Article: **Cloud cover detection combining high dynamic range sky images and ceilometer measurements**

Authors: **R. Román, A. Cazorla, C. Toledano, F. J. Olmo, V. E. Cachorro, A. de Frutos and L. Alados-Arboledas**

Journal: **Atmospheric Research**

Volume: **196**

Pages: **224-236**

Year: **2017**

DOI: **10.1016/j.atmosres.2017.06.006**

Cloud cover detection combining high dynamic range sky images and ceilometer measurements

R. Román[1,2,3], A. Cazorla[1,2], C. Toledano[3], F. J. Olmo[1,2], V. E. Cachorro[3], A. de Frutos[3] and L. Alados-Arboledas[1,2]

[1]{Department of Applied Physics, University of Granada. 18071, Granada (Spain)}

[2]{Andalusian Institute for Earth System Research (IISTA-CEAMA), University of Granada, Autonomous Government of Andalusia. 18006, Granada (Spain)

[3]{Grupo de Óptica Atmosférica (GOA), Universidad de Valladolid. Paseo Belén, 7, 47011, Valladolid (Spain)}

Correspondence to: R. Román (robertor@ugr.es)

Abstract

This paper presents a new algorithm for cloud detection based on high dynamic range images from a sky camera and ceilometer measurements. The algorithm is also able to detect the obstruction of the sun. This algorithm, called CPC (Camera Plus Ceilometer), is based on the assumption that under cloud-free conditions the sky field must show symmetry. The symmetry criteria are applied depending on ceilometer measurements of the cloud base height. CPC algorithm is applied in two Spanish locations (Granada and Valladolid). The performance of CPC retrieving the sun conditions (obstructed or unobstructed) is analyzed in detail using as reference pyranometer measurements at Granada. CPC retrievals are in agreement with those derived from the reference pyranometer in 85% of the cases (it seems that this agreement does not depend on aerosol size or optical depth). The agreement percentage goes down to only 48% when another algorithm, based on Red-Blue Ratio (RBR), is applied to the sky camera images. The retrieved cloud cover at Granada and Valladolid is compared with that registered by trained meteorological observers. CPC cloud cover is in agreement with the reference showing a slight overestimation and a mean absolute error around 1 okta. A major advantage of the CPC algorithm with respect to the RBR method is that the determined cloud cover is independent of aerosol properties. The RBR algorithm overestimates cloud cover for coarse aerosols and high loads. Cloud cover obtained only from ceilometer shows similar results than CPC algorithm; but the horizontal distribution cannot be obtained. In addition, it has been observed that under quick and strong changes on cloud cover ceilometers retrieve a cloud cover fitting worse with the real cloud cover.

Keywords

Sky camera; ceilometer; cloud cover; HDR; clouds; aerosols

Acronyms

AE: Angström Exponent.

AEMet: State Meteorological Agency of Spain (Agencia Estatal de Meteorología).

AERONET: Aerosol Robotic NETwork.

AOD: Aerosol Optical Depth.

CBH: Cloud Base Height.

CC: Cloud Cover.

CEI: CEIlometer.

CPC: Camera Plus Ceilometer.

FOV: Field Of View.

HDR: High Dynamic Range.

MABE: Mean Absolute Bias Error.

MBE: Mean Bias Error.

MDBE: Median Bias Error.

RBR: Red-Blue Ratio.

RGB: Red Green Blue.

SZA: Solar Zenith Angle.

SD: Standard Deviation.

SONA: Automatic Cloud Observation System.

WMO: World Meteorological Organization.

1 Introduction

Clouds play a critical role in the Earth’s radiative budget, since they backscatter to space a portion of incoming solar radiation but also reemit back to the surface a fraction of Earth infrared radiation. Hence, changes on cloud properties like lifetime (and subsequently cloud cover: CC), or albedo could dramatically impact on Earth’s climate (Boucher et al., 2013). From the energy production point of view, solar energy systems are strongly affected by cloud presence. Especially in the case of concentration solar power plants and concentration photovoltaic systems, that strongly depend on direct beam solar irradiance, their energy output is highly reduced when the sun is obstructed by clouds (Beyer et al., 1994; Frederick and Steele, 1995; Bartlett et al., 1998; Antón et al., 2011; Cazorla et al., 2015). Both climate and solar energy issues motivate the need for cloud observations.

Cloud cover can be determined by different ways. Visual observations of CC, made by a human observer, are hemispheric “instantaneous” observations that depend on the visible horizon and are subjective observations, prone to human effects due to differences between observers (WMO, 2012). Some authors such as Sánchez-Lorenzo et al. (2009) used this kind of measurements to study long-term CC data series due to the availability of these data in the past. These measurements cannot be automatically done and the time resolution is limited.

It is feasible replacing some of these observations by automated and continuous measurements from a ceilometer, which is an active instrument that emits pulsed laser signals and records with a receiver telescope the vertical signal based on the backscatter of the atmosphere (Tapakis and Charalambides, 2013). Cloud base height (CBH) and CC can be estimated from these measurements due to the strong backscatter of clouds (Martucci et al., 2010; Mittermaier, 2012; Costa-Surós et al., 2013, 2014), however this methods are only based on the vertical information, ignoring the spatial dimensions and excluding some clouds which are not in the vertical line of the ceilometer.

Satellite images can be used to retrieve cloud cover (e.g., Arking and Childs, 1984; Rossow and Schiffer, 1999; Gao et al., 2002; Zhao and Di Girolamo, 2006), but on a global scale, which is not useful to determinate if sun is obstructed by clouds in a particular place. Radiometers, radars, and radiosondes are also used in the retrieval of cloud properties (for a review see: Tapakis and Charalambides, 2013).

Sky cameras are devices that usually provide a hemispherical image of the full sky in the visible range, typically at red-green-blue (RGB) channels. Ghonima et al. (2012) simulated sky images under cloud-free conditions and detected cloudy pixels comparing the measured image with the cloud-free simulated one; to this purpose an aerosol correction is included in the simulations because aerosols can change significantly the sky radiance distribution. Yabuki et al. (2014) presented an algorithm based on the spectral contrast between the RGB channels of the sky image and various constrains. Cazorla et al. (2008) and Linfoot and Allis (2008) applied neural networks to sky images for detection of cloudy pixels after a previous training. Liu et al. (2015) used the superpixel segmentation technique to locate cloudy pixels in sky images. Some authors combined sky imagery with radiometric data (shortwave or longwave) in order to obtain cloud cover and classification (Martínez-Chico et al., 2011; Alonso et al., 2014; Wacker et al., 2015).

However, most of the sky camera algorithms for detection of cloudy pixels (e.g., Koehler et al., 1991; Long et al., 2006; Calbó and Sabburg, 2008; Kreuter et al., 2009) are based on the whiteness of the pixels quantified by the RBR value: ratio of the red to the blue channel; a threshold value is chosen and the pixels with a ratio below threshold are considered cloud-free (pixel too blue) and for ratio above the threshold the pixel is cloudy (high red channel). This method presents some problems since the commercial cameras usually do not have linear pixel sensitivity (although it can be mitigated by gamma correction), then RBR is not linear, and if their white balance is not fixed it could vary the RBR value. In addition, under large loads of coarse aerosol the cloud-free pixels look whiter than under low aerosol loads; pixels near the circumsolar are usually saturated looking also whiter (Long et al., 2006; Olmo et al., 2008; Heinle et al., 2010; Román et al., 2012); in this way the cloud-free pixels are erroneously classified as cloudy. Saturation and non-linearity problem can be solved taking high dynamic range (HDR) images, which are a composition of various images taken with different exposure times (Debevec and Malik, 1997; Stumpfel et al., 2004; Cazorla et al., 2015; Román et al., 2017).

The main aim of this paper is to develop an algorithm to retrieve the cloud cover and sun condition but removing, or at least reducing, the problems obtained with other algorithms. To this purpose, a sky camera is used and configured to take HDR images (avoiding saturation). HDR images are synergistically combined with information from a ceilometer to improve the cloud detection algorithm. The basis of our cloud detection is partially based on the consideration that a cloud-free sky image presents high symmetry relative to the solar principal plane. Other concepts like variation between two consecutive images or edge detection are considered. The proposed algorithm retrievals are compared with other algorithms and tested against suitable references like those based in trained meteorological observers.

Section 2 presents the locations and data from different instruments used in this work. Relevant information about the configuration of the sky camera to take HDR images can be found in Section 3. The new algorithm developed in this work is explained in detail in Section 4 and Appendix A and B. Section 5 shows the main results of the comparison of the new algorithm and others against reference values and, finally, Section 6 summarizes the main conclusions.

2 Location and Instrumentation

Data used in this paper was recorded at stations sited at Valladolid (41.66ºN; -4.71ºW; 705 m a.s.l.) and Granada (37.16ºN; -3.61ºW; 680 m a.s.l.), both cities located in Spain. The predominant aerosol at Valladolid, sited in North-Central Iberian Peninsula, can be considered as “clean continental” (Bennouna et al., 2013, Román et al., 2014b) but some Saharan dust episodes also happen, especially in summer (Cachorro et al., 2016). These Saharan episodes are more frequent at Granada due to its proximity to North Africa, since this city is located at the South-East of the Iberian Peninsula (Valenzuela et al., 2012). Both stations are equipped with a “CHM-15k Nimbus” ceilometer (*Lufft* manufacturer), a “SONA” sky camera (*Sieltec Canarias S.L.*) and a “CE318-N” sun/sky-photometer (*Cimel Electronique*).

Both ceilometers belong to the Iberian Ceilometer Network (Cazorla et al., 2017). They provide CC and CBH measurements every 15 seconds. The cloud cover determined by the ceilometer is described on the Jenoptik CHM15k user manual (Jenoptik, 2013). It is determined using the previously calculated cloud bases heights. First, a time interval is considered and its length depends on the cloud base height, being longer for higher clouds creating a “temporal cone of influence”. The frequency of cloud bases is calculated for each time interval. Peaks in the frequency distribution are separated and all cloud bases in the space of a peak will be clustered to one cloud layer. The calculation of the total cloud cover value is done within a rectangle depending on time and altitude. For this purpose the mentioned time interval (“truncated cone of influence”) will be divided in a fixed number of small truncated cones. Parts containing cloud bases are counted against the total number of cone parts and the cloud cover value is expressed as a percentage value from this comparison. Finally the percentage value is expressed in oktas.

SONA ("Sistema de Observación de Nubosidad Automático": Automatic Cloud Observation System) sky camera takes hemispherical sky images along the whole day but in this work we only use daytime images. It consists of a surveillance CCD camera, which provides three channels (RGB) images with 8 bit-digitalization providing 256 counts per channel (Cazorla et al., 2015). This camera with a fisheye lens is inside a waterproof case which has a quartz dome and a shadow band in order to block the sun (González et al., 2012). The field of view (**FOV**), zenith (**ZEN**), and azimuth (**AZI**) matrices,representing the angles viewed by each pixel, were obtained correlating the pixel positions of celestial bodies whose coordinates are well known (Román et al., 2017).

The sun/sky-photometers used in this work are integrated in AERONET network (AErosol RObotic NETwork; Holben et al., 1998). The AERONET processed data used in this work are the daily average of *AOD* at 440 nm and the daily Angström Exponent (*AE*) obtained in the spectral range 440-870 nm. All these cloud-screened data (level 1.5) correspond to the AERONET Version 3 algorithm, and are available at http://aeronet.gsfc.nasa.gov.

CC measured visually by trained meteorological observers at two AEMet stations (State Meteorological Agency of Spain) is also available. These measurements are taken three times at day: 07:00, 13:00 and 18:00 UTC, and are given in oktas with a resolution of 1 okta. These AEMet stations are 3.75 km and 4.75 km far away in a straight line for Granada and Valladolid stations, respectively.

Finally, direct beam solar shortwave (SW_b) irradiance data was obtained at each minute as the difference between global and diffuse components recorded by two CM-11 pyranometers (Kipp & Zonen); diffuse is measured using a shadow-ball in a sun-tracker. This kind of data is not available at Valladolid, at least near to the sky camera used in this work. Both pyranometers at Granada presents a relative uncertainty of 1.9% and are frequently calibrated using a reference instrument at the site (Antón et al., 2012).

3 HDR imagery

In order to avoid saturated pixels and to linearize the radiometric pixel response of the sky images, both SONA sky cameras were configured to take HDR images. First, an image with the lowest exposure time (~4 µs) is taken and then, another 11 images are taken doubling the exposure consecutively (the last image taken with ~4 ms); the whole process requires about 5 seconds. As a result, just one HDR linear image is obtained applying the method of Debevec and Malik (1997) to the 12-image set using the pixel sensitivity calculated by Román et al. (2017). Every 5 minutes two consecutive HDR images, I_1 and I_2, are taken, expending about 10 seconds in the process. The availability of HDR images used in this work spans from 10th March 2015 to 24th September 2015 at Valladolid and from 16th November 2015 to 19th September 2016 at Granada.

Some differences between the use of individual direct images and the use of HDR images are shown in Figure 1 for Granada. This figure shows the individual sky image (left panels) which is usually used in CC detection algorithms; the HDR composition for the same images are shown (right panels) using a tone map (Reinhard et al., 2002). In the first and third cases the direct image shows cloudy pixels, especially near the sun, that are saturated; it does not give information about those pixels because it is impossible to know to what degree are these parts more intense than the non-saturated ones or which are the most intense pixels inside a saturated area. This problem does not appear for the HDR images, where every pixel can be discerned. An interesting case is the last image of Figure 1, which was taken under cloud-free conditions but with a high coarse aerosol load (*AOD* at 440 nm around 0.5 and AE ~ 0.15) corresponding to a Saharan dust episode. As can be observed, in the direct image the pixels around the sun are saturated at the three channels, however the HDR image avoids this problem that frequently appears under dusty conditions.

4 CPC algorithm

The CPC (Camera Plus Ceilometer) algorithm is designed in order to detect if the sun is obstructed or unobstructed by clouds, and also to discern if the portion of sky viewed by each pixel is cloud-free or not. This algorithm is based on individual pixel information (like red to blue ratio of each pixel) and also on information of the full HDR image as a whole (like the average of red-blue ratio of all pixels) as has been previously used by Calbó and Sabburg (2008), Heinle et al. (2010) and Kazantzidis et al. (2012) to classify clouds. CPC discern cloudy pixels to cloud-free pixels with high aerosol load using the concept that sky under cloud-free conditions shows symmetry. In addition: the temporal change between consecutive images is considered to observe cloud conditions near the sun; some criteria are applied or not depending on the presumable sky condition which is identified from ceilometer information; and image edge detection is used to detect cloud borderlines.

CPC algorithm first detects if the sun is obstructed or not by clouds and then applies some criteria, which vary if sun is obstructed or not, to classify pixels as cloudy or cloud-free. Most of the decisions of CPC algorithm are based on the use of threshold for some variables derived from the sky images. All these threshold values, shown in Table B.1 in Appendix B, were selected tuning manually these values and observing the output of the algorithm for a large image-set with multiple observations of different sky conditions. Some thresholds are site dependent because of the differences in the configuration of the cameras at Granada and Valladolid, since they have some differences in properties like the white balance, gamma correction and others.

The CPC method is based on deriving some variables from the sky HDR images. The Appendix A describes these variables, usually matrices directly derived from the HDR images, but also scalar variables derived from these matrices. The workflow of CPC algorithm is described in detail in Appendix B, where RBR algorithm is also explained. RBR needs only the I_1 HDR image as input while CPC uses I_1 and I_2 HDR images and also some ceilometer information. Both algorithms return as output the sun condition and the identification of every single pixel of the HDR image (I_1) as cloudy or cloud-free.

5 Results

5.1 Individual cases

RBR and CPC algorithms were applied to the HDR images of the particular cases of Figure 1, and results are shown on Figure 2. Black pixels are the masked pixels, white pixels the cloudy pixels and blue pixels the cloud-free. Regarding sun condition, the white/yellow disk on the sun position represents obstructed/unobstructed condition. In the first case, the clouds are well detected by RBR algorithm, but it considers that the sun aureole, which is whiter, is cloudy and hence the sun is obstructed; it does not happen with CPC algorithm, which determines that sun is unobstructed, however, some cloudy pixels are considered as cloud-free due to the symmetry criteria. RBR under a cloud-free and clean atmosphere, like the second case of Figure 2, both algorithms work well although CPC shows some cloud-free pixels as cloudy in the right side due to the symmetry criteria; however for this case the ceilometer cloud cover algorithm (CEI) provided a CC of 7 oktas since the sky was too cloudy half hour before the image was taken, which evidences that CEI algorithm presents some problems at least when the cloud cover rapidly varies.

A similar problem happens in the third case of Figure 2, where the clouds started to appear in the sky and hence CEI considered a CC value of zero oktas; RBR and CPC perfectly detect the cloudy pixels in this case. Finally, in the case of cloud-free sky with high dust aerosol load, RBR selects all pixels as cloudy because under these conditions the sky is whiter (AE tends to zero giving a similar scatter radiance at all wavelengths); CPC also shows some pixels as cloudy like the pixels with azimuths below -178º and above 178º (symmetry criteria is not applied at these angles), but the most of them are detected as cloud-free and at least the sun is considered as unobstructed; CEI provides a CC value of 1 okta, more in agreement with the real sky conditions.

All these examples indicates that RBR works fine when sun is obstructed or under very clean conditions but it fails under not clean conditions, the sky being considered overcast when the atmosphere is cloud-free but with high coarse aerosol load. CPC presents a better performance but it also shows cloudy pixels as cloud-free and vice versa. CEI algorithm provides CC values far to the real scenario when the CC quickly and strongly changes.

5.2 Comparison of sun condition

In order to compare the obtained sun condition by RBR and CPC algorithms, the measurements of SW_b at Granada were used as reference, since SW_b is very sensitive to clouds obstructing the sun. To this purpose, for all available measurements of SW_b, the same SW_b but under cloud-free conditions (SW_b^{cf}) were simulated using the UVSPEC tool of the libRadtran 1.7 package (Mayer and Kylling, 2005) in a similar way as in Román et al. (2014a, b). The used radiative transfer solver was the two-stream fluxes (“twostr”) solver developed by Kylling et al. (1995), and the inputs were: the daily *AOD* at 440 nm, AE and water vapour column from AERONET (version 3, level 1.5), daily total ozone column from OMI onboard Aura satellite (TOMS algorithm version 8.5).Monthly climatological values are used as input for days without availability of daily input.

Once the simulations were performed, the ratio (R_{sw}) from the measured SW_b to the simulated SW_b^{cf} was calculated for all measurements. The sun was considered as unobstructed if R_{SW} is between 0.87 and 1.13, giving and error margin around 13% between the simulations and the measurements, in a similar way to Cazorla et al. (2015). Considering as reference of unobstructed sun the R_{SW} values between 0.87 and 1.13, P_{sc} was calculated as the percentage of data that camera algorithms determinate the same sun conditions that R_{sw} criterion. The images taken with SZA above 80º were rejected and, as a result, 34504 pairs of coincident data of R_{SW} and camera sun condition were available. 85% of sun conditions obtained from CPC algorithm are in agreement with the reference, while only 48% of sun conditions obtained from RBR algorithm fit with the reference marked by R_{sw} criterion. For the cases considered as unobstructed (20204 data), 88% of sun conditions from CPC algorithm are in agreement with the reference, but this percentage is only 12% for the RBR algorithm. P_{sc} is 81% and 99% for CPC and RBR algorithms, respectively, when the obstructed cases are only taken into account. These results indicate that RBR method tends to consider the sun obstructed in the most of the cases.

Figure 3 shows the values of P_{sc} for both algorithms at different SZA intervals. CPC algorithm presents P_{sc} values near 90% for SZA below 60º independently on sun condition, but they slightly decrease, especially for obstructed conditions, to around 75% for SZA between 70º and 80º. The worse agreement at high SZA values, especially for obstructed conditions, could be caused because clouds obstructing the sun at high zenith angles are closer to the horizon and they are more difficult to be identified. On the other hand, the P_{sc} values obtained from RBR algorithm shows a strong dependence on the reference sun condition because they are always near 100% for obstructed conditions but never above 20% for unobstructed conditions; the increase with SZA of P_{sc} for all conditions case in the RBR algorithm is caused by the increase with SZA of the number of obstructed cases.

A similar study appears in Figure 4 but for P_{sc} as a function of *AOD* daily mean intervals instead of SZA. CPC fits better the sun condition for the unobstructed cases, ranging from 75-95%, and does not present any dependence on *AOD*. P_{sc} from RBR algorithm shows that RBR considers sun obstructed for all cases when *AOD* at 440 nm is above 0.3.

Figure 5 shows the P_{sc} data versus AE intervals and considering only data which daily *AOD* at 440 nm was between 0.1 and 0.2, in order to remove any dependence on *AOD*. CPC does not show a significant dependence on AE, showing high P_{sc} values (between 75-95%) except for the obstructed conditions at low AE values (coarse particles), where P_{sc} ranges from 34% to 52%. RBR algorithm clearly shows that considers the most of time the sun obstructed except when the size of particles decreases (high values of AE) because P_{sc} increases with AE under unobstructed conditions.

5.3 Comparison of cloud cover

Cloud cover from camera usually is calculated, in %, as the sum of pixel detected as cloudy divided by the sum of all available pixels (not masked), and then this ratio is multiplied by 100%. However, this way does not take into account that the solid angle of the sky viewed by each pixel is different. Hence, in order to give more weight to the pixels viewing larger solid angle, the CC was calculated from HDR sky images as:

$$CC(\%) = 100\% \frac{\sum_{i,j} c_{i,j} FOV_{i,j}}{\sum_{i,j} FOV_{i,j}}, \quad (1)$$

being $c_{i,j}$ equal to 1 if the i,j-pixel is cloudy and 0 if is cloud-free; masked i,j-pixels are not taken into account. The differences between CC (retrieved by CPC algorithm) obtained without weighting solid angle and by Eq. (1) were calculated for 61008 available cases at Valladolid and Granada with different sky conditions. Maximum and minimum values of these differences were 8.3% and -7.5%, respectively; the mean and standard deviation values were 0.6% and 1.5%, respectively. 68% and 96% of the absolute differences were below 1% and 4%, respectively. However, if only the cases with CC, obtained by Eq. (1), above 10% and below 90% are chosen (17324 data) the mean and standard deviation of the difference increase to 1.3% and 2.4%, respectively. These results indicate that there is not a significant difference between the two methods, but CC obtained as the sum of cloudy divided by the sum of all pixels slightly overestimates the CC values calculated by Eq. (1), especially for partially cloudy conditions.

In this work Eq. (1) was used to obtain CC for both CPC and RBR algorithms. CC was transformed to oktas multiplying by 8 oktas and dividing by 100%. This CC was not rounded to integer like in ceilometers and human observations.

In the case of Valladolid, images were taken every 5 minutes except at the full hour (this camera was configured to be rebooted every hour), then, the CC at the full hour was considered as the average of the CC obtained 5 minutes before and after. To be coherent in this location the CC from CEI at the full hour was also averaged from the CC measured 5 minutes before and after.

In order to know the performance of the different methods to obtain CC (CPC, RBR and CEI), the CC measurements performed manually by the AEMet observers are used as reference. Therefore, the ΔCC distribution was calculated for each algorithm by:

$$\Delta CC = CC_a - CC_r, \tag{2}$$

where CC_a is the CC of the algorithm (CPC, RBR or CEI) and CC_r the reference (AEMet). Table 1 shows some statistical parameters of ΔCC that helps to quantify the agreement between the algorithms and the reference: N is the number of ΔCC data used; MBE (Mean Bias Error) is the average of ΔCC, with positive values indicating that algorithms overestimate the reference and viceversa; MDBE (Median Bias Error) is the percentile 50 (median) of ΔCC and also is related to the accuracy quantifying the over/under-estimation. MDBE is less affected by ΔCC outliers than MBE; MABE (Mean Absolute Bias Error) is the average of the absolute values of ΔCC and gives information about the average difference in absolute value between algorithm and reference; SD is the standard deviation of ΔCC and provides information about the deviation of ΔCC.

As can be observed in Table 1, the MBE for CPC is zero at Granada while CEI MBE is small at both sites. The MDBE indicated that the accuracy of CEI and CPC are similar, with differences of zero oktas. MABE and SD are similar for CEI algorithm while RBR algorithm presents larger values. For the two stations together the averaged absolute difference with the reference is around 1okta for CPC and CEI algorithms, while the precision quantified by the SD is around 2 oktas for both algorithms too.

The previous results could be affected by the fact that reference values (AEMet) were recorded 4-5 km far away from the sky cameras, especially for low clouds. Therefore, in order to see the influence of cloud altitude, *ΔCC* values have been represented in Figure 6 as a function of different CBH intervals. The CBH used for this purpose was obtained as the 1-hour averaged CBH, from ceilometer, half hour before and after the measurement. Against the expected, *ΔCC* presents, for the three algorithms, mean and median values near to zero and the standard deviation for low clouds with CBH below 2 km. In fact the standard deviation of *ΔCC* for these clouds is lower than for other higher clouds. Both CPC and RBR algorithms presents mean values near to zero for CBH below 9 km a.g.l., while for clouds with CBH above 9 km a.g.l. these values are near to -1, indicating that CPC and RBR underestimates the cloud cover of the highest clouds, like cirrus, likely due to this kind of clouds cannot be easily appreciated by the sky camera. On the other hand, the CEI algorithm overestimates around 1 okta the human observations for clouds which CBH ranges from 3 to 9 km a.g.l., like mid-level clouds; however CEI presents mean *ΔCC* values near to cero for clouds with CBH below 3 km a.g.l. and above 9 km a.g.l.. In general CEI presents the lowest standard deviation for all CBH intervals, and RBR the highest one.

In order to see any dependence of the agreement between the algorithms and human observations on cloud cover, Figure 7 shows the *ΔCC* distributions for different CC values and for the three algorithms. For cloud-free conditions (CC=0 oktas) the most of CC values retrieved by CEI and CPC algorithms are zero, while RBR presents more deviation. The deviation is higher for CEI in the case of overcast conditions (CC=8 oktas), where inter-quartile range of *ΔCC* is very low for CPC and RBR algorithms. For partially cloudy conditions all algorithms have a similar behaviour not showing a clear dependence on CC, and being the highest deviation for CC between 3 and 5 oktas. In general, the standard deviation and the inter-quartile range are lower for the CPC algorithm, and the absolute value of the *ΔCC* mean is below 1 for all algorithms.

Looking for a dependence on AOD, Figure 8 shows the *ΔCC* as a function of AOD at 440 nm intervals. CEI algorithm shows an agreement with the reference similar for all AOD values, with mean and median of *ΔCC* close to zero in all AOD intervals. The mean and median of *ΔCC* of CPC algorithm slightly increase with AOD, being always below 2 oktas; however, these values increase with AOD for RBR algorithm, showing values higher than 2 oktas for AOD above 0.3. CEI and CPC present similar standard deviation and inter-quartile range for all AOD intervals, but RBR shows increasing deviation as AOD increases. It confirms that RBR algorithm does not retrieve a good CC value under turbid conditions.

In addition, trying to observe the effect of aerosol size, *ΔCC* is represented in Figure 9 as a function of AE intervals. In this case only data with AOD (440 nm) between 0.1 and 0.2 were used in order to avoid any AOD dependence. CEI and CPC show similar behaviour, without any dependence on AE, presenting CPC the best accuracy and precision against the reference for the coarsest particles (AE~0.3). On the other hand, the CC is overestimated by RBR algorithm for coarse aerosol, being the mean of *ΔCC* near 2 oktas for AE below 0.4.

Finally, all the calculations of this section have been also done (not shown) with CPC algorithm but removing the nodes classified as potential candidates to be removed (Figure B1). The obtained results have been basically the same that with these nodes, which indicates that for CC calculation the CPC algorithm can be rewritten in a simpler way. However, in this work these nodes have been remained in the CPC algorithm in order to identify a few cloudy/cloud-free pixels which could be interesting although their influence on the total cloud cover is very small.

6 Conclusions

Regarding the detection of sun condition (obstructed/unobstructed), the Camera Plus Ceilometer (CPC) algorithm has shown to be in good agreement with the reference, fitting around 85% of estimations with the reference. CPC does not show any strong dependence on the sun condition, solar zenith angle (at least for values below 70º), aerosol size nor optical depth (AOD). However, the Red-Blue camera ratio (RBR) algorithm usually classifies the sun condition as obstructed independently of the reference condition, except for low AOD values or fine particle predominance (high Angstrom exponent) when RBR algorithm fits better with the reference. The ceilometer does not provide this kind of information.

Cloud cover retrieved by the ceilometer algorithm fits in general better with the reference, but CPC algorithm shows a similar agreement, being the mean absolute differences with respect to the reference for CEI and CPC algorithms equal to 1 okta (around the resolution of the reference). CPC and RBR algorithms underestimate the cloud cover of high clouds above 9 km, but CEI overestimates the cloud cover of clouds with cloud base height between 3 and 9 km. RBR algorithm overestimates the cloud cover under the presence of coarse aerosol (low Angstrom exponent), like desert dust, and also under high aerosol load. In fact, RBR has detected as overcast sky a full cloud-free sky with moderate-high dust load. CPC and CEI algorithms do not present any dependence on aerosol.

These results could indicate that only a ceilometer is enough to retrieve cloud cover since the agreement with the reference is similar to the CPC algorithm, however CPC algorithm, instead of CEI, has the power to locate the clouds in their position inside the celestial vault; in addition CPC is capable to discern if the sun is obstructed or not. In a next step, the CPC algorithm could be also combined with cloud base height information from ceilometer in order to retrieve the cloud type. The authors encourage researchers to test the CPC algorithm in other stations, even rejecting the nodes assumed as potential candidates to be removed in order to reduce complexity. Finally, CPC algorithm has been designed to take some actions or not depending on the presumable sky conditions, which are obtained from a ceilometer; however, there are a lot of sky camera systems being not near to a ceilometer; hence, although it is beyond the scope of this paper, authors propose to find another way to obtain the presumable sky conditions in the CPC algorithm for these cases without a ceilometer. Some possible ideas for this purpose could be, for example, using solar radiation measurements as proxy, or maybe considering the variation in the sky images in a time window.

Acknowledgements

This work was supported by the Andalusia Regional Government (project P12-RNM-2409) and by the “Consejería de Educación" of "Junta de Castilla y León" (project VA100U14); the Spanish Ministry of Economy and Competitiveness and FEDER funds under the projects CGL2013-45410-R, CMT2015-66742-R, CGL2016-81092-R, and “Juan de la Cierva-Formación” (FJCI-2014-22052) program; and the European Union H2020-INFRAIA-2014-2015 project ACTRIS-2 (grant agreement No 654109).

Appendix A: CPC variables

- **R**: This matrix is the red channel of the image I_1 divided (right-array division) by **FOV**. Each $R_{i,j}$ element represents the relative radiance viewed by the i,j-pixel of the image I_1in the red channel.
- **B**: Similar than **R** but for the blue channel. $B_{i,j}$ values (and also $R_{i,j}$ values) are higher for cloudy pixels and pixels near sun aureole, especially under high coarse aerosol load.
- **R2**: Similar than **R** but for the image I_2.
- **DIS**: This represents the distance of each i,j-pixel to the sun (scattering angle). It is calculated in degrees using the Solar Zenith Angle (SZA) value and the **ZEN** and **AZI** matrices as in Ghonima et al. (2012).
- **RT**: This matrix is the ratio of the elements of **R** to the elements of **B**. $RT_{i,j}$ represents the whiteness of the i,j-pixel, since the sky is less blue when $RT_{i,j}$ increases. $RT_{i,j}$ values are higher for cloudy pixels, but also for cloud-free pixels under high coarse aerosol load.
- **SRT**: is a smoothed matrix of **RT**, being each i,j-element of **SRT** the average of the square of 5x5 pixels of **RT** centred in the i,j-element. This matrix also gives information about sky whiteness, and it is more useful than **RT** for dark skies since, under these conditions, **R** and **B** present low values, which provides higher deviations to **RT** values.
- **CI**: This matrix represents the relative difference (pixel by pixel) between the matrix obtained by smoothing **B** (like in the **SRT** case) using a 5x5 pixels window and the matrix obtained smoothing **B** using a window of 10x10 pixels. The high values of **CI** usually indicates the presence of cloud borderlines but also cirrus (hence the name CI since it helps to detect cirrus)

and even contrails from aviation, which Red-Blue ratio is not too high but their frequently filamentous shape leads to higher **CI** values.

- **SR**: This is a symmetrical matrix of **R** calculated using as symmetry axis the solar principal plane.
- **DSR**: It represents the absolute value of the relative difference between the red relative radiance of two symmetric pixels. Each element of this matrix is calculated as:

$$DSR_{i,j} = 100\% \frac{|R_{i,j} - SR_{i,j}|}{\min(R_{i,j};SR_{i,j})} \tag{A.1}$$

Low values of DSR usually indicate cloud-free pixels, although it is also possible to have low values for cloudy skies. This matrix is finally smoothed by a 5x5 pixels window. DSR is near to zero when one pixel and its symmetric are similar, even if both pixels are too white.

- **SDSR**: It represents the signed relative difference between the red relative radiance of two symmetric pixels. Each element of this matrix is calculated as:

$$SDSR_{i,j} = 100\% \frac{R_{i,j} - SR_{i,j}}{SR_{i,j}} \tag{A.2}$$

This matrix helps to discern if two pixels that are not symmetric are both cloudy or one of them is cloud-free. **R** in a cloud-free pixel must be lower than in a cloudy pixel and, therefore, high negative values of **SDSR** could be related with cloud-free pixels.

- **DR2**: This matrix represents the absolute value of the relative differences of the red channel of the image I_1 and image I_2. High values of **DR2** indicate changes on sky in a short time, which can be attributed to cloud presence.

This matrix can help to discern cloudy pixels near the aureole. **DR2** is calculated as:

$$DR2_{i,j} = 100\% \frac{\left|R_{i,j} - R2_{i,j}\right|}{\min(R_{i,j}; R2_{i,j})} \quad \text{(A.3)}$$

This matrix is finally smoothed by a window of 3x3 pixels. The values of this matrix usually are high for borderlines of clouds.

In addition to these matrices, some scalar variables can be used as a proxy for the sky condition, and they are the next:

- *mrt*: This is the median value of **RT** computed for all non-masked elements. It represents the averaged whiteness of an image.
- *sdr*: This values is defined as:

$$sdr = \frac{\mathrm{std}(\mathbf{R})/\mathrm{mean}(\mathbf{R})}{\mathrm{std}(\mathbf{B})/\mathrm{mean}(\mathbf{B})} \quad \text{(A4)}$$

  where std(**R**) and mean(**R**) are the standard deviation and average, respectively, of all elements of **R** which are not masked. *sdr* represents the ratio of the relative variation of red channel to the variation of the blue channel; this ratio shows lower values for overcast conditions and dark skies, then it will be useful to determine if sun is obstructed.
- *cns*: It represents the percentage of cloudy pixels near the sun (**DIS**<10º). If this percentage is low likely the sun will be unobstructed by clouds.
- *cis*: Once pixels are identified as cloudy/cloud-free, the cloud pixel conditions is interpolated to the sun position because the sun is blocked by the shadowband. *cis* is the percentage of cloudy pixels (interpolated) inside a disk with a radius of 7 pixels and centred in the sun position. Sun is likely obstructed if *cis* value is high.

- *cir*: is defined as the percentage of CBH data that is not null and higher than 6000 m a.g.l. in the time interval of half hour before and after (length of 1 hour) the image was taken. High *cir* values indicate the presence of high clouds like cirrus, which are more difficult to identify with a RGB camera than with a ceilometer.
- *clo*: is the percentage of CBH data that is not null and above 600 m a.g.l. in the time interval of 5 minutes before and after (length of 10 minutes) the image was taken. The mean of *clo* is similar to *cir* but for all clouds and in a shorter time interval.

Appendix B: CPC workflow

Figure B.1 shows the workflow diagram of CPC algorithm. In order to test the diagram conditions, 549 cases at Granada, when cloud cover visually measured by AEMet staff was available, have been chosen. The main inputs of CPC are the **R, B** and **R2** matrices of the recorded HDR images. Other inputs are SZA, **ZEN** and **AZI** (in order to locate some elements as the sun and its aureole), and *cir* and *clo* from ceilometer in order to obtain the presumable sky condition. The first step is to apply a mask to the HDR image, removing all pixels with zenith angles above 80º and also all pixels corresponding with the shadowband.

If *sdr* is below a threshold value T_{sdr1} (dark sky) and in addition *mrt* is above other threshold T_{mrt} (most of pixels enough white), then CPC algorithm will consider that the sun is obstructed (left branch) by clouds and the i,j-pixels which $SRT_{i,j}$ value is below/above T_{SRT1} will be chosen as cloud-free/cloudy. These conditions are usually satisfied by overcast skies, therefore the threshold for **SRT** was defined dynamic (Table B.1) considering as cloud-free the pixels with a **SRT** below the three quarters of the averaged Red-Blue ratio. 7% of the chosen cases satisfy the first condition, and the averaged cloud cover from AEMet for them is 7.5 oktas, which indicates that this node really identifies overcast skies.

If the previous condition (left branch) is not satisfied but *sdr* is below a second threshold (T_{sdr2}) and also *mr* higher than T_{RT1} (Table B.1), then the sun is considered obstructed by clouds (middle branch) and the i,j-pixels which $SRT_{i,j}$ value is below the threshold T_{SRT2}, will be considered as cloud-free. 13% of chosen cases cross this middle branch, being the averaged cloud cover in these cases around 6.6 oktas, which still corresponds to high amount of clouds and, therefore sun should be likely obstructed.

If the last condition is still not satisfied (80% of selected cases; averaged cloud cover about 1.7 oktas), then the red-blue ratio (RBR) algorithm, which considers as cloudy all i,j-pixel which $RT_{i,j}$ value is above T_{RT1} (too white pixels) and cloud-free the rest (too blue pixels), is applied (Figure B.1). After RBR is applied, the pixels are temporally classified as cloudy/cloud-free following the RBR method. Cloud-free pixels which show **CI** values higher than T_{CI} (threshold value in Table B.1) will be considered as cloudy, indicating the presence of cloud borderlines, cirrus or contrails which are not detected by RBR method. On average, for cloud cover of 0-1 oktas, around 10% of cloudy pixels detected after this node are due to the criterion based on **CI** matrix; however this percentage is below 4% for cloud cover above 3 oktas. These results indicate that the node based on **CI** matrix has not too much influence in the cloud cover detection, hence it could be removed from the algorithm (the three red nodes are potential candidates to be removed), however it has not been removed because it is useful to identify a few cloudy pixels which could be interesting.

After that, CPC algorithm considers the ceilometer information in order to establish the presumable sky condition. If *cir* is higher than T_{cir} or *clo* above T_{clo} (both threshold shown in Table B.1), CPC will consider that sky is presumable cloudy and it will not apply a symmetry criteria to transform pixels classified as cloudy to cloud-free; it occurs for the 18% of chosen cases that reach the node, which present an averaged cloud cover of 3.7 oktas. On the other hand, if the sky is not considered presumable as cloudy (82% of cases; averaged cloud cover of 1.2 oktas), then CPC will classify as cloud-free all cloudy i,j-pixels (except i,j-pixels which $AZI_{i,j}$ is below -178º or above 178º since the distance between symmetric pixels is too short) that present a value of $DSR_{i,j}$ below a threshold named T_{DSR1} (shown in Table B.1). This criterion assumes that pixels which are similar to their symmetric pixels can be assumed as cloud-free like in the AERONET almucantar observations (Holben et al., 1998), at least under a presumable cloud-free conditions (given by ceilometer information). Regarding T_{DSR1}, this threshold is higher for pixels near to the sun aureole (**DIS**<10º) because the sky radiance gradient in this area is stronger and little uncertainties in camera geometry characterization can introduce higher differences between symmetric points. The average percentage of cloudy pixels transformed in cloud-free by this criterion is 63%, 55%, 47% and 39% for 0, 1, 2, and 3 oktas respectively. This percentage increases with AOD at 440 nm, being below 50% for AOD values lower than 0.15, and around 70% for AOD above 0.25. These results indicate that this node is, at least, one of the main responsible to avoid the aerosol influence on cloud cover detection.

However, even if symmetrical criteria turn a pixel to cloud-free, and independently on the presumable sky conditions, the i,j-pixels which $RT_{i,j}$ values are above T_{RT2} (shown in Table B.1) will be assumed as cloudy at this section of the algorithm (pixels are too much white). After that, the cloud-free pixel satisfying that **DSR** is above the threshold T_{DSR2} (pixel does not show symmetry) and also **RT** is higher than T_{RT3} (likely too white) will also be considered as cloudy. It identifies pixels which do not show enough symmetry and are whiter than an established threshold T_{RT3}, which was dynamically defined in Table B.1 as a value 10% whiter than the mean of all pixels. The last criteria are included in only one node. The averaged percentage of cloud-free pixels turned to cloudy by this node is 8% for 8 oktas, which indicates that this node can identify around half okta of cloudy pixels, previously considered as cloud-free, under overcast conditions. This node is useful to detect cloudy pixels showing symmetry in overcast conditions.

The next steps of the algorithm are focused on estimating the cloud conditions near the sun aureole, and the sun conditions. In the case of presumable cloud-free conditions (*cir* below T_{cir} and *clo* below T_{clo}) CPC will turn to cloud-free all the cloudy pixels near the sun (**DIS**<10º) which show **DR2** values below an established threshold T_{DR2}. This assumes that cloud-free pixels at sun aureole do not present changes in a short time (few seconds) higher than a threshold which was manually determined as 4% (Table B.1). The last criterion based on **DR2** is not applied under presumable cloudy conditions since sometimes under stable cloudy conditions, especially overcast conditions, the differences between I_1 and I_2 are too low, **DR2** showing values near to zero. The averaged percentage of cloudy pixels turned to cloud-free by the criterion based on **DR2** matrix is 99% for zero oktas and it monotonously decreases to 66% up to 5 oktas; it does not show any dependence on AOD.

With the temporal classification of all pixels in this point, CPC calculates *cns* and *cis* in order to identify the sun condition. CPC will identify the sun as unobstructed by clouds if *cns* is below T_{cns} or *cis* is below T_{cis}, and as obstructed on the other cases. The RBR algorithm considers that sun is unobstructed by clouds if *cis* is below T_{cis}. 80% of chosen cases reaching this point have been classified as sun unobstructed.

Once the sun condition is defined, the algorithm shows two branches, one for each sun condition. If the sun is considered obstructed by clouds then the symmetry criteria are rejected, and the only i,j-pixels classified as cloudy will be the pixels which $RT_{i,j}$ value is above T_{RT1}, or $DSR_{i,j}$ is above T_{DSR2} at the same time that $RT_{i,j}$ is higher than T_{RT3}(as explained above), or $CI_{i,j}$ value is higher than T_{CI}. Symmetry criteria for cloud-free pixels detection are not applied when the sun is considered as obstructed by clouds because it has been observed that this kind of criteria does not work well under this sun condition.

If the sun is considered unobstructed then a last criterion based on **SDSR** is applied in order to identify cloud-free pixels that are considered cloudy because their symmetric pixels are cloudy, which provides high **DSR** values. To this purpose, CPC turns to cloud-free all cloudy i,j-pixels which $SDSR_{i.j}$ is lower than T_{SDSR}, $DSR_{i,j}$ is higher than T_{DSR1} (symmetry criteria for cloud-free pixels detection applied above is not satisfied), and $RT_{i,j}$ is below T_{RT4} (pixels too white are still considered as cloudy). The averaged percentage of cloudy pixels converted to cloud-free by this last node ranges from 1% to 6% for the different cloud cover values. The influence of this node in the final cloud cover is small, which indicates that it is also a potential candidate to be removed from the CPC algorithm.

References

Alonso, J., Batlles, F.J., López, G., Ternero, A. (2014): Sky camera imagery processing based on a sky classification using radiometric data. Energy, 68, 599-608.

Antón, M., Gil, J.E., Cazorla, A., Fernández-Gálvez, J., Foyo-Moreno, I., Olmo, F.J., Alados-Arboledas, L. (2011): Short-term variability of experimental ultraviolet and total solar irradiance in Southeastern Spain. Atmos. Environ., 45, 28, 4815-4821.

Antón, M., Valenzuela, A., Cazorla, A., Gil, J.E.,Fernández-Gálvez, J.,Lyamani, H., Foyo-Moreno, I., Olmo, F.J., Alados-Arboledas, L. (2012): Global and diffuse shortwave irradiance during a strong desert dust episodeat Granada (Spain). Atmos. Res.,118,232-239.

Arking, A., Childs, J. D. (1985): Retrieval of cloud cover parameters from multispectral satellite images. J.Clim. Appl.Meteorol., 24, 4, 322-333.

Bartlett, J.S., Ciotti, Á.M., Davis, R.F., Cullen, J.J. (1998): The spectral effects of clouds on solar irradiance. J.Geophys. Res.: Oceans, 103, C13, 31017-31031.

Beyer, H.G., Hammer, A., Luther, J., Poplawska, J., Stolzenburg, K.,Wieting, P. (1994): Analysis and synthesis of cloud pattern for radiation field studies. Sol. Energy, 52, 5, 379-390.

Bennouna, Y.S., Cachorro, V.E., Torres, B., Toledano, C., Berjón, A., De Frutos, A.M., Coppel, I. A. F. (2013): Atmospheric turbidity determined by the annual cycle of the aerosol optical depth over north-center Spain from ground (AERONET) and satellite (MODIS). Atmos. Environ., 67, 352-364.

Boucher, O., Randall, D.,Artaxo, P., Bretherton, C., Feingold, G., Forster, P.,Kerminen, V.M., Kondo, Y., Liao, H., Lohmann, U., Rasch, P., Satheesh, S.K., Sherwood, S., Stevens, B., Zhang, X.Y. (2013): Clouds and Aerosols. In: Climate Change 2013: The

Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Cachorro, V.E., Burgos, M.A., Mateos, D., Toledano, C., Bennouna, Y., Torres, B., de Frutos, A., Herguedas, A. (2016). Inventory of African desert dust events in the north-central Iberian Peninsula in 2003–2014 based on sun-photometer–AERONET and particulate-mass–EMEP data. Atmos. Chem. Phys., 16(13), 8227-8248.

Calbó, J., Sabburg, J. (2008): Feature extraction from whole-sky ground-based images for cloud-type recognition. Journal of Atmospheric and Oceanic Technology, 25(1), 3-14.

Cazorla, A., Olmo, F.J., Alados-Arboledas, L. (2008). Development of a sky imager for cloud cover assessment. J. Opt. Soc. Am. A, 25(1), 29-39.

Cazorla, A., Husillos, C., Antón, M., Alados-Arboledas, L. (2015): Multi-exposure adaptive threshold technique for cloud detection with sky imagers. Solar Energy 114, 268–277.

Cazorla A., Casquero-Vera, J.A., Román, R., Guerrero-Rascado, J.L., Toledano, C., Cachorro, V.E., Orza, J.A.G., Cancillo, M.L., Serrano, A., Titos, G., Pandolfi, M., Alastuey, A., Hanrieder, N., Alados-Arboledas, L. (2017): Near Real-Time processing of ceilometer network data: characterizing an extraordinary dust outbreak over the Iberian Peninsula. Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2017-151, 2017 .

Costa-Surós, M., Calbó, J., González, J.A., Martin-Vide, J. (2013): Behavior of cloud base height from ceilometer measurements. Atmos. Res., 127, 64-76.

Costa-Surós, M., Calbó J., González, J A., Long, C.N. (2014): Comparing the cloud vertical structure derived from several methods based on radiosonde profiles and ground-based remote sensing measurements. Atmos. Meas. Tech., 7, 9, 2757–2773.

Debevec, P.E., Malik, J. (1997): Recovering high dynamic range radiance maps from photographs. Proceedings of SIGGRAPH 97, 369–378. ISBN 0-89791-896-7.Held in Los Angeles, California.

Frederick, J. E., Steele, H.D. (1995): The transmission of sunlight through cloudy skies: An analysis based on standard meteorological information. J. Appl. Meteorol., 34, 12, 2755-2761.

Gao, B.C., Yang, P., Han, W., Li, R.R., Wiscombe, W.J. (2002):An algorithm using visible and 1.38-μm channels to retrieve cirrus cloud reflectances from aircraft and satellite data. IEEE T.Geosci. Remote, 40, 8, 1659-1668.

Ghonima, M.S., Urquhart, B., Chow, C.W., Shields, J.E., Cazorla, A., Kleissl, J. (2012): A method for cloud detection and opacity classification based on ground based sky imagery. Atmos. Meas. Tech., 5(11), 2881-2892.

González, Y., Lopez, C., Cuevas, E. (2012): Automatic observation of cloudiness: analysis of all sky images, TECO-2012, In: WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observation, Brussels, Belgium, 16–18 October 2012.

Heinle, A., Macke, A., Srivastav, A. (2010): Automatic cloud classification of whole sky images. Atmos. Meas. Tech., 3, 3, 557-567.

Holben, B.N., Eck, T.F., Slutsker, I., Tanré, D., Buis, J.P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y. J., Nakajima, T., Lavenu, F., Jankowiak, I., Smirnov A.

(1998): AERONET – A federated instrument network and data archive for aerosol characterization. Remote Sens. Environ., 66, 1–16, 1998.

Jenoptik (2013).CHM15k – Nimbus Ceilometer User Manual. Revision P0, September 2013.

Kazantzidis, A., Tzoumanikas, P., Bais, A.F., Fotopoulos, S., Economou, G. (2012): Cloud detection and classification with the use of whole-sky ground-based images. Atmos. Res., 113, 80-88.

Koehler, T.L., Johnson, R.W., Shields, J.E. (1991): Status of the whole sky imager database. In Proceedings of the Cloud Impacts on DOD Operations and Systems, 1991 Conference, 77-80.

Kreuter, A., Zangerl, M., Schwarzmann, M., Blumthaler, M. (2009): All-sky imaging: a simple, versatile system for atmospheric research. Appl. Optics, 48, 1091–1097.

Kylling, A., Stamnes, K., Tsay, S.C. (1995): A reliable and efficient two-stream algorithm for spherical radiative transfer: Documentation of accuracy in realistic layered media. J. Atmos. Chem., 21, 2, 115-150.

Liu, S., Zhang, L., Zhang, Z., Wang, C., Xiao, B. (2015): Automatic cloud detection for all-sky images using superpixel segmentation. IEEE Geosci. Remote S., 12, 2, 354-358.

Linfoot, A., Alliss, R.J. (2008): A cloud detection algorithm applied to a whole sky imager instrument using neural networks. American Meteorological Society, 88th Annual Meeting, Chantilly, Virginia, 1-13.

Long, C.M., Sabburg, J.M., Calbó, J., and Pagés, D. (2006): Retrieving cloud characteristics from ground-based daytime color all-sky images. J. Atmos. Ocean. Tech., 23, 633–652.

Martínez-Chico, M., Batlles, F.J., Bosch, J.L. (2011): Cloud classification in a Mediterranean location using radiation data and sky images. Energy, 36, 7, 4055-4062.

Martucci, G., Milroy, C., O'Dowd, C.D. (2010): Detection of cloud-base height using Jenoptik CHM15K and Vaisala CL31 ceilometers. J. Atmos. Ocean. Tech., 27, 2, 305-318.

Mittermaier, M. (2012): A critical assessment of surface cloud observations and their use for verifying cloud forecasts. Q. J. Roy. Meteorol. Soc., 138, 668, 1794-1807.

Mayer, B.,Kylling, A. (2005): Technical note: The libRadtran software package for radiative transfer calculations-description and examples of use. Atmos. Chem. Phys. 5, 7, 1855-1877.

Olmo, F. J., Cazorla, A., Alados-Arboledas, L., López-Álvarez, M. A., Hernández-Andrés, J., & Romero, J. (2008):Retrieval of the optical depth using an all-sky CCD camera. Appl.Optics, 47,34, H182-H189.

Reinhard, E., Stark, M., Shirley, P., Ferwerda, J. (2002). Photographic tone reproduction for digital images. ACM Transactions on Graphics (TOG), 21(3), 267-276.

Román, R., Antón, M., Cazorla, A., de Miguel, A., Olmo, F.J., Bilbao, J., Alados-Arboledas, L. (2012): Calibration of an all-sky camera for obtaining sky radiance at three wavelengths. Atmos. Meas. Tech., 5, 2013–2024.

Román, R., Bilbao, J., and de Miguel, A. (2014a): Reconstruction of six decades of daily total solar shortwave irradiation in the Iberian Peninsula using sunshine duration records. Atmos. Environ., 99, 41–50.

Román, R., Bilbao, J., and de Miguel, A. (2014b): Uncertainty and variability in satellite-based water vapor column, aerosol optical depth and Angström Exponent, and its effect on radiative transfer simulations in the Iberian Peninsula. Atmos. Environ., 89, 556–569.

Román, R., Torres, B., Fuertes, D., Cachorro, V.E., Dubovik, O., Toledano, C., Cazorla, A., Barreto, A., Bosch, J.L., Lapyonok, T., González, R., Goloub, P., Perrone, M.R., Olmo, F.J., de Frutos, A., Alados-Arboledas, L. (2017): Remote sensing of lunar aureole with a sky camera: Adding information in the nocturnal retrieval of aerosol properties with GRASP code. Remote Sens. Environ., 196,238-252.

Rossow, W.B., Schiffer, R.A. (1999): Advances in understanding clouds from ISCCP. B. Am.Meteorol. Soc., 80, 11, 2261.

Sánchez-Lorenzo, A., Calbó, J., Brunetti, M., Deser, C. (2009): Dimming/brightening over the Iberian Peninsula: Trends in sunshine duration and cloud cover and their relations with atmospheric circulation. J. Geophys. Res., 114, D00D09.

Stumpfel, J., Tchou, C., Jones, A., Hawkins, T., Wenger, A., Debevec, P. (2004): Direct HDR capture of the sun and sky. In Proceedings of the 3rd international conference on Computer graphics, virtual reality, visualisation and interaction in Africa, 145-149.

Tapakis, R., Charalambides, A.G. (2013): Equipment and methodologies for cloud detection and classification: A review. Sol. Energy, 95, 392-430.

Valenzuela, A., Olmo, F.J., Lyamani, H., Antón, M., Quirantes, A., Alados-Arboledas, L. (2012): Aerosol radiative forcing during African desert dust events (2005–2010) over Southeastern Spain. Atmos. Chem. Phys., 12, 21, 10331-10351.

Wacker, S., Gröbner, J., Zysset, C., Diener, L., Tzoumanikas, P., Kazantzidis, A., Vuilleumier, L.,Stöckli, R.,Nyeki, S., Kämpfer, N. (2015): Cloud observations in Switzerland using hemispherical sky cameras. J. Geophys. Res., 120, 2, 695-707.

WMO (2012): Recommended Methodsfor Evaluating Cloud andRelated Parameters. World Weather Research ProgrammeWWRP 2012 – 1, WWRP/WGNE Joint Working Group on Forecast Verification Research (JWGFVR), April 2012.

Yabuki, M., Shiobara, M., Nishinaka, K., Kuji, M. (2014): Development of a cloud detection method from whole-skycolor images. Polar Sci., 8, 4, 315-326.

Zhao, G.,DiGirolamo, L. (2006): Cloud fraction errors for trade windcumuli from EOS-Terra instruments. Geophys. Res. Lett., 33,L20802.

Tables

Table 1: Statistical parameters comparing the cloud cover obtained by various methods against the cloud cover visually measured by AEMet.

Location	Method	N	MBE (oktas)	Median (oktas)	MABE (oktas)	SD (oktas)

Granada	CPC	549	0.0	0.0	1.1	1.8
	RBR	549	0.6	0.1	1.6	2.3
	CEI	549	-0.1	0.0	1.1	1.7
Valladolid	CPC	469	-0.4	0.0	1.1	1.4
	RBR	469	-0.4	0.0	1.1	1.5
	CEI	469	0.2	0.0	1.0	1.5
All	CPC	1018	-0.2	0.0	1.1	1.6
	RBR	1018	0.2	0.0	1.4	2.1
	CEI	1018	0.0	0.0	1.0	1.6

Table B.1: Threshold values of the CPC algorithm. These values were obtained by the observations of multiple HDR images under different sky conditions.

Threshold	Condition	Threshold value
T_{sdr1}	All	1.05
T_{mrt}	All	0.48

T_{SRT1}	All	0.75 *mrt*
T_{sdr2}	All	1.18
T_{RT1}	Valladolid	0.48
T_{RT1}	Granada	0.65
T_{SRT2}	Valladolid	0.47
T_{SRT2}	Granada	0.64
T_{CI}	All	1.5%
T_{cir}	All	20%
T_{clo}	All	75%
T_{DSR1}	**DIS**>10º	15%
T_{DSR1}	**DIS**≤10º	20%
T_{DSR1}	**AZI**>178º	-1%
T_{DSR1}	**AZI**<-178º	-1%
T_{RT2}	Valladolid	0.9
T_{RT2}	Granada	1.18
T_{DSR2}	All	30%
T_{RT3}	All	1.1 *mrt*
T_{DR2}	**DIS**>10º	-1%
T_{DR2}	**DIS**≤10º	4%
T_{cns}	All	10%
T_{cis}	All	70%
T_{SDSR}	All	-15%
T_{RT4}	All	0.6

Figure captions

Figure 1: Direct sky image (left panels), tone mapped HDR sky image (right panels) for four cases at Granada: 15th May 2016 13:15 UTC (first row); 24th January 2016 13:25 UTC (second row); 15th June 2016 14:25 UTC (third row); 9th June 2016 08:40 UTC (fourth row).

Figure 2: Cloud cover detected by RBR method (left panels) and by CPC method (right panels) for four cases at Granada: 15th May 2016 13:15 UTC (first row); 24th January 2016 13:25 UTC (second row); 15th June 2016 14:25 UTC (third row); 9th June 2016 08:40 UTC (fourth row). White/blue pixels are considered cloudy/cloud-free, while yellow/white circle in the sun position means that sun has been identified as unobstructed/obstructed.

Figure 3: P_{sc} values at different SZA intervals for sun obstructed, unobstructed and all conditions, using the CPC (upper panel) and RBR (middle panel) methods. The number of data used under unobstructed and obstructed conditions is also given at different SZA intervals (bottom panel).

Figure 4: P_{sc} values at different AOD at 440 nm intervals for sun obstructed, unobstructed and all conditions, using the CPC (upper panel) and RBR (middle panel) methods. The number of data used under unobstructed and obstructed conditions is also given at different AOD at 440 nm intervals (bottom panel).

Figure 5: P_{sc} values, calculated only with data which AOD ranges from 0.1 to 0.2, at different AE intervals, for sun obstructed, unobstructed and all conditions, using the CPC (upper panel) and RBR (middle panel) methods. The number of data used under unobstructed and obstructed conditions is also given at different AE intervals (bottom panel).

Figure 6: Box plots for the distribution of ΔCC (using as reference the CC AEMet values) as a function of various CBH intervals for three methods: CPC (upper panel), RBR (second panel) and CEI (third panel). The box limits are the 25 and 75 percentiles, the error bar is the standard deviation, the circle is the mean, the red line inside the box is the median, the crosses are the 5 and 95 percentiles, and the triangles are the 1 and 99 percentiles. The number of data used is shown in the bottom panel for different CC intervals.

Figure 7: Box plots for the distribution of ΔCC (using as reference the CC AEMet values) as a function of CC (measured by AEMet) for three methods: CPC (upper panel), RBR (second panel) and CEI (third panel). The box limits are the 25 and 75 percentiles, the error bar is the standard deviation, the circle is the mean, the red line inside the box is the median, the crosses are the 5 and 95 percentiles, and the triangles

are the 1 and 99 percentiles. The number of data used is shown in the bottom panel for different CC intervals.

Figure 8: Box plots for the distribution of ΔCC (using as reference the CC AEMet values) as a function of AOD at 440 nm for three methods: CPC (upper panel), RBR (second panel) and CEI (third panel). The box limits are the 25 and 75 percentiles, the error bar is the standard deviation, the circle is the mean, the red line inside the box is the median, the crosses are the 5 and 95 percentiles, and the triangles are the 1 and 99 percentiles. The number of data used is shown in the bottom panel for different AOD at 440 nm intervals.

Figure 9: Box plots for the distribution of ΔCC (using as reference the CC AEMet values) as a function of AE for three methods: CPC (upper panel), RBR (second panel) and CEI (third panel). The box limits are the 25 and 75 percentiles, the error bar is the standard deviation, the circle is the mean, the red line inside the box is the median, the crosses are the 5 and 95 percentiles, and the triangles are the 1 and 99 percentiles. The number of data used, where only AOD from 0.1 to 0.2 were selected, is shown in the bottom panel for different AE intervals.

Figure B.1: Diagram of the CPC method. The three red nodes are potential candidates to be removed.