RC1: 'Comment on amt-2022-80', David Winker, 26 Apr 2022

This paper examines whether the impacts of sparse sampling from a nadir-viewing satellite lidar varies with the revisit time of the satellite orbit. Orbits of several different existing satellite lidars are chosen as examples. Parameters of interest are cloud amount, cloud top height, and cloud optical depth. SEVIRI cloud retrievals are taken to be truth. Lidar sampling errors are then simulated by sampling SEVIRI retrievals along the ground tracks of the various lidars. Lidar sampling error is measured by the difference between statistics based on SEVIRI retrievals sampled along the orbit track of each of the modeled lidars and SEVIRI retrieval statistics sampled by a hypothetical lidar with one day revisit time.

The paper is well organized and clearly written, for the most part. I have one major concern and a few specific comments.

My major concern is that Section 3.3 and Section 4 (Table 4) seem to come to opposite conclusions. Section 3.3 shows that shifting the initial day of the CALIPSO 16-day orbit cycle (essentially, shifting the orbit tracks observed on a given day) can be a major source of uncertainty. On the other hand, Table 4 shows that at the annual scale, with 10x10 grid cells, accuracy requirements can be met for most locations. Are all the results in Table 3 for 1x1 degree grid cells? Figure 5 shows that sampling uncertainties decrease when size of the grid cells increases but the uncertainties seem to be larger than what is indicated by the results in Table 4. But the metrics shown in the two sections are different and difficult to compare. Are results in the two sections consistent or do results in Table 4 ignore uncertainties due to initial day of the cycle? Please explain.

Reply: Thank you for your comments. In my opinion, there are two issues that need clarification.

First, Section 3.3 does indeed only consider uncertainties that resulted from the shift of the initial day of the cycle, while all other sources of uncertainty are ignored. Similarly, Section 3.1 ignores the shift of the initial day of the cycle, and only considers one particular execution of the revisit scheme (i.e., the actual calendar date of a satellite pass). The idea was to evaluate sources of uncertainty one by one. Second, Table 3 and Figure 5 give information about a range of CA/CTH, between the highest and the lowest estimates within a 16-day period (more precisely, within 16 sampling scenarios; lines 304-310). On the other hand, Table 4 gives information about how many grid cells have reported a difference that is below a threshold. Data in Table 4 were derived from differences between *n*-day and 1-day climatologies. This difference ('error') is a different measure, and is not the same as the difference between maximum and minimum values ('range') reported in Section 3.3. As the Reviewer has correctly noted, these two metrics address different aspects, and therefore are difficult to compare directly. A direct point of reference for Table 4 is Figure 1 (or Figure 2), but not Figure 5.

Following the Reviewer's comment, I have added further details, and clarification in the Discussion (see lines: 409-417).

Minor comments:

- I did not find the latitudinal extent of SEVIRI CLAAS dataset in the text. Figure 2 seems to show the CLAAS data extends from about 70S to 70N. This is important to mention in the text, to make clear that lidar sampling of the high Arctic is not being evaluated in this study.

  Reply: Meteosat, like any geostationary satellite, covers Earth's disc up to ~81° N/S. However, at the most extreme angles, image distortion is significant (eventually the line of sight becomes tangential to the sphere). For that reason, the study only considered Meteosat locations that were within ±70° latitude and longitude. This information has been added to the manuscript, as requested (see lines: 181-183).

- Line 133 states that Aeolus is in an equatorial (0-degree inclination) orbit. This is not correct. Aeolus is in a 97-degree inclination orbit.

  Reply: Thank you for pointing this out, naturally it should be 'polar' instead of 'equatorial'; now corrected.

- In Section 3.3 it is not clear what grid cell size is used in generating the statistics which are reported. Other than Figure 5, do all statistics refer to 1 degree grid boxes? What grid cell size is shown in Figure 5 c, d, g, and h?

  Reply: In Section 3.3, statistics have been calculated for all resolutions, but for purely technical reasons only selected results can be shown in Table 3 and Figure 5. As requested, I have added information on resolutions to the text and captions, to make it clear which data are discussed.

- Line 432: "spatial resolution above 10 degrees" is ambiguous. Does this mean "spatial resolution better than 10 degrees" ?

  Reply: Indeed, 'better than' is more appropriate; changed as suggested.

- Line 438. Please explain why confidence intervals are preferred over means and medians in this circumstance. Also, provide a reference on how to compute confidence intervals.

  Reply: Mean and median alone provide no information about the uncertainty level related to a specific sampling scheme. This information can only be represented by a confidence interval for a mean or median. If a variable follows a normal distribution, intervals can be found using the mean and standard deviation. However, in the case of cloud amount, where the distribution is frequently non-Gaussian, a (nonparametric) method like bootstrapping is much more appropriate. This information (and a reference) has been added to the manuscript, as suggested (see lines: 459-463).