Response to Reviewers

The authors greatly acknowledge the anonymous reviewer for carefully reading the manuscript and providing constructive comments. This document contains the author's responses. Each comment is discussed separately with the following typesetting:

Reviewer's comments

Authors response

Changes in the manuscript

Reviewer #1

The manuscript provides a summary of ambient measurements of F11 (phase function) and F12 made in Granada, Spain by the Airphoton/GRASP-Earth PIN100 as well as measurements of PM10 and a suite of other optical quantities. Two cases with very high concentrations of advected Saharan dust are explored in detailed followed by a summary of several cases with more moderate dust loading. Results are contextualized by single scattering calculations for spheres and spheroids showing theoretical F11 and F12 for hypothetical urban, mixed, and pure dust aerosols.

The complex shape of desert dust particles limits our ability to accurately predict the optical properties required in aerosol modeling and remote sensing but direct, ambient measurements of these quantities, especially phase function and other scattering matrix elements has been very limited. This paper presents some of the only ambient dust scattering matrix measurements to date and the results have great potential to help improve on these past limitations. The identification of the strong dependence in F12 to the fine mode, particular at shorter visible wavelengths is also a valuable contribution. I do however feel the manuscript fails to properly contextualize results, particularly in its accounting for potential measurement errors. At times it also seems to jump to very strong conclusions based on what seems to be limited or inconclusive evidence. To be suitable for publication in ACP, I feel that these issues, outlined in detail below, must be addressed.

We thank the reviewer for his/her time and efforts to review our work, and we believe that the feedback received is very positive to improve the manuscript.

The question about the uncertainties of the Airphoton/GRASP-Earth PIN100 uncertainties were also raised by the other referees. We did not give many details about the uncertainties because in Bazo et al., (2024) we performed an exhaustive characterization of the PI- Neph and of its uncertainties. Here, we want to give an overview that will help in this public discussion.

PUBLIC DISCUSSION OF Airphoton/GRASP-Earth PIN100 UNCERTAINTIES FROM BAZO ET AL., (2024) MANUSCRIPT

Appropriate calibration of the instrument

The instrument follows an exhaustive calibration procedure that consists of two different steps: The first is a geometric correction that corrects from the different light paths to the different pixels in the CMOS camera used as detector. To do so we used gases of known phase functions that follow Rayleigh theory (e.g. CO_2 or particle free air), particularly the perpendicular signal that follows the shape of $\cos^2\theta$ (see Eq. 13 in Bazo et al., for details). Measurements are fit to the theoretical shape by using an 8th order polynomial. The following Figure (Fig. 5 in Bazo et al., (2024)) illustrate how the application of the computed geometric correction serves to accurately reproduce theoretical phase functions of CO_2 and clean air

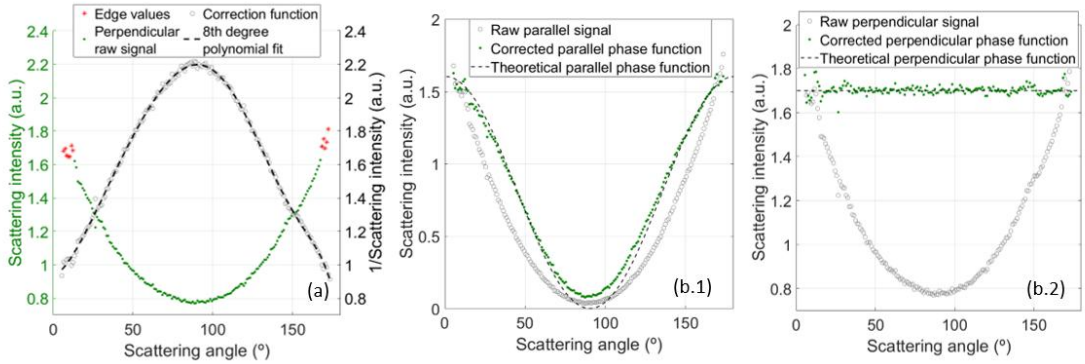

Figure 5 in Bazo et al., (2024). (a) Geometric correction procedure: raw perpendicular signal for 515 nm (green dots); red asterisks represent the edge values used to calculate the correction function (gray circles); the correction function is fitted to an 8th degree polynomial (dashed line), that corresponds to the geometric correction. (b) Application of the geometric correction to parallel (b.1) and perpendicular (b.2) particle free air signals.

The second step is to perform an absolute calibration that permits to obtain phase matrix in physical units. We did it by measuring the parallel and perpendicular signals of pure CO_2 and particle free air, since their phase functions are well-known. The sum of both components corresponds to the phase function in digital counts $C_{11}(\theta) = C_{output,||}(\theta) + C_{output,\perp}(\theta)$. Therefore, it is possible to obtain a linear relationship between the counts recorded by the CMOS and the scattered light pattern: $F_{11}(\theta) = K \cdot C_{11}(\theta) + W$ where K and W are the coefficients of a linear fitting between theoretical and experimental data. These coefficients of the linear relationship between the phase functions in physical units $F_{11}(\theta)$ and in digital counts can be computed following the same procedure used in common integrating nephelometers by fitting the theoretical scattering coefficients of pure CO_2 and particle free air to the integral of the phase function in digital counts ($\tilde{\sigma}_{counts}$):

$$\tilde{\sigma}_{\text{counts}} = \frac{1}{2} \cdot \int_{0°}^{180°} C_{11}(\theta) \cdot \sin\theta \cdot d\theta$$

where the values of $C_{11}(\theta)$ between 0 and 5° and 175 and 180° are extrapolated by the closest neighbor method, since the theoretical Rayleigh $F_{11}(\theta)$ follows the $(1 + \cos^2\theta)$ shape. For the computation of the theoretical scattering coefficients the methodology

proposed in Bodhaine et al. (1991) was used, applying corrections of pressure and temperature from PI-Neph sensor measurements. We can then compute $\sigma_{\text{theoretical}} = K \cdot \tilde{\sigma}_{\text{counts}} + W$. Note that the geometric correction is first applied to the measurements. The evaluation of the calibration constants with time was also done and minimum differences (below 5%) were obtained.

Our evaluation of the different ways of calibrating the PI-Neph instrument was even further by performing an angle-to-angle calibration that consists of associating each angle with a different calibration constant. To do so, we use again well-known scatters that follow the Rayleigh scattering matrix (again CO_2 and particle free air). Then, an angle-to-angle comparison of measurements with the theoretical scattering matrix elements is done, obtaining a calibration constant for each data point in the measurements. That approach can serve to mitigate measurements affected by stray light. Figure below (Fig. S1 in Bazo et al., (2024)) shows the results in F_{11} and $-F_{12}/F_{11}$ using the above calibration method and the angle-to-angle calibration. Differences in F_{11} and $-F_{12}/F_{11}$ are negligible.

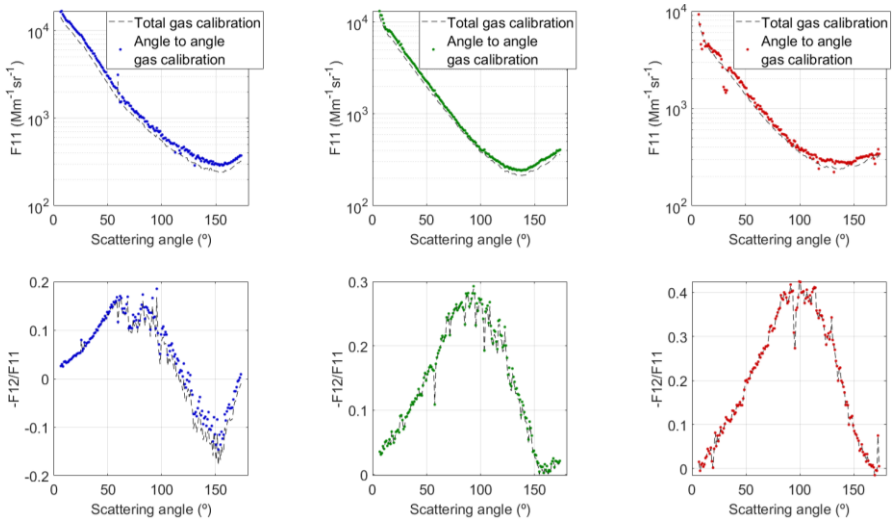

Figure S1 in Bazo et al., (2024): F_{11} (up) and $-F_{12}/F_{11}$ (bottom) for 405 (left), 515 (middle) and 660 nm (right) calculated with total gas calibration (dashed line) and angle to angle gas calibration (colored dots).

Evaluation of Instrument Stability

Figure below - Figure S2 in Bazo et al., (2024) - shows the evolution of the total scattering coefficient during 15 min both for particle free air and CO_2 measured at a constant flow rate of 10 $Lmin^{-1}$ versus theoretical values (corrected of temperature and pressure). In general, no significant changes in the scattering coefficients are observed with time, which reveals very good stability of the instrument. Mean differences between the theoretical and calculated particle free air scattering coefficient are 0.0608 Mm^{-1} for the 660 nm channel (1.07% difference), 0.0341 Mm^{-1} for 515 nm (0.25% difference), and 0.3152 Mm^{-1} for 405 nm (0.73% difference). Similar values are found for the CO_2 scattering coefficients, with mean differences of 0.0567 Mm^{-1} for 660 nm (0.38% difference), 0.0952 Mm^{-1} for 515 nm (0.17% difference), and 0.4258 Mm^{-1} for 405 nm (0.40% difference)

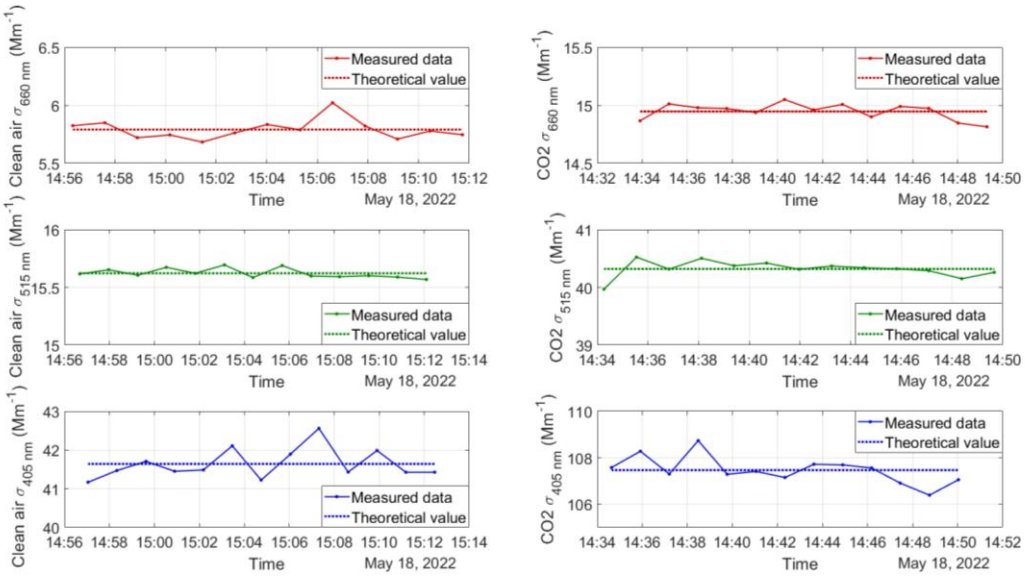

Figure S2 in Bazo et al., (2024): Changes in the scattering coefficients of particle free air (left) and CO_2 (right) with time, over a 15 min measurement period.

Errors Associated with the Imaging Technique

The first one is the different background pixel counts due to changes in the temperature of the CMOS detector, but this error is minimized because background images are taken at the beginning of each sequence when differences in temperature are negligible. The other source of error in the imaging technique comes from the selection of the pixels associated with the scattering of the laser beam. The PI-Neph software is capable of appropriately detecting saturated pixels and those with low SNR.

Natural fluctuations of the sample can lead to noisy measurements being particularly important in conditions of high aerosol variability and enhanced with very short exposure times. To characterize these effects, we carried out measurements with very low aerosol loads and different exposure times, ranging from 1s to 20s. Figure below - Figure S4 in Bazo et al., (2024) - shows instantaneous F_{11} and $-F_{12}/F_{11}$ (in arbitrary units) for 515 nm for the different exposure times, revealing an improvement in both parameters as the exposure time increases. Selecting the appropriate exposure time for every measurement makes the continuous operation of the instrument difficult, and although exposure times must be revised by the user it is fixed during long periods. That is why noisy measurements plus saturated and low SNR measurements are present and must be adequately treated.

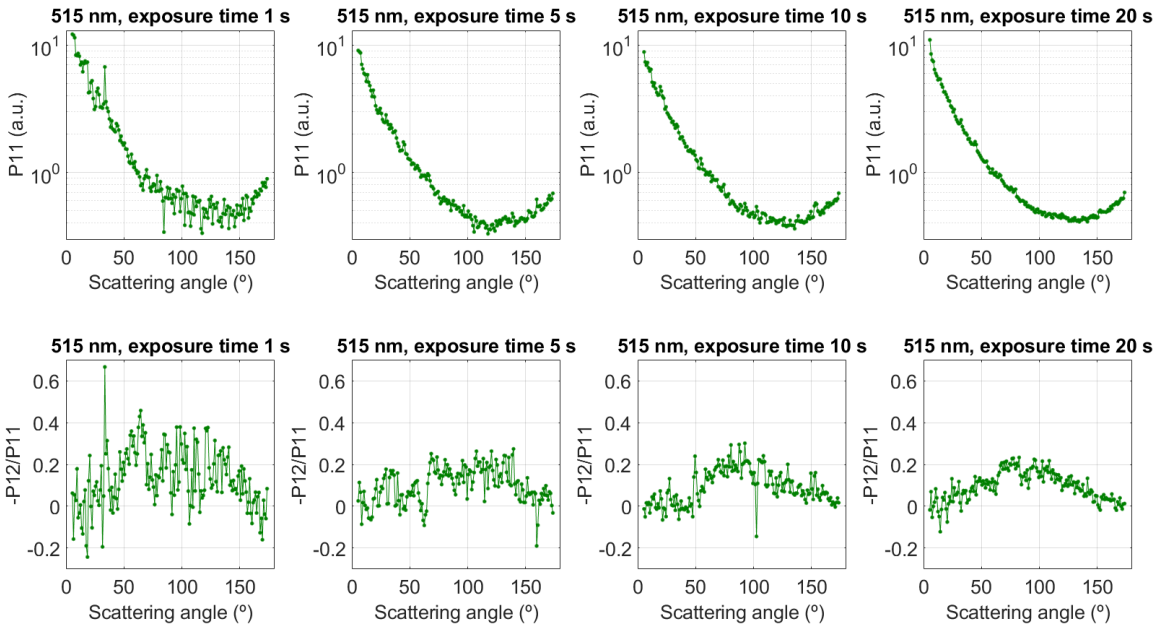

Figure S4 in Bazo et al., (2024): For 515 nm, P11 and -P12/P11 with different exposure times.

Instrument Validation versus Monodisperse Aerosol

For the validation of PI-Neph data, monodisperse aerosol measurements were carried out. In particular, measurements of PSL of 0.75 μm and 1 μm diameter have been performed in the laboratory. For the generation of the aerosol sample, an Aerosol Generator (TSI 3076) was used, followed by a Diffusion Dryer (TSI 3062) that was connected to the PI-Neph's inlet. The set of measurements was done for 30 minutes, with a total of 30 matrix elements measurements. Exposure times were adjusted depending on the wavelength and polarization state. Results presented here correspond to temporal averages of the entire period. Figure below (Figure 10 in Bazo et al., (2024)) shows experimental data and fits with the GRASP algorithm. Generally, the agreements in F_{11} are excellent, reproducing all resonances, while for $-F_{12}/F_{11}$ there are some deviations. If we assume that errors in the instruments are of 5-10% and 10-20% for F_{11} and $-F_{12}/F_{11}$, respectively, and associated with issues associated with calibration and data acquisition of the CMOS, plus other errors associated with the imperfections and preparations of the PSL sample, then the differences observed in the Figure can be explained by errors propagation. Nevertheless, all resonances are well reproduced confirming that the pixel to angle relation in the PI-Neph is accurate and therefore serves as validation of PI-Neph measurements. The agreement between measured and retrieved value is supported by the high correlation coefficients and low residuals. GRASP retrievals were able to reproduce particle diameters and refractive index. Similar results were obtained when using a Mie code instead of GRASP algorithm.

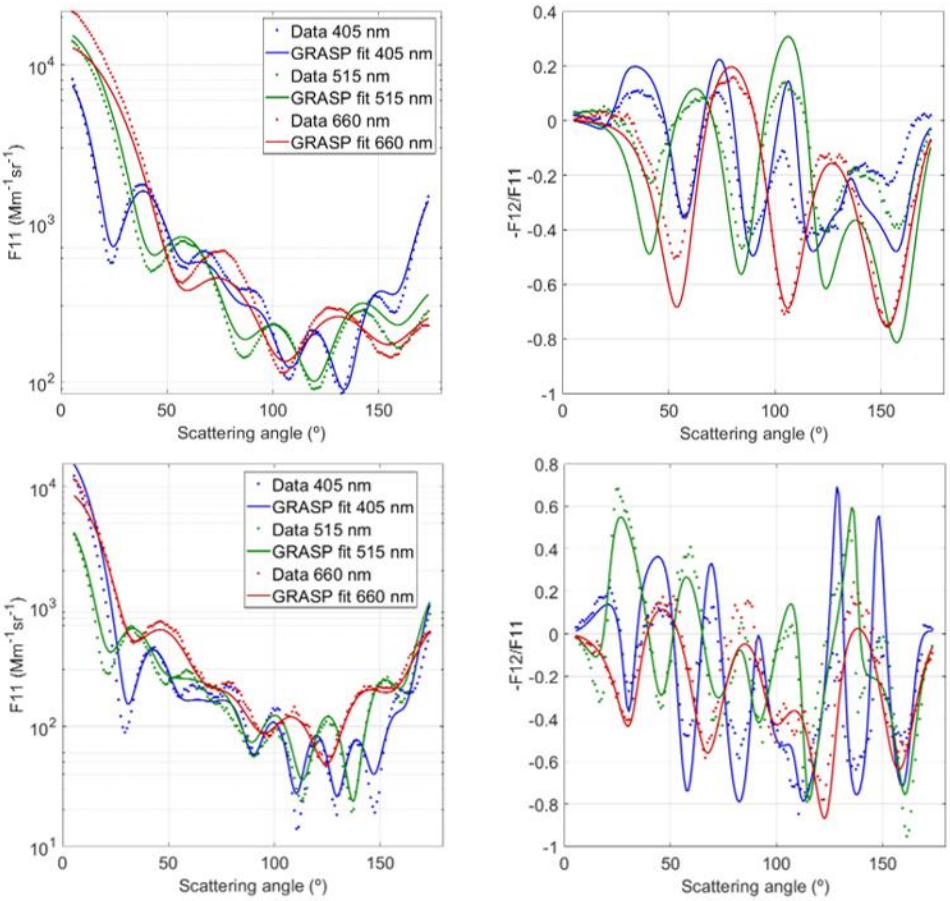

Figure 10 in Bazo et al., (2024). PSL PI-Neph measurements of F_{11} (left) and $-F_{12}/F_{11}$ (right) and GRASP fits for 0.75 μm (top) and 1 μm (bottom).

Data Quality Check

For analyzing a large dataset where the atmospheric conditions are variable, we developed a data quality procedure illustrated in the Figure below – Fig. 6 in Bazo et al., (2024). The idea behind this data quality check is to minimize errors associated with the measurement process. The complete procedure is applied to each channel and polarization individually. The first step is to select reliable data from parallel and perpendicular signals using the quality check parameter that serves to mark the data point as non-reliable if the averaged counts reach the saturation threshold. Since the number of counts at a specific angle is averaged through a group of pixels, the value of the quality check parameter of that angular measurement will depend on the number of pixels within the group that reached saturation. Specifically, it can vary from 0 to 1, where 1 means that all CMOS counts per pixel that have been average do not reach the saturation threshold. On the other hand, a value of 0 implies that all pixels in the region were saturated. Any value between 0 and 1 gives information about the ratio of CMOS counts per pixel that did or did not reach the saturation threshold. More details about the quality check parameter can be found in Section 2.2 of Bazo et al., (2024).

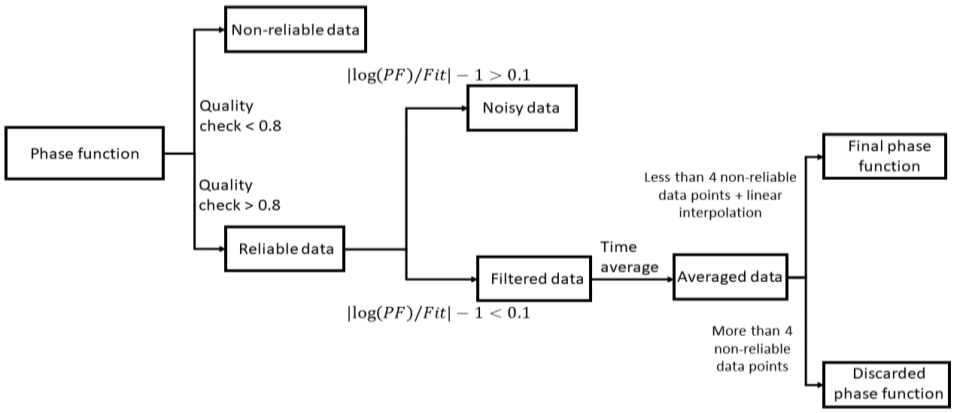

Figure 6 in Bazo et al., (2024). Flow-chart for the filtering procedure for the PI-Neph data.

In the data quality check procedure data with quality check parameter below 0.8 are discarded. Later, the procedure evaluates any possible peaks in individual measurements caused by a particle or by stray light. The filtering is based on the fit to a Legendre polynomial of the natural logarithm of the phase functions, which can be expressed as an expansion of Legendre polynomials. This method serves to detect outliers and assumes that if individual (angle dependent) values of the natural logarithm of the phase functions differ from the values of the fit to Legendre polynomial by more than a 10%, these points are discarded, since it is most likely that they come from a large particle that crossed the sample chamber or from stray light. The 10% threshold was selected after appropriate optimizations (see Bazo et al., 2024 for details).

After applying the filtering process, temporal averages are carried out. Temporal resolution of phase functions typically depends on the concentration of aerosols being sampled. Low aerosol concentrations lead to noisier averaged phase functions, hence larger averaging times are needed to have smoother patterns of the matrix elements, especially of $-F_{12}/F_{11}$. After performing averages, if more than 75% data points are qualified as non-reliable, all measurements are discarded to avoid any biases caused by the filtering process. If after performing the time averages there are more than 4 values in consecutive angles that are considered unreliable by the filtering procedure, the entire phase function is discarded as well. For cases with fewer values that do not comply with the filtering, a linear interpolation is performed (see Bazo et al., 2024 for details).

The Figures below – Figure 7 and 8 in Bazo et al., (2024) - show examples of the final F_{11} and $-F_{12}/F_{11}$, respectively, after applying the data-quality check algorithm. Data corresponding to rows a, b, c to instantaneous, 30-minute average and 60-minute average, respectively; whereas columns 1, 2, 3 correspond to non-filtered, quality filtered, and Legendre filtered data (previously quality filtered), respectively. Data was acquired for ambient aerosol samples. In general, quality filter mostly affects the 405 nm channel in the angular range of 70 to 100°, since it is the most sensitive wavelength to low SNR data. Legendre filter removes data in the three channels for instantaneous data. Note that peaks in the three phase functions at 120° get filtered. However, there is lack of data in many regions in the angular range, so temporal averages (rows b, c) must be done to have a continuous phase function. It can be observed that temporal averages reduce noise in the phase functions but there are some outliers that remain. Therefore, applying a Legendre filter to the instantaneous phase function components (parallel and perpendicular signals)

and then performing temporal averages is the best option to have reliable data, especially for the $-F_{12}/-F_{11}$ element, which is affected by more noise because of the subtraction of the parallel and perpendicular signals measured by the PI-Neph.

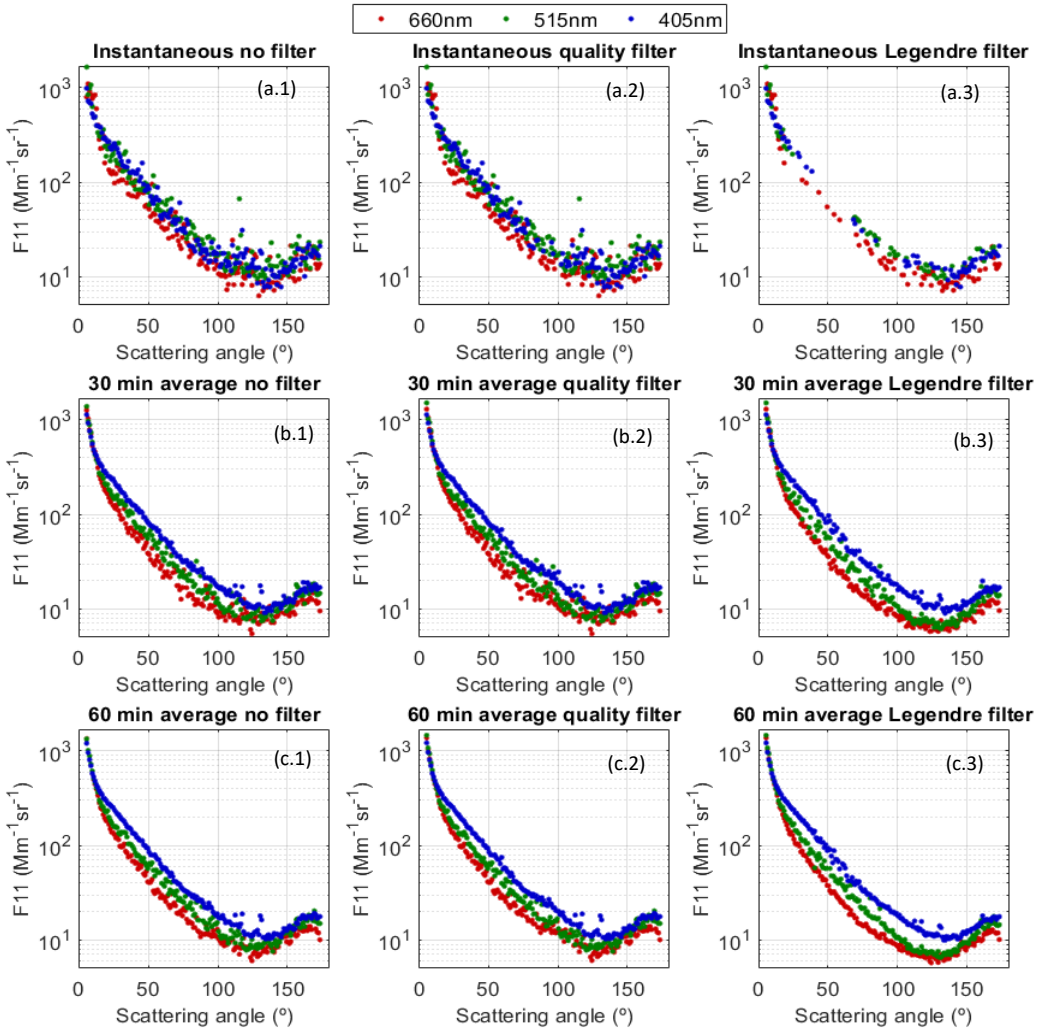

Figure 7 in Bazo et al., (2024). Phase function F_{11} for different stages of the filtering process. Rows (a), (b), (c) show instantaneous, 30 min average and 60 min average phase functions, respectively. Columns (1), (2), (3) show non-filtered, quality filtered, and Legendre filtered data, respectively.

The discussion presented serves to illustrate the different sources of errors that can affect PI-Neph measurements. We note that for laboratory measurements under controlled conditions uncertainties are in the range of 5-10% for F_{11} and 10-20% for F_{12}. For ambient aerosol samples, the situation is more complex because of the natural variability of the aerosol. The large variability of particles that can be present in ambient aerosols make natural fluctuations in the phase matrix to appear. For example, a very coarse particle can be in the air sample and depending on the orientation, it can affect critically the shape of the phase functions. The data quality approach tries to avoid these fluctuations, although they are present in nature. We therefore believe that PI-Neph uncertainties are typically 5% for F_{11} and 10% for F_{12} but for real aerosol samples the fluctuations in aerosol particles plus the complexity of their mixture make the standard deviations higher than the claimed uncertainties in most cases.

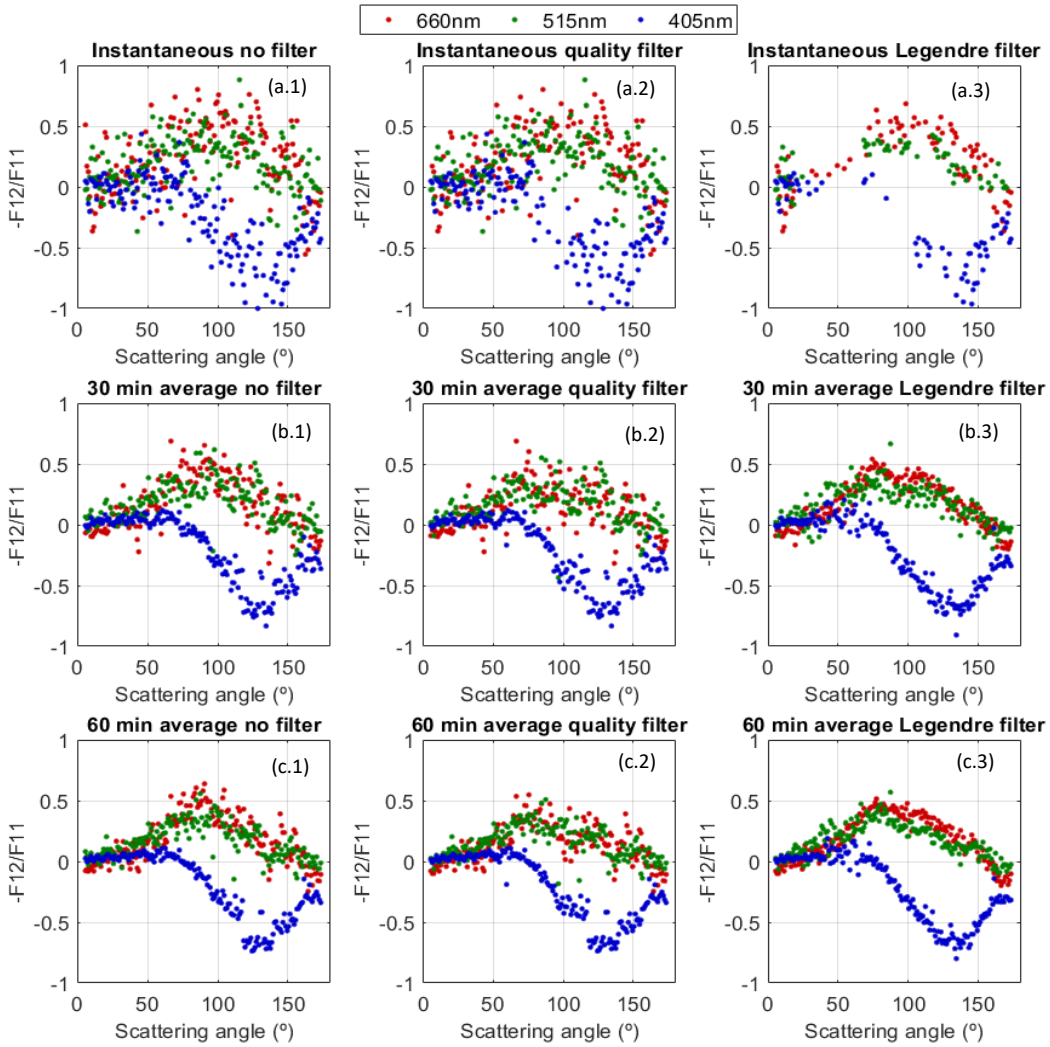

Figure 8 in Bazo et al., 2024. Polarized phase function $-F_{12}/F_{11}$ for different stages of the filtering process. Rows (a), (b), (c) show instantaneous, 30 min average and 60 min average polarized phase functions, respectively. Columns (1), (2), (3) show non-filtered, quality filtered, and Legendre filtered data, respectively.

Validation versus TSI Integrating Nephelometer

The PI-Neph was measuring continuously in April-September 2022 and a with a collocated integrated nephelometer (TSI 3563). That allowed to evaluate the scattering coefficient defined as:

$$\sigma_{sca}(\lambda) = \frac{1}{2}\int_0^{180} F_{11}(\theta,\lambda) sin\,\theta \cdot d\theta\,, \qquad (1)$$

The integrating nephelometer measured in the range 7º-170º while the PI-Neph measured F_{11} in the range 5º-175º. Therefore, to avoid any hypothesis about corrections we evaluated the truncated scattering coefficient in the range 7-170º that was measured directly by both instruments. The results of the general fits were very good with slopes around 0.93 and R^2 above 0.81. Differences could be associated with the uncertainties of each instrument. Computed RMSE were of 3.79, 4.57 and 8.51 for 660, 515 and 405 nm channels. More details are in Bazo et al., (2024)

****

We understand referee concerns about the uncertainties of the instrument, and therefore we have included an overview of instrument uncertainties and data quality check procedure in Section 2.2.1 (L199-225):

"An extensive analysis of the error sources in the PI-Neph was performed in Bazo et al., (2024) but an overview is given here: an exhaustive calibration of the instrument is performed consisting of two different steps. The first is a geometric correction that corrects from the different light paths to the different pixels in the CMOS camera. Later the absolute calibration permits to obtain phase matrix elements in physical units. In each step we used known scatterers (CO_2 and particle free air) whose parallel and perpendicular signals can be computed analytically using the Rayleigh theory (Anderson et al., 1996). Evaluation of the calibration with time did reveal great stability (variations around 3%). Instrument stability was evaluated with CO_2 measurements at a constant flow rate of 10 $Lmin^{-1}$ during 15 min. These measurements revealed constant values of scattering coefficients with differences below 1% versus theoretical values from (Bodhaine et al., 1991). Finally, inherent aspects of the imaging technique were evaluated such as the impact of the exposure time. The largest noise is found for exposure times below 5 s, while the smoother values are obtained for exposure times of 10-20 s. However, large exposure times can yield to more angles that are saturated, and the software must find a compromise between noise and saturation. Thus, the typical exposure time is of 10 s and with that we estimate that uncertainties in measured parallel and perpendicular signals are around 5% in laboratory conditions. The evaluation of the instrument versus known scattered (monodisperse polystyrene spheres - PSL) showed good agreements with RMSE around 0.10 for both F_{11} and $-F_{12}/F_{11}$.

The uncertainties in direct measurements of the instrument (parallel and perpendicular signals) under laboratory conditions imply uncertainties below 10% in F_{11} and below 20% in $-F_{12}/F_{11}$. However, in-situ measurements present natural variability of the aerosol sampled and the differences can be enhanced because of the short exposure times (~ 10s). Effects during the measurements such as saturation or low signal to noise ratios (SNR) of some pixels can happen. Other issues such as the passage of an individual super-coarse particle can have an impact on certain angles of the phase matrix. Therefore, we apply a data-quality check procedure that accounts for all these issues and provide an effective phase matrix representative of an average time of 30 min or 1 hour, depending on the specific conditions of natural aerosol variability. Note that standard deviations during these periods might be larger than the uncertainties of the instrument. Details of this quality check procedure are in Bazo et al., (2024)."

To clarify better the main results of our study, we have modified the abstract that now is given by (L16-42):

"This work investigates the scattering matrix elements during different Saharan dust outbreaks over Granada (South-East Spain) in 2022 using the Polarized Imaging Nephelometer (PI-Neph PIN100, GRASP-Earth). The PI-Neph is capable of measuring continuously the phase function (F_{11}) and the polarized phase function ($-F_{12}/F_{11}$) at three different wavelengths (405, 515 and 660 nm) in the range 5° - 175° with 1° resolution for ambient aerosol samples. Extreme dust events (PM_{10} > 700 μgm^{-3}) occurring in March 2022 are compared with more frequent and moderate events registered in summer 2022

(PM_{10} between 50 and 100 μgm^{-3}). These intercomparisons allow the evaluation of F_{11} and $-F_{12}/F_{11}$ when dust particles predominate in the aerosol sample, but there is a possible mixture with other anthropogenic particles. For F_{11} there are no remarkable differences between extreme and moderate events. However, results of $-F_{12}/F_{11}$ show large differences between extreme and moderate events: for 660 nm the $-F_{12}/F_{11}$ pattern is characterized by a bell-shape with a positive maximum in the 90º-120º scattering region, and this pattern is observed both in the extreme and moderate dust events. However, there are remarkable differences in $-F_{12}/F_{11}$ at 405 nm showing a very similar pattern with 660 nm during the peaks of the extreme dust events while for moderate events it shows a different pattern characterized by values around zero up to ~ 50°, decreasing later to negative values ~120° and increasing again to values close to zero in the backward scattering region. For 515 nm we found out intermedia patterns. The temporal evolutions during the extreme dust events reveal that $-F_{12}/F_{11}$ is very sensitive to the particle concentration at 405 nm. For the peak of the events, F_{11} and $-F_{12}/F_{11}$ agree with the laboratory measurements available in the Amsterdam-Granada Light Scattering database at all wavelengths. The combination of PI-Neph measurements with additional in-situ instrumentation allowed to obtain scattering (SAE) and absorption (AAE) Ångström exponents and to conduct a typing classification that revealed extreme dust events as pure dust, while moderate dust events were classified as a mixture of dust with urban background pollution. In addition, simulations with the Generalized Retrieval of Atmosphere and Surface Properties (GRASP) code explain the different patterns in $-F_{12}/F_{11}$ with changes in the refractive indexes and the contributions of the fine and coarse mode. Therefore, our results confirm that differences in the phase matrix elements of Saharan dust outbreaks of varying intensity can be explained by the mixing conditions of dust with the background particles."

And we have also modified part of the introduction to contextualize better the study and objectives of our manuscript. The main modifications are in Lines 82-158:

"Remote sensing techniques are widely used to infer dust properties. For example, passive remote sensing techniques such as sun-photometry by the Aerosol Robotic Network (AERONET – (Holben et al., 1998)) or star/moon photometry (i.e Pérez-Ramírez et al., 2008, Pérez-Ramírez et al., 2011;Berkoff et al., 2011) allow to have a representation of column-integrated values, particularly aerosol optical depth (AOD). But to infer other aerosol optical (e.g. aerosol complex refractive index and single scattering albedo) and microphysical (e.g. aerosol size distribution) properties it is necessary to solve ill-posed problems where the information content is low (Dubovik & King, 2000; King et al., 1978; Nakajima et al., 1996; Olmo et al., 2006, 2008; Pérez-Ramírez et al., 2015). These algorithms use the Mie theory for the internal computation of particles phase functions, but in the case of dust particles more complex approaches such as T-Matrix are needed because of the non-sphericity of dust particles (Mischenko & Travis, 1994, 1997) . Nevertheless, several inversion algorithms have been developed incorporating T-Matrix modeling, being one of the most popular algorithms developed within the AERONET network (Dubovik et al., 2006).

Ground-based remote sensing techniques are only representative of the measurement site, and to face these limitations satellite measurements are ideal because they can cover wide regions of the world. However, passive remote sensing space platforms deal with

additional complexity in the retrieval of aerosol properties because of the influence of surface reflectance (Kahn et al., 1998; Levy et al., 2007). The simplest retrievals use look-up tables with a priori aerosol types with great success in obtaining AOD, but limited capacity for obtaining other aerosol parameters because of the difficulties to separate the signals corresponding to the atmosphere and surface (Dubovik et al., 2019). To solve these limitations, the use of multiwavelength and multi-angle polarization measurements is ideal to improve the information content (Mishchenko et al., 2007). Some of the first polarized-based measurements for aerosol studies were carried out by the POLDER instrument (POlarization and Directionality of the Earth's Reflectances – (Deuzé et al., 1993)) that acquired 9 years of data. These measurements were used as inputs in the Generalized Retrieval of Atmosphere and Surface Properties algorithm (GRASP – (Dubovik et al., 2014, 2021)) for obtaining extended aerosol optical and microphysical properties. Algorithms such as GRASP are becoming the operational algorithms in new satellite missions (Remer et al., 2019; Fuertes et al., 2022; Hasekamp et al., 2024), but these algorithms need phase matrix measurements that allows the optimization of the kernels used internally, particularly for non-spherical particles.

The main difficulties for measuring aerosol phase matrix of ambient air are in the design and development of appropriate polar nephelometry capable of measuring light scattered with appropriate angular resolution. The first polar nephelometry developments were based on moveable detectors, but they must be mechanically stable and require a constant population of aerosol particles that does not change appreciably during the detector sweep (Holland & Gagne, n.d.; Hovenier et al., 2003; Jaggard et al., 1981; Kuik et al., 1991; Perry et al., 1978; Volten et al., 2001a). Other polar nephelometry designs use arrays of many detectors placed on representative scattering angles (Barkey et al., 1999; Gayet et al., 1998; Pope et al., 1992; West et al., 1997; Wyatt et al., 1988), but this technique requires careful calibration of the detectors and generally suffers from low angular resolution (~2º). Those instrumental limitations have implied that the usual study of scattering matrix elements of dust particles were done in the laboratory for synthetic samples minerals that compose dust particles (Curtis et al., 2008; Huang et al., 2020; Meland et al., 2010; Muñoz et al., 2010a; J. B. Renard et al., 2014; J.-B. Renard et al., 2010) or with collected dust samples (Muñoz et al., 2007a; J. B. Renard et al., 2014; J.-B. Renard et al., 2010, 2024). Actually, the parametrizations of mineral dust phase matrix used for AERONET algorithm were calculated by fitting the laboratory measurements of different non-spherical particles samples (i.e. Dubovik et al., 2006). Such measurements were performed at a few wavelengths, and what is more important, they might be non-representative of real aerosol measurements because of the different transformations and interactions of dust particles since they are emitted in their source regions. There is therefore a current challenge in having an extended database of measurements of dust phase matrix elements for different dust types and mixtures.

The latest developments in polar nephelometry use imaging techniques (Bian et al., 2017; Curtis et al., 2007; Dolgos & Martins, 2014) to determine phase matrix with single detector and relatively compact design that does not require moveable parts. The Polarized Imaging Nephelometer (PI-Neph) was one of the first designs of a polar nephelometer that used imaging techniques, developed by the University of Maryland, Baltimore County (UMBC). This first prototype of the PI-Neph could acquire aerosol phase matrix at 473, 532 and 671 nm with 0.5º resolution. The instrument was deployed

on the NASA DC8 aircraft and operated during special field campaigns (Espinosa et al., 2018; Reed Espinosa et al., 2017). Other PI-Neph instruments based on the first UMBC design are operated by NOAA (Ahern et al., 2022; Manfred et al., 2018). The main novelty of these prototypes is that they measure phase matrix elements of ambient air, where conditions can be very different to laboratory measurements. However, to date none of these instruments have been operating continuously and reported any multiwavelength measurements of Saharan dust. The imaging technique is being expanded worldwide with further designs although limited to laboratory operation yet (Moallemi et al., 2023). All designs in polar nephelometry present physical limitations that limit the measurements to the range 3º-178º, but synthetic tests have revealed that multi-wavelength polarimetric PI-Neph measurements improve the information content for the retrieval of aerosol optical and microphysical properties (Moallemi et al., 2022) . Therefore, measurements of dust phase matrix elements for ambient aerosol samples in the atmosphere will serve to further advance in the understanding of mineral dust absorption properties and chemical composition (Di Biagio et al., 2017, 2019).

This work presents phase matrix measurements of ambient Saharan dust particles by the GRASP-Earth's (https://www.grasp-earth.com/) multi-wavelength PI-Neph. The instrument was developed using the heritage of previous PI-Neph developments made by UMBC and can provide aerosol phase matrix elements at 405, 515 and 660 nm of ambient samples in the range 5° - 175° with 1° resolution. Measurements were acquired in the urban background station (UGR) of the Andalusian Global ObseRvatory of the Atmosphere (AGORA) located in the Southeast of the Iberian Peninsula where the main source of natural particles is the Sahara Desert's transported particles (Querol et al., 2019). We present the results for extreme outbreaks that occurred in March 2022 (Rodríguez & López-Darias, 2024) with PM_{10} (particulate matter with diameter < 10 μm) concentrations over 700 μgm^{-3}, and for more typical situations of moderate dust events with PM_{10} concentrations ~100 μgm^{-3}. The measurements presented of the phase matrix for Saharan dust are unique and are and step forward from the ancillary measurements performed in the region by Horvath et al., (2018) with a single wavelength polar nephelometry (no polarization was available)."

GENERAL COMMENTS

The exact method for obtaining several key optical quantities is not specified. Four aspects of the data were particularly unclear:

(1) The PI-NEPH does not measure F_{11} at exact backscattering, but LR is provided throughout. How was $F_{11}(180°)$ estimated given the available data?

For the computation of the scattering coefficient, the objective is not to compute an exact value for $F_{11}(180°)$. Indeed, we want to estimate F_{11} for the entire range 175º - 180º. To do so we used the methodology based on linear extrapolation using the neighbors measurements. According to studies of Horvath et al., (2015), these linear extrapolations only imply 5% uncertainty in the computation of σ_{scat}, g and B_s. We have modified the text to clarify this point (L233-235).

"…where data from 0 to 5° and from 175 to 180° have been linearly extrapolated to obtain the complete phase function which according to Horvath (2015) only implies uncertainties up to 5% in the computations of the σ_{scat}, g and B_s."

Once computed the scattering coefficients, we use measurements of the absorption coefficient to obtain single scattering albedo – note that the absorption coefficient at the exact wavelengths of the PI-Neph is computed using absorption Angström exponent (see reply to comments below). However, the computation of extinction-to-backscattering ratio (LRs) requires exact values of $F_{11}(180°)$, and we use values of $F_{11}(180°)$ computed following the linear extrapolation method. More robust methods could have been used being the differences between up to 20-30 % (i.e. Gomez-Martin et al., 2021). These differences are close to the uncertainties of the PI-Neph and therefore LRs estimations will serve only as an illustration of how this parameter varies under different conditions. The text has been modified accordingly (L265-268):

"The computation of $F_{11}(180°)$ has been made using the interpolation method used for completing the entire angular range in σ_{sca}. Other more robust methods can be used (i.e. Gomez-Martin et al., 2021), that can imply differences in $F_{11}(180°)$ of up to 20-30%. Therefore, LRs estimations will serve as an illustration of how this parameter varies under different conditions"

We have also added the reference Horvath (2015) to the reference list:

"Horvath, H.: Extrapolation of a truncated aerosol volume scattering function to the far forward and back region, J. Aerosol Sci., https://doi.org/10.1016/j.jaerosci.2015.08.001, 2015."

(2) The TSI integrating nephelometer is described and it is stated that "σ_{sca} can be also obtained with the integrating nephelometer" but it's not clear if this data is ever used, or all σ_{sca} values are obtained via the PI-Neph. If the latter is the case, then there is no need to introduce the TSI instrument.

We have used the scattering coefficients from the TSI integrating nephelometer to obtain the Angstrom matrix in section 5.1. We could have done this with the scattering coefficients by the PI-Neph, but we wanted to maintain the same wavelength range for the calculation of the scattering Angstrom exponent to apply the classification proposed by Schmeisser et al. (2017). We have added the following statements to the manuscript (L623-625):

"For the SAE, we have used the σ_{sca} measured with the TSI integrating nephelometer since it directly provides measurements at the same wavelengths than those required in the Schmeisser et al., (2017)."

(3) Equations 5 and 6 show general definitions for SAE and AAE but I don't think the exact wavelengths used are ever stated. Please specify the wavelengths used in these calculations.

-Thanks for this point. We included this information in the results section, whenever both variables were introduced. But to make this clear, we specified the wavelengths used for SAE and AAE when defining these parameters in Equations 5 and 6 (L257-L258).

"… The wavelengths used in this work to calculate both SAE and AAE have been 405-660 nm and 450-700 nm, respectively."

(4) How were the particles sampled and did any aspect of the process have the potential to produced size dependent biases relative to the ambient aerosol. This is particularly import for the dust dominated scenes since a significant portion of dust particle volume concentration occurs in particles several microns in diameter that are easily lost to inlets and tubing.

-We agree with the reviewer's comment on the importance of sampling losses during dust events. For this reason, the experimental set-up used in this study was designed to minimize any potential losses of large particles. Measurements were performed with a total inlet (no size-cut) consisting of 5 m long stainless-steel tube displayed vertically to minimize deposition losses, and a flow splitter that provided the appropriate flow to each instrument. The final connection to all instruments, including the PI-Neph, was performed with conductive tubing (TSI), avoiding bends. No aerosol dryer, that might lead to additional particle losses, was used due to the low ambient relative humidity at the measurement site. Further details of the inlet system can be found in Lyamani et al. (2008). We have added the following statement to the manuscript (L177-L181):

"In this study, particles were sampled using a total inlet (no size cut) that consist of a 5 m long stainless-steel tube with a 20 cm diameter (Lyamani et al., 2008). Inside the stainless-steel tube there are several pipes that split the aerosol flow into the different instruments. The inlet system is completely vertical to minimize deposition losses. The final connection to the instruments is performed with conductive tubing avoiding bends. Additionally, all the measurements were at ambient conditions (no aerosol dryer was used)."

And we have added the reference Lyamani et al. (2008) to the reference list between

"Lyamani, H., Olmo, F.J., Alados-Arboledas, L. Light scattering and absorption properties of aerosol particles in the urban environment of Granada, Spain. Atmos. Environ. 42, 2630–2642. https://doi.org/10.1016/j.atmosenv.2007.10.070, 2008"

<> At times measurements seem give nonphysical results. For example, $-F_{12}/F_{11}$ is unrealistically low in Figure 9b.2 and trends in AAE do not always match spectral trends in SSA (see specific comment – LN 467). Measurement errors are unavoidable, but I think these issues further emphasize the need to include well characterized uncertainties on the measurements, as the other reviewers have noted.

We agree with the referee, and we have added a more detailed description of the instrument including uncertainties of the instrument (see above comment). About the specific case in Figure 9.b, we highlight the variability in $-F_{12}/F_{11}$ due to the variability of aerosol samples in this period. Particularly, standards deviations at that moment were around 0.2, above the uncertainty of the instrument. This is now clarified in the revised manuscript (Lines 568-570):

"The standard deviations were 20-30% for F_{11} and around 0.2 in $-F_{12}/F_{11}$, which are larger than the uncertainties of the instruments for all cases and explained by the large variability of aerosol samples during the measurement process."

<> In my view, some conclusions are not well supported, especially regarding relationships between optical features and the underlying microphysical properties of the aerosol (see specific comments for examples). The inclusion of additional measurements like those of size distribution or chemical composition, if available, could go a long way here towards proving or disproving some of these claims. Additionally, GRASP could be a very powerful tool for interrogating some of the relationships between optics and microphysics hypothesized, at least in the context of spheroidal particles. This seems to have been started in Section 3.2.4 but it was not clear to me why this analysis was not extended to actual inversions of the phase matrix (and possibly absorption) measurements.

The referee is right in all the issues raised to complement our study, but there were many shortcomings that avoided adding those points to our analyses.

The AGORA observatory has an Aerosol Particle Sizer (APS TSI 3321) and an Aerosol Chemical Speciation Monitor (ACSM - Aerodyne). However, during the two extreme Saharan dust outbreaks in March 2022 these two instruments were not available to operate at the station due to maintenance issues, that got extended until summer 2022. That is why these important measurements were not available for this study.

Concerning the issue pointed out of using GRASP, we would like to mention that currently GRASP is not optimized to perform inversion of bimodal size distributions representative of mixture of mineral dust (coarse particles) and anthropogenic pollution (fine mode particles) using F_{11} and $-F_{12}/F_{11}$ as inputs. The simulations carried out only illustrate that the patterns observed in section 3.2.4 can be obtained for a mixture of particles– now section 5.2 in the revised manuscript. We have highlighted these issues in the revised manuscript (Lines 763-7678):

"However, studying the relationships between measured F_{11} and $-F_{12}/F_{11}$ with other aerosol optical and microphysical properties requires further analyses because F_{11} and $-F_{12}/F_{11}$ ultimately depend on the size distribution, refractive indexes, and particle shapes. The problem is even more complex if we differentiate optical properties between fine and coarse mode. Future optimization in GRASP will permit the retrieval of aerosol refractive indexes between fine and coarse mode separately using as inputs F_{11} and $-F_{12}/F_{11}$, and thus permitting further analyses of the different study cases discussed in this work."

Moreover, given the importance of the points raised by the referee, we have decided to modify the conclusion section to emphasize the need to acquire correlative measurements of PI-Neph with additional instrumentation that provide information of particles size distribution and of particles chemical compositions (Lines 860-864).

"However, going further in understanding the interaction of dust with these anthropogenic particles requires further analyses that provide the chemical composition and size distribution of the ensemble of particles and the final composition and shape of the particles after interacting. This is planned in future studies that will allow a more complete comprehensive analysis."

<> I found the structure of the text to be a bit difficult to follow. Perhaps moving some sections towards to beginning may help provide a clearer "landscape" for the reader as they progress through the text. Moving section 3.2.3 higher up, or at least providing a clear and consistent definitions of the aerosol types observed/hypothesized earlier on would help a lot.

Due to comments from Reviewer 2, we have modified the numbering of the results and have divided them into three different sections. We have divided section 3 'Results and Discussion' in three different sections:

Section 3 'Overview of extreme dust events during March 2022' that initially was sub-section 3.2.1.1. This section gives an overview of the meteorological conditions associated with these two extreme events, plus some satellite observations that give an overview of the intensity of the dust plume.

Section 4 'Results of aerosol phase matrix from different dust scenarios': Here we include now sub-section 4.1 'Extreme events' that was initially subsection 3.2.1.2, and sub-section 4.2 'Moderate dust events during spring/summer 2022' that was initially subsection 3.2.2

Section 5 'Discussions': Here we include the sub-section 5.1 'Comprehensive assessment of the different dust events' that initially was sub-section 3.2.3 and sub-section 5.2 'Phase matrix simulations for different aerosol mixture scenarios' that initially was sub-section 3.2.4

SPECIFIC COMMENTS

Eq 1: I see the reason 2π and 4π were kept separate, but I wonder if reducing it to 1/2 would make this equation a little easier to read, especially given the improved consistency with Eq 4.

Thank you for pointing this out. We have changed the expression of Eq. 1 so it is consistent with Eq. 4.

LN 140: This sentence is confusing. It is not clear that the single beam novelty they are referring to is the PIN-100. Also, it should be "University of Maryland, Baltimore County".

-The referee is right, it should be noted University of Maryland, Baltimore County. We have modified this throughout the text. Concerning the text about the PIN-100, we have modified this part of the introduction (L189-191):

"The instrument uses previous heritage in PI-Neph developments in the University of Maryland, Baltimore County (Dolgos and Martins, 2014), where the novelty in the PIN-100 is the use of one beam instead of a mirror system to fold the laser beam, as was in previous models."

Eq 7: How is $F_{11}(180°)$ obtained? F11 has a lot of structure in the backscattering direction and linear extrapolation is probably not very accurate, especially for nonspherical particle like dust. The method for extrapolating should be some discussion of the potential resulting uncertainties should be provided.

As explained discussed in the general comment above, $F_{11}(180°)$ is only used in the computations of lidar ratios, that is a variable of interest for the lidar community. We used the linear extrapolation method, and we highlight the possible differences versus other more robust methods. We have corrected the text highlighting that the values given in the text serve to illustrate the LRs variability (L265-L268):

"The computation of $F_{11}(180°)$ has been made using the interpolation method used for completing the entire angular range in σ_{sca}. Other more robust methods can be used (i.e. Gomez-Martin et al., 2021), that can imply differences in $F_{11}(180°)$ of up to 20-30%. Therefore LRs estimations will serve as an illustration of how this parameter varies under different conditions."

LN 197: "pression" -> "pressure"

-Done.

Figure 2/3: If all four subplots needed the color bars should be made consistent for all subplots and some discussion should be added to the text explaining what features differentiating the four subplots should be provided in the text. As it is, there is no significant difference that stands out to me between the four model times shown.

-We have modified the color bars of Figures 2 and 3 and now the same scale is shown for all cases. We plotted the evolution of each event with the CAMS model with the objective of illustrating the temporal evolution of the event and what regions are affected. The following discussions have been added to the text (L290-L292):

"The wind field at the surface is also represented in these Figures, and different times selected serve to understand how the dust was transported and affected different regions…"

And Lines 296-298:

The wind pattern reveals how dust enters through the southeast of the Iberian Peninsula reaching later the northwest latitudes reaching even southern France.

Figure 4: It would be good to include the name of the specific satellite/sensor that obtained these images.

-Done, now the caption reads (L346-L347):

"Figure 4. (a) Satellite image from 15th March 2022. (b) Satellite image from 25th March 2022. Images from https://wvs.earthdata.nasa.gov, obtained with MODIS (Moderate Resolution Imaging Spectroradiometer)."

LN 273: I assume this is a regulatory limit? This should be made clearer in the text.

-The reviewer is right, it is a regulatory limit. We have added the following information in the text to make it clearer (L358):

"…regulatory daily limit value of 50 μgm^{-3} established by the Ambient Air Quality Directive (2008/50/CE European Directive)"

LN 284: It is well established that nonspherical particles can have weaker scattering near scattering angles of 180° but that only contributes a very small amount to the

total Bs which integrates from 90° to 180°. If there is a specific reference stating Bs decreases with particle nonsphericity that should be included here, otherwise I would suggest this sentence be removed or rephrased.

The text has been modified as follows (L367-L369):

"The low values of the Bs observed at the three wavelengths are associated with small values of F_{11} at the backward scattering angles, which is common for non-spherical particles (Nousiainen & Kandler, 2015)."

And we have added the reference Nousiainen and Kandler (2015) to the reference list between

Nousiainen, T., & Kandler, K. (2015). Light scattering by atmospheric mineral dust particles. In *Light Scattering Reviews 9: Light Scattering and Radiative Transfer* (pp. 3–52). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-37985-7_1

Figure 5: The time series for Event 1 should be extended back to at least 7:00 on March 15th to be consistent with F11 and F12 shown in Figure 6 and to give better context to the background conditions prior to the dust intrusion.

-We agree with the reviewer on this. However, most of the data between 7:00 and 12:00 did not pass the quality check procedure explained above, due to the intensity of the episode. We tried to monitor the dust event by adjusting the exposure time of the images to prevent saturation of the detector, but it was still impossible to obtain quality data so we had to lower the gain of the camera. That is why we do have a continuous quality-checked data set from 12:00 until the end of the episode. In Figure R1, the reviewer can find the extended time series where the is clearly a lack of data in the PI-Neph products. For this reason, we did not include this period in the figure but included Table 1 which contains the properties of Figure 5 at 7:00.

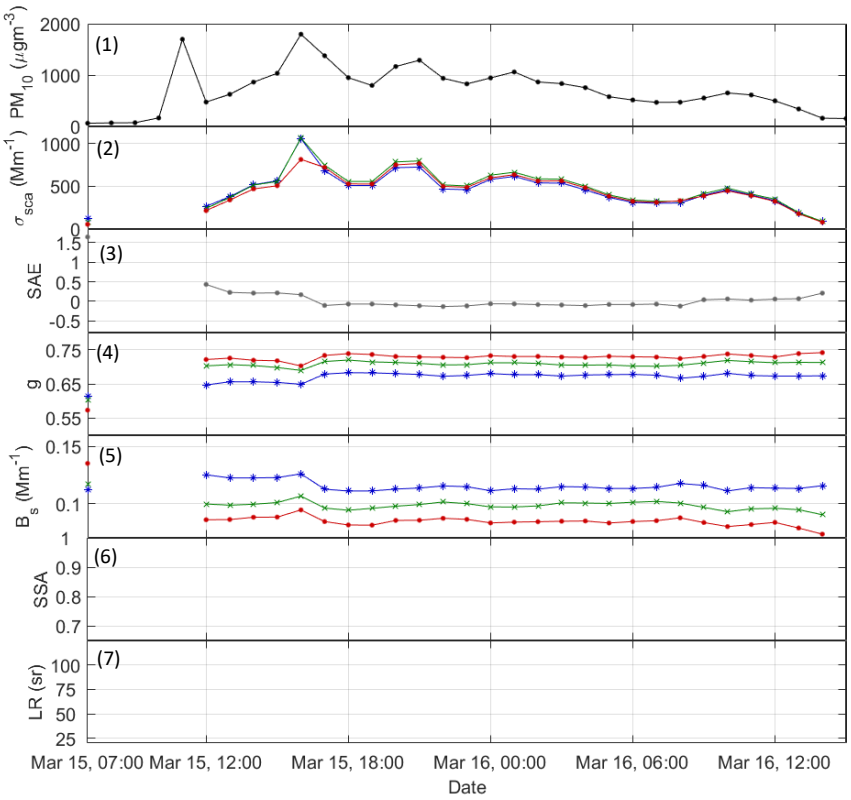

Figure R1. Time series of the PM_{10} (1), σ_{sca} – scattering coefficient (2), SAE – scattering Angström exponent (3), g – asymmetry parameter (4), Bs – fraction of backscattered light (5), SSA – single scattering albedo (6) and LR – lidar ratio (7) for the extreme dust events of 15th March.

Section 3.2.1.2: Because data at each angle in a PI-Neph is based on the sum of many individual pixels, saturation tends to happen gradually and result in a steadily increasing bias rather than a sharp, obvious upper limit in F11. The challenge of avoiding such saturation given an aerosol with a combination of a strong forward peak and very high concentration is not discussed in the text. Was each image checked to ensure that the counts of all pixels were below the upper limit of the detector's dynamic range? The text should include at least a brief explanation of the steps taken to ensure saturation did not bias the measurements.

-This is a very interesting point. As stated in Bazo et al. (2024) the software of the PI-Neph PIN-100 has a quality check parameter that marks the point as non-reliable if saturation happens. Whenever this happens, we discard the affected point and if it happens for many angles, we directly discard the complete measurement. Therefore, since we have the quality check parameter to account for saturation, we have not checked the images every image. Moreover, we had to change the gain of the camera due to the extremely high scattering that happens in the extreme dust events. For these reasons, and to also clarify why there were some gaps of measurements early in the morning on 15th March, we add the following text to the manuscript (L408-L413):

"Phase matrix elements were exhaustively monitored with the use of the PI-Neph during both extreme events. Given the high aerosol concentrations, the usual configuration of the measurements could lead to saturation of many angles in the forward scattering.

Therefore, it was necessary to reduce the gain of the PI-Neph's camera, changing the dynamic range of the camera for obtaining non-saturated measurements at such high concentrations. After this was done, some instantaneous measurements might still present saturation at some angles that were filtered out by the data quality criterion of the instrument (Section 2.2.1)."

LN 341: Again, it is not clear to me how the values of g and Bs on their own–which are generally accepted to have the strongest dependence to particle size–are able to clearly suggest the predominance of spherical vs nonspherical particles.

-We agree with the reviewer, and we are sorry for the misunderstanding in the original manuscript. What we wanted to remark is that the obtained values of g and Bs are typical of that observed for fine mode particles predominance. After carefully reading the comments of the three referees, we agree that we must be very cautious about making conclusions on particle size, shape and chemical composition from our phase matrix measurements. They have encouraged us to re-read the manuscript and make corrections to avoid any ambiguity. Specifically, for the line mentioned here we have simplified and avoided any mention to spherical vs non-spherical particles.

LN 343: It's worth noting that holding intensive properties constant but increasing the aerosol concentration will increase F11 at all angles. Given that the authors seem to be using the PI-Neph to interrogate changes in microphysics here, I wonder if P11 would be a better variable to directly analyze? Total scattering and spectral dependencies between the three wavelengths can easily be captured in σscat.

We agree with the reviewer, and we have re-written these sentences including a better discussion of F_{11}. Now, between lines 431 – 442 is given by:

"Figure 6 shows a general pattern in F_{11} characterized by strong predominance of forward scattering up to two orders of magnitude greater than backward scattering. However, there are significant changes in both magnitudes and spectral dependence over time, that is, with the intensity of the dust outbreak passage. At the beginning of the dust event (Fig. 6a), the values of F_{11} in the forward scattering region are around 1000 $Mm^{-1}sr^{-1}$ for all three wavelengths, which is even one order of magnitude lower when compared with the cases at the other moments of the event (i.e. 50000 $Mm^{-1}sr^{-1}$ for the three channels during the peak). Also, at the beginning of the event (Fig. 6a) notable spectral separation in F_{11} is observed while such spectral separation is negligible during the rest of the event when coarse mode particles largely predominate. All F_{11} show the minimum in the region 120º-140º but the magnitude of that minimum varies between the different stages. Also, around that minimum is the region where some spectral difference is observed during the cases of strong predominance of coarse mode (Fig. 6c-d). A recovery from that minimum is also observed, being more pronounced in cases close to the peak of the event."

The hourly temporal evolutions of F_{11} and $-F_{12}/F_{11}$ are given in the supplementary material for more details.

LN 348: "constant tendency" needs to be made more precise. It looks to me like F11(~180°) is an order of magnitude higher in panel c.1 relative to a.1 in Figure 6.

-The reviewer is right about the values of F_{11} from panels a.1 and c.1. By constant tendency we only refer that in the backward region the different F_{11} show similar pattern in the backward scattering. In the revised manuscript we avoid these misunderstandings (see previous comment with the revised paragraph).

LN 349: While the spectral dependence is less, the red channel, and maybe the green too, appears lower than blue to me in Figure 6a.1. The language here should also be made more precise.

-The reviewer is right. We have avoided ambiguities in the revised manuscript and re-write the whole paragraph – see previous comment.

LN 381: Should this say, "increase in F11" rather than "decrease"?

-Following the suggestion of all referees, we have avoided ambiguities in our interpretation of the results. The whole paragraph where we discussed F_{11} on 15th-16th March was rewritten, and now is given as (L481-491):

"Figure 7 shows that F_{11} patterns are very similar to the previous extreme event on 15th - 16th March, with strong predominance of forward scattering (~25000 $Mm^{-1}sr^{-1}$), being two orders of magnitude above the backscattering (~100 $Mm^{-1}sr^{-1}$) at the peak of the event on 25th March 9:00 UTC. There are no significant spectral differences, as also happened for the other extreme event on 15th-16th March. These patterns in F_{11} agree with laboratory measurements of dust samples (i.e. Muñoz et al., 2007; Renard et al., 2014; Volten et al., 2001). Nevertheless, there are some features in F_{11} with different situations: the slope in F_{11} in the forward scattering region becomes sharper when the PM_{10} concentrations are higher (Figs. 7b.1 and 7c.1). For the backward region F_{11} shows a flatter behavior for high PM_{10} concentrations (Figs. 7b.1 and 7.c.1), while for the cases with lower PM_{10} concentrations there is a sharp increase in scattering from 150° to 180°. During the previous extreme dust outbreak, we observed flat patterns for the backward scattering region during the peaks of the dust intrusions."

LN 383: Many things can increase the forward peak of F_{11} and it is known to be particularly sensitive to aerosol size. However, it seems to be implied here that change in shape is the most probable cause. This should be explained.

-The reviewer is right. As we have commented in the previous referee´s question, we have re-written the entire paragraph to avoid ambiguities. In the conclusion section we mention that the patterns in F_{11} and $-F_{12}/F_{11}$ depend on size distribution, particle shapes and refractive indexes. More measurements are needed to draw further conclusions. Also, optimization of GRASP inversion will help in the future.

Figure 7: I think the caption should say "four different moments".

-Thank you for identifying this typo, we have corrected it.

LN 425: I found this a bit hard to follow. Is the statement that some studies suggest that SAE is a useful discriminator between the pure mineral dust and dusty mixtures with urban background pollution?

-We use SAE as a proxy that helps to identify possible presence of mineral dust, which agrees with what's stated in the bibliography (i.e. Lyamani et al., 2008; Teri et al., 2024).

Measurements of PM10 plus the Angstrom matrix classification proposed by Cazorla et al. (2013) and Schmeisser et al. (2017) are later used to confirm the presence of mineral dust.

To clarify that SAE is used as a proxy based on bibliography, we have modified the manuscript that now reads as (L524-L527):

"For the identification of cases with influence of mineral dust particles, data shown in Fig. 8 are filtered out and correspond only to values of SAE< 1, which are used as a proxy to identify possible presence of dust particles in the atmosphere (i.e. Lyamani et al., 2010; Teri et al., 2024).

Table 1: Aside from the ± variability windows, is there any information in this table that is not already provided in Figure 5? If not, I wonder if those could be represented by bars in Figure 5 and the table could be moved to the supplement.

-We thank the referee for this suggestion. We consider that adding error bars to Figure 5 would make the graph more difficult to read, since there are several subplots that show many variables. Also, the objective of Figure 5 is to show a detailed temporal evolution. Table 1 is devoted to showing the mean values of the different variables at the exact times of PI-Neph data in Figures 6 and 7. We therefore believe that the current configuration of the Figures and Tables is appropriate.

LN 443: "absorbing particles" should be made more precise. Is this referring specifically to black carbon?

-We agree with the referee. We have changed 'absorbing particles' to 'black carbon particles'

LN 463: AAE=1.71 is significantly higher than pure black carbon so it may be too strong to say that "absorption is mostly explained by black carbon".

That value is for the case on 16th April that is the case with more complex aerosol mixture. Black carbon is a possible contributor in the mixture. We have modified the text to remark that flat dependences of SSA might suggest the presence of black carbon, but the high AAE suggests the presence of other particles such as mineral dust (L575-L581):

"The first event on 16th April is the one of the most complex in terms of mixture of particles, with the lowest PM_{10} but with the largest SAE and AAE. The values of g suggest the lower contribution of dust particles when compared to the other cases (Horvath et al., 2018). The flat spectral pattern in SSA is the typical observed when there is contribution of fine mode particles of anthropogenic origin during a Saharan dust outbreak (Valenzuela et al., 2014). Nevertheless, the high AAE is not typical of black carbon and thus remarks the complexity of the mixture of aerosol particles in that day."

The reference Valenzuela et al., (2014) was added:

Valenzuela, A., Olmo, F. J., Lyamani, H., Granados-Muñoz, M. J., Antón, M., Guerrero-Rascado, J. L., Quirantes, A., Toledano, C., Perez-Ramírez, D., & Alados-Arboledas, L. (2014). Aerosol transport over the western mediterranean basin: Evidence of the contribution of fine particles to desert dust plumes over alborán island. *Journal of Geophysical Research*, *119*(22), 14,028-14,044. https://doi.org/10.1002/2014JD022044

LN 465: Again, it is not clear to me that g can be used as a proxy for non-sphericity.

-We have changed this sentence to (L576-577):

"The values of g suggest the lower contribution of dust particles when compared to the other cases (Horvath et al., 2018)"

LN 467: It should be noted that while the SSA measurements on July 5 may show stronger spectral dependence, the AAE measurement is actually lower on July 5 than Apr 16.

-The referee is right, we have added this information in L581-583.

"The more pronounced spectral SSA when compared to 16th April also relates to the influence of dust in absorption (Dubovik et al., 2002), although the event on 5th July shows a lower AAE than for the previous case."

LN 469: It should probably be emphasized here that eBC is only equivalent BC for a given amount of absorption. Even at the longer wavelengths at which eBC is derived, very high levels of dust could produce significant absorption.

-We agree with the reviewer, so we have added the following statement to the manuscript (L583-586):

"Finally, the case on 30th August is the one with the largest PM_{10} concentrations but also with the largest eBC, which can make a very complex mixture. However, it should be noted that when affected by high concentrations of dust, the dust particles might interfere with the eBC measurements."

Figure 9: Blue -F_{12}/F_{11} is much lower than any prior theoretical calculations (including those presented here in Figure 12) or measurements have shown. There are also some unnatural looking "kinks" in F_{11} around the same angular region. This emphasizes need for good uncertainty estimates and careful data screening.

-We agree with the referee. Uncertainties of the instrument are ~ 10% for F11 and ~20% for -F12/F11 but in laboratory conditions. For ambient samples, such as those presented in this work, natural variability of the samples during the 1 hour of data acquisition to fulfill the data quality requirements can lead to standard deviations larger than the uncertainties of the instrument. This is explicitly mentioned in the revised manuscript:

(L415-L419): "Computed standard deviations were larger than instrument uncertainties and they are associated with the variability of the different parcels of air sampled throughout the hour of measurements. These standard deviations were of ~ 20% for F_{11} and ranging between 0.1 and 0.2 for -F_{12}/F_{11} (minimums for the forward region and maximums in the middle region around 90º)."

L463-L465: "Again, the computed standard deviations are larger than the uncertainties of the instruments and they represent the variability of the samples. As for the previous extreme event, these standard deviations are ~20% for F_{11} and between 0.1- 0.2 for -F_{12}/F_{11}. "

L568-L570: "The standard deviations were 20-30% for F_{11} and around 0.2 in $-F_{12}/F_{11}$, which are larger than the uncertainties of the instruments for all cases and explained by the large variability of aerosol samples during the measurement process"

LN 504: I'm not following this sentence. Why would high levels of noise mean that the measurement is "very sensitivity to aerosol mixtures"?

-When comparing to Figure 7, phase functions in red and green channels for the stage of high dust concentrations (panel c) show a clear bell-shape typical of mineral dust. However, when the concentration of dust is not so high, this shape is not that clear, which is more like what we find in the phase functions of Figure 9. This is what we wanted to emphasize with this sentence, however the use of 'noisy' might not be very accurate. Therefore, we have re-phrased the sentence as (L605-L607):

"For 515 and 660 nm the $-F_{12}/F_{11}$ patterns are characterized by a bell-shape with large variability that might be associated with the complexity of the mixture of mineral dust and anthropogenic particles."

LN 506: This also left me confused on first read. Is the idea here that the blue channel of -F12/F11 is depressed in anthropogenic (or absorbing?) aerosol and the feature shows up at low dust concentrations when the dust is not optically dominate?

-Yes, the reviewer is right and that is the point we want to highlight. We have modified the text accordingly (L608-616):

"However, the $-F_{12}/F_{11}$ pattern in 405 nm shows a very different behavior, having $-F_{12}/F_{11}$ positive values until ~ 70° and negative $-F_{12}/F_{11}$ values for the following angles. Minimum $-F_{12}/F_{11}$ values are in the region around 120°. Therefore, $-F_{12}/F_{11}$ measurements can be potentially used for investigating the mixture of particles in the sample. That pattern with negative values has been observed in the UGR station for cases with no influence of Saharan dust particles (Bazo et al., 2024), and also for biomass-burning at 473 nm (Espinosa et al., 2017). Nevertheless, there are differences between the three different cases that might be associated with the differences in the mixtures of aerosol particles."

The reference Espinosa et al., (2017) was added:

Espinosa, W. R., Remer, L. A., Dubovik, O., Ziemba, L., Beyersdorf, A., Orozco, D., Schuster, G., Lapyonok, T., Fuertes, D., & Martins, J. V. (2017). Retrievals of aerosol optical and microphysical properties from Imaging Polar Nephelometer scattering measurements. *Atmospheric Measurement Techniques*, *10*(3), 811–824. https://doi.org/10.5194/amt-10-811-2017

LN 517: I think this sentence is redundant with the first sentence of this paragraph.

-We have removed the last part of the sentence.

LN 540: I believe it is important to emphasize here that utility of 405 nm P_{12} as a marker for smaller, likely anthropogenic particles may be unique to the region/circumstances studied here. For example, because of the size invariance rule for single scattering properties, an aerosol that is identical but with diameters only ~27% (1 - 515nm/405nm) smaller would produce scattering patterns in the blue that

are identical to what was measured in the 515 nm channel, which did not show this unique anthropogenic "fingerprint". This is not meant to diminish the finding, but instead to recommend that it be clearly caveated by the statement that this fingerprint may not be generalizable to all regions/seasons with dust intrusions.

-The reviewer is right. We wanted only to highlight the potential of polarization measurements to distinguish different mixtures of mineral dust with other types of particles (L643-L646):

"However, for the cases classified as mixtures, there was a different pattern $-F_{12}/F_{11}$ for the 405 nm channel. Thus, the typing classification explains the differences in the phase matrix for the temporal evolution of the extreme dust events during 15th-16th and 24th-25th March 2022 in the UGR station…."

In the conclusion section we have highlighted that more measurements (i.e. chemical composition, size distribution) and optimized GRASP retrievals are needed to further understand the interaction of dust with anthropogenic particles (L860-L864):

"However, going further in understanding the interaction of dust with these anthropogenic particles requires further analyses that provide the chemical composition and size distribution of the ensemble of particles and the final composition and shape of the particles after interacting. This is planned in future studies that will allow a more complete comprehensive analysis."

LN 603: Should this say, "blue wavelength"? I generally consider 400 nm to be the upper limit of the UV. Either way, the 473 nm measurements of Espinosa et al. (2017) were certainly not in the UV.

-The referee is right, we have replaced UV with 405 nm channel.

LN 655: I'm not sure what is meant by "optical properties" here. Either way, GRASP is certainly capable of simulating separate fine and coarse single scattering properties and refractive indices, as is evident in the forward calculation results show in Figure 12. I strongly suggest the authors include such inversions here.

-As the reviewer says, GRASP can do the forward calculations, which were done for studying how different mixtures can produce different patterns in F_{11} and $-F_{12}/F_{11}$. However, at the moment of writing the manuscript GRASP configuration using F_{11} and $-F_{12}/F_{11}$ as inputs is not optimized for the retrieval that provides different optical properties between fine and coarse mode. Indeed, GRASP retrieves effective optical parameters (i.e. refractive index) that are the same for fine and coarse mode. It is true that GRASP retrieval can separate between fine and coarse mode for the configuration lidar + sun photometry sky radiances and based on that experience we are working on GRASP optimization for retrieving two modes with different optical properties using F_{11} and $-F_{12}/F_{11}$. But this will be the scope of a future publication where we plan to analyze with GRASP the entire PI-Neph database at AGORA, with the use of additional in-situ instruments that provide measurements of absorption and size distribution. In that study we plan to add other cases such as biomass-burning that were measured in AGORA.

We have re-phrased this sentence to highlight that the forward simulation with GRASP helps us to understand that the negative values in $-F_{12}/F_{11}$ can be explained by a mixture

of dust with anthropogenic particles, but that going further needs future optimizations with GRASP. We would like also to note that future GRASP optimizations also need to implement super-coarse mode as suggested by Dr. Jean-Baptiste Renard.

L758-770: "Figure 12 results show how the presence of anthropogenic particles (fine mode) can alter the spectral dependencies in $-F_{12}/F_{11}$ when compared with only dust particles (coarse mode) in the sample, particularly in the blue channels. However, changes in the F_{11} patterns were not so evident. These results help to understand the different phase matrix elements discussed in this manuscript, and their temporal evolutions during the extreme dust events (Supplementary Material). However, studying the relationships between measured F_{11} and $-F_{12}/F_{11}$ with other aerosol optical and microphysical properties requires further analyses because F_{11} and $-F_{12}/F_{11}$ ultimately depends on the size distribution, refractive indexes, and particle shapes. The problem is even more complex if we differentiate optical properties between fine and coarse mode. Future optimization in GRASP will permit the retrieval of aerosol refractive indexes between fine and coarse mode separately using as inputs F_{11} and $-F_{12}/F_{11}$, and thus permitting further analyses of the different study cases discussed in this work. It is important to mention that the super-coarse mode can also affect the behavior of F_{11} and $-F_{12}/F_{11}$ and the presence of this mode is also observed for long-range transport (i.e. Renard et al., 2010). Future GRASP developments also need the consideration of this super-coarse mode."

LN 707: I still am not following how instrument noise can be linked with particle properties.

We thank the referee for pointing this out. We wanted to remark that the large variability in $-F_{12}/F_{11}$ at 515 and 660 nm can be associated with the different mixtures of mineral dust and anthropogenic particles. We have re-phrased this sentence to make this point clearer (L826-830):

" …. while for the other channels at 515 and 660 nm they show the bell-shape pattern but with larger variability if compared to the extreme dust cases, probably associated with the complexity of the mixture of fine (anthropogenic) and coarse (dust) mode particles. These patterns in F_{11} and $-F_{12}/F_{11}$ agreed with those observed in the intense Saharan dust outbreaks in March for the instants when the dust was entering and withdrawing"