Response to the Anonymous Referee #4 comments for the manuscript "Retrieval of aerosol properties from zenith sky radiance measurements" By Sara Herrero-Anta et al. in AMT

First of all, we would like to thank the time and effort of the referee for their detailed review of the manuscript. Reviewer comments (RC) are in black font and author comments (AC) are in red.

Author's answer to Anonymous Referee #4

RC: This manuscript describes the retrieval of aerosol properties with the synergetic use of zenith sky radiance measurements at 4 wavelengths and GRASP.

The instrument used ZEN-R52 radiometer, is a simple radiometer with uncertainties arising mainly from the temperature dependencies. The authors have done a laboratory characterization of the instrument, introduced corrections and accessed their uncertainties, which were further used in the inversion algorithm.

This paper provides detailed information for the normalization, validation and inversion strategies followed.

This work provides new insights in the possibility of retrieving aerosol properties (AOD, VCT, VCC and VCF) using a simple measurement geometry and skipping the laboratory radiance calibration (or using it for stability monitoring purposes). The results are very promising, and the authors provide information to the community of possible improvements through their validation against AERONET products and uncertainty budget.

The manuscript is clear, but it would improve, if a more concisely writing style was used.

I propose that this article is accepted for publication, after minor revisions.

RC: L131 Silicone diode sensitivity range: 180 nm to 1100 nm

AC: The clarification of the sensitivity range is irrelevant for the paper and it has been removed in the new version of the manuscript.

RC: L134 is this the plateau of the FOV or the FWHM?

AC: It is the FOV.

RC: L142-143 the software name is misleading. I would suggest to rename it to combined variability since it describes both the atmospheric variability and the noise of the ZEN.

AC: The software does not have a specific name, it is just the software of the instrument. This software gives the parameter as ZEN error, so we used it like that, but we agree that

ZEN variability is more appealable. To clarify it we have modify the description of the parameter as follows:

L144-146: "*For each measurement, it is also provided a variability parameter (ZEN variability) that describes both the atmospheric variability and the noise of the ZEN-R51 within the minute of measurement, which is calculated as the standard deviation of the 30 samples.*"

RC: L157 and?

AC: We think "Therefore" fits better than "and".

RC: L164 (NO2 and O3)

AC: The GOD includes the gases given by AERONET: O3, NO2, CO2, CH4 and Water Vapor.

RC: L207 (Section 3 calibration) A comparison to a RT model is not a calibration procedure. This section should be renamed to "Normalization to GRASP forward model"/ "Responsivity to GRASP forward model radiance" to. A calibration would be the comparison of co-located, synchronous zenith radiance measurements to a laboratory calibrated instrument (eg section 3.5), accounting for the uncertainties of differences in wavelength, FOV, extrapolation,...

AC: Although this calibration method may not be conventional, we think it can be considered a calibration method since, in this case, model data are used as a reference or standard instead of another instrument, and these model data do are co-located and synchronous. In fact, as demonstrated in this study, by applying this methodology, the raw signal from the ZEN instrument is transformed into physical units of radiance ($Wm^{-2}nm^{-1}sr^{-1}$), which is equivalent to what is achieved with a traditional calibration. Therefore, we consider the method well-defined as calibration, even though we understand it may generate potential confusion due to its more unconventional nature as a calibration method.

RC: L208-212 This normalization methodology requires dark, temperature corrected and quality assured signals over the analysed period. There is limited effort in the laboratory characterization, since the radiance calibration is replaced by the RT model. However, laboratory test are done and presented in 3.1, 3.3.

AC: Laboratory measurements have been conducted only for Section 3.1, but it is clarified that it can also be realized using nigh-time measurements or even measurements during day using a dark cover for the instrument. No laboratory measurements have been conducted in Section 3.3 for the temperature correction, it is used the normalized ZSR_{DSC}/ZSR_{SIM} ratio.

To clarify that, the next sentence has been modified:

L238-240: "*In this work, the DS has been characterized in the laboratory to cover a wide range of temperatures, but it could be calculated from the night-time measurements (dark sky) or even from day-time measurements (covering the instrument with a black piece), when a thermal chamber is no available.*"

RC: L223 negligible temperature dependency (<1%)

C: It has been replaced by "negligible dependence on temperature".

RC: L224 steep?

AC: "Steeped" has been replaced by "staggered".

RC: L228-230 Is the dark signal level of the instrument constant over time? Is this monitored through the night-time measurements?

AC: Here it can be seen a residuals graph for the differences between the modelled (using the laboratory data) and the raw real ZEN signal for night-time measurements for the whole study period, April 2019 to September 2021. The residuals are consistent for the whole dataset, being mainly within +-1 DC, so we can consider the modelled dark signal represents well the dark signal for the all the period. It means that dark signal did not vary over the period.

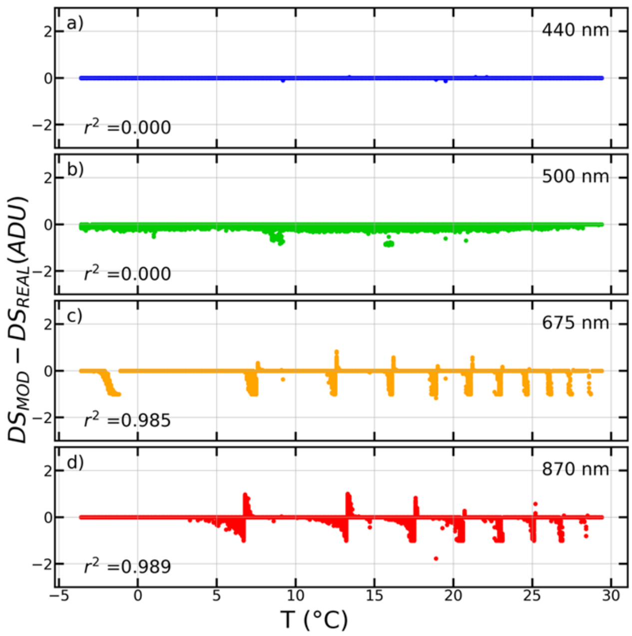

Figure: Residuals graph for the dark signal (DS) correction calculated for the whole study period, April 2019 to September 2021. The differences between the modelled dark signal (DS_{MOD}) and the correspondent raw real dark signal recorded by the ZEN-R52 (DS_{REAL}) are plotted against the temperature at a) 440nm, b) 500nm, c) 675 nm and d) 870 nm. The determination coefficient (r^2) obtained for the direct comparison of the DS_{MOD} against the DS_{REAL} is also included.

To point out that, the next sentence has been added in the new version:

L236-238: '*It has also been verified that the dark signal behaviour has remained constant over time, comparing the modelled DS against the nigh-time measurements*'

RC: L234 λ is used later as wavelength, please consider changing he letter or even skipping the information about the smoothing.

AC: This has been removed to avoid confusion.

RC: L241-242 It would be easier for the reader if the physical parameters are described here and explained later. e.g. atmospheric variability, stray-light, uncertainties in temperature correction.

AC: These sentences have been rewritten following the reviewer's comment as follows:

L250-255: '*The ZEN-R52 measurements can be affected in different ways. For example, the possible sun stray-light intromission when sun is very elevated can increase the measured signal, clouds presence can also alter it, or the variation in temperature can introduce some dependency. To identify and reject the cloud-contaminated or wrong measurements, different thresholds have been identified after the visual analysis of some parameters in scatter plots. For the SZA, the signal of the instrument is higher than expected for SZA values below 30º, which could be explained by sun stray-light intromission.*'

L263 typos

This has been corrected changing 'Despise outliners' by 'disregard outliers'.

RC: L266-267 Please rewrite it clearer eg: are used for the temperature dependency correction following Equation 2

AC: Done

RC: L274-276 I would suggest simplification of the sentence. eg the "calibration" factor are obtained by comparing the dark and temperature corrected QA signals of ZEN to ...

AC: It has been simplified by:

L287-288: '*The calibration factors can be directly obtained by comparing the dark and temperature corrected ZSR from the ZEN-R52 against the values simulated by GRASP*'.

RC: Section 3.5.1 The geometry of the "Cloud Mode" is identical to the ZEN, introducing less uncertainty in the comparison with respect to PPL method. However, the high variability of the data set used results in high uncertainty in the comparison. It would be worth while to perform the analysis using the strict QA criteria or limit the discussion in the validation of the cloud screening thought the std of the 30.

AC: If we applied the QA criteria at this comparison, we would eliminate most of the ZEN data, since we are under the presence of clouds and that results in a high ZEN variability associated to those measurements. Furthermore, the CIMEL measurement is quasi-instantaneous, while the ZEN measurement represents a one-minute average, so in the presence of clouds, conditions can vary significantly within that minute. Therefore, it is expected to observe a high level of uncertainty in the comparison between both datasets.

In order to clarify that, the next paragraph has been added:

L338-344: "*This comparison against the cloud mode measurements will not be used to quantify the uncertainty of the ZEN measurements; it is because clouds are very variable and, therefore, the recorded signal. Therefore, we should need to compare both measurements carried out at exactly the same time; but this is not the case since ZEN measurements are 1-min averages while CE318 photometer measurements are quasi-instantaneous. In addition, for the retrieval of aerosol properties, it is necessary to employ measurements under cloud-free conditions, therefore, the results obtained in following comparison will be the reference ones.*"

RC: L348 it seem that the distributions start deviating from the assumed normal one for 675 nm and especially 870 nm. A better stimulation of the distribution would give more representative values for these wavelengths.

AC: It is true that these distribution looks like a distribution with positive skewness, especially for 675 and 870 nm. But we consider the values in the tail are due to occasional malfunctions of this instrument at those wavelengths. Then, if we neglect these outlier values, we can assume a normal distribution. In fact, the use of the median instead of the mean is to neglect these values. Regarding the standard deviation, it is calculated with these outliers, which provides a higher value, but we prefer to be conservative assuming a lower precision than the expected.

RC: L349-350 when was the IARC calibration performed and at which temperature? it is erroneous to apply a calibration factor applying different corrections than those in the calibration procedure. This paragraph doesn't add some information since the calibration and "normalization" factors can be directly compared.

AC: We are comparing the calibration factors obtained using two independent calibration methods, the one proposed here, and the one described by Almansa et al. (2020) which has been called IARC in this work. IARC calibration method has no additional corrections (no temperature correction, no dark signal removal), the calibration factors obtained by IARC are directly applied to the ZEN-R52 raw signal. Maybe the IARC calibration could be improved considering dark and temperature corrections, but it is out of the scope of the paper.

Anyway, the similarity of the results between our calibration method and the IARC method demonstrates the quality of the calibration method proposed.

RC: L358-365 Information of the impact of the uncertain of each channel, is repeatedly mentioned. Please consider simplifying the paragraph.

AC: The paragraph has been simplified as follows:

L384-390: "*These results indicate that the ZEN-R52 measurements are more reliable at shorter wavelengths and, therefore, should be given more importance than those corresponding to longer ones in the retrieval of aerosol properties. The inversion module from GRASP code considers the importance of each measurement through the so-called 'noises'; allowing to associate a different 'noise' or reliability to each channel. The standard deviations collected in Table 2 (using the calibration proposed in this work), associated with the ZSR_{ZEN} uncertainty, are used to this end in the GRASP-ZEN method.*"

RC: L384-385 of ZEN-R52 for these scenarios (ZSRSYN).

AC: This sentence has been modified as follows:

L409-411: "*For both tests, synthetic aerosol scenarios have been created and used as input to the GRASP forward module to simulate the ZSR of the ZEN-R52 under these scenarios (ZSR_{SYN}).*"

RC: L527-531 Not relevant to the publication

AC: It has been removed.

RC: L534 'overstimates'

AC: It has been replaced by 'higher than'.

RC: L537-538 'a match-up has been done. In this case, the GRASP-ZEN values closest to the AERONET values within 5 minutes are chosen,'

AC: The next sentence has been added instead:

L559-561: "*For a more quantitative analysis of the correlation between VCF, VCC and VCT from GRASP-ZEN and AERONET datasets a synchronization with a time window of ±5 min was done, obtaining a total of 4356 coincident points for each volume concentration.*"

RC: L552 'offered by'

AC: It has been changed by “of the AERONET products”.

RC: L554 ‘issue’

AC: It has been changed by ‘study’.

RC: L568 retrieve a normalization factor converting the ZEN-R52 signal to radiance

AC: We prefer to maintain the calibration word as expressed above.

RC: L573 a substantial amount of

AC: Done

RC: L601 lower

AC: It is low, since we are referring to the fact they are indeed low, not a comparison with anything else.