Atmos. Meas. Tech. Discuss., 4, C1596–C1602, 2011
www.atmos-meas-tech-discuss.net/4/C1596/2011/

Atmospheric Measurement Techniques Discussions

Interactive comment on "Where do we need additional in situ aerosol and sun photometer data?: a critical examination of spatial biases between MODIS and MISR aerosol products" *by* Y. Shi et al.

Anonymous Referee #7

Received and published: 14 September 2011

General Comments ==========

This paper describes an excellent concept for assessing the shortcomings of different aerosol retrievals from space by (a) comparing satellite retrievals to AERONET data and (b) cross-comparing MISR vs. the MODIS Dark Target and MODIS Deep Blue products. It is an inspiring read and a very important contribution to the field. It starts out with the more traditional approach of a satellite product validation using AERONET, and continues with a consistency analysis of MISR vs. MODIS products.

Other than previous studies where, e.g., level-2 products or level-3 products from different satellites are compared regardless of the different data sampling, filtering, cloud mask etc., the comparison here makes an effort of a more quantitative comparison by (1) separating MODIS DT (the 'standard' MODIS algorithm) and MODIS DB and (2) by doing comparisons within 6-hour data windows per 0.5x0.5degree grid-box only if both MODIS and MISR provided a valid aerosol retrieval. This is a different, and probably more promising, strategy compared to a comparison of climatologies derived from individual satellite instruments because it allows a closer look at the details. Another compelling aspect about this study is the effort to engage the community. This is done by offering, e.g., kml files such that the user can visualize the study's results in different ways. This could be a model for future publications of this kind.

Despite the excellent concept and novelty of this paper, there are a few shortcomings which can be addressed in minor to moderate revisions.

(1) Using rˆ2 as an indicator for the goodness of a correlation in Figure 1, Table 1, and Figure 2a/b is the appropriate tool in statistics. However, and this is a common misconception, it does not necessarily reveal the level of confidence in the products without taking into account the number of data points and their distribution across the parameter space. To put it bluntly: Fitting a line to two data points gives rˆ2=1, but this is meaningless. Conversely, fitting a line to a lot of data points that are quite well correlated may produce a lower rˆ2 the more points are added. rˆ2 is only the first step in the regression analysis. In order to make a statement about the consistency of two datasets, this needs to be translated into a a confidence level of some sort within the framework of statistical analysis. Since the authors introduce such a nice new sampling concept, it is somewhat disappointing that they then resort to the rather rudimentary approach of using rˆ2, and only that, as indicator, since the aggregated data offers much more with little effort.

(2) The paper argues in numerous places that somehow, a large intercept (sometimes wrongly referred to as 'interception') between retrievals is indicative of "problems with

AMTD

4, C1596–C1602, 2011

surface reflectance", whereas a slope different from unity is indicative of "microphysics" (i.e., aerosol model) problems. While this makes sense intuitively, and is probably (hopefully) described in one of the cited papers, it should at least be mentioned on what grounds this assertion can be justified.

(3) The language is highly 'heterogeneous' across the manuscript. Some of the sections, especially, the introduction, are brilliant; others are, in the worst places, sometime incoherent and are crowded with awkward sentences or plain mistakes (examples given below).

In sum, while the paper shows the pathway toward a new way of looking at aerosol retrievals from space (and how they can be improved with ground-based observations), it is still a little bit away from introducing a quantitative metric for the goodness of satellites retrievals. It may be beyond the scope of this manuscript to introduce a more formal way of comparing retrievals, in which case the weakness of the approach here (i.e., rely on simple r^2, point (1) above, and point (2) above) should at least be mentioned in the revised version.

Specific Comments in sequential order: ========== p4297,l25:"The more complicated land surface also reduces the degrees of freedom in available microphysics models." - This needs to be explained. Why should that be?

p4299,l7: '0.03+-0.15*AOD' - unclear what this means - is 0.03 an absolute error, or 15% a relative error? Should there be a bracket around 0.03+-0.15? In essence, why are there two numbers for the error? On the other hand, "0.05 or 0.2*AOD" on line 3, page 4300 makes much more sense.

p4301: Fig 1: It is probably explained somewhere, but what's the date range of the AERONET-satellite comparisons? The section before only talks about satellite data, not AERONET.

AMTD

4, C1596–C1602, 2011

l23: "Scores" need to be introduced - slopes? intercepts? correlation coefficients?

l27: What is a 'reasonable' correlation?

p4302,l2: 'areas dominated by different aerosol species' - What are those areas?

p4302,l6: 'The influence of lower boundary conditions is less evident in MISR-AERONET than the MODIS-AERONET comparisons.' – How so? Which parameter distinguishes the impact of surface vs., say, aerosol type on the bias between satellite vs. AERONET? At some point later in the manuscript, it is mentioned in a side note that this is done by looking at slope and intercept as indicators (this should be better explained). On that note, what is the explanation for MISR performing sometimes "better" and sometimes "worse" than MODIS when compared to AERONET in Table 1? Since MISR does better with surfaces (even if they are bright), it should, at least following the theory brought forward later, have smaller intercepts, provided that the surface around AERONET sites is perfectly well known. Another question on that note: Was any segregation of data by type done, based on the AERONET data? And if so, where is it shown? Such a segregation would (arguably) be a better way of deciding whether 'the surface' or 'the aerosol type' has the largest impact on MODIS/MISR/AERONET discrepancies.

l11: What is an 'insufficient number of data points'?

p4306, l2-5; l6-8: Here is the core description of 'high intercept' = 'surface problems' vs. 'slope' = 'microphysics'. (major point number (2) above). This needs to be explained - not even a reference is given here. If this is indeed true, demonstrate how this show differently with MISR/AERONET vs. MODIS/AERONET because under several circumstances, MISR should perform better than MODIS because MISR does better for bright surfaces than MODIS, and it actually retrieves aerosol type along with other aerosol parameters, as opposed to MODIS where it is fixed and relies on a climatology - or is this a misunderstanding. It might be worthwhile pointing out these differences in the operational MODIS and MISR algorithms were explained briefly in the beginning.

AMTD

4, C1596–C1602, 2011

Even if most reader are familiar with it, it doesn't hurt to repeat.

p4307,l12-18: How were the regions segregated into the different 'problem groups' defined above? Certain parameters, a combination of parameters, or by empirical arguments?

p4310,l24-25: How so - how were the regions identified? (Relates to a number of previous comments as well.)

p4311, l10-12: If problems are related to surface reflectance, how would AERONET sites help then?

l17: 'Our regressions show that...are robust.' - What is robust? What does 'robustness' mean. Avoid statistical slang and say what parameter this relates to - i.e.: high rˆ2 etc. (although, as pointed out above, rˆ2 alone is not necessarily sufficient!)

Technical/Language Comments: ==========

p4296: What is "greater south america"?

p4297:l7: 'surface-observtion-data-poor regions' - excessive use of hyphenation looks awkward

l10: Check 'Complicating...bias' - something missing? word order?

l20: conditionS –> conditions

p4298: l8: bias –> biases?

l14: 'causes of the discrepancies should be collected' –> awkward, collecting not the appropriate word.

l15: 'which aims to help direct' –> awkward

l29: 'heterogeneity ' should be 'discrepancy' or 'inconsistency'

p4299,l4: hyphenate near-UV, near-IR, fine-more, near-surface (farther down)

AMTD

4, C1596–C1602, 2011

l2: can be very effectively used: word order?

l14: retrieving process –> retrieval process

p4301: This is where most language problems begin.

p4301, l1: "our results begin by presenting..." does not appear to be correct language; the entire paragraph should probably be reworded, it sounds awakward.

l19,20: replace 'wavelengths' with 'channels'?

p4302,l25: 'Satellite products correlate well' - with each other? with AERONET? What's a 'good' correlation?

l27: 'The question now becomes' –> 'The question then/therefore becomes'?

p4303:l3-l8: Bad explanation of what's shown in the Figure. It took many times to understand. Also, the caption of the figure is not helpful: What's 'MISR AOD that corresponding to operational MODIS DT' - this is grammatically incorrect and doesn't make sense. During the first read-through I thought that Figure 2a is the *average* of the MISR and MODIS retrieval - which obviously is not true. But that's the way it is described somewhere in the text.

l7-8: 'Shown in Fig. 2a, the commonly ... are visible.' awkward

p4304: l29: 'ratio values' –> 'ratios'

p4305,l18: 'The uncertainties...due to ratios from small values...' Sounds awkward, also needs to be expanded.

l22: 'data have MISR AOD values' - sounds like the data own the values - there should always be a better word than 'have'

l23: Please change all occurrences of 'interception' to 'intercept'.

p4307,l4: 'In summation' - one of the avoidable English mistakes – it should be 'In sum'

AMTD

4, C1596–C1602, 2011

l29: 'at the visible spectrum' –> over? throughout?

p4308,l3-7: run-on sentence

l17-19: 'The use of...retrievals.' does not make sense. What does it mean?

p4310,l2: 'indexes' –> 'indices'

p4311, l1: 'The AERONET has data' - fix English

l2: 'helps address the Arabian Peninsula' - what is addressed? need to state a problem, not a region.

l3: 'invite further experiments' - awkward; people invite, data are not alive.

l2: Break out 'UAE' (United Arabic Emirates?)

Interactive comment on Atmos. Meas. Tech. Discuss., 4, 4295, 2011.