RC1: 'Comment on amt-2023-259', Anonymous Referee #1

General Comment:

The authors apply the dark target aerosol optical depth algorithm to six satellite instruments (3 geo, 3 leo). They produce a quarter degree gridded product with statistics for each instrument and an ensemble average. They use the gridded product for intercomparison and validation against AERONET. The manuscript is a significant contribution that fits well within the scope of AMT. I have only minor comments that focus on methodological clarity and minor editorial comments. The resultant data product will likely be extremely valuable to multiple air pollution disciplines.

Response:

We thank the reviewers for insightful comments that have improved the manuscript significantly. We have provided response to each comment below.

Specific Comments: (original comments are in black and responses are in blue)

pg1, 22-23, I suggest moving the correlation before the percent within EE because it makes it could be read that the correlation is related to that subset.

modified as suggested.

pg1, 44: grid[d]ed

checked and corrected throughout the manuscript.

pg2, 16: my copy shows a strikeout that should be addressed.

identified and fixed

pg2, 16: SNPP and Aqua seem close in time, but Terra seems like a meaningfully different overpass time.

modified as follow:

"Having multiple polar-orbiting views of the same scene might increase data product availability, but not much if the two instruments pass close in time, such as do Aqua, and S-NPP in North America."

pg2, 20-21: As written, this excludes the main reasons for missing pixels and then concludes nearly complete… The no clouds *and otherwise retrievable* seems weird.

text revised to clarify it.

"In fact, because a scene is rarely continually cloudy from sunrise to sunset, we see that geostationary sampling can find at least one cloud-free opportunity to make an aerosol retrieval on any day, This increases the probability of at least one aerosol retrieval sometime during the day to nearly 100% (Remer et al., 2012)."

pg3, 30: (ATBD, 2023[])

edited

pg4, 18: Can you be more specific about "after some time"? Are we talking about Phase F or something earlier?

text revised to clarify it.

"At first they produced very similar results (Remer et al. 2006), but after the implementation of Collection 5 calibration and data processing that began in September 2006 the DT aerosol results from the two sensors began to diverge (Levy et al. 2018)."

pg4, 25: Section 3 really only addresses LUT updates. Are algorithm adjustments always LUT updates? Or are there any more substantial updated?

This is covered in section 3.0. Table 2 provide specific differences in each sensor. We had added text to point out additional aspects more specifically such as pixel size, cloud masking, coverage, etc.

pg5, 14: It would be good for Table 2 or the text to explicitly mention overpass times.

Added information on equatorial overpass time in Table 2

pg5, 34-35: Are any of the AERNET not collocated with leo orbits?

both LEO and GEO sensors are collocated with AERONET stations in their respective coverage area.

pg6, 18: viewing "angle" will vary by product.

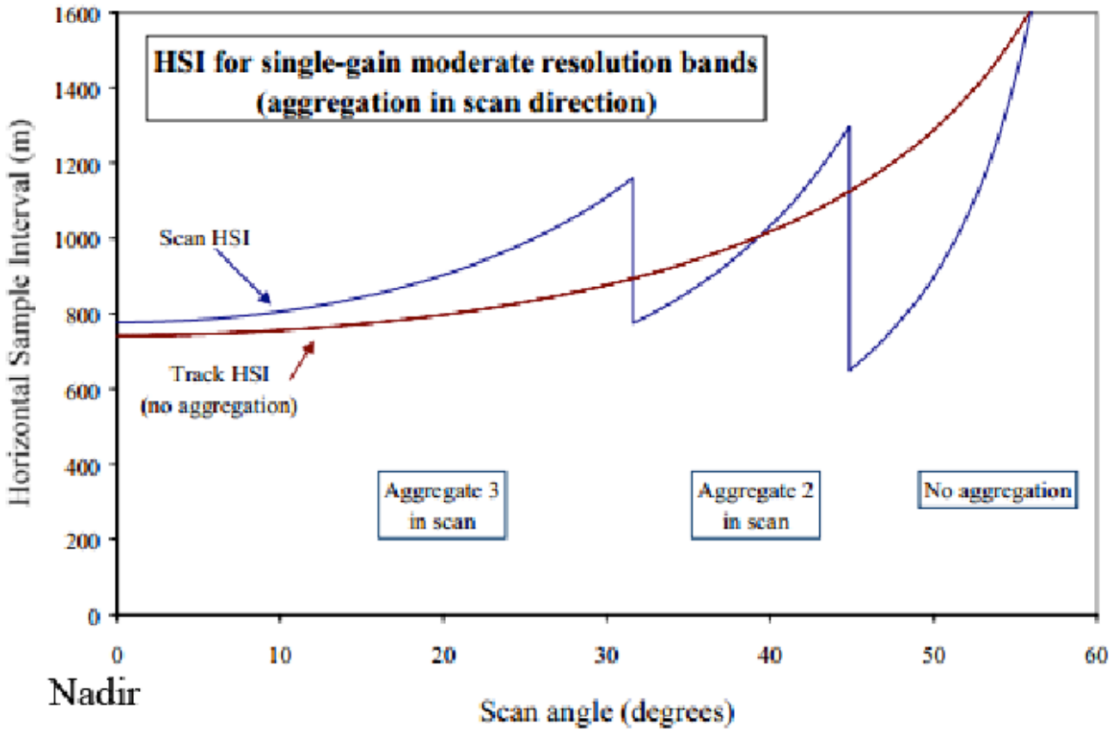
Yes, view angle varies for each sensor. For each GEO sensor, the viewing geometry is fixed, while for LEO it changes for every orbit and only repeat once every 16 days.

pg6, 34: Are you saying finer pixel measurements at nadir are aggregated so that the pixel size range is smaller? Is that what the jumps are in Figure 2?

The text is revised, and following reference is added to further explain VIIRS pixel aggregation:

Elvidge, C.D.; Zhizhin, M.; Hsu, F.-C.; Baugh, K.E. VIIRS Nightfire: Satellite Pyrometry at Night. *Remote Sens.* **2013**, *5*, 4423-4449. https://doi.org/10.3390/rs5094423

The figure below (from Elvidge et al., 2013) and caption explains VIIRS different aggregation regimes.



Horizontal sample interval chart shows how the growth of VIIRS M-band ground field of view from nadir to edge of scan is constrained by switching the number of pixels that are aggregated [12]. In aggregation zone 1, from nadir to 31.72 degrees, the signal from three pixels are averaged. In aggregation zone two, the signal from two pixels are averaged. In aggregation zone three, signal from a single pixel is recorded.

pg6, 38: "box gridding" is not a term I am used to. Is this referring to binning pixels based on their centroids being within a quarter degree cell (nearest neighbor based on centroids)**?**

The box griding is discussed in our previous publication (Gupta et al., 2020) where averaging method is explained in great detail. In simple words, “box gridding” here refers to everybody’s simple concept of an average of all the pixels with center lat-lon that fall within the 0.25x0.25 degree latitude and longitude box for each grid. Because some MODIS pixels are 40 km apart, we end up with grid boxes having no MODIS

centers within the box. This is compounded by bow tie effects that can place two 40 km pixel centers closer together and another two 40 km pixels further apart than the nominal 40 km. Modified text to clarify this.

pg6, 39: "spatial filling method" as described sounds like "averaging pixels whose footprint overlaps a grid cell".

Some sensor's pixels at the edge of the swath can cover multiple quarter degree grids of measurements due to their large size, therefore, simple averaging can create artificial gaps in gridded data due to 0.25x0.25 restriction. Therefore, to avoid these gaps, a simple gap filling method based on viewing angle and pixel size is adopted so that grids represent actual measurement footprints without any gaps in data. Gupta et al., 2020 describe this in more details with examples.

pg7, line 21: Visible discontinuity at the scale displayed seems like an unreasonable metric. We'd expect the discontinuity to be larger for a single scene when zoomed in.

Yes, agreed. Therefore, we specify it as qualitative and later we demonstrate quantitative differences among sensors.

pg7, 37-38: This seems like a weird choice. I agree that it likely doesn't change the conclusions, but a 1 in 30 sample seems like an unnecessary simplification.

One day per month is selected to ensure sampling represents all the seasons as well as data presented are manageable. As we increase to number of days, volume of data become challenging to put on scatter plot. But we agree that the conclusion will not change. We further clarify this in the text.

pg8, 4: The g17 also looks at the arid west where aod comparisons have revealed higher uncertainty. I think it is important to note that it isn't just US vs Asia, but within countries as well.

Agreed. We added text to further clarify this.

Meanwhile, ABI-G16 covers the generally wetter and darker eastern North and South America, while ABI-G17 covers mostly ocean and the dryer and brighter western North America. The significant differences in surface type will affect the accuracy of the retrievals.