relative sampling of AOD = 0.000 is enormous (a factor of 40x), the absolute sampling from the AOD = 0.000 bin is only about  $2 \times 106$ . It is clear from the figure that there are increases in every bin for AOD < 0.050 and decreases in every bin for AOD > 0.050, which is not consistent with the explanation that including more AOD = 0.000 (with high QAC) in calculating the mean value is the driver for this large change in the global mean AOD over ocean. It would be easier to parse out the underlying causes if a second histogram was generated using just matched C5/C6 data, which would make it clearer how the values in the bins changed in C6 relative to the C5. A regression plot of the ocean retrievals from the two collections would also likely be extremely informative.

With the revision of this paper, we see that the change over ocean is not quite as large, now more like 0.018. Nonetheless these are large changes (>10%) of the global mean. Again, following the land discussion above, these changes are smaller than the EE. We have revised the histograms (new Fig 14) to show absolute frequency rather than relative, and for each of the four months during 2008. We see that there is an across-the-board increase in all bins having AOD < 0.05, and somewhat fewer cases in the moderately clean (<0.2) range. The heavy AOD frequency remains the same. We are not sure why a second histogram would be necessary, as the patterns are clear from the combination of plots and tables. Lower AODs are retrieved in midlatitude storm tracks (higher winds), and higher AODs are retrieved near the ITCZ (lower winds). Cloud masking plays a role, and so does the new gas correction (increasing AODs).

From Figure 12 it is clear that the inclusion of the wind speed has had a significant (likely positive) impact on

the MODIS AOD retrievals, particularly noticeable in the storm track regions, and the monsoon region of the Indian Ocean in July. However, I also note an increase in the AOD in the warm pool region, in particular, associated with a decrease in the AE, which could indicate increased cloud contamination.

Finally, in this section, the authors claim that the 2% reduction in the AERONET col- location in the AE comparison “is due to not reporting size parameters when the AOD signal is too small to retrieve size robustly” (p. 195). Except in this specific section, I did not find any mention of a change in the threshold from C5 to C6 where size parameters are reported, which I assume is based on the discussion in Section 3.3.3. If a change was indeed made, then it would be appropriate to include it in Section 3.3.3, rather than having it appear in this section.

This figure is now Fig. 1 and 2, where Fig 2 shows step by step changes and their impacts. Lower AODs are retrieved in storm track regions, that is partially offset by a new gas absorption correction. Cloud masking provides both positive and negative changes to the retrieved ocean AOD. We cannot completely rule out cloud contamination for the warm pool region, but the overall change can be explained by combination of wind speed and gas absorption corrections. Size parameters are not reported for the cases when TOA reflectance (at 0.86  $\mu\text{m}$ ) is less than 1.5 \* Rayleigh.

#### 4 New MODIS 3 km product (MxD04\_3K)

The authors state that, “the new MxD04\_3KM file will provide only a subset of the SDSs offered by the standard MxD04\_L2 file” (p. 203). It would be useful to summarize what SDSs will be made available.

This information is in the Appendix C as Table A4.

#### 8 Conclusions

Once the paper concludes, the big unanswered question is when the MODIS algorithm team expects C6 to be publically released, for both Terra and Aqua.

As soon as possible! At this point, publication of this paper should be close to public release of C6 aerosol products. Note that Level 1B and other “upstream” files are already available via <http://ladsweb.nascom.nasa.gov>.

#### Appendix A

It would be helpful to have a reference to the solar spectral database used to perform the weighting of the response functions. Clearly, small differences in this database can lead to differences in the spectral response functions.

We added this information about solar spectral database. We also added information about the slight differences between Terra and Aqua CW, and that we took the middle ground.

From the discussion here it appears that a single Rayleigh optical depth is calculated and applied globally based on a calculation performed for 45°N latitude. Is this correct, or is Rayleigh included in the radiative transfer?

This is pretty much what we did. We used Bodhaine et al., to calculate ROD corresponding to MODIS filter function. Rayleigh optical depth is included in the radiative transfer that creates LUTs.

The discussion of the derivation of the ozone correction in this section is difficult to follow. While the formula for the calculation of the water vapor transmission is provided (Equation A5), no similar formula is provided for ozone. Also, it would be somewhat more useful to provide the ozone correction factors, which are multiplied by the column ozone abundance to yield ozone optical depths (see Kahn et al., 2007, Table 2a).

Formulas have been added.

#### References

The following references do not appear anywhere (that I could find) in the paper:

Anderson et al. (2003), Kinne et al. (2003), Kishcha et al. (2007), Schafer et al. (2008), and Takemura et al. (2002). A citation for Chu et al. (1998) from page 167 is missing.

These references are either deleted from the list or added within the text.

## Tables

Table A2: The correction factors for carbon dioxide appear to be missing.

CO2 is now lumped in with “other gases”.

Table A4 (p. 239): I believe the parameter should be “precipitable water” not “perceptible water,” this could be a spell checking error.

Fixed

Table A5 (p. 241): Again, “precipitable water” not “perceptible water.” Figures

Fixed

Figure 5: There’s a typo in the figure caption. It should be “0.55  $\mu\text{m}$ .” As discussed above, it would be helpful to have both a histogram of matched data and the histogram plotted with equal sized bins. It is interesting to note that this histogram has the opposite behavior seen in Figure 11, the histogram for the ocean retrievals. In the land histogram the sampling in all the bins less than the modal (peak) value decreases, while the sampling in all the bins greater than the modal value increases. The biggest driver in the reduction of sampling appears to be in the first (-0.050) bin.

Typo fixed. It is now Fig 10. We agree it would be helpful to have different forms of histograms and matching, but this became too complicated. The message remains the same. Now that we have plotted four months of data, separately, in some months the frequency increases smaller than modal peak, other months it is the opposite.

Figure 7: There’s a minor inconsistency in the figure caption and the axis labels. The caption gives the wavelength in  $\mu\text{m}$ , while the axis labels use nm. It would be useful to have additional plots showing the C5/C6 values regressed against one another, and the AERONET regression for the matched dataset (i.e., times and places where both C5 and C6 retrievals are available).

Sorry about the minor inconsistency. This seems to remain no matter how many times we redo figures! We agree it would be helpful to have different forms of C5/C6 matching, but this complicates the message. There are a few extra co-locations for C6, and the overall comparability is not hindered.

Figure 8: The NCEP wind speed is provided at  $1^\circ$  resolution, so contouring the data in the way it is done in this figure makes it difficult to see how the values are changing, and the contour labels are hard to read. It would make more sense to provide an additional panel showing the NCEP winds by themselves.

Point taken. This figure has been revised: It is now Fig 12.

Figure 9: If it were true that “the largest differences are near the glint mask,” why is the glint exclusion region not “outlined” in the difference plot? Instead, differences appear near the latitudinal extremes of only a few of the glint exclusion regions. It appears that there may be a relationship with the AOD as well, so a map of the percent change in AOD might also be informative.

This figure has been deleted, because we like Fig 12. (and Fig 2 panel H)

Figure 10: As the point of this figure is to demonstrate improved screening of high cloud, it would be more compelling if there were a panel showing the presence of high cloud (e.g., from the MYD35 product).

Now, Fig 13. Yes, maybe, but resolution is different, and the point is that the MODIS aerosol retrieval does not always feel comfortable with the MYD35 product for high cloud.

Figure 11: There's a typo in the figure caption. It should be "0.55  $\mu\text{m}$ ." See comments related to Figure 5 and the earlier discussion of Section 3.3.5.

Fixed typo. Now Fig 14. Same comments as related to old Fig 5 (now Fig 10).

Figure 13: There's a minor inconsistency in the figure caption and the axis labels. The caption gives the wavelength in  $\mu\text{m}$ , while the axis labels use nm. See comments related to Figure 7.

Now Figs 15 (AOD) and 16 (AE). Sorry, the inconsistency remains. Note that we have revised the EE from pre-C005 ( $\pm 0.03 \pm 5\%$ ) to an asymmetric ( $+0.04 + 10\%$  and  $-0.02 - 10\%$ ). We added Fig 15 panels C and D, to show the asymmetry and that 5% wouldn't satisfactorily cover the larger AOD bins.

Figure 14: This figure would be more compelling if it included at least an RGB image for visual comparison.

Now Fig 18: Point taken!

Figure 17: There's no color scale for the figure (it looks like it is cut off). The color scale itself is poorly chosen, as it is difficult to see the changes against the black background. Perhaps a white or gray background, consistent with some of the previous figures, would be appropriate.

Now Fig 24: This figure has been revised as recommended.

#### Technical Corrections

The technical corrections are presented in terms of page and line number in the manuscript as much as possible. Note that there are a number of places where comma usage for the last item in a list is done inconsistently. Also, punctuation with regard to quotation marks appears to be done incorrectly in most cases. In the interest of space, these are not noted here.

Page 160, Line 1: I'm pretty sure it's the "Moderate Resolution Imaging Spectroradiometer."

Fixed

Page 160, Line 2: The word "incredible," used to describe the dataset, is somewhat subjective and could perhaps be replaced with "extensive."

Okay, done

Page 160, Line 12: The abbreviation "DT" is not introduced, but it can be inferred from the context.

Now introduced

Page 161, Line 2: The abbreviation "DB," which appears in the "merged DT/DB product," is not introduced.

Deep blue now spelled out

Page 161, Lines 2-3: "In addition to changes to the aerosol retrieval, C6 will include ..."

Not sure what is wrong, but changed retrieval to parameters

Page 161, Line 11: The reference should be "Koren and Feingold, 2011"

Ok

Page 162, Line 8: Should it be a "constant effort," rather than a "consistent effort"?

How about "Continued"?

Page 163: Line 1: "C6 will represent a continuous, consistent data record. . ."

Ok

Page 163, Line 4: "...retrievals in urban and suburban. . ."

Ok

Page 163, Lines 24-26: The list should be separated by commas.

Ok

Page 164, Line 2: "... or updated in the C6 aerosol retrieval. . ."

Ok

Page 165, Line 8: "... in the case of the aerosol retrieval. . ."

Revised to "... is true for aerosol retrieval..."

Page 165, Line14: I think the sentence should read: "MxD04 retrievals require input L1B files. . ."

ok

Page 166, Lines 16-17: "non-cloudy"

Ok

Page 166, Line 27: The Kaufman et al., 1997 reference needs an "a" or "b" to specify which one.

Done

Page 167, Line 6: The expression "more central visible wavelengths" is ambiguous. I believe the intention is to say "longer, visible wavelengths. . ."

Yes

Page 167, Line 13: Again, the Kaufman et al., 1997 reference needs an "a" or "b."

Done

Page 168, Line 2: I appreciate the sentiment, but I question the ability of MODIS to resolve the presence of "puddles."

Changed

Page 168: Line 5: There doesn't need to be a comma in "Martins et al. (2002). . ."

Ok

Page 168, Line 7: What is the meaning of the term "arbitrarily" in this sentence? The implication is that there is no particular reason behind the choice of the thresholds over land and ocean.

Understood. Removed the term.

Page 168, Line 14: "...represent conditions under which an aerosol retrieval is possible. . ." Is the word "possible" the appropriate choice in this context, or is the intention rather to choose conditions under which an aerosol retrieval using the MODIS DT retrieval algorithm is likely to succeed?

Yes, that was the idea. Now text reads: "representative of conditions that DT aerosol retrieval can succeed"

Page 169, Line 1: Don't need a comma after "This means. . ."

Ok

Page 169, Line 26: "This means that, for both DT algorithms, the primary. . ."

Changed text to read "Thus both DT algorithms retrieve"

Page 170, Line 27: ". . . extend the range of retrievals. . ."

Changed to: "expand retrieval coverage"

Page 171, Line 1: I think the meaning of this sentence is "Although minor changes were made under the same Collection number."

Yes, but it has been removed in the revised version.

Page 171: Line 19: The acronym for "AERONET" was already spelled out on the previous page.

Removed

Page 171, Line 23: Is it important to mention that the SPs come from "moving" ships?

Probably not, although collocating with the satellite is difficult!

Page 172, Line 11: No comma after "Yet. . ."

Removed

Page 172, Line 24: Are "local computing machines" equivalent to "local computers"? What does "local" refer to in this context? My assumption is that it is NASA's Goddard Space Flight Center (GSFC).

We meant our machines, not MODAPS'.

Page 174, Lines 15-16: The sentence should either read: "Subtle differences between 'layers' and 'levels' when computing transmission functions were confused" or "The subtle difference between 'layers' and 'levels' when computing transmission functions was confused."

Yes

Page 174, Lines 21-22: "In fact" is used twice in this sentence.

True, so it is fixed.

Page 176, Line 30: "These studies show that, on a global basis, the MODIS. . ."

Not sure the problem, but sentence has been revised.

Page 180, Line 25: As written, this sentence does not make sense. I believe the issue is that the first "conservative" refers to "cloud conservative," and the second refers to "clear sky conservative."

Yes

Page 183, Line 15: ". . . the QAC reduced to 0."

Noted

Page 183, Line 18: The word “cases” appears twice in this sentence. The second instance could likely be replaced with “observations.”

Replaced with “detection”

Page 183, Line 19: “Figure 4. . .” (typo)

Not sure what is, but the text has been revised.

Page 183, Line 21: It might be more correct to say, “. . . were potentially cirrus contaminated pixels that would have been tagged. . .”

Ok

Page 183, Line 22: “. . . AOD stay farther from. . .”

Fixed

Page 186, Line 20: “. . . a larger relative error for low AOD cases.”

Sentence now reads: “Uncertainties in water leaving radiance, glint, and white foam properties would introduce an error, that is larger relative to low AOD cases, but also may be non-negligible even when AOD is high”

Page 188, Line 20: There should be a reference (Cox and Munk, 1954) to the statement that increasing wind speed diffuses the glitter pattern.

ok

Page 189, Line 20: Should probably read, “. . . (Levy et al., 2009), but we describe them here as well.”

Now write: “However, we have made changes, so we briefly describe here”

Page 190, Line 22: The sentence should probably read: “To undo the extra cirrus contamination. . .”

Yes is should

Page 193, Line 6: As discussed in the Specific Comments, the NCEP data packaged in the GDAS format for MODIS may more properly be termed an “analysis” as opposed to a “re-analysis.”

Ok

Page 194, Line 12: “. . . the large decreases (~0.04 or more) are in the midlatitudes of both. . .”

Noted

Page 194, Line 17: “. . . from 0.55 and 0.86  $\mu\text{m}$ ” (typo in units)

Ok

Page 196, Line 26: “This information is carried along in an array. . .” (no comma)

Ok

Page 198, Line 7: “. . . transition regions with comparatively low vegetation cover, but which are sufficiently. . .”

Fixed

Page 199, Line 7: “(although by definition these products should be at least as good. . .”

Removed this parenthetical statement

Page 199, Line 12: The standard NASA definition of Level 2 (L2) products is “swath” as opposed to Level 3 “gridded” products. Using “along orbit” to describe the L2 products here is confusing.

Ok

Page 199, Line 20: Instead of “Greenwich Mean Time,” should probably just use “UTC”.

Ok

Page 200, Line 3: “Analogous to **the** selection process. . .”

Noted

Page 202, Lines 2-3: This sentence reads awkwardly. To keep the overall intention, it might be rephrased as, “Since before the launch of Terra, the MODIS aerosol algorithms were designed to retrieve aerosol information at 10 km resolution (at nadir).”

Ok

Page 202, Line 6: “. . . while producing manageable volumes. . .”

Ok

Page 202, Line 28: The punctuation around the “(Remer et al. 2012)” reference is confusing.

Fixed

Page 203, Line 22: “Note that, initially for C6, the 3 km. . .”

Ok

Page 204, Line 8: “. . . requires information as to whether the ‘ocean’ or ‘land’ fork. . .” (missing “the”)

**This whole section has been moved up as Section 4.2. But the sentence has been fixed.**

Page 204, Line 16: “This means that in any area. . .”

Page 205, Line 8: Should probably read, “. . . in MODIS-Terra’s blue channel calibration to lead to artificial trends. . .”

**Discussion has been revised**

Page 206, Line 25: I’m not sure if a CDR is a “quality” or a “quantity” or something else.

**You are right, got rid of the statement about elusive, and went right to the definition.**

Page 206, Line 26: There’s a space missing before the open quotes.

**Found it!**

Page 207, Line 17: “. . . at least as good an aerosol product as the MODIS. . .”

Ok

Page 207, Lines 21-22: “. . . than that provided by MODIS.”

Ok

Page 210, Line 9: The abbreviation “ROD” was introduced two sentences previously.

Page 211, Equation (A4): Missing a closing parenthesis.

Page 212, Lines 20-21: I think what is meant by this sentence is “. . . and applied the values from table 9.5 of Jacobson (2005).”

**This entire appendix has been revised. This reference has been removed. The LBL code includes NO<sub>2</sub>.**

#### Used References

Anderson, T. L., Wu, Y., Chu, A. D., Schmid, B., Redemann, J., and Dubovik, O.: Testing the MODIS satellite retrieval of aerosol fine-mode fraction, *J. Geophys. Res.*, 110, D182004, doi:10.1029/2005JD005978, 2005.

Cox, C. and Munk, W.: Measurement of the roughness of the sea surface from photographs of the sun's glitter, *J. Opt. Soc. Am.*, 44, 838–850, 1954.

Kahn, R. A., Garay, M. J., Nelson, D. L., Yau, K. K., Bull, M. A., Gaitley, B. J., Martonchik, J. V., and Levy, R. C.: Satellite-derived aerosol optical depth over dark water from MISR and MODIS: Comparisons with AERONET and implications for climatological studies, *J. Geophys. Res.*, 112, D19205, doi:10.1029/2006JD008175, 2007.

Kahn, R. A., Nelson, D. L., Garay, M. J., Levy, R. C., Bull, M. A., Diner, D. J., Martonchik, J. V., Paradise, S. R., Hansen, E. G., and Remer, L. A.: MISR aerosol product attributes and statistical comparisons with MODIS, *IEEE T. Geosci. Remote Sens.*, 47, 4095–4114, 2009.

Kleidman, R. G., O'Neill, N. T., Remer, L. A., Kaufman, Y. J., Eck, T. F., Tanré, D., Dubovik, O., and Holben, B. N.: Comparison of Moderate Resolution Imaging Spectroradiometer (MODIS) and Aerosol Robotic Network (AERONET) remote-sensing retrievals of aerosol fine mode fraction over ocean, *J. Geophys. Res.*, 110, D22205, doi:10.1029/2005JD005760, 2005.

Lebsack, M. D., Stephens, G. L., and Kummerow, C.: Multisensor satellite observations of aerosol effects on warm clouds, *J. Geophys. Res.*, 113, D15205,

Levy, R. C., Leptoukh, G. G., Kahn, R., Zubko, V., Gopalan, A., and Remer, L. A.: A critical look at deriving monthly aerosol optical depth from satellite data, *IEEE T. Geosci. Remote Sens.*, 47, 2942–2956, 2009.

Nakajima, T., Higurashi, A., Kawamoto, K., and Penner, J. E.: A possible correlation between satellite-derived cloud and aerosol microphysical parameters, *Geophys. Res. Lett.*, 28, 1171–1174, 2001.

Shi, Y., Zhang, J., Reid, J. S., Hyer, E. J., Eck, T. F., Holben, B. N., and Kahn, R. A.: A critical examination of spatial biases between MODIS and MISR aerosol products – application for potential AERONET deployment, *Atmos. Meas. Tech.*, 4, 2823–2836, 2011.

Tomasi, C., Vitale, V., Petkov, B., Lupi, A., and Cacciari, A.: Improved algorithm for calculations of Rayleigh-scattering optical depth in standard atmospheres, *Appl. Optics*, 44, 3320–3341, 2005.

Young, A. T.: Revised depolarization for atmospheric extinction, *Appl. Optics*, 19, 3427–3428, 1980.

Yu, H., Chin, M., Remer, L. A., Kleidman, R. G., Bellouin, N., Bian, H., and Diehl, T.: Variability of marine aerosol fine-mode fraction and estimates of anthropogenic aerosol component over cloud-free oceans from the Moderate Resolution Imaging Spectroradiometer (MODIS), *J. Geophys. Res.*, 114, D10206, doi:10.1029/2008JD010648, 2009.

Wu, A., Xiong, X., Doelling, D. R., Morstad, D., Angal, A., and Bhatt, R.: Characterization of Terra and Aqua MODIS VIS, NIR, and SWIR spectral bands' calibration stability, *IEEE T. Geosci. Remote Sens.*, in press, 2012. doi:10.1029/2008JD009876, 2008.

## Response to: Interactive comment by Anonymous Referee #2

Received and published: 25 March 2013

Thank you for the valuable comments, suggestions and commentary. We understand we have a product that many people use, and we have tried to produce an accessible and user-friendly document. We hope that we have succeeded in answering your concerns.

We want to note that since this paper was originally submitted in November, 2012, there have been some updates and bug fixes implemented into the C6 aerosol retrieval code. Therefore, in the revised manuscript, some tables and figures will include updated information. In general, there are no major changes to the overall conclusions.

The manuscript provides a detailed description of the transition from MODIS C5 to C6 aerosol products. It is well written and systematically structured with sections describing general considerations, over land and over water improvements and the merging of dark target (DT) and deep blue (DB) data to provide better coverage. Also the provision of a higher resolution product (3x3 km) is described. The MS documents problems that have emerged with previous versions of MODIS aerosol products and how these were handled in C6, either by providing a better solution or giving justification to continue with the same solution as in C5. This is an excellent overview of the various steps that have been made and how MODIS aerosol products have evolved. This is very useful for both MODIS aerosol data users and scientific algorithm developers. In particular an overview is given of current and previous products and motivation for the deletion of previous products from C6 is provided. Usually such changes and descriptions are provided as a report, i.e. an ATBD. However, the risk is that such ATBD's are published on a website and not easy to trace, or eventually may get lost. Furthermore, the wide use of the MODIS aerosol data requires an easily accessible and traceable reference and AMT offers an excellent forum for this. The MS reads indeed a bit like a report but this makes it also easy to follow the different steps. The paper has become very long. However, splitting it in two or more different papers would lead to duplication in the various papers and since the authors have done a good job in outlining the various topics it fits very well together, so I recommend keeping it all together. For all these reasons the manuscript deserves publication in a peer-reviewed journal like AMT with an open discussion which can be followed by the scientific and user's communities. I noted that I was asked to provide a review at a later stage when 2 reviews had already been provided. Apparently a third opinion was needed. However, I chose to provide an independent evaluation rather than commenting on the earlier two reviews and leave it to the editor to balance these with mine. They are very good reviews and raise a number of interesting points which should be addressed. Certainly the paper will improve from these reviews.

I have only few comments. The authors have done validation for January and July in a few years. I think this is sufficient in the present context of outlining changes to the algorithm. However, certainly they will have tested each change individually and they are indeed reporting on the results. In this respect it would be good to see some more detail on testing the effects of using different aerosol models for dust or biomass (or absorbing smoke) aerosol which have been identified as problem areas (e.g. p. 186, lines 15-16). For instance an evaluation versus a limited number of AERONET data for selected sites.

Yes, we agree this would be a good idea. However, as the changes are mostly in setting geographical boundaries, we will provide a more complete evaluation at a later date.

p. 187-188: I am very happy to see that the effect of wind speed is addressed and that there is a significant effect on the over-ocean AOD. However, with only three additional LUTs with simulations for 3 different wind speeds, I wonder how these are interpolated and how effects of wind speeds exceeding 14 m/s are dealt with. Whitecap fraction increases with the cube of the wind speed (U10^3.4, Monahan, E.C., and I.G. ÓMuircheartaigh (1980), Optimal power law description of oceanic whitecap coverage dependence on wind speed, *J. Phys. Oceanogr.*, 10, 2094–2099, doi:10.1175/1520-0485(1980)010<2094:OPLDOO>2.0.CO;2).

For wind speeds much above 14 m/s, the “dark-target” aerosol retrieval becomes impossible. Fortunately, there are very few cases where the wind speed is that high, and yet the cloud fraction is low enough to retrieve. We chose not to worry about only a few cases. As for wind speed interpolation, we chose to interpolate upon wind speed using the same methodology when interpolating upon geometrical angle. The AOD error for 1 m/s error in wind speed is not large. Plus, using 1° GDAS data is crude anyway. We added this reference in the paper.

p. 192, line 9: I can hardly imagine that any case could exist with "no aerosol" or AOD=0. At the end of the page is indicated that in C6 cases with no retrieval are given QAC=1. However, if there is no retrieval possible, and AOD is set to 0, this data should not be used and QAC=1 would be misleading. QAC should be either 0, or maybe even a -1 should be introduced for cases where no retrieval has been made. Such values should certainly be discarded from any further processing. I emphasize that I am not advocating to discard all AOD values < or = 0, when they indeed results from a retrieval. Only those cases where retrieval is not possible, for whatever reason, should be flagged and discarded from further analysis.

Not "no aerosol", but can be very low. And the uncertainty in retrieving 0.01 is really  $\pm 0.05$ . We added a small discussion here (based on reviewer #2), in that cases of  $1.1 * \text{Rayleigh}$  (at  $0.86 \mu\text{m}$ ) are really of order 0.01 AOD (at  $0.55 \mu\text{m}$ ). Fortunately, these retrievals are rare over the ocean (as indicated by histograms).

p. 194 raises an interesting question as to the cause of the large decrease of the order of 0.04 or more in AOD in the Roaring Forties. Increasing the wind speed from the previous default value of 6 m/s to the actual values (which I presume are generally much higher than that) would increase AOD (more production of sea spray aerosol) while at the same time increasing whitecap fraction. But in current sea spray source function models the production is linear with whitecap fraction, and hence it must be mixing which reduces sea spray aerosol concentrations to a less than cubic function of wind speed (as often observed). Does the new cloud screening also affect the retrieval in these areas? Have tests been made to this effect to see the effect of each of these changes (i.e. cloud screening vs wind speed)?

There is absolutely no statistical difference in retrieved AOD due to new cloud screening in the Roaring 40s. Using the wrong wind speed was one reason for the high bias in this region, which has been fixed for C6. The other was that too many of the low AOD retrievals were thrown out, which has also been fixed for C6. It would be interesting to have a conclusive study on the effect of sea spray on AOD.

I have a general question on the merging of different aerosol products provided from different algorithms but the same sensor. Apparently the QAC is used to determine which product is best. Have tests been made to determine the continuity of the aerosol products between adjacent pixels from different algorithms, and how this affects the various quality assessments?

The QAC product (at least for the Dark Target retrieval) is designed to measure the success of the aerosol retrieval, not necessarily its accuracy. However, they are correlated. Unfortunately, there are few AERONET sites in the semi-arid "merge" regions of the world, so it is impossible to tell whether one product is better or another in these regions. All we have right now is subjective understanding of how the two products behave. Fortunately, even on a granule scale, the merge product looks continuous. The averaging smooths the transition. When C6 is produced operationally, there will be enough data to make a quantitative evaluation.

Figure 1: the lower left shows the difference new-old; however, new goes to larger sza, and hence further south, to -60. However, the difference plot should then also extend to -60 and show the actual AOD from blue as plotted in top right. Why is that not the case here? Figure 11 shows a histogram over ocean that goes up to AOD= 5. However, over land (Figure 5) the maximum AOD is 3. How can AOD over ocean be higher than over land where the aerosol sources giving rise to high-AOD sources are situated? And why are there no negative AODs over ocean?

The difference was plotted only for where both old and new retrieve. Now (new Fig. 1), there are separate panels that also show "new" coverage for C6.

The maximum AOD over both land and ocean is 5. However, we rarely see this, so doesn't show up on histograms. AOD can be higher over ocean because the contrast over ocean may be better for doing retrieval. Over land, high AOD may be correlated with a source that is so spatially variable (in reflectance units) that it is not retrievable. The new cloud mask over land helps detect a few more high AOD cases than before. There are no negative AODs over ocean, because of a "retrievability constraint" (reflectance in  $0.86 \mu\text{m}$  must be greater than  $1.1 * \text{Rayleigh}$ ).

I noted a few typos, but there may be more; and some minor questions: (page, line no) 171, 1; had been made

179, 1: should be SSA > 0.95 for low absorption

Yes

180, 23 what do you mean with 'heavy' aerosol: high concentrations or AOD? Or large large particles?

Heavy aerosol loading

183, 19: Figure 190, 22: I guess that continuation should be contamination?

Yes, fixed

192, 3: C5 over ocean MODIS . . .

200, 2: I think the reference is Levy et al., 2009a

Yes

207, 16-17: the sentence reads a bit strange, could it be written as "the VIIRS algorithm should provide at least as good an aerosol product as the MODIS algorithm"

Noted

207, 21: should it really be ensure, or something weaker like "the VIIRS aerosol record may be somewhat different from that has been . . ."

Ok, "that suggest that the VIIRS aerosol record will not exactly follow MODIS"

208, 1: something missing, should it something like: ". . . VIIRS shows up in the global climate data record"?

210, eq A1: does atmospheric pressure, i.e. gas pressure, play a role here?

It could, but a 5% variability in nominal sea level pressure is not significant.