Response to Short Comment

Collocation Mismatch Uncertainties in Satellite Aerosol Retrieval Validation

Timo H. Virtanen, Pekka Kolmonen, Larisa Sogacheva, Edith Rodríguez, Giulia Saponaro, and Gerrit de Leeuw
Finnish Meteorological Institute, Helsinki, Finland

We thank Dr. Schutgens for the positive feedback and the interesting questions.

- *The authors conclude that there is an optimal sampling distance of 0.3-0.4 degrees for such validations. That should give some confidence in 1 by 1 degree L3 products.*

  Indeed, the AOD comparison results indicate that a larger sampling area, approximately the size of a L3 pixel, gives better agreement between the satellite and AERONET than a single-pixel comparison. Also, the AOD variability is better captured at the $\sim 1^\circ$ scale, which gives some credibility to L3 uncertainty estimates.

- *I wonder though why correlations drop off much faster for MODIS data (Fig 9) than for AATSR data (Fig 4)? According to the authors, this drop-off is due to the natural variability in an AOD field (i.e. not related to retrieval errors). Wouldn't you expect MODIS and AATSR data to show similar results?*

  The main problem here may be the different scaling of the y-axis in Figs. 4 and 9. In Fig. 1 below we have rescaled Fig. 7 a) from the manuscript so that the y-axis has the same scale as in Fig. 9 a). Note that the y-axis position is still different for AATSR and MODIS, but in both panels below the y-axis covers a correlation coefficient range of 0.35 (from 0.92 to 0.955 for AATSR and from 0.935 to 0.97 for MODIS). On this scale, the drop in AATSR correlation with increasing sampling distance is better visible, although it is slower than for MODIS. However, the correlation coefficients remain slightly higher for MODIS than for AATSR. The main difference between the instruments seems to be that for the smallest sampling distances AATSR does not agree so well with AERONET.

  We will now point out the difference in the y-axis scales in the paper.

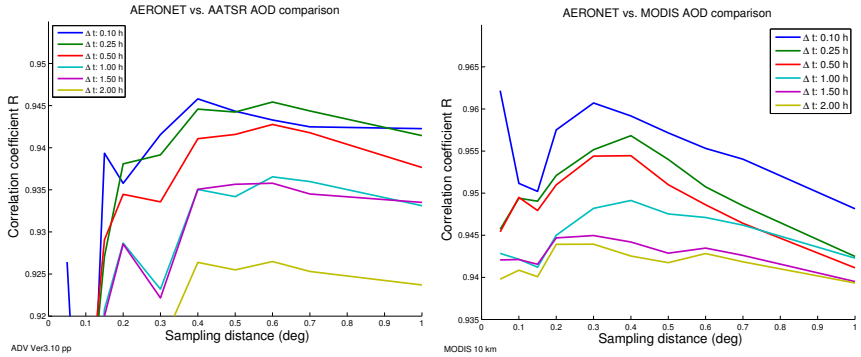

Figure 1: Comparison of AATSR (a) and MODIS (b) correlation coefficients with AERONET for various sampling parameters. Here the y-axes have the same scale in both panels.

- *We recently studied the representativity of observations using high resolution models: https://www.atmos-chem-phys.net/17/9761/2017/ . One interesting aspect is that modelled fields show high correlation among neighbouring points while your satellite-AERONET comparison shows very low correlation amongst neighbouring points (i.e. short sampling distance). As the authors themselves conclude this points to large random errors in the satellite product that average out when using larger sampling distances.*

  The high noise in AATSR data which averages out for larger sampling distances is certainly a point we want to address in future studies, possibly utilizing the high resolution (1 km) AATSR retrievals. The issue may be related to cloud screening: a fragment of cloud missed by the cloud screening algorithm will cause an erroneous high AOD value in the satellite data. As discussed in the manuscript, the number of valid AATSR pixels (N_{ADV}) in the sampling area is an indirect measure of the cloudiness in the area. Fig. 2 shows that by requiring at least three AATSR pixels in the sampling area improves the comparison considerably for the small sampling distances.

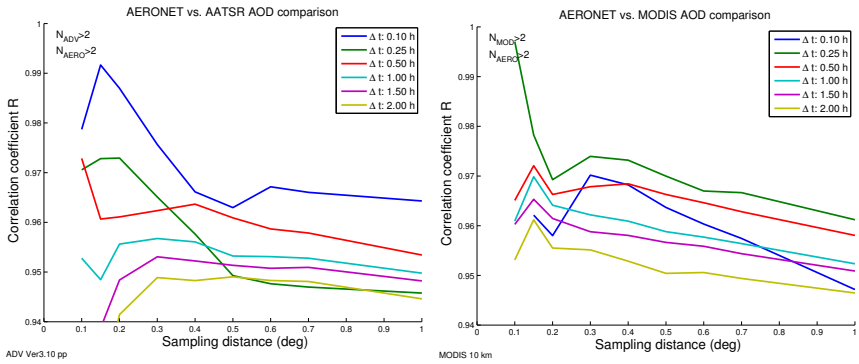

Figure 2: Comparison of AATSR (a) and MODIS (b) correlation coefficients with AERONET. Here we have required at least 3 satellite pixels in the sampling area and at least 3 AERONET observations in the sampling time window.