Cloud screening and quality control algorithm for star photometer data: assessment with lidar measurements and with all-sky images

D. Pérez-Ramírez1,2,3, H. Lyamani1,2, F. J. Olmo1,2, D. N. Whiteman3, F. Navas-Guzmán1,2, and L. Alados-Arboledas1,2

1Centro Andaluz de Medio Ambiente (CEAMA), Universidad de Granada, Junta de Andalucía, Av. del Mediterráneo s/n, 18006 Granada, Spain
2Departamento de Física Aplicada, Universidad de Granada, Campus de Fuentenueva s/n, 18071 Granada, Spain
3Mesoscale Atmospheric Processes Laboratory, NASA Goddard Space Flight Center, Greenbelt, 20771, Maryland, USA

Correspondence to: D. Pérez-Ramírez (dperez@ugr.es, daniel.perezramirez@nasa.gov)

Received: 6 January 2012 – Published in Atmos. Meas. Tech. Discuss.: 17 February 2012
Revised: 6 June 2012 – Accepted: 11 June 2012 – Published: 11 July 2012

Abstract. This paper presents the development and set up of a cloud screening and data quality control algorithm for a star photometer based on CCD camera as detector. These algorithms are necessary for passive remote sensing techniques to retrieve the columnar aerosol optical depth, $\delta_{Ae}(\lambda)$, and precipitable water vapor content, $W$, at nighttime. This cloud screening procedure consists of calculating moving averages of $\delta_{Ae}(\lambda)$ and $W$ under different time-windows combined with a procedure for detecting outliers. Additionally, to avoid undesirable $\delta_{Ae}(\lambda)$ and $W$ fluctuations caused by the atmospheric turbulence, the data are averaged on 30 min. The algorithm is applied to the star photometer deployed in the city of Granada (37.16° N, 3.60° W, 680 m a.s.l.; South-East of Spain) for the measurements acquired between March 2007 and September 2009. The algorithm is evaluated with correlative measurements registered by a lidar system and also with all-sky images obtained at the sunset and sunrise of the previous and following days. Promising results are obtained detecting cloud-affected data. Additionally, the cloud screening algorithm has been evaluated under different aerosol conditions including Saharan dust intrusion, biomass burning and pollution events.

1 Introduction

The importance of atmospheric aerosol is remarked in atmospheric sciences because the atmospheric aerosol particles modify Earth-atmosphere radiative balance and also act as cloud condensation nuclei and thus modify cloud droplet size and cloud albedo. Passive remote sensing techniques allow the retrieval of aerosol properties, taking advantage of the modification of the radiative field due to the atmospheric aerosol. Until now the scientific community has mainly used the solar radiative field, measuring direct solar radiation, sky radiance and polarization via the well-know sun photometry. During the last decades, sun photometry has been developed considerably, and the international network AERONET (http://aeronet.gsfc.nasa.gov/) has been established to study daytime columnar aerosol properties worldwide. This last network has established and standardized the measurement protocol, data processing and quality control of data (Holben et al., 1998; Smirnov et al., 2000).

Currently, columnar aerosol properties at nighttime are poorly known. However, the study of these properties are taking an ongoing interest (Zhang et al., 2008), which include the development of appropriate instrumentation to fill the nighttime gaps in large temporal observations that rely on the sun (Tomasi et al., 2007; Stone et al., 2008; Eck et al., 2009). In particular, it is interesting to know the aerosol behavior in the pre-convection and pre-photochemistry effects, and thus study the dynamic of the aerosol particles during the whole day and their relationship with the boundary layer evolution. In addition, aerosol optical depth measurements at nighttime can be used as constraints for lidar systems, both at airborne or ground level. In this sense, different approaches have been developed using the moon (Esposito et al., 1998; Berkoff et al., 2011) or the stars (Leiterer et al., 1995; Ansmann et
As for sun photometry, one of the critical points in star photometry measurements for aerosol characterization is the separation between cloud-affected and cloud-free data. In principle, human observers can detect clouds based on subtle textural and spatial patterns. But the development of automatic instruments posed the problem of defining an effective and automatic cloud-screening procedure.

The most common algorithm for cloud screening used in sun photometry is the well-known algorithm developed by Smirnov et al. (2000), which is the standard algorithm in AERONET network. Basically, this algorithm applies a triplet stability criterion and assumes that the spectral aerosol optical depth varies less than 0.02 within one sequence of triplet measurements (acquired in less than 1 min) if the atmosphere is considered stable and cloud-free. The star photometer EXCALIBUR takes 3–5 min to acquire one sequence of measurements including all the filters, and thus the cloud screening algorithm developed by Smirnov et al. (2000) is not appropriate for this instrument. This fact made us look for an alternative automatic cloud screening technique for the purpose of retrieving aerosol properties and water vapor content at nighttime.

This paper presents a cloud screening procedure for star photometry that consists basically of calculating moving averages at different time-window intervals combined with a procedure for detecting outliers. Additionally, a quality control scheme is proposed. Section 2 describes the experimental site and the instrumentation used. Moreover, a brief description of the methodology used to retrieve aerosol optical depth and precipitable water vapor is given. Section 3 describes in detail the procedure developed for cloud screening and quality control of star-photometry data. In Sect. 4 the cloud screening procedure is tested versus lidar measurements and it is evaluated under Saharan dust intrusions, biomass burning influence and pollution cases. Additionally, all-sky images at sunset and sunrise are used to check the cloud conditions before and after star-photometer measurements. Concluding remarks are given in Sect. 5.

2 Experimental site and instrumentation

2.1 Experimental site

The data used in this work were acquired between March 2007 and September 2009 in the city of Granada (37.16° N, 3.60° W, 680 m a.s.l.) located in the South-East of Spain. Granada is a medium-sized city (around 250,000 inhabitants and twice that including its metropolitan area) situated in a natural basin surrounded by mountains. The near-contontinental conditions prevailing at this site are responsible for large seasonal temperature differences. Most rainfall occurs during spring and winter. Summers are normally very dry with few rainfall events. Due to its location in the Iberian Peninsula and depending on the prevailing synoptic situation, the study area can be affected by air masses originated in North Africa, the Atlantic, the Mediterranean and Europe (Lyamani et al., 2006, 2010). Moreover, the city is affected by local anthropogenic emissions, mainly from domestic heating and vehicles (Lyamani et al., 2008, 2010, 2011; Titos et al., 2012). Thus, this area is suitable for determining the cloud screening procedure due to the presence of different aerosol types and different meteorological conditions during the year.

2.2 Star photometer EXCALIBUR

The star photometer EXCALIBUR (iTec. Astronómica S. L., Spain) allows measurements of direct flux from a given star. Details about this instrument are given in Pérez-Ramírez et al. (2008a). Basically, the instrument consists of a telescope (CELESTRON CGE 1100) that collects the parallel incident light rays from the required star. Later, the starlight passes through a 10-filter wheel where six narrow band filters are set up with central wavelengths at 380, 436, 500, 670, 880 and 1020 nm (nominal wavelengths) for aerosol characterization, and an additional filter at 940 nm for retrieving precipitable water vapor. The FWHM ranges between 7.7 and 11.2 nm for the different filters. Once the starlight is spectrally separated, it reaches the CCD camera detector (model SBIG ST8-XME, Santa Barbara, USA), whose linear response and high quantum efficiency makes this device ideal for our photometry purposes. On the other hand, an external wide field CCD camera is employed to assure a correct pointing for a given star. The star photometer EXCALIBUR also has software that is able to minimize errors associated with a CCD camera, and also to focus the telescope and to calculate the background light associated with the measurements (Pérez-Ramírez et al., 2008b).

Due to the quantum efficiency of the CCD camera and to the different spectral types of the stars, the exposure time varies among the different filters and also for the different stars. Table 1 shows the average exposure time for each filter. In addition, the CCD camera takes approximately 4 s to process one particular measurement (technical specifications of SBIG ST8 CCD are in http://www.sbig.com/ST-8XME-C2.html). As a result, the star photometer EXCALIBUR takes 3–5 min approximately to make a sequence of measurements that includes all the filters.
Table 1. Mean, maximum and minimum exposure times for each filter of the star photometer EXCALIBUR.

<table>
<thead>
<tr>
<th>Nominal wavelength (nm)</th>
<th>Average exposure time (s)</th>
<th>Minimum exposure time (s)</th>
<th>Maximum exposure time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>380</td>
<td>23 ± 30</td>
<td>0.3</td>
<td>124</td>
</tr>
<tr>
<td>436</td>
<td>1.4 ± 1.0</td>
<td>0.3</td>
<td>4.9</td>
</tr>
<tr>
<td>500</td>
<td>0.7 ± 0.4</td>
<td>0.3</td>
<td>2.5</td>
</tr>
<tr>
<td>670</td>
<td>0.53 ± 0.3</td>
<td>0.3</td>
<td>1.7</td>
</tr>
<tr>
<td>880</td>
<td>2.5 ± 1.9</td>
<td>0.3</td>
<td>10.3</td>
</tr>
<tr>
<td>940</td>
<td>13 ± 11</td>
<td>0.3</td>
<td>62</td>
</tr>
<tr>
<td>1020</td>
<td>49 ± 40</td>
<td>0.3</td>
<td>140</td>
</tr>
</tbody>
</table>

2.3 Sun photometer CIMEL CE-318-4

The CIMEL CE-318-4 (CIMEL electronique, France) radiometer used in this work is within the frame of AERONET network (Holben et al., 1998). This instrument makes direct Sun irradiance measurements with a 1.2° full field of view at 340, 380, 440, 670, 870, 940, and 1020 nm (nominal wavelengths), with FWHM between 2 and 10 nm. A sequence of three such measurements is taken 30 s apart, creating a triplet observation per wavelength. The solar extinction measurements can be used to compute aerosol optical depth at each wavelength except for the 940 nm channel, which is used to retrieve precipitable water vapor. More details about this instrument are given by Holben et al. (1998).

2.4 All-sky imager and Raman lidar system

The all-sky imager was developed by the Atmospheric Physics Group at the University of Granada (GFAT) to provide images of the whole sky dome during the daytime for cloud and aerosol characterization (Cazorla et al., 2008a,b; Olmo et al., 2008). The all-sky imager registers full color images at three channels using a fish-eye lens: one is centered in red, another in green and the last one in blue. This instrument captures an image every 5 min. Further details about the all-sky imager can be consulted in Cazorla et al. (2008a,b). The Raman lidar system model LR321D400 (Raymeters s.a., Greece) is used to derive range corrected signals at 532 nm. More details about this instrument can be found in Guerrero-Rascado et al. (2008).

2.5 Aerosol properties and precipitable water vapor retrievals

Attenuation of light passing through Earth’s atmosphere follows the Beer-Bouguer-Lambert law, for an average star-Earth distance, is given by the following expression:

\[ V(\lambda) = V_0(\lambda) \exp(-m_t \delta_{atm}(\lambda)) \]  

where \( V(\lambda) \) is the measured signal by the photometer, \( V_0(\lambda) \) is the extraterrestrial signal (that is known as calibration constant), \( \delta_{atm}(\lambda) \) is the total atmospheric optical depth and \( m_t \) is the relative optical air mass which is the ratio between the optical air mass in the slant path and that associated with the vertical path. Its expression is given by Kasten and Young (1989):

\[ m_t = \left( \cos \theta + 0.15(93.885 - \theta)^{-1.253} \right)^{-1} \]  

where \( \theta \) is the zenith angle of the sun or star. The calibration constants of the star photometer EXCALIBUR was determined using the Astronomical Langley Method (Mitchell and Forgan, 2003) during special calibration campaigns at high mountain site once a year (Pérez-Ramírez et al., 2011).

The total atmospheric optical depth is calculated from Eq. (1), and thus aerosol optical depth, \( \delta_{Ae}(\lambda) \), is retrieved by subtracting Rayleigh scattering by molecules and ozone and nitrogen optical depths to \( \delta_{atm}(\lambda) \). The optical depths associated with those scattering and absorption processes are calculated as by Gueymard (2001). The ozone content is obtained from an OMI instrument on board of AURA satellite (http://aura.gsfc.nasa.gov/instruments/omi.html), being nighttime values computed by interpolating the values obtained the day after and before. The NO\(_2\) content used is 0.00017 cm-atm given by Kneizys et al. (1988) for mid-latitudes atmosphere. AERONET network uses monthly fixed values of NO\(_2\) content retrieved from meteorological databases, with an annual mean value of 0.00017 for our station, which is equal to the NO\(_2\) content value used in this work. The work of Pérez-Ramírez et al. (2011) explained in detail the retrieval procedure of \( \delta_{Ae}(\lambda) \) for the star photometer, and showed that changing NO\(_2\) content by 50% produced uncertainties in \( \delta_{Ae}(\lambda) \) below 0.001. Pérez-Ramírez et al. (2011) have also shown that the uncertainties in \( \delta_{Ae}(\lambda) \) retrieved by the star photometer (considering all sources of errors) were ±0.02 for \( \lambda < 800 \) nm and ±0.01 for \( \lambda > 800 \) nm. Thus, the uncertainty induced by using a fixed NO\(_2\) content is negligible. For the sun photometer, the \( \delta_{Ae}(\lambda) \) values of level 2.0 provided by AERONET are used.

Using the Angström turbidity formula \( \delta_{Ae}(\lambda) \propto \beta \lambda^{-\alpha} \), the least-squares fits of \( \delta_{Ae}(\lambda) \) (in a log-log scale) has been applied to determine the Angström parameters \( \alpha \) and \( \beta \). The Angström parameter \( \alpha \) characterizes the spectral features of aerosols and is related to the size of the particles, while the \( \beta \) parameter is related to particle concentration and represents the aerosol optical depth at 1 \( \mu \)m. In the solar spectrum, \( \alpha \) is a good indicator of the atmospheric particles size: cases with \( \alpha > 1.5 \) are mainly dominated by submicron particles, while those with \( \alpha < 1.5 \) are dominated by coarse particles (Dubovik et al., 2002; Gobbi et al., 2007). The Angström parameters determined from star photometer measurements were computed in the wavelength range 436–880 nm and those obtained from sun photometer measurements were calculated in 440–870 nm.

On the other hand, in the near infrared spectrum, water vapor presents a large number of strong wavelength
Fig. 1. Flow diagram of the cloud screening and quality control algorithm for star photometry. The parameter $\Delta t$ refers to the time-interval used for calculating the moving-averages. “$\Delta t = \text{all-night}$” is the time between the first and the last measurement in each night. The $\delta_{\text{Ae}}(\lambda)$ data point is not included in the computation of its corresponding moving-average $\delta_{\text{Ae}}^{\overline{M}}(\lambda)$.

dependent absorptions. In this sense, the response of the instrument $V(\lambda)$ at 940 nm to radiance attenuation through the atmosphere, at medium Earth–Sun/star distance, is given by Halthore et al. (1997):

$$V(940) = V_0(940) \exp(-m_r(\delta_{\text{Ae}}(940) + \delta_R(940))) T_w(940)$$  (3)

where $T_w(940)$ is the water vapor transmittance at 940 nm. For determining the calibration constant at 940 nm channel of the star photometer, we follow the methodology described in Pérez-Ramírez et al. (2012). Aerosol optical depth at 940 nm is computed by extrapolation from the values at 880 and 1020 nm, while Rayleigh scattering at 940 is computed as in Guemard (2001). Once the instruments are calibrated from the corresponding measurements at 940 nm, we calculate $T_w(940)$. Using this value of $T_w(940)$ and the corresponding relative optical air mass, the precipitable water vapor ($W$) is computed from a look-up table (Pérez-Ramírez et al., 2012).
Fig. 2. Frequency histograms of the absolute differences between two consecutive values of: (a) aerosol optical depth at each wavelength, $\Delta \delta_{\text{Ae}}(\lambda)$, and (b) precipitable water vapour, $\Delta W$.

3 Cloud screening and quality control of star photometer data

3.1 Cloud screening procedure

Figure 1 shows the flow diagram of the cloud screening algorithm for star photometer data proposed in this work. The first step of Fig. 1 is devoted to calculate $\delta_{\text{Ae}}(\lambda)$ and $W$, eliminating all $\delta_{\text{Ae}}(\lambda)$ and $W$ that are smaller than their absolute uncertainties. On the other hand, measurements acquired at $m_t > 3$ require large exposure time, in particular for filters with low quantum efficiency, and thus $m_t$ can experience significant changes during this time. To avoid this effect, the star photometer measurements are limited to $m_t < 3$.

The second step of the algorithm deals with different cloud screening procedures. The first one consists of analyzing the differences between two consecutive $\delta_{\text{Ae}}(\lambda)$ and $W$ data (obtained always in less than 5–6 min). When temporary clouds obstruct the field of view of the star photometer telescope, usually, sudden and large increases in $\delta_{\text{Ae}}(\lambda)$ and $W$ are observed. In fact, high differences between two consecutive values of $\delta_{\text{Ae}}(\lambda)$ and $W$ data are usually associated with passing clouds. Thus, in order to detect and eliminate data affected by these type of clouds, the absolute differences between two consecutive values of spectral aerosol optical depth, $\Delta \delta_{\text{Ae}}(\lambda)$, or precipitable water vapor, $\Delta W$, are analyzed.

Figure 2a and b shows the histograms of $\Delta \delta_{\text{Ae}}(\lambda)$ and $\Delta W$ for one and half years of measurements (March 2007–September 2009) obtained by the star photometer EXCALIBUR at Granada. From Fig. 2a, it can be observed that more than 90% of the differences $\delta_{\text{Ae}}(\lambda)$ at 380, 436, 500, 670, 880 and 1020 nm are below 0.03. These values are close to the absolute uncertainties in $\delta_{\text{Ae}}(\lambda)$; 0.02 for $\lambda < 800$ nm and 0.01 for $\lambda > 800$ nm (Pérez-Ramírez et al., 2011). Thus, the threshold value of 0.03 is established for every wavelength for eliminating $\delta_{\text{Ae}}(\lambda)$ data affected by clouds. Similarly, Fig. 2b shows that 90% of the $\Delta W$ data are below 0.2 cm, which is established as the threshold for $W$. Therefore, if any pair of consecutive data presents absolute difference between their values larger than the thresholds just mentioned, the largest value of the pair is eliminated.

The second procedure assumes that the sensitivity of the cloud variability in their optical depths during a short period of time is much larger than those expected for $\delta_{\text{Ae}}(\lambda)$. With this assumption, Harrison and Michalsky (1994) developed an algorithm to eliminate cloud-affected data for a multifilter rotating shadow-band radiometer based on analyzing their measurements in different time intervals. However, this algorithm was developed to make Langley calibrations. Using this concept, the cloud screening procedure for star photometer data consists basically of calculating the moving average, $\delta_{\text{Ae}}^{i \cdot \text{M}}(\lambda)$, for every aerosol optical depth, $\delta_{\text{Ae}}(\lambda)$, retrieved from star irradiance measurements:

$$\delta_{\text{Ae}}^{i \cdot \text{M}}(\lambda) = \frac{1}{n-1} \sum_{j=1}^{n} \delta_{\text{Ae}}^{j}(\lambda) \quad (4)$$

where “$n$” is the number of data included in a temporal interval, $\Delta t$, which can vary from some minutes up to an all night measurement period. In our case, we computed the moving averages using 1 h and all-night period temporal intervals. It is important to note that the value $\delta_{\text{Ae}}^{i}(\lambda)$ is not included in the computation of its corresponding moving average. Moreover, the procedure also generates the standard deviation $\sigma^{i \cdot \text{M}}$ of each moving average as

$$\sigma^{i \cdot \text{M}}(\lambda) = \sqrt{\frac{1}{n-2} \sum_{j=1}^{n} \left( \delta_{\text{Ae}}^{j}(\lambda) - \delta_{\text{Ae}}^{i \cdot \text{M}}(\lambda) \right)^2} \quad (5)$$
where we have replaced \( n-1 \) of the standard deviation definition by \( n-2 \) because the point “\( i \)” is not included in the moving-average computation. Finally, we compute the differences \( \Delta \delta_{\text{Ae}}^i(\lambda) \) between each \( \delta_{\text{Ae}}^i(\lambda) \) and its corresponding \( \delta_{\text{Ae}}^M(\lambda) \):

\[
\Delta \delta_{\text{Ae}}^i(\lambda) = \delta_{\text{Ae}}^i(\lambda) - \delta_{\text{Ae}}^M(\lambda).
\]

Once all computations are done, the procedure flags each \( \delta_{\text{Ae}}^i(\lambda) \) that presents \( \Delta \delta_{\text{Ae}}^i(\lambda) \) larger than three times its \( \sigma_{\text{M}} \). From these outliers of \( \delta_{\text{Ae}}^i(\lambda) \) data, the \( \delta_{\text{Ae}}^i(\lambda) \) with the largest positive deviation is eliminated from the database. The previous procedure is repeated until no outlier is detected. Finally, we obtain a final data series where each \( \delta_{\text{Ae}}^i(\lambda) \) does not present strong variability versus its surrounding values.

### 3.1.1 All-night window averages

To detect \( \delta_{\text{Ae}}^i(\lambda) \) data affected by persistent clouds on time, the cloud screening procedure just described above is firstly applied to the data set using an all-night window. An all-night window indicates that, for each night, the time \( \Delta t \) used in Eq. (4) is the time between the first and the last measurement in this night. It is worth noting that the \( \delta_{\text{Ae}}^i(\lambda) \) data point is not included in the computation of its corresponding night average \( \delta_{\text{Ae}}^M(\lambda) \).

As an example, Fig. 3a shows the evolution of \( \delta_{\text{Ae}}(436 \text{ nm}) \) and the Angström parameter \( \alpha(436–880 \text{ nm}) \) obtained by the star photometer during the night 18–19 March 2009 after applying a time interval equal to an all-night window. Full circles and triangles correspond to the data obtained as cloud-free, while empty circles and triangles are the determined as cloud-affected data. As can be seen before 19 March 2009 at 02:00 UTC, the \( \delta_{\text{Ae}}(436 \text{ nm}) \) was almost constant with values around 0.1 (Fig. 3a). However, after 02:00 UTC there was a large increase in \( \delta_{\text{Ae}}(436 \text{ nm}) \), reaching values up to 1.0 in less than 30 min and maximum value of 1.2 thereafter. There was also a great change in \( \alpha(436–880 \text{ nm}) \) on 19 March 2009 at 02:00 UTC (Fig. 3a). Before 02:00 UTC, \( \alpha(436–880 \text{ nm}) \) values were above 1.0 with mean value around 1.5, suggesting a predominance of fine particles. After 02:00 UTC, the \( \alpha(436–880 \text{ nm}) \) values were quite different with a mean value around 0.3, suggesting a predominance of large particles.

In our study area, for cloud-free conditions, \( \delta_{\text{Ae}}(440 \text{ nm}) \) larger than 0.7 are usually associated with extreme Saharan dust intrusions (Lyamani et al., 2006; Guerrero-Rasco et al., 2009). Additionally, large \( \delta_{\text{Ae}}(440 \text{ nm}) \) can be also associated with strong pollution (Lyamani et al., 2006) or biomass burning events (Alados-Arboledas et al., 2011). In other conditions in our study area, \( \delta_{\text{Ae}}(440 \text{ nm}) \) is usually lower than 0.5 (Alados-Arboledas et al., 2009). Moreover, over the study region Saharan dust intrusions usually present low values of the Angström parameter (<0.5) (Lyamani et al., 2006; Guerrero-Rasco et al., 2009; Valenzuela et al., 2012), while during pollution and biomass burning cases, this parameter presents large values (>1.5) (Pérez-Ramírez et al., 2008a; Alados-Arboledas et al., 2011). On the other hand, Smirnov et al. (2000) reported that clouds present high optical depths and low values of the Angström parameter. Thus, the data showed in Fig. 3a after 02:00 UTC have to be either cloud affected-data or associated with long-range transport of dust particles from North Africa.

Strong Saharan dust outbreaks affecting our study area present a large horizontal spreading affecting a relevant part of the Iberian Peninsula (Pérez et al., 2006; Toledano et al., 2007; Cachorro et al., 2008; Guerrero-Rasco et al., 2009; Córdoba-Jaboner et al., 2011). So, for checking if the data presented in Fig. 3a after 02:00 UTC are possibly associated with the presence of dust particles over the Iberian Peninsula, we investigated the time evolutions of \( \delta_{\text{Ae}}(440 \text{ nm}) \) and the Angström parameter \( \alpha(440–870 \text{ nm}) \) obtained by sun photometer CIMEL at Granada and Malaga AERONET stations (36.5°N, 4.7°W, 40 m a.s.l.), from 18 to 21 March 2009.
Table 2. Mean coefficient of variation (COV) of $\delta_{Ae}(\lambda)$ and $W$ at different averaging time intervals for the whole data series obtained by star photometer EXCALIBUR.

<table>
<thead>
<tr>
<th>Averaging time interval</th>
<th>$\delta_{Ae}(380)$</th>
<th>$\delta_{Ae}(436)$</th>
<th>$\delta_{Ae}(500)$</th>
<th>$\delta_{Ae}(670)$</th>
<th>$\delta_{Ae}(880)$</th>
<th>$W$</th>
<th>$\delta_{Ae}(1020)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instantaneous data</td>
<td>0.16</td>
<td>0.18</td>
<td>0.23</td>
<td>0.25</td>
<td>0.23</td>
<td>0.10</td>
<td>0.25</td>
</tr>
<tr>
<td>5 min</td>
<td>0.15</td>
<td>0.17</td>
<td>0.21</td>
<td>0.24</td>
<td>0.22</td>
<td>0.09</td>
<td>0.25</td>
</tr>
<tr>
<td>15 min</td>
<td>0.15</td>
<td>0.16</td>
<td>0.20</td>
<td>0.22</td>
<td>0.21</td>
<td>0.08</td>
<td>0.23</td>
</tr>
<tr>
<td>30 min</td>
<td>0.14</td>
<td>0.15</td>
<td>0.18</td>
<td>0.20</td>
<td>0.19</td>
<td>0.07</td>
<td>0.22</td>
</tr>
</tbody>
</table>

The values for $\delta_{Ae}(440 \text{ nm})$ were lower than 0.3 on 18 and 19 March at both stations, while on 20 March slightly larger values were obtained (up to 0.4). Thus, no very large values of $\delta_{Ae}(436 \text{ nm})$ (>1.2) were expected for the night 18–19 March. On the other hand, $\alpha(440–870 \text{ nm})$ presented values above 1.0 at both stations during this study period. Therefore, the presence of dust particles over the South of the Iberian Peninsula is not expected during this period. Additionally, MODIS satellite images (http://modis.gsfc.nasa.gov/) on 18 and 19 March do not reveal the presence of dust particles over our study area on these days (images not shown). Moreover, five-days backward-trajectories computations of air mass ending at Granada at the altitudes of 500, 1500, 2500, 3500, 4500 and 5000 m a.g.l. by HYSPLIT model (Draxler and Rolph, 2003) indicated that the air masses affecting our study area on the night 18–19 March 2009 came from the Mediterranean Sea (graphs not shown). In addition, MODIS satellite images for 18 March 2009 reveals cumulus formation in the Mediterranean Sea just beside the South-East coast of the Iberian Peninsula, which moves to Gibraltar on the following day (graphs not shown).

Therefore, according to the discussions in the above paragraphs, $\delta_{Ae}(436 \text{ nm})$ larger than 0.5 are not expected on the night 18–19 March 2009, and $\delta_{Ae}(436 \text{ nm})$ values larger than 0.5 are very probably cloud-affected data. These data were successfully detected as cloud-affected data after applying the cloud screening procedure with $\Delta t$ equal to all-night, as can be seen in Fig. 3a.

### 3.1.2 One-hour window moving-average procedure

Later, to the remaining data of the algorithm, the cloud screening procedure is applied again, but using a one-hour window moving-average. For example, as can be seen in Fig. 3b, after applying the cloud screening with a $\Delta t$ equal to all-night, there are data (outliers) that still deviate from the rest, which might be associated with cloud-affected data. But these deviations are not as strong as in the case illustrated in Fig. 3a. As can be observed, the moving-average procedure with one-hour window eliminates these outliers data (Fig. 3b). Therefore, the moving-average procedure with one-hour window can be useful to detect cloud-affected data that have not been detected in the previous steps.

### 3.1.3 Cloud screening of precipitable water vapor data

For the retrieval of $W$ cloud-free data, the previous moving-average procedures are not applied. The retrieval of $W$ implies the determination of $\delta_{Ae}(940 \text{ nm})$ which is calculated by linear interpolation between $\delta_{Ae}(880 \text{ nm})$ and $\delta_{Ae}(1020 \text{ nm})$ (Pérez-Ramírez et al., 2012). In this sense, if either $\delta_{Ae}(880 \text{ nm})$ or $\delta_{Ae}(1020 \text{ nm})$ is detected as cloud-affected data, then the corresponding $W$ measurement is also considered as cloud-affected data.

### 3.2 Data quality control

Nights with few measurements are not reliable for a large database. Additionally, few measurements make it more difficult to detect and eliminate cloud-affected data. Thus, only the nights that have at least 2 hours of measurements are considered. In addition, if after applying moving-average procedures to data obtained in a particular night, the number of data eliminated is more than one third of the original data, then all data obtained during this night are eliminated from the database.

The filters of the star photometer EXCALIBUR that have not been detected in the previous steps.

The filters show the largest decrease as the average-time interval increases. These filters have the shortest exposure times (less than 1 s) and thus they are more affected by atmospheric turbulence effects. Moreover, COV of $\delta_{Ae}(436 \text{ nm})$ and $\delta_{Ae}(880 \text{ nm})$ present similar decrease with increasing...
average-time-interval because the filters have similar exposure times. In spite of the large exposure times of the filter at 1020 nm, $\delta A_e(1020 \text{ nm})$ presents large COV due to its low signal-to-noise ratio. Finally, $\delta A_e(380 \text{ nm})$ and $W$ present the lowest COV because the filters at 380 and 940 nm have larger exposure times (> 10 s) and larger signal-to-noise ratio than the other ones.

Table 2 reveals that the lower COV of $\delta A_e(\lambda)$ and $W$ are found for 15 and 30 min averaging-time intervals. However, slight fluctuations in $\delta A_e(\lambda)$ can induce large variations in the Angström parameter $\alpha$. In this sense, it is important to note that if there are no rapid changes in $\delta A_e(\lambda)$, high fluctuations in $\alpha(436–880 \text{ nm})$ are not expected. Thus, the time evolution of $\alpha(436–880 \text{ nm})$ is used to select the most appropriate averaging-time interval. As an example, Fig. 4 shows the day- and nighttime evolution of the Angström parameter $\alpha$ from 22 to 25 June 2007 (Fig. 4a) and from 30 July to 2 August 2007 (Fig. 4b). It is important to note that $\alpha(436–880 \text{ nm})$ obtained at night time have been calculated from the corresponding average values of cloud-free $\delta A_e(\lambda)$ at 15 or 30 min. From Fig. 4 can be seen that averaging the data either at 15 or 30 min provides good COV and reasonable time evolution of $\alpha(436–880 \text{ nm})$, with an apparent coherency and continuity with daytime values. However, much better smoothness of nighttime $\alpha(436–880 \text{ nm})$ is obtained by averaging the data in 30 min. Therefore, to minimize the fluctuations associated with the atmospheric turbulence effects, a 30 min average-time interval is assumed.

Usually, low and highly variable values of Angström parameter $\alpha$ are associated with clouds (Smirnov et al., 2000). However, in our study area strong Saharan dust outbreaks also present low Angström parameter (Lyamani et al., 2006; Guerrero-Rascado et al., 2009). Moreover, several aerosol types are present (Alados-Arboledas et al., 2003), with large variability of $\alpha$ during the year. Thus, we have not imposed any limitation on the Angström parameter.

**Fig. 4.** Temporal evolution of day- and nighttime Angström parameter $\alpha$. (a) From 22 to 25 June 2007. (b) From 30 July to 2 August 2007. The Angström parameter $\alpha (436–880 \text{ nm})$ at nighttime was obtained from the average values of cloud-free aerosol optical depths at 15 or 30 min while Angström parameter $\alpha (440–870 \text{ nm})$ values obtained at day time are instantaneous cloud-free data after applying Smirnov et al. (2000) algorithm. The averaging times and the coefficient of variation, COV, are also included in the figure.
Fig. 5. (a) Temporal evolution of the aerosol optical depth at 380, 670, 880 and 1020 nm during the night 27–28 August 2007. Empty symbols correspond to cloud-affected data and full symbols are cloud-free data. (b) Range corrected signal measured by lidar system at 532 nm between 27 August 2007 at 20:00 UTC and 28 August 2007 at 02:15 UTC. (c) All-sky images during the early morning of 28 August 2007 (times expressed in UTC).

4 Evaluation of the cloud screening procedure

4.1 Assessment of the cloud screening algorithm with lidar measurements

The cloud screening algorithm proposed in this work is now checked versus simultaneous data obtained by lidar system at our station. Until now the channel at 436 nm has been used as representative of the visible range. However, it is important to remark that the proposed cloud screening algorithm also provides similar results for the rest of $\delta A_{\epsilon}(\lambda)$, and thus this section makes use of $\delta A_{\epsilon}(\lambda)$ at 380, 670, 880 and 1020 nm.

Figure 5 shows the data obtained after applying the cloud screening algorithm to the star photometer data acquired on 27–28 August 2007 night (Fig. 5a). Empty symbols correspond to cloud-affected data, while full symbols are cloud-free data. Range corrected signal at 532 nm obtained between 27 August 2007 at 20:00 UTC and 28 August 2007 at 02:00 UTC is also shown (Fig. 5b).

From Fig. 5a can be observed that in this night, the aerosol load was relatively high (mean $\delta A_{\epsilon}(436 \text{ nm})$ of 0.49 ± 0.03; after applying the cloud screening algorithm). The range corrected signal from lidar system revealed a strong backscatter signal at 5.5 km between 22:15 and 23:30 UTC, which has been associated with clouds. The $\delta A_{\epsilon}(\lambda)$ values obtained during this period were successfully identified as cloud-affected data by the cloud screening algorithm proposed in this work. On the other hand, the all-sky images for the early morning on 28 August 2007 (Fig. 5c) reveal partly cloudy skies. This could explain the data eliminated late at night. Unfortunately, due to technical problems the all-sky imager did not acquire images on evening of 27 September.

As another example, Fig. 6 shows the $\delta A_{\epsilon}(\lambda)$ for the night 7–8 July 2008 (Fig. 6a) and the range corrected signal obtained by the lidar system between 20:00 and 22:00 UTC of 7 July 2008 (Fig. 6b). During this night the aerosol load was low (mean $\delta A_{\epsilon}(436 \text{ nm})$ of 0.10 ± 0.02; after applying the cloud screening algorithm). In addition, a strong backscatter range-corrected signal, at the height of 10 km between 20:05 and 20:40 UTC was observed, which has been associated with cirrus clouds. During this period, the $\delta A_{\epsilon}(\lambda)$ measured by the star photometer presented outliers data, which were successfully detected as cloud-affected data by the cloud screening algorithm (Fig. 6a).

The all-sky images (Fig. 6c) reveal the presence of cirrus clouds between 18:00 and 18:15 UTC on 7 August, which could also explain the data eliminated by the cloud screening algorithm during the early night. In the early morning of 8 August the all-sky images shows cloudless skies and therefore no clouds are expected at the end of this night. This agrees with the cloud-free data obtained late at night.
4.2 Cloud screening algorithm evaluation under different aerosol types

This section is devoted to check the proposed cloud screening algorithm under different aerosol types. Five-day air-masses backward-trajectories calculated by HYSPLIT model and MODIS sensor images were used to characterize each event, although graphs are not shown.

Figure 7 shows the temporal evolutions of $\delta_{Ae}(\lambda)$ and $\alpha$ from 18 to 19 July 2008 and all-sky images on the late evening of 18 and the early morning of 19 July. Backward-trajectories analysis revealed that the air masses affecting our study area during this period came from North Africa. Satellite observations by MODIS showed the presence of Saharan dust particles over the study area on 18 and 19 July 2008. During daytime, mean values of $\delta_{Ae}(440\text{ nm})$ were $0.26 \pm 0.08$ and $0.43 \pm 0.02$ on 18 and 19 July 2008, while $\alpha(440–870\text{ nm})$ mean values were $0.35 \pm 0.14$ and $0.10 \pm 0.02$ respectively (Fig. 7b). During the night 18–19 July 2008, original $\delta_{Ae}(436\text{ nm})$ values ranged from 0.38 to 1.07 (Fig. 7a). Values of $\delta_{Ae}(436\text{ nm})$ larger than 0.8 were detected as cloud-affected data by the proposed cloud screening algorithm (Fig. 8a). It is worth noting that although Saharan dust outbreaks at the Iberian Peninsula can present large spatial and temporal variability (Escudero et al., 2005; Pérez et al., 2006; Lyamani et al., 2006; Cachorro et al., 2008; Rodríguez et al., 2011; Córdoba-Jabonero et al., 2011), the algorithm eliminates data taking into account the non-stability of the data series in a very short period less than 1 h (Fig. 7a) as compared to the spatial and temporal variability registered in the bibliography. Thus, for this night cloud-free data mean values of $\delta_{Ae}(436\text{ nm})$ and $\alpha(436–880\text{ nm})$ were $0.48 \pm 0.06$ and $0.23 \pm 0.08$, showing good continuity and agreement with daytime data (Fig. 7b). The values of $\delta_{Ae}(\lambda)$ and $\alpha$ obtained during this period are typical values for desert dust particles at our study area (Lyamani et al., 2006; Toledano et al., 2007; Guerrero-Rascado et al., 2009).

The all-sky images (Fig. 7c) reveal no presence of clouds late at evening on 18 July, which is in agreement with the lack of cloud-affected data at the beginning of the night. However, during early morning on 19 July cloudy skies were observed, which could explains the presence of cloud-affected data above 0.85 registered between 02:15 and 03:25 UTC on 18 July and the lack of measurements from 03:00 UTC (Fig. 8b).

Figure 8 shows $\delta_{Ae}(\lambda)$ and $\alpha$ obtained from 11 to 12 July 2007, when our study area was under the influence of air-masses coming from the Mediterranean basin as revealed by backward-trajectories analyses. These air-masses can transport anthropogenic particles to our study area (Lyamani et al., 2006; Cachorro et al., 2008; Rodríguez et al., 2011; Córdoba-Jabonero et al., 2011).
Fig. 7. (a) Time evolution of aerosol optical depth during a Saharan dust outbreak over the city of Granada during the night 18–19 July 2008. (b) Day-and-night time evolutions of aerosol optical depth and Angström parameter $\alpha$ from 18 to 19 July 2008, after applying the cloud screening algorithm presented in this work to the nighttime data. (c) All-sky images during late 18 and early 19 July 2008 (times expressed in UTC).

Fig. 8. (a) Nighttime evolution of aerosol optical depth during Mediterranean air masses influence over the city of Granada on the night 11–12 July 2007. (b) Day-and-night time evolution of aerosol optical depth and Angström parameter $\alpha$ from 11 to 12 July 2007, after applying the cloud screening algorithm presented in this work to the nighttime data. (c) All-sky images during the early morning of 12 July 2007 (times expressed in UTC).
During daytime, mean values of $\delta_{\text{Ae}}(440\text{ nm})$ were $0.28 \pm 0.04$ and $0.38 \pm 0.06$ on 11 and 12 July 2007, while $\alpha(440–870\text{ nm})$ mean values were $1.09 \pm 0.08$ and $1.22 \pm 0.06$, respectively (Fig. 8b). These values are similar to those obtained under Mediterranean air-masses influence in the same study area by Lyamani et al. (2006). During the nighttime, original $\delta_{\text{Ae}}(436\text{ nm})$ values ranged between 0.25 and 0.96, with many data detected as outliers (Fig. 8a). After applying the cloud screening algorithm, mean values of $\delta_{\text{Ae}}(436\text{ nm})$ and $\alpha(436–880\text{ nm})$ were $0.33 \pm 0.02$ and $0.85 \pm 0.12$, respectively, showing as in the previous case a good continuity and agreement with daytime data (Fig. 8b).

The all-sky images (Fig. 8c) reveal partly cloudy skies on 12 September 2007 early morning, which could explain the lack of data from 02:00 to 09:00 UTC on 12 September, and also some outliers reported during the rest of the night (Fig. 8b).

Figure 9 shows the time evolution of $\delta_{\text{Ae}}(\lambda)$ and $\alpha$ from 30 to 31 July 2007. In the previous days, some forest fires were detected in the South-West of the Iberian Peninsula (http://firefly.geog.umd.edu/firms/). Backward-trajectories analysis showed that the air-masses reaching our area on the night 30–31 July 2007 have passed over this forest fires area. During nighttime, on 30 and 31 July 2007 mean values of $\delta_{\text{Ae}}(440\text{ nm})$ were $0.29 \pm 0.03$ and $0.30 \pm 0.06$, while mean values of $\alpha(440–870\text{ nm})$ were $1.24 \pm 0.08$, $1.22 \pm 0.06$ on 11 and 12 July 2007, respectively. During the night 30–31 July 2007, original $\delta_{\text{Ae}}(436\text{ nm})$ data ranged between 0.25 and 0.85 (Fig. 9a). However, the maximum value is reduced up to 0.5 after applying the cloud screening algorithm (Fig. 9b), being the mean value of $0.32 \pm 0.08$. The $\alpha(436–880\text{ nm})$ for cloud-free data ranged between 1.95 and 1.0 with mean value of $1.7 \pm 0.4$ (Fig. 9b). The time evolution of aerosol properties presented in Fig. 10b remarks the ability of the star photometer to detect aerosol plumes at nighttime, which could be ignored if a linear interpolation is made between the end of the previous day and the beginning of the following day. Moreover, biomass burning aerosol present $\alpha(440–880\text{ nm})$ similar to those values just reported, being the aerosol load proportional to the strength of the fires and to the distance to the source (O’Neill et al., 2002; Reid et al., 2005; Schafer et al., 2008; Eck et al., 2009; Alados-Arboledas et al., 2011).

5 Conclusions

This work has dealt with the development and set up of a cloud screening and data quality control algorithm for a star photometer based on a CCD camera as detector and used for obtaining columnar aerosol properties at nighttime. The main challenges for developing these algorithms are that the instruments take around taking 3–5 min to acquire a sequence of measurements for the seven narrowband filters installed (central wavelengths at 380, 436, 500, 670, 880, 940 and 1020 nm) and the variation of the exposure time of each filter inducing fluctuations on the signals. The algorithm developed includes a routine that study the variations between consecutive measurements. The analysis performed has shown that establishing thresholds values that limit maximum variations, cloud-affected data caused by temporary clouds that obstruct the field of view of the telescope during a few minutes are detected successfully. After that, the next step has been to study moving averages under different time-intervals. In this sense, the algorithm generates the series for the aerosol optical depth and another for their corresponding standard deviation. Later, it checks each retrieved aerosol optical depth and its corresponding moving average and those whose difference is larger than three times the standard deviation are considered as outliers. After evaluating each aerosol optical depth, the maximum outliers is eliminated and the previous processes are repeated until no outlier is detected.

The ability of the previous procedures to detect cloud-affected data has been successfully studied. For cases with strong variations in aerosol optical depth during the night...
(more than one order of magnitude), the outliers’ values were successfully eliminated. The evaluation of these outliers as cloud-affected data has been performed by analyzing the evolution of the Angström exponent and ancillary information about the synoptic conditions and aerosol model predictions. However, special attention has been paid to correlational measurements of star-photometer and lidar, being cloud-affected data detected by both instruments. These results have given more reliability to the cloud-screening algorithm proposed.

To minimize the effects of the fluctuations of the signals due to atmospheric turbulence effects, the cloud-free data were averaged over 30 minutes. This interval has been selected as the most appropriate after analyzing the fluctuations on aerosol optical depth at the different filters and mainly on the Angström exponent. Therefore, a standard procedure for high quality data of star photometry has been obtained. Particularly, good continuity and agreement in the day- and nighttime evolution of aerosol optical depth and Angström exponent have been obtained for a Saharan dust outbreaks, Mediterranean air-masses and biomass burning events. The quality of the procedures proposed for obtaining cloud-free data is also indicated by all-sky images at late evening and early morning.

To finalize these discussions, we would like to make few final comments. The work deals with the quality control and cloud screening tool for a specific place, database and a star photometer based on a CCD camera that has a concrete measurement sequence. Therefore, application of the method proposed to other instruments must be done carefully.

Acknowledgements. This work was supported by the Spanish Ministry of Science and Technology through projects CGL2008-01330-E/CLI (Spanish Lidar Network), CGL2010-18782 and CSD2007-00067; by the Andalusian Regional Government through projects P10-RNM-6299 and P08-RNM-3568; by the EU ACTRIS project (EU INFRA-2010-1.1.16-262254); and by the Postdoctoral Program of the University of Granada. The authors express gratitude to the NOAA Air Resources Laboratory (ARL), Naval Research Laboratory for the HYSPLIT transport and dispersion model. We gratefully acknowledge the MODIS mission scientists and associated NASA personnel for the production of the data used in this publication. We also would like to thanks AERONET team and the corresponding PI of the stations used in this work for the deployment and maintenance of the instruments and for providing the data.

Edited by: V. Amiridis

References


