We wish to thank the reviewer for the careful examination of the manuscript. Please find our responses below in red, with proposed amendments or revisions highlighted with underlined text.

Minor issues:

Method: the retrieval method is well known to be robust to generate good retrievals but it does not account for coupling between the surface and the scattering atmosphere close to it. Could the authors expand a bit on this point?

We assume that the comment refers to the simplified treatment for multiple scattering processes within the SMAC atmospheric correction algorithm employed in CLARA albedo retrieval. It is certainly true that multiple scattering is treated simply through the spherical albedo of the atmosphere (so not fully decoupled either), and an explicit multiple-layer atmosphere is not a component of the SMAC calculations; these were intentional original design choices to enable the very fast correction calculations in SMAC which are inherently necessary for the very large data volumes present in the 40-year CLARA timeseries (Rahman and Dedieu, 1994).

Rahman, H., & Dedieu, G. (1994). SMAC: a simplified method for the atmospheric correction of satellite measurements in the solar spectrum. Remote Sensing, 15(1), 123-143.

It is difficult to estimate the degree of additional error resulting from the simplicity of the multiple interaction accounting, given the overall simplicity of the SMAC approach. Certainly, we know that the additional error is a function of viewing and illumination geometry, which is yet another reason for our relatively conservative angular cutoffs in processing when compared to e.g. MODIS albedo products. We propose revising the text around line 110 to include a cautionary note on the underlying limitations in SMAC, noting that the effects will manifest mostly in cases of low Sun and/or satellite zenith angles.

Explain the meaning of the acronyms SAL, WAL, BAL (pag 3 line 63). The three albedos are explained with equations in the relative paragraph, but it is missing a sentence explaining them. Those variables are just not only mathematical terms but have a meaning and I think this must be added, in particular for non-expert readers or data users.

We propose to expand the text around line 63 as follows to explain the physical interpretation of the albedo quantities, particularly for non-expert readers:

"Conceptually, black-sky albedo would be observable in the absence of an atmosphere, when all solar illumination comes from a single direction. Conversely, white-sky albedo would be observable only in cases where the incoming illumination is fully diffuse, i.e. evenly distributed from all directions in the sky. In real-world situations on Earth, neither extreme case is achievable, and the incoming illumination is a combination of direct and diffuse radiation fluxes. The blue-sky albedo is the parameter that seeks to estimate these cases."

Table 1: add reference for ERA-Interim and ERA5.

Revised as requested.

Pag. 4 Add a similar plot (Figure 1 a) for a pentad product. This is necessary to give a complete example of the different products provided to the users. This would show if specific difference can be due to the amount of days accumulated.

We will revise Figure 1 to include the map and zonal means of monthly example as a/b, and map plus zonal means of one pentad within that month as c/d.

Pag. 12, Table 2 add a column to define the surface type of each site. It would be of great help for the readers to have a plot showing the location on Earth of the sites.

We will add the land covers of the sites to the table. We will also add the site locations as markers to Figure 1a.

Pag. 24, Figure 11. This figure has potential for me but the re no clear meaning on what is the message in it. Could the authors spend a couple of sentences to add a deeper explanations?

Figure 11 is meant as a comparison to a similar figure produced for the publication describing the first CLARA edition (Riihelä et al., 2013). The underlying message is rather simple, the colors indicate zonal mean anomalies against 1992-1998 mean albedo (of land and snow) for each month and latitude in the CLARA-A3 time series. While in general low values are desirable particularly for the low latitudes where snow is not present and land surfaces change relatively slowly, the figure illustrates well e.g. the impact of the Mt. Pinatubo eruption in 1992 as rapid anomalies in the albedo data, partly natural and partly due to imperfect atmospheric descriptions following this dramatic event.

Another message to be obtained here is the consistently negative anomaly over the “desert belt” of 20 N latitude for 2014-2020, most likely following from the use of the AOD climatology which would mask some of the natural variability seen during the prior decades of the record. Finally, the anomalies are generally now lower than for CLARA-A1, demonstrating that the retrievals are more stable due to advances in algorithms and supporting data, such as cloud screening. We will add these notes to the text.

Major issues:

Pag 3 line 62,63: You introduce the white, black, blue sky albedos. You miss to explain their meaning and their relation. This should be add. See for instance (https://doi.org/10.1175/JAS3479.1)

The comment appears to refer to the same item seen above in minor comments; please see the proposed revision there for additional information to the reader.

Pag. 10 Figure 4. I suggest to add a figure similar to (b) but for one region over the tropics. This will show how different geographical regions are impacted in the retrieval and provide the users with a clear view of the potentialities of the products in different areas.

We propose to add a subplot into Figure 4, illustrating sampling over e.g. the rainforests of central Africa as a counterpoint to the polar region of subplot b.

Pag. 9 Line 219. The authors say that being the retrieval deterministic rather than probabilistic but the measure of uncertainty is left to statistical values. The albedo variables are estimated using clear mathematical relations. Could the authors explain what the uncertainty are not also mathematically estimated using the well known procedures. See for instance the method used in https://doi.org/10.1029/2006JD007313. I find this section personally to be the only one a bit weak in the paper, otherwise well structured.

The formalism in the paper referenced (Govaerts & Lattanzio, 2007) is designed for geostationary imagers. A central idea there is that since the viewing geometry is fixed, variability in derived reflectivity shall correlate with errors in the instrument, the radiative transfer calculus or the atmospheric description relevant for the scattering/absorption processes along the pathway. The continuously variable solar/viewing geometry of AVHRR and the diurnally limited sampling of any one place on the Earth would imply violation of those assumptions if we were to attempt implementing the same approach for CLARA error estimation.

The geostationary viewpoint also allows for sub-daily characterization of e.g. AOD variability for the uncertainty characterization, which is not possible for polar orbiters (unless the imager constellation is very large). Finally, the method assumes non-variability of the surface albedo over an extended sampling period to maximize the robustness of the uncertainty estimate. While this is certainly achievable and a valid assumption for low-latitude vegetated surfaces, high-latitude snow and ice cover will cause sudden dynamic changes in reflectivity that would invalidate uncertainty estimates using this approach.

This being said, we are not at all opposed towards improved characterizations of uncertainty for the CLARA retrievals. One approach being considered is to use the level 2 validation data gathered during the CLARA-A3 production process as a source for training machine learning methods to create a retrieval error prediction algorithm for the next edition, that is foreseen to keep the present core retrieval algorithms. It could also be possible to release separate per-pixel uncertainties for CLARA-A3 *a posteriori*, although this would depend on the robustness of the machine learning-based uncertainty estimates, given that the available (reference) validation data is spatially limited on the global scale.

Section 3 on validation. The authors have a very long list of validation sites. Even if they clearly worked a lot and well on this part, I have several issues with this paragraph. The authors includes a lot of sites (in red italics) not used in the analysis being “spatially” unrepresentative. Could you explain why they have been included if not sued? For the remaining sites, I would have expected some full time range (4 decades or the time covered by the ground measurements at least) plots (selecting some representative among all, covering the different surface types). Only basing the analysis on monthly mean could be misleading for me. There are some summary statistics in Table 4. The table shows interesting figures. I would recommend to provide the numbers for each valid site and not as merged statistic.

Thank you for the comment, we considered this carefully. <u>We propose to first remove</u> the unused sites from tables 2 and 3, but we note the excluded sites in the text regarding the representativeness analysis for clarity. <u>Second, we propose to add a selection of time series plots</u> showing the CLARA estimates and the reference albedos, in style of Figure 9. The question of how to clearly illustrate both monthly and pentad data over decadal scale has been considered for each CLARA release validation, and unfortunately no optimal choice has been found. We suggest showing two sites from BSRN and two sites from PROMICE, with separate subplots for pentads and monthly means vs. reference to illustrate the variability. This implies 8 panels, which is already a large figure, but experience has shown that mixing monthly and pentad data into the same subplots degrades readability considerably, as unfortunately does portrayal of multi-decadal data in a single plot.

Regarding the site-specific results, Figures 7 and 8 are specifically designed to illustrate the bias and other quality indicators on a site-to-site basis. We thus suggest not to add additional large tables to the already rather lengthy manuscript.

Section 4 on intercomparison to MODIS. This is the only comparison against satellite data. It is a pity not to intercompare against a geostationary product but I can see the difficulty even if it would have made the section more complete. The comparison against MODIS is a bit weak. There is only one figure, again showing means over months. It would be really important to select some ground sites and show time series including also MODIS and CLARA-3.

We will expand the visualization here to include a time series-based comparison of MCD43 and CLARA over a small selection of ground sites.