

Atmos. Meas. Tech. Discuss., referee comment RC1
<https://doi.org/10.5194/amt-2022-303-RC1>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on amt-2022-303

Anonymous Referee #1

Referee comment on "Aerosol Optical Depth Retrievals over Thick Smoke Aerosols using GOES-17" by Zhixin Xue and Sundar Christopher, Atmos. Meas. Tech. Discuss.,
<https://doi.org/10.5194/amt-2022-303-RC1>, 2022

This paper shows case studies for a retrieval of smoke aerosol optical depth (AOD) from GOES observations during intense wildfire events in the western USA in September 2020. The motivation is that the operational GOES AOD product (here, OAOD) often misses these events (the implication is that they are flagged as cloudy). The motivation is important and the topic is relevant to the journal, because extreme events are important for a variety of purposes and are not well-captured by the journal. The results from the proposed technique are compared to OAOD, ground-based AERONET sites (at two locations), and the MODIS MAIAC product.

The method is a simple one which relies on using the minimum reflectance technique to create a "background" surface albedo map at two wavelengths (470 and 650 nm) from days prior to the fires. Then, AOD is retrieved independently at each wavelength using an assumed aerosol optical model, and interpolated to 550 nm using one of two methods. These are not scientifically new ideas – it is the same concept as the initial versions of the MODIS/SeaWiFS Deep Blue algorithms (e.g. Hsu et al 2004, <https://doi.org/10.1109/TGRS.2004.824067> and 2006, <https://doi.org/10.1109/TGRS.2006.879540>), with a shorter compositing period. The authors do not cite these papers but do cite Knapp (2005) which uses basically the same method as well. So I struggle to recommend this paper on grounds of scientific novelty. It is not clear to me what is new here. Minimum reflectance is a simple technique that works ok for extreme events because the neglected "background" optical thickness which is present in the reference images becomes a fairly small component of the total AOD during the intense event. But there's no sensitivity analyses backing up any of the assumptions or justifications for choices made at all. It seems guided by intuition.

The authors use only GOES observations from 20 UTC “due to the maximum solar forcing”. That sentence does not make sense to me and I’m not sure that it is relevant. I suggest it is rephrased. Do they mean “minimum solar zenith angle” or something else? Also, why is this geometry preferred? I would have thought that minimum solar angle would mean sampling geometries closer to the surface backscattering hotspot, which normally hinders AOD retrieval. This should be better justified. Also, how much does solar zenith angle vary over the 45 day compositing period? What implications does this have for the surface albedo obtained?

It would be nice to process more than just this 20 UTC time slot. That way the authors could assess at times e.g. comparable to both Terra and Aqua overpasses, and maybe another time which polar-orbiting sensors don’t sample (an advantage of geostationary). This would also increase the data volume available for validation, and may make that part of the analysis more robust. An additional option would be keeping one time slot but expanding beyond September 2020 (looking on Worldview one sees smoke beginning mid-August and continuing at fairly large scales until late October). I understand that the relevant computational power may not be available, but if not, that makes strengthening the motivation for choosing 20 UTC more critical.

In comparing their results to OAOD, the authors make much of coverage improvements. However these discussions seem to me to be mixing two things up. There are two components here: pixel selection and retrieval algorithm (i.e. pixel-level inversion). The coverage difference likely comes from pixel selection differences between the algorithms. However the discussion seems often to be framed in terms of the retrieval algorithm itself. I think it would be more useful to focus on why OAOD is throwing out pixels as invalid first, and establish that. From Figure 6 and Table 4, OAOD doesn’t appear to perform any worse than the authors’ algorithm (though the limited data volume and short analysis make it hard to say). That makes it sound like the main benefit here is the pixel selection – so maybe it is better to focus on assessing and improving pixel selection for extreme aerosol events in OAOD. If I am misunderstanding something here, then I think the authors should rewrite the manuscript to make things clearer.

Technical statistical terms are routinely used incorrectly. The word “accuracy” appears in the manuscript 21 times, but I don’t think it is the correct word in most cases. The authors seem to be using it as a synonym for “error”. “Accuracy” should refer to how biased a measurement is with respect to the true value. “Uncertainty” is an estimate of the dispersion of repeated measurements from the true value (combining systematic and random error sources). “Error” is the mismatch between a specific measurement and

specific instance of the truth. These terms are not the same. Table 3 also has a column "Accuracy within +/-0.5" measured in percent. There isn't a charitable interpretation of the use of "accuracy" here – I think the authors might mean "Percent agreeing within +/-0.5". Finally, the MAIAC comparison is referred to as a "validation" – this is not the case as MAIAC is not a ground truth. Only AERONET here can be reasonably assumed to be a ground truth. Correctness of terminology is important in scientific writing; the fact that a lot of published papers are sloppy is not an excuse.

Two methods are explored to estimate AOD at 550 nm from 470 and 650 nm. One takes the retrieved AODs at those wavelengths and calculates the Ångström exponent (which the paper calls angstrom – this should also be corrected). The other takes the 470 nm AOD and an assumed AE of 2. (A third method, not tested, would be to take the 650 nm AOD and an assumed AE of 2). This felt exploratory and hand-wavy to me, and I don't think that the AERONET data volume available for this case study is sufficient to judge what is best. Really the authors seem to be trying to see whether using both wavelengths adds value over just using one, and the answer will depend on the magnitude and spectral correlation of error sources. Without a thorough uncertainty analysis (which is missing) and larger-scale validation, I don't think we can tell, so I didn't find this aspect useful.

The figures need improvement. In Figures 3, 8, and 9 "no retrieval" and "AOD=0" seem to be the same color. This should be fixed. The color palette is also not color blind safe, which I think is needed now by AMT submission guidelines: <https://www.atmospheric-measurement-techniques.net/submission.html#figurestable> Figure 4 caption does not explain horizontal/vertical bars. Figures 5 and 7 need substantial font size increases, and the background data might be better presented as a heat map. Figure 6, there are clearly nonlinearities in the relationships here, so a linear regression is inappropriate and should be removed (perhaps also true for 4). Also not clear that correlation is the right metric to highlight here for the same reason (it measures the degree of linear association between two quantities so isn't so relevant when the underlying relationship isn't linear).

Table 1 would be improved by listing AERONET site names here. I'm not sure of the relevance of the FRM/FEM column as it is "No" in both cases. It looks like this is listing only the two sites with a PM monitor – why not list all 15 sites used for the validation? That would also be an improvement as currently this information does not seem to be provided.

As the scientific novelty is not clear and there are numerous technical, language, and presentation shortcomings, I recommend rejection. I would encourage a resubmission if the scope were expanded to more directly illustrate and understand the reasons for coverage differences between this and OAOD, and make some actionable recommendation. In my opinion, this would require more work than the traditional scope and time frame of major revisions. I know this seems like a negative review but I do really think the idea is important – extreme events are poorly-retrieved (both in terms of pixel masking and quantitative retrieval) and these new geostationary sensors are great tools to tackle this problem. It just feels to me like this paper was submitted prematurely.