Response to Review from Michael Fromm

We thank Dr. Fromm for his careful review of our paper. Our responses are embedded below.

In my report I use "auth" to refer to Colarco et al.

Auth focus their study on satellite observations of aerosol optical thickness (AOT), particularly the question of strategic spatial sampling reduction, with respect to a MODIS- like full swath, and its impact on the value of the resultant AOT data for climate studies. They choose AOT for its relevance as a primary metric in the assessment of direct aerosol radiative forcing (DARF) of climate. They adopt MODIS as the standard for global satellite remote sensing of AOT because of the existing decade-long data set, large swath width (synonymous with global spatial coverage), and the maturity level of aerosol retrieval. Auth undertake a statistical analysis of 10 years of Aqua MODIS level 2 AOT data aggregated onto regular lat/lon grids of various size. From those grids they are able to perform global/regional 1-year (2010) analyses, time series, and trend analyses for defined regions. The limited-area swaths they test are intended to mimic MISR-like narrow swaths or to approximate the real estate covered by CALIOP (what they refer to as "curtain" mode). Auth conclude that spatial sampling matters in the qualification of an AOT data set for DARF assessment. In particular they conclude that a MODIS-like AOT in curtain mode is inadequate for climate studies, however they're less definitive on the limitations imposed by MISR-like narrow swaths.

Spatial sampling of AOT is without doubt an important consideration for qualification as a climate data record. Hence it is an appropriate topic for AMT. Moreover, considering the remarkable MODIS record dating to 2000, it is compelling to evaluate its position as a possible standard for global AOT data. However, my assessment is that this paper has approached their science question in a largely unsupportable way, significantly limiting its value. Unfortunately, I cannot recommend publication of this paper in its present form. My major reasons are given below.

We would like to clarify that it is not the intention of this paper to establish MODIS "as a possible standard for global AOT data." Our choice of the MODIS dataset is pragmatic: it is the best available long time series, wide-swath, nearly global aerosol data set. However, see later, we took your suggestion to look at this problem from the point of view of an aerosol transport model seriously, and so we present as well that complementary analysis.

Additional comments and suggestions are provided as comment bubbles in the pdf of the manuscript.

Motivation of Study

A singular motivation for auth's paper is explicitly given in the introductory paragraph starting on line 57. "Spatial coverage is among the primary considerations for any future satellite instrument designed to measure aerosols. Given technological and budgetary constraints, trade-offs are made between spatial coverage (i.e., measure ment swath width) and other instrument measurement characteristics, including the number of spectral and polarized channels, relative precision and accuracy, angular and temporal coverage, and pixel size." Since the entire justification of this research flows exclusively from this statement, auth are obligated to reinforce this statement with citations. Without objective reinforcing information, this statement amounts to informal speculation. I.e. auth need to build a compelling case here, identifying real assessments of future satellite concepts w.r.t. spatial sub-sampling as a "primary consideration" for decision-making. If indeed they cannot cite firm evidence of swath width as an acknowledged trade-off against other instrument factors, the sole motivation for this research evaporates.

This study derives from our involvement in a science working group (SWG) that NASA empanelled to respond to the NRC Decadal Survey recommended ACE (Aerosol-Clouds-Ecosystems) mission. The Decadal Survey description of ACE called repeatedly for across-track scanning capabilities. The NASA ACE SWG document attempted to refine the requirements on that question, but there was not a large body of literature to draw on. Hence our initial work. The topic has as well been taken up in Geogdzhayev et al. (2013). We have substantially rewritten the introduction to make this motivation more explicit and cite appropriate literature more completely.

Even if auth can convince that spatial coverage is a factor in competition with other instrument design elements, they must acknowledge that the trade-offs add value to the whole content of aerosol measurement (e.g. MISR's stereographic views of scattering and aerosol height). This fact further diminishes the motivation to study only swath width and the importance of any results so derived.

We make clear the trade space to be explored, and that spatial coverage is but one issue, writing:

"Given technological and budgetary constraints, trade-offs between the benefit of spatial coverage and other desired instrument characteristics must be considered, such as the number of spectral and polarized channels, relative precision and accuracy, angular and temporal coverage, and pixel size."

And in our conclusions we write:

"Other measurement approaches would enhance retrieval of aerosol properties over bright land surfaces. Future aerosol instruments will undoubtedly improve upon MODIS in these and other respects, such as providing enhanced information about aerosol single scattering albedo and particle size, other important drivers DARF (e.g., Loeb and Su, 2010). It was not our intention here to demonstrate the benefit of those enhanced capabilities, but rather to investigate the limits imposed on the measured AOT by one aspect of any future measurement strategy, its spatial coverage."

We are of course not impugning the utility of narrow swath but enhanced capability instruments. However, we confine ourselves to one aspect of that trade, to address the specific issue raised in the introduction.

Auth need to strengthen their stated motivation and add additional background to justify the research plan and results restricted to swath-width considerations. Otherwise this work comes across as an arbitrary exercise with results having minimal impact.

See the above responses.

Method

Auth are looking strictly at artifacts introduced via reduced spatial sampling of AOT. In the Introduction they state: "If spatial sampling artifacts introduce sufficient uncertainty in the satellite-derived AOT, we will not be able to meaningfully improve estimates of DARF." So they undertake to isolate this factor (narrower swaths with respect to a full MODIS image) to answer the question of whether certain image-swath subsets are insufficient for climate sensitivity studies. They test the subsamples on actual MODIS AOT, with all of its well documented issues. An example of just one issue is illustrated in their Figure 1c. Ms. Lisa's picture is significantly disfigured by clouds. The ubiquitous issue of clouds adds a large unknown to the puzzle of true AOT and trends thereof. As far as I can tell from what is presented here, auth do not acknowledge that the grid-point AOT averages are only representative of the clear-sky areal fraction (as do Remer and Kaufman (2006)). With or without this acknowledgement, it is essential (in my assessment) to discuss the implications of cloud cover and regional/temporal changes thereto in the context of the "true" vs. MODIS areal/temporal AOT variations. This and the many other MODIS issues, some discussed by auth, impose a high level of uncertainty regarding a true representation of AOT that may well dominate spatial sampling considerations.

We do; in Section 2.2:

"AOT is retrieved in the daytime portion of the MODIS orbit under cloud-free and glint-free conditions using separate aerosol retrieval algorithms for ocean (Tanré et al., 1996, 1997) and land (Kaufman et al., 1997; Levy et al., 2007a, 2007b)."

Furthermore, we make the following qualification:

"The AOT dataset so obtained is representative of the clear-sky (cloud-free) aerosol distribution, and it should be understood that a portion of our subsequent spatial sub-sampling (see next section) is entwined with the spatial sampling of clouds."

Possibly some next-generation sensor can resolve total columnar AOT in the presence of clouds (i.e., aerosol embedded within or layered above and below clouds), but this has not been demonstrated yet (OMI retrievals of AOT above clouds and CALIOP retrievals of aerosols near clouds and above clouds do get at this issue, but both have issues retrieving aerosols below clouds, for example). The presence of clouds is implied in our analysis of the aerosol transport model results in subsequent sections.

It is fundamental from a statistical standpoint that reduced sampling will increase uncertainty. This can be formulated and demonstrated one theoretical grounds.

Auth have already decided that they are basing their work on an imaging instrument identical to MODIS. MODIS by no means has been established objectively as a satisfactory AOT data set either for short-term, regional, or climate applications. The strongest defendable affirmation is that MODIS AOT is the best we have (and there might not even be consensus on that point). As valuable as MODIS aerosol data are, they can by no means be generalized as "high quality" as auth claim in the Discussion and Conclusion section (line 632). Their own discussion, and papers they cite, fully reveal the several weaknesses inherent in the MODIS data; scan-angle systematics, cloud contamination, cloud clearing, minimal diurnal sampling, land/ocean differences to mention some. Hence, to achieve auth's stated goals, applying a statistical technique to real MODIS data does not produce convincing results. To illustrate that contention I refer to the discussion of Figure 12 after line 445. Here auth compare their map of AOT trends with Figure 7a of Zhang and Reid (2010). Auth point to only one region of differing result (lines 449-451), the Pacific Ocean west of Mexico. However, there is a much larger area of opposite sign of AOT trend in the region encompassing In- donesia, New Guinea, and the western Pacific northeast to Japan. This is not meant to criticize their focus, but rather to point out that indeed large, coherent regions have full-swath-derived MODIS AOT trends of opposite sign between two thoughtful, rigorous, independent AOT analyses. My conclusion, after comparing auth's and Zhang and Reid's maps, is that factors other than swath width likely dominate an analysis of AOD trends, regional and global.

The choice of the MODIS dataset for our analysis is based on the pragmatic consideration that it is, as you say, the best we have: a long time series, wide-swath dataset of global aerosol properties. Indeed there are issues with the dataset, and we deal explicitly with one that is relevant to the current study (the view angle biases) as best we can. With the MODIS data, the question we are addressing is how well narrower-swath instruments can reproduce even the MODIS AOD pattern and trend results, which represent only the best available representation of the actual AOD distribution for this purpose. But we also address this concern using another dataset, derived from an aerosol transport model simulation (see below).

The comparisons of our derived trends to those from Zhang and Reid (2010) need to be understood in the context of looking at two somewhat different (though not unrelated) datasets: theirs is their "data assimilation grade" dataset, which attempts to corrected biases present in the MODIS dataset using AERONET and MISR data, and is for a different period of time and is from the Terra MODIS instrument; ours is derived from the Collection 5.1 Aqua MODIS dataset.

In my view, this type of approach to answering the science questions auth pose would be substantially more convincing if it invoked synthetic MODIS data. Here all the unknowns and hard-to-disentangle issues (like scan-angle AOT artifacts) could be eliminated, and replaced with a synthetic model of AOT that replicates regional sources and introduces a specified trend (regional or global). This model AOT atmosphere, sampled by real MODIS swaths (and strategic subswaths), would give the analyst the most defendable tool for making a convincing assessment of spatial sampling artifacts on AOT patterns/trends.

We took this suggestion to heart, and it is the main reason for the lengthy delay in our response to the initial reviews!

We had of course considered this idea, but it is also not straightforward. After all, we are trying to make some assessment of the behavior of a real observing system looking at real aerosols. Whatever its flaws, that's what MODIS is doing. On the other hand, a synthetic dataset derived from an aerosol transport model has to hurdle the tall pole of how realistically it captures aerosol variability, so that conclusions drawn from it are relevant to real-world applications. As such models are notoriously poor at representing natural variability, this is not a trivial task, and goes far beyond making simple comparisons of monthly mean model output to MODIS.

So we have added to the paper a complementary analysis from a recently made aerosol reanalysis. The model results presented are based on the NASA GEOS-5 Earth System Model running a version of the GOCART aerosol module. Aerosol information is assimilated in the model based on a "data assimilation grade" AOT product derived from MODIS. We analyze model output that is sampled on the Aqua track, but extracted from the model immediately prior to the AOT assimilation step, so that the model fields shown in this work have not seen the satellite data for about 24 hours. A lot of this is new stuff, and a full evaluation of this system is far beyond the scope of the current paper, though the Supplementary Material includes the publication track record for the model itself. The aerosol assimilation is a relatively newer piece of work. We show comparisons of the model to AERONET data by way of providing some initial tests (Figure S6), but further validation of the assimilation system is a (separate) work in progress. Some preliminary evaluation of the aerosol assimilation can also be found in references cited in the Supplementary Material.

The results of the model analysis complement and extend the analysis of the satellite dataset. Spatial sampling artifacts in AOT are found in all the regions analyzed, and are comparable in magnitude to those found in the satellite dataset (see Figures 9, 10, and S1). Furthermore, we analyze the model-computed aerosol radiative forcing (we show specifically SW TOA all-sky aerosol forcing), and find a spatial-sampling artifact in forcing (that is, the uncertainty in seasonal-regional mean aerosol forcing due to spatial sampling) that is about 1 W m-2, and sometimes as high as 2 – 3 W m-2.

So, the spatial sampling artifacts we identify in the MODIS dataset are also present, and similar in scale, in the model fields.

Results

My perception of auth's results is that there is one clear signal: curtain measurements won't cut it for DARF applications. However, their rendition of curtain measurements is nothing more than MODIS AOT along a very narrow swath. They argue against some hypothetical curtain measurement techniques with presumptive statements about a lack of information on surface reflective qualities (line 630). Referring back to the above-mentioned concern about unsubstantiated future satellite missions, statements like the one on line 630 echo as empty speculation/argumentation. This and other curtain-centric concerns are already seen to be mitigated when considering that an existing curtain instrument, CALIOP, has synchronous visible and IR imagers, and also makes aerosol measurements day and night, twice as often as a MODIS curtain would. Hence the simple conclusions stated online 630 come across to me as prejudicial and speculative.

We address this concern in part by inclusion of the model simulation analysis. The conclusions have also been substantially rewritten, and we hope this addresses the issue raised regarding the tone of this section.

As elaborated on in the comment bubbles imbedded in the manuscript, there is also some unnecessary difficulty in evaluating curtain-analysis map panels in Figures 4 and 13. Auth's interpretation of these figures suggests that the curtains completely miss large or numerous areas. However, close-up viewing shows that these findings are not accurate. It still may be true that the curtain sampling considerably degrades the detection of regional/temporal AOT features, but it is also true that the color schemes employed and the representation of the whole globe in a small plot area smears out features that argue against the written interpretation. I suggest that auth experiment with

zoom views and other color schemes to give more contrast and detail than is presented here. In addition, I suggest that a clear and objective test be applied, over and above eyeballing maps, to evaluate whether or not curtain sampling sufficiently captures regional/temporal AOT variations.

Most of the figures in the paper have been revisited with the reviewer's comments in mind, and we hope their clarity is enhanced.

Embedded Comments

Line 11: I don't know why auth chose a curtain sub-sample alternative. Except for the fact that its footprint mimics CALIOP, what is its significance? Presumably an instrument that would be operating in that mode would be vastly different than MODIS in terms of its aerosol measurement modus operandi, so this comparison seems to be a lot of wasted space.

The Glory APS instrument would have provided a single pixel wide, along-track scan of the global aerosol field. Whatever its capabilities, the question of the impact of such reduced spatial sampling in deriving statistics about global or regional aerosol remains. Geogdzhayev et al. (2013) addressed this same question, and we are offering an alternative analysis. From the first sentence of their abstract: "We examine likely effects of pixel-wide along-track sampling on climatological means of aerosol optical thickness (AOT) derived from observations with satellite instruments such as CALIOP and APS by sub-sampling AOT retrievals from a wide-swath imaging instrument (MODIS)." So this seems to us to be a reasonable sampling to consider.

Line 24: This statement presumes that a true trend signal is achievable and available to this study. No evidence of that "truth" data set is presented in this paper.

Several authors have already published AOT trends derived from the MODIS data set. Zhang and Reid (2010) do a careful job of this. Others are less considered. In any case, we don't claim that the trends derived truly represent Earth, but instead are assessing whether even those trends could be derived from a reduced dataset.

Line 25: Maybe there weren't any.

The curtain-like sub-sample is a subset of the full-swath dataset. The full-swath dataset has the widest coverage, and so has the most information from which to derive and assign significance to trends. The point is that statistically meaningful decadal-scale regional trends can be derived from the full-swath data, whereas with the reduced spatial coverage of the curtain subset, there are not enough observations to assign statistical significance to the trends.

Line 58: This paragraph is meant to motivate this paper's analysis. But it has no weight (above auth's speculation) due to a total absence of citations or other quantifiable documentation showing spatial sampling as a dilemma for future instruments. Auth need to present evidence of real budgetary and/or technological hurdles confronting planners of future satellite missions.

We have bolstered the motivation for the paper in the introduction.

Line 59: Auth give no accounting for the fact that we have over a decade of satellite AOT data that is still absent an objective, accurate benchmark for AOT. Space- and ground-based AOD are both inherently confounded by cloud, the cloud-aerosol boundary, and optically thick aerosol.

AERONET is the best benchmark available. MODIS (and other instruments) are necessarily tied to AERONET because of its quality. How MODIS compares to AERONET is well documented in the references included in our text.

Line 128: What is the rationale for using different criteria?

The assigned QA flags provide a means to further subset the MODIS retrievals. The assessment of which QA flags to use where is derived from comparisons of ocean and land retrievals to relevant AERONET sites. It is found that QA = 1, 2, or 3 compare equally well to AERONET sites for the ocean retrieval, whereas for land, QA = 3 compares the best. So these become the QA recommendations. We have added the relevant citations for this issue.

Line 206: I disagree with this characterization. I still see the Russia AOT, but it is simply more speckled, and the gray background mutes the contrast. But at a suitably large zoom the AOT shows up clearly.

Agreed. We write: "Other features, such as the 2010 Russian fires (Witte et al., 2011), still exhibit similar magnitude of AOT across sampling strategies, but their spatial coherence is quite different. "

Line 219: It is important for auth to qualify "large." As I see it the largest differences are on the order of 10% of the mean AOT. How does this compare with the true geophysical variability and/or the AOT uncertainty?

This section has been rephrased in the new text.

Line 434: "decreasing trends" describes, or at least implies, a second derivative. The figures show a first derivative. Here and elsewhere, it would be preferable to use something like "negative trend" and "positive trend."

Thank you for the suggestion, we have cleaned up the text accordingly.

Line 450: There is a more substantial area of opposite-sign difference, Indonesia east to New Guinea and north all the way to Japan.

We write: "Our trends differ from theirs primarily in the Pacific west of Mexico, where we show a slight positive trend and they show a weak negative trend, and as well in the western Pacific and near the Maritime Continent where they find a weak positive trend and we find a weak negative trend."

Line 450: "positive" instead.

Corrected (see above).

Line 490: Of course the number of blue pixels is largely reduced in the C1 map. However, the characterization given here is incorrect. Certainly that is the impression portrayed by this small map of the whole world, with pastel blue and gray competing for visibility. If one zooms in on the map one sees that blues are speckled across the globe, over land and ocean. It would be highly desirable to make the color of the >95% pixels truly distinctive, e.g. red, or otherwise reduce the competition for viewing these pixels.

We have revised the figures in question to select a more visible color scheme.

It is conceded that the coverage of statistically significant pixels is reduced, but it can be made readily apparent (by the above suggestion) that these are spaced well enough around the globe that every portion outside the gray areas of Figure 13a is represented. This will then raise the important question of how well this coverage captures regional aerosol phenomenologies. It is not clear to me that even this sparse (but regular) spacing is insufficient to achieve regional aerosol monitoring.

We have revisited this point. In addition to missing some regional-scale aerosol patterns, substantially reduced sampling fails to capture the extremes of the AOT distribution, and does not provide statistically significant trends on decadal timescales. In our conclusions we write:

"Spatial sampling affects the derived magnitude and assignment of statistical significance in aerosol trends. Along-track curtain sampling results in reduced trend magnitude and essentially eliminates statistical confidence in the derived, decadal-scale trends when the data are aggregated at high spatial resolutions. Trend magnitudes and statistical significance were more similar to full swath values for the narrow-swath sampling.

"Aerosol trends and statistical significance were found to be similar across sampling strategies when the trends were composed from coarsely gridded aggregates of the sub-sampled MODIS AOT data, suggesting that single pixel width sampling may be sufficient to detect and attribute trends at spatial scales of order 1000 km."

Line 497: A "Discussion and Conclusion" section should not contain new analysis. The cross-swath analysis belongs in another section prior to this one.

We have rearranged the paper as suggested. The across-track sampling is presented in Section 2.5, and no new analysis is presented in Section 4 (Discussion and Conclusions).

Line 541: The Geogdzhayev et al. work has little do do with auth's paper, which is motivated exclusively by the idea that future satellite mission designs would employ narrower (along-track) swaths than MODIS. There is no evidence given in this paper for a future aerosol measurement strategy along the lines of the Geogdzhayev subsampling. It seems that Geogdzhayev's intention was specific to MODIS and its inherent aerosol data quality issues. Assessing their approach and findings does not pertain to the sole motivation of this paper.

We disagree. Although Geogdzhayev et al. (2013) certainly do address issues inherent in the MODIS retrievals, the first line of their abstract makes clear that their intention is to assess single pixel wide sampling further on its own merit. They write: “We examine likely effects of pixel-wide along-track sampling on climatological means of aerosol optical thickness (AOT) derived from observations with satellite instruments such as CALIOP and APS by sub-sampling AOT retrievals from a wide-swath imaging instrument (MODIS).” The point of their across-track sampling strategy was to (as much as possible) reduce the scan angle bias to answer their question. Their conclusions about the global, annual mean AOT are supported in our analysis (see our Figure 4). But we go further, and look at the seasonal and regional AOT, and the significance of the associated trends.

Line 609: It does nothing of the sort. It only calls into question what a single-scan-angle MODIS data item would provide. Even the existing CALIOP "curtain" measurement provides double the MODIS sampling per day.

We investigated the seasonal-regional AOT and trends from several perspectives, including taking the reviewer’s suggestion to incorporate results from an aerosol transport model. What we have found consistently is that there are spatial sampling artifacts; that is, uncertainty in the seasonal-regional mean AOT which results from the spatial coverage of the observing system. This uncertainty would impact assessments of loading, forcing, or trends.

Line 616: This isn't necessarily true. Why does a bias have to result when sampling is decreased?

Our use of the word “bias” was imprecise. We state now “sampling related uncertainty.”

Line 618: This is not a new finding, nor is it disputable. But what is the point?

This sentence is removed.

Line 620: Of what relevance is that? Is there some instrument being pushed that will be a curtain-limited MODIS?

This sentence is removed.

Line 628: Where does this number come from?
For the reader to know how to interpret this, it is essential to know what percent is deemed sufficient.

This sentence is removed.

Line 630: On what do auth base this statement that any hypothetical curtain instrument would not observe the underlying surface? They give no evidence.

This sentence is removed. But the point, made elsewhere, is that to calculate DARF, you need at least AOT, aerosol type, and surface albedo (possibly also aerosol vertical distribution), and as surface albedo varies on kilometer spatial scales, having wide coverage (e.g., more than ~10% of the planet from curtain sampling) is required to significantly reduce DARF uncertainty.

Line 632: What is meant by "high quality" and how is that determined? What paper(s) show conclusively that MODIS AOD is of high quality?

This sentence is removed.

Figure 5: It would really help the reader to have the ordinates consistent between a and b.

Fixed.

Figure 12: Label "-0.3" has the wrong sign in each panel.

Fixed.

Figure 13: The caption and color bar need more explanation. E.g. what is the meaning of the "0" and "2" on the color bar? What is the meaning of gray vs. white?

Fixed.