Real time automatic cloud detection using a low-cost sky camera

Joaquín Alonso-Montesinos[a,b]

[a]Department of Chemistry and Physics, University of Almería, 04120 Almería, Spain
[b]CIESOL, Joint Centre of the University of Almería-CIEMAT, 04120 Almería, Spain

Correspondence: Joaquín Alonso-Montesinos (joaquin.alonso@ual.es)

Abstract. Characterizing the atmosphere is one of the most complex studies to undertake due to the non-linearity and phenomenological variability. Clouds are also amongst the most variable of the atmospheric constituents, changing their size and shape over a short period of time. There are several sectors in which the study of cloudiness is of vital importance. In the renewable field, the increasing development of solar technology and the emerging trend for constructing and operating solar plants across the earth's surface requires very precise control systems that provide optimal energy production management. Likewise, airports are hubs where cloud coverage is required to provide high-precision periodic observations that inform airport operators about the state of the atmosphere. This work presents an autonomous cloud detection system, in real time, based on the digital image processing of a low-cost sky camera. The system's overall success rate is approximately 94% for all types of sky conditions.

1 Introduction

With the expansion in solar plant development [Chapman et al., 2018; Li et al., 2020,], comprehensive knowledge of the events that might affect plant production quality is required. Solar technologies involving energy transformation generally have inherent issues that must be overcome. Knowing when clouds will appear in the solar field is essential information for solar plant operators [Alonso and Batlles, 2014,]. With this knowledge, operators can perform a range of actions to optimize solar plant operation.

At airports, a daily meteorological report is provided showing the state of cloud cover; this is usually carried out by a human observer with experience in cloud visualization. It is therefore necessary to develop a real-time meteorological cloud detection system capable of repeatedly providing concise information on the state of the atmosphere.

Clouds have been detected using a wide variety of tools, one of which is satellite imagery. A simple classification method was developed based on the split-window technique. This system provided a detection accuracy of 44%, with an underestimation error of 56%, correctly classifying the areas in 88% of cases [Paszkuta et al., 2019,]. Another work presented two machine-learning methods to determine cloud masking using Spinning Enhanced Visible and Infrared Imager (SEVIRI) images that measured the reflectance values obtained from the IR3.9 channel. In general, the determination coefficient (r^2) presented results higher than 75% for the cases analyzed with MODIS (Moderate Resolution Imaging Spectroradiometer) and CLM (Cloud Mask product of EUMETSAT) images [Hadizadeh et al., 2019,]. Also, several authors have used Landsat 8 data to deal with the problem of detecting clouds in visible and multispectral imagery from high-resolution satellite cameras [Francis

Atmospheric
Measurement
Techniques
Discussions
Open Access

et al., 2019,]. Focusing on the importance of monitoring cloud cover over solar power plant areas, satellites have been used for cloud estimation, thus making it possible to track clouds and forecast their future position to predict when the sun will be blocked [Escrig et al., 2013,]. Other authors developed the Function of mask (Fmask) 4.0 algorithm for the automated cloud

30   and cloud shadow detection of Landsats 4-8 and Sentinel-2 images, in which the computational time needed to process a single Landsat or Sentinel-2 image generally takes between 1.5 and 10min (depending on the percentage of cloud cover and the land surface characteristics) [Qiu et al., 2019,]. However, space satellites have certain drawbacks, such as spatial and temporal resolution. Geostationary satellites provide images at a frequency of about 15 minutes and a temporal resolution of several square kilometers. Polar satellites have a higher spatial resolution, in the order of meters, but usually take only one image per

35   day. Added to this, the process for performing certain tasks with these images (matrix calculations) involves more time and may not be very efficient depending on the context in which the obtained information is applied.

Sky cameras are a way of providing a vision of the sky that complements the satellites, where clouds can be identified more accurately and at a higher temporal resolution. In some works, dust episodes have also been studied with this technology, as in the case of the Saharan dust intrusion over southern Spain in 2017. This appeared as though it were an overcast image that

40   consequently affected the Direct Normal Irradiance (DNI) [Alonso-Montesinos et al., 2017,]. The study of aerosol optical properties is another important field that uses sky cameras [Román et al., 2017b,]. Some authors have created their own sky cameras for cloud cover assessment without using conventional solar occulting devices [Fa et al., 2019,]. Furthermore, to counteract the temporal resolution limitations of satellite images, a digital camera was used for night-time cloud detection, detailing the percentage of cloud cover at 5-min intervals over the Manila Observatory [Gacal et al., 2018,]. In addition,

45   sky cameras have been used to record cloud detection from solar radiation data [Alonso et al., 2014a,] and then to predict the solar resource over the short term using digital levels and maximum cross-correlation [Alonso-Montesinos and Batlles, 2015; Alonso-Montesinos et al., 2015,]. Other authors have compared cloud detection data using satellite imagery and sky cameras [Luiz et al., 2018; Werkmeister et al., 2015,]; in [Luiz et al., 2018,], the comparison was made in south-eastern Brazil over a period of approximately three months. Good agreement was obtained for clear sky and overcast conditions,

50   with detection probabilities of 92.8% and 80.7%, respectively. For partially cloudy skies, the agreement was around 40%. In that article, the authors cited problems with the sky camera, for example, very bright areas around the sun, which were sometimes identified as clouds, leading to cloud cover overestimation. A similar observation was cited in [Román et al., 2017a,]. In [Werkmeister et al., 2015,] the observations took place in Germany and New Zealand, over time frames of 3 and 2 months, respectively. For a clear sky, the authors found detection probabilities of between 72 and 76%. They recommended

55   that more automated ground-based instruments (in the form of cloud cameras) should be installed as they cover larger areas of the sky than less automated ground-based versions. These cameras could be an invaluable supplement to SYNOP observation as they cover the same spectral wavelengths as the human eye. It is also common to use the R/B, R-B or (R-B)/(R+B) ratios to obtain better cloud characterization, where R is the red channel pixels and B the blue channel pixels [Yang et al., 2015,]. Nevertheless, these ratios lack precision when the image is processed, especially in the solar area where problems with pixel

60   brightness tend to overestimate the presence of clouds.

This work presents an automatic and autonomous cloud detection system using a low-cost sky camera (Mobotix Q24). The system mainly uses digital image levels and the solar height angle calculated at minute intervals. For each image, the system generates a processed image representing the original image; this is accomplished by identifying the clouds in white and the sky in blue.

65   The article is structured as follows: Section 2 presents the properties of the image acquisition system. Section 3 gives a step-by-step description of the methodology used to identify clouds (due to its size, it has been divided into subsections). Section 4 presents the most important results, both visual and numerical, after comparing this system with another consolidated cloud detection system. Finally, Section 5 presents the main conclusions.

2 Sky camera acquisition and optical properties

70   This work used images from a hemispheric low-cost sky camera, model Mobotix Q24, placed on the rooftop of the Solar Energy Research Center (CIESOL) at the University of Almería, Spain (36.8^oN, 2.4^oW; at sea level). Figure 1 shows the camera installation on the CIESOL building.

Figure 1. Mobotix Q24 camera installed in the University of Almería.

The facility has a Mediterranean climate with a high maritime aerosol presence. The images produced are high resolution from a fully digital color CMOS (2048 x 1536 pixels). One image is recorded every minute in JPEG format, the optimal time

75   to identify clouds in the sky. The three distinct channels represent red, green and blue levels. Each image pixel is made up of 8 bits, obtaining values between 0 and 255.

For this work, images were selected from all possible sky types, spanning the earliest times of the day to the latest, just before sunset, and at different times of the year. The period studied was from 2013 to 2019.

3 Methodology

80   This section describes the particular tasks undertaken to find the different clouds that can appear in a sky capture, defining the methodology and the steps involved to process an image. All developments have been carried out in the Matlab environment because it is an optimal platform to operate with matrices in record time. Its efficiency is the main reason we used this software for the methodology. To consider the starting point, one raw image is presented in Figure 2, in which the sky is represented as a common image.
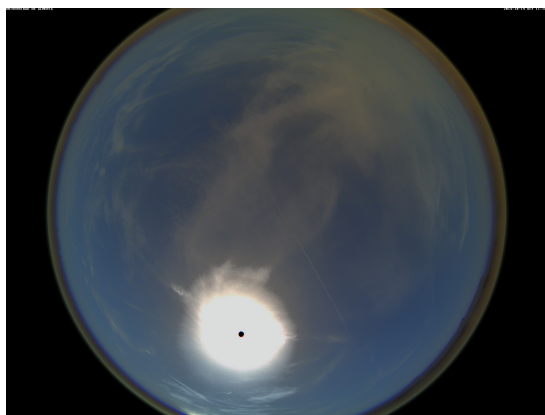

Figure 2. Sky image acquired from the low cost sky camera on 19 October 2014 at 11:18 (Universal Time Coordinated - UTC).

85   Here, we see the circular representation of the sky appearing over a black background. Therefore, the first step is to determine the area of interest from the raw image by applying a white mask, thus obtaining the image in Figure 3, where the sky is perfectly identified in the circular area.
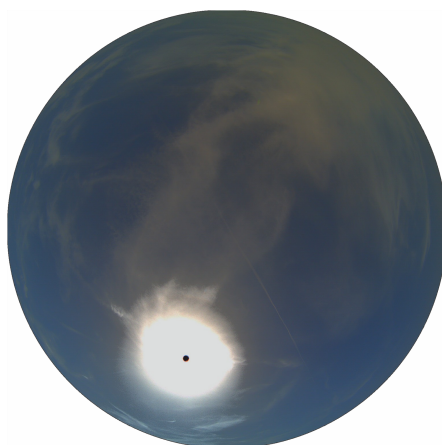

Figure 3. Application of a white mask to the raw sky image on 19 October 2014 at 11:18 (UTC).

After applying this customized mask, the image is ready to be used in the developed algorithm to detect clouds. The Matlab environment allows us to make changes in the color space of the images so that we can study specific properties. This is the case with HSV (Hue, Saturation, Value), NTSC and yCbCr color spaces. The first gives information about the gray color scale, where pixels vary between 0 and 1. The second is an RGB color space, where the first component, luminance, represents grayscale information and the last two components make up the chrominance (color information). The last one is used to digitally encode the color information in the computing systems: Y represents the brightness of the color, Cb is the blue component relative to the green component and Cr is the red component relative to the green component. Figure 4 shows the image presented beforehand in the three color spaces.
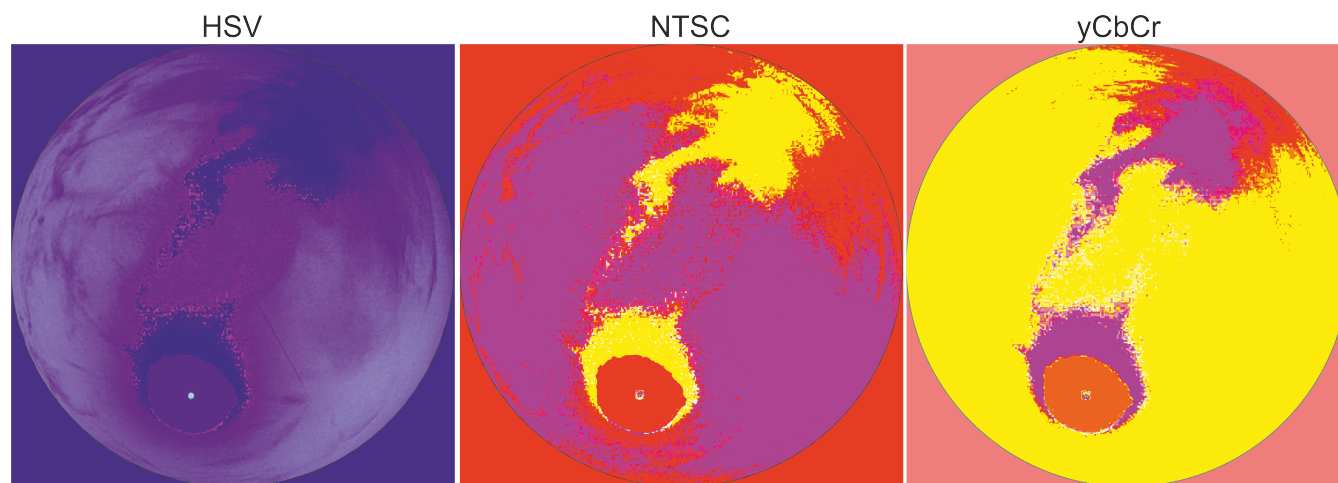

Figure 4. Representation of different color spaces for the image acquired from the low cost sky camera.

The different colors of the three images represent the main image characteristics. Focusing solely on the inside of the circle, the blue color identifies the more saturated areas of the HSV image. These areas are represented by red and yellow in the NTSC image, and by orange, rose and red in the yCbCr image. Normally, in the NTSC and yCbCr color spaces, the pixel values acquire inflexible static values in each color space channel whereas in the HSV color space, the pixels (represented by the three bands/channels) provide values with better precision for the purpose of cloud detection. Moreover, the representation of clouds is not perfectly represented in the three color spaces, so it is important to define the most significant color space to work with. Given that the HSV color space represents cloudy pixels better, and more clearly, it has been used together with the RGB color space to identify clouds in the image processing procedure. For the complete image processing procedure, the developed algorithm was structured into different parts, as described in the following subsections. Different tables appear to define specific criteria for image processing. Basically, each table has been made following the purpose of each subsection and fitting the intensity levels of the channels, obtaining a precise detection of the zones (clouds or sky). For that, different images have been analysed at different times of the day and times of the year. Afterwards, a general initial state has been assumed, to be precise in the adjustment of intensity values. Normally, the general state of each table starts analyzing the R value and the

comparison with G and B channels. If the intensities of these channels are enclosed in particular values, the HSV channels are
110 also analyzed. Therefore, this algorithm has been gradually formed by parts, finally remaining connected and sequential, in the
same order as it appears in the manuscript.

3.1   Recognition of the solar area: classification of pixels

The first step for detecting clouds in the whole sky image is to determine the solar area. Being able to recognize this area is
fundamental for establishing the Sun's position in the image. To track the angle of solar altitude each minute, the Cartesian
115 coordinates are obtained, with the south being represented by the center-bottom pixels and the east by the center-right pixels.
Subsequently, the original image (in JPG format), defined by the RGB color space, is also converted into the HSV color space.
As seen in the previous images, the sun appears as bright pixels, so one needs to consider the position of the pixels to determine
the bright solar pixels. To do this, after locating the sun pixel, a matrix is created to determine the distance of the other pixels
from it. This operation allows us to classify whether a bright pixel is a 'solar pixel' or not (based on its position). As a general
120 rule, when the value of the red, blue and green pixels is greater than 160, the pixels are identified as being in the sun area.
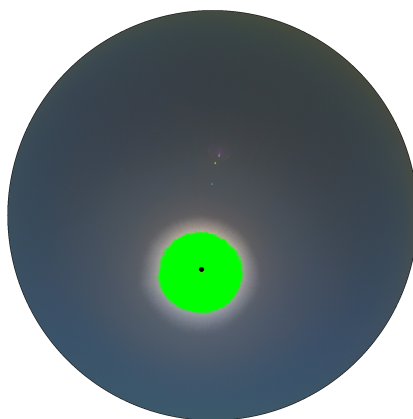Figure 5 shows the general detection of the sun pixels.

Figure 5. Identification of sun pixels applying a green mask.

The main step consists of applying a green mask to pixels that are placed in the sun area. After that, the idea is to detect if
these pixels are cloudless or overcast. Table 1 shows the rules for determining cloudless pixels in the solar area.

Different strategies are carried out to determine cloudless pixels in the sun area according to the pixel intensity in each image
125 channel. Figure 6 shows the general detection of cloudless pixels in the sun area (represented in red) after this filter has been
applied.

Subsequently, the algorithm looks for cloudy pixels in the same area in case some cloud are present. Table 2 shows the
condition for classifying the pixels in the solar area as cloudy.

Table 1. Criteria for selecting cloudless pixels in the solar area.

| Red (R)	Green (G)	Blue (B)	Hue (H)	Saturation (S)	Special Condition
[171,180]	-	-	0	0	R=G & G=B
[225,235]	-	R-2	(0.166,0.168)	(0.0070,0.0100)	R=G
≥ 200	-	R-2	(0.166,0.168)	(0.0070,0.0100)	R=G
228	-	R-2	(0.166,0.168)	(0.008,0.009)	R=G
≥ 230	-	-	> 0	-	R=G & G=B
≥ 210	-	-	≥ 0.416	≤ 0.005	-
[173,218]	-	-	[0.416,0.418]	(0.0112,0.1130)	Dis < 135
[173,178]	≥ 178	-	(0.166,0.168)	(0.0110,0.1120)	Dis ≥ 135
[172,177]	-	-	(0.166,0.168)	(0.0100,0.1113]	R=G
≥ 176	-	R-2	$\neq 0.167$	-	Dis < 135 & R=G
≥ 176	-	R-2	(0.166,0.168)	(0.0103,0.0107)	Dis < 135 & R=G
≥ 172	-	-	(0.416,0.418)	(0.0109,0.0113)	R=(G-2) & R=(B-1)
≥ 180	-	-	(0.166,0.168)	(0.0100,0.0110)	R=G & R=(B+2)
≥ 172	-	-	(0.416,0.418)	(0.0105,0.0107)	R=(G-2) & R=(B-1)
[172,185]	-	-	(0.166,0.168)	(0.0112,0.0113)	R=G & R=(B+2)
[170,174]	R+2	R+1	(0.416,0.418)	(0.0107,0.0109)	-
[190,194]	R+2	R+1	(0.166,0.168)	(0.0112,0.0113)	R=G & R=(B+2)
194	R+2	R-1	(0.277,0.279)	(0.0152,0.154)	-
≥ 225	≥ 225	≥ 225	0	0	$\alpha \leq 14$
≥ 200	≥ 200	≥ 200	[0.416,0.418]	≥ 0.0080	$\alpha \leq 14$
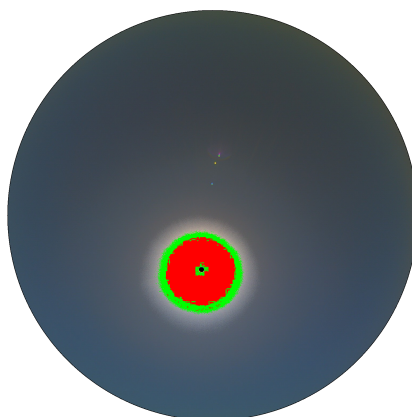
Figure 6. Determination of cloudless pixels in the solar area of the sky cam image.

Only one sentence is applied for detecting cloudy pixels in the solar area. In these situations, when a cloud is identified by means of a pixel, the mask applied is also green. When the solar area has been fully treated, the algorithm focuses on the rest of the image, starting with the solar area periphery.

Table 2. Criteria for selecting cloudy pixels in the solar area.

Red (R)	Green (G)	Blue (B)	Hue (H)	Saturation (S)	Special Condition
≥ 172	-	-	< 0.168	≥ 0.013	R=G

3.2 Detection of bright zones around the solar area

The pixels located around the solar area have an intermediate bright characteristic. In other words, the pixels present values lower than the solar area pixels and higher than those in the rest of the image. The size of this area varies according to the day and the atmospheric conditions at each moment. Table 3 shows the adjusted criteria for determining these pixels.

Table 3. Criteria for detecting bright pixels around the solar area.

Red (R)	Green (G)	Blue (B)	Hue (H)	Saturation (S)	Special Condition
> 125	-	-	-	-	(B-R > 8) & (B-4 > 4)
130	≥ 140	≥ 150	-	≥ 0.1900	Dis < 650
(130,140)	≥ 150	≥ 159	-	≥ 0.1900	Dis < 650
[140,150)	≥ 160	≥ 169	-	≥ 0.1900	Dis < 650
[150,165]	≥ 165	≥ 175	-	≥ 0.1900	Dis < 650
>130	-	-	-	-	(Dis < 650) & (B-R > 30)
>130	$\neq$ R	-	< 0.1	≥ 0.1000	(Dis < 650) & (R-G $\geq$ 8)
-	> 110	-	-	-	(Dis < 650) & (G > R) & (G > B) & (G-R < 8)
> 140	-	-	-	-	(Dis < 650) & (R-G $\leq$ 10) & (R-B $\geq$ 25)
($\neq$ 257) & ($\neq$ 260)	$\geq$ 130	-	[0.400,0.580]	$\leq$0.0800	(Dis < 650)
($\neq$ 257) & ($\neq$ 260)	-	-	$\leq$0.090	$\leq$0.0900	(Dis < 650)
($\neq$ 257) & ($\neq$ 260)	-	-	$\leq$0.090	$\leq$0.0900	(Dis < 650) & (G < R) & (R-G < 8) & (|G-B| < 3)
($\geq$100) & ($\neq$ 257) & ($\neq$ 260)	-	-	-	-	(Dis < 650) & (G-R > 20) & (B-R > 35)
($\geq$100) & ($\neq$ 257) & ($\neq$ 260)	-	-	-	-	(Dis < 650) & (G-R > 20) & (B-R > 35)

One of the most important tasks is to locate each pixel. The Dis variable almost always appears because the pixel emplacement is very important in this process. Therefore, to distinguish previously classified areas in subsequent processes, the yellow color is used to mark the new area (Fig. 7).

The new pixels classified as yellow do not represent a homogeneous area; they are dispersed across the image but at a distance of less than 650 units from the central solar pixel. In this new preprocessing, there are gaps between the yellow and green pixels that need to be classified beforehand. With the solar and surrounding area processed, the algorithm looks for cloudy pixels in the rest of the image.

3.3 Detection of cloudy pixels in the rest of the image

In general, clouds present several characteristics that allow us to identify the most common cloud types (white or extremely dark clouds). Table 4 shows the general pattern for detecting the clouds in the complete image by characterizing the digital levels of these common cloud types.
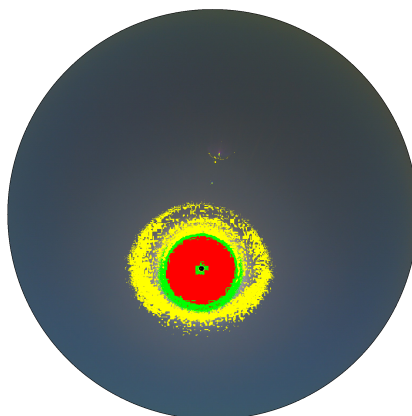
Atmospheric
Measurement
Techniques

Open Access

Discussions

EGU

Figure 7. Detection of the pixels with an intermediate value of bright around the solar area.

Table 4. Criteria for cloud detection excluding the solar area.

Red (R)	Green (G)	Blue (B)	Hue (H)	Saturation (S)	Special Condition
B	B	-	-	-	-
$\neq$ [259,261]	-	-	-	> 0.2500	(V $\neq$ 256) & (R $\geq$ B)
$\neq$ [259,261]	-	≥ 100	-	≤ 0.1800	(V $\neq$ 256) & (R/B $\leq$ 0.90)
$\neq$ [259,261]	-	-	≤ 0.200	≤ 0.2000	(V $\neq$ 256)
$\neq$ 45	-	≥ 80	< 0.600	≥ 0.3500	-

If a cloudy pixel is detected, it is marked in white. There are many cases in which some pixels are identified as cloudy although no clouds are present in the sky. This is caused by the similarity in the range of channel values, whereby dark skies can be confused for dark clouds. However, this mistake can be remedied during the algorithm's next steps. An example is presented in Figure 8, where a few pixels are classified as cloudy in white color.

Only a few pixels are classified as cloudy near the sun area. The first picture for this day showed no clouds appearing in the image so no cloudy pixels could be generated. In spite of this, a few pixels are interpreted as cloud. When solar area pixels and cloudy pixels are evaluated, the process continues to detect the pixels as unclassified.

3.4 Detection of cloudless pixels in the image excluding the solar area

After the solar area has been classified, the rest of the image is analyzed to identify if a pixel represents a cloud or not. Table 5 represents the set of sentences implemented to detect the cloud-free pixels in the parts of the image not including the solar area.

On can see that the Dis variable was not used even though we have presented the criteria to identify pixels that are outside the solar area. This is because, for cloudy pixels, the digital pixel levels never appear as in the range levels shown in the table. For this reason, it was not necessary to include the aforementioned variable in the sentences used (Fig. 9).
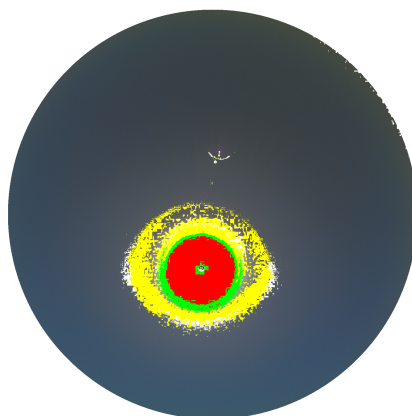
Figure 8. Detection of cloudy pixels in the total area of the sky cam.

Table 5. Criteria for detecting cloudless pixels in the image excepting in the solar area.

Red (R)	Green (G)	Blue (B)	Hue (H)	Saturation (S)	Special Condition
-	-	≤ 83	≥ 0.250	≥ 0.3000	-
≥ 55	-	≤ 83	≥ 0.580	≥ 0.5700	-
≤ 52	-	≥ 100	≥ 0.570	≥ 0.5700	-
≤ 50	≤ 60	≥ 83	≥ 0.610	-	-
-	-	-	≥ 0.500	≥ 0.4500	-
[50, 60]	[57,63]	[66,75]	[0.610,0.650]	[0.2000,0.3000]	-
≤ 70	-	≥ 75	≤ 0.700	≤ 0.3100	-
≥ 55	-	-	≥ 0.600	-	R/B ≤ 0.80
[50,68]	-	-	-	≥ 0.4600	R/B ≤ 0.80
> G	-	-	≥ 0.660	-	(R/B ≥ 0.94) & (G/B ≥ 0.88)
-	-	≥ 90	≥ 0.620	-	-
-	-	≥ 115	≥ 0.550	≥ 0.3000	-
-	-	-	≥ 0.800	-	-
[65,72]	[80,90]	≥ 108	-	-	-
[49,54]	[60,69]	[86,93]	-	-	-
G	-	-	-	-	B > G
≥ 100	≥ 110	≥ 130	-	-	-
-	B	> 70	-	-	R < B
[70,75]	-	[84,89]	-	-	G-R < 4
-	-	[90,100]	-	-	(G-R ϵ [6,7]) & (G-R ϵ [9,12])
-	-	-	-	-	(|R-B| ≤ 3) & (G>B) & (R/G ϵ [0.90,0.92])
[80,85]	[82,90]	[88,94]	-	-	(G-R ≤ 3) & (B-R ≤ 7)
[59,64]	[69,75]	-	-	-	(G-R ≥ 6) & (B-R ≥ 12)

In the image, the sun area and the surrounding area are processed, along with a small part of the remaining image. Therefore, at this point in the algorithm, it is possible that a large part of the image has still not been processed. Consequently, a further step is necessary to conclude the algorithm and classify all the pixels.
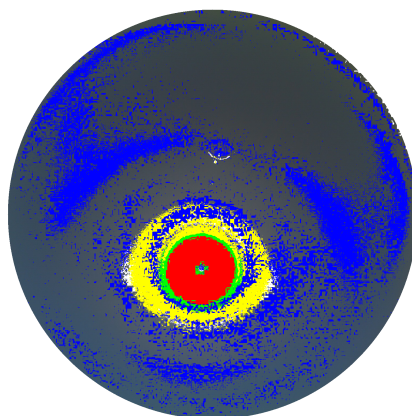
Figure 9. A representation of the sky cam image processing where the sun area and a part of the sky have been processed.

3.5 Determination of non-classified pixels

165   The final steps for classifying the pixels in a complete image establishes a statistical criterion which depends on the pixels that
have already been classified. Knowing the amount of pixels for each color, we determine those pixels that do not have a label.
To do this, there are different strategies for establishing the classification criteria for these, as yet, unclassified pixels. Table 6
shows the steps to determine whether the pixels should be classified as cloudless; if not, they will be classified as cloudy.

Table 6. Criteria for the determination of cloudless pixels for non-labelled pixels.

Main condition	Other conditions
$\alpha < 5$	(SkyPixels > 1000000) & (RedPixels $\geq$ 7000)
$\alpha \geq 5$	CloudPixels $\leq$ 25000
$\alpha \geq 5$	(CloudPixels > 25000) & (GreenPixels $\leq$ 1.6 RedPixels)
$\alpha \geq 5$	(CloudPixels > 25000) & (RedPixels $\geq$ 0.7 (GreenPixels + YellowPixles))
$\alpha \geq 5$	(CloudPixels > 25000) & (NonClass $\geq$ 1100000 & (RedPixels $\geq$ 2 GreenPixels) & (SkyPixels > CloudPixels))
$\alpha \geq 5$	(CloudPixels > 25000) & (NonClass $\geq$ 800000 & (SkyPixels $\geq$ 400000) & (CloudPixels $\leq$ 140000))
$\alpha \geq 5$	(CloudPixels > 25000) & (NonClass $\geq$ 600000 & (SkyPixels $\geq$ 700000) & (CloudPixels $\leq$ 120000))

In the table, different expressions appear. $SkyPixels$ are pixels that have been classified as cloudless, whereas $CloudPixels$
170   are those that have been labelled as cloudy. Red, green and yellow pixels have been obtained in the previous processes and
$NonClass$ is used to refer to the pixels that remain unclassified. For these operations, the Matlab environment allows us to
perform matrix operations in an efficient way. This part of the algorithm results in a matrix in which all the pixels have been
labelled, as shown in Figure 10.

As one can see, all the pixels have been assigned a color: blue, yellow, red or green. Now, the mission is to finish the
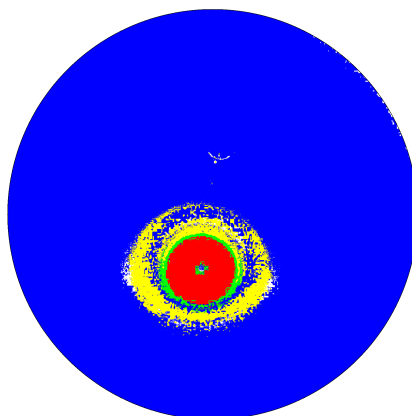175   classification process according to a common criterion.

Figure 10. Graphical representation of the identification of the pixels in the sky cam image.

3.6 Final step in the sky cam image classification

To finish the sky cam image processing, a final step is needed in which the differently colored pixels are converted to determine whether they are cloudless or cloudy. This process has been determined following the experience gathered working with a great number of images and scenarios. Table 7 shows the specific criterion for assigning the final pixel classification.

Table 7. Criteria for the last classification of pixels in the sky cam image processing.

Label	Classification conditions	Final classification
Green	(Dis $\geq$ 700)	Cloud
Yellow	(Dis $\geq$ 700)	Cloud
Green/Yellow	($\alpha \geq 5$) & (RedPixels $\geq$ 0.5 GreenPixels) & (GreenPixels > 25000)	Cloudless
Green/Yellow	($\alpha \geq 5$) & (RedPixels $\geq$ 0.3 GreenPixels) & (SkyPixels $\geq$ 1000000)	Cloudless
Green/Yellow	($\alpha \geq 5$) & (RedPixels $\geq$ 0.3 GreenPixels) & (RedPixels $\leq$ 15000)	Cloud
Green/Yellow	($\alpha \geq 5$) & (RedPixels $\geq$ 0.3 GreenPixels) & (CloudPixels $\geq$ 500000) & (RedPixels $\leq$ 0.5 GreenPixels)	Cloud
Green/Yellow	($\alpha \geq 5$) & (RedPixels $\geq$ 0.3 GreenPixels) & (CloudPixels $\geq$ 500000) & (GreenPixels $\leq$ 25000)	Cloud
Green/Yellow	($\alpha \geq 5$) & (RedPixels $\geq$ 0.3 GreenPixels) & (CloudPixels $\geq$ 500000) & (RedPixels ϵ [15000,30000])	Cloud
Green/Yellow	($\alpha \geq 5$) & (RedPixels $\geq$ 0.5 GreenPixels) & (GreenPixels > 25000)	Cloudless
Green/Yellow	($\alpha \geq 5$) & (RedPixels $\geq$ 0.3 GreenPixels) & (SkyPixels $\geq$ 1000000)	Cloudless
Green/Yellow	($\alpha \geq 5$) & (SkyPixels $\geq$ 1000000) & (RedPixels $\leq$ 15000)	Cloud
Green/Yellow	($\alpha \geq 5$) & (SkyPixels $\geq$ 1000000) & (CloudPixels $\geq$ 500000) & (RedPixels $\leq$ 0.5 GreenPixels)	Cloud
Green/Yellow	($\alpha \geq 5$) & (SkyPixels $\geq$ 1000000) & (CloudPixels $\geq$ 500000) & (GreenPixels $\leq$ 25000)	Cloud
Green/Yellow	($\alpha \geq 5$) & (SkyPixels $\geq$ 1000000) & (CloudPixels $\geq$ 500000) & (RedPixels ϵ [15000,30000])	Cloud
Green/Yellow	($\alpha < 5$) & (RedPixels $\geq$ 7000)	Cloudless
Green/Yellow	($\alpha < 5$) & (RedPixels $\geq$ 3 GreenPixels)	Cloudless
Green/Yellow	($\alpha < 5$) & (RedPixels $\geq$ 10000) & (SkyPixels $\geq$ 1000000)	Cloudless
Green/Yellow	Other-case	Cloud
Red	(Dis $\geq$ 700)	Cloud
Red	-	Cloudless

Following the assignations in the table, the final image can be generated, the result of which is shown in Figure 11.

The criteria presented in the above tables have been carefully defined, so, changes in the correlations represent alterations in the final image processed. Each criteria has an associated sensitivity according to the amount of pixels involved (the pixels that comply the criteria determined). Therefore, the sensitivity associated to each criteria affect the pixels that comply the condition
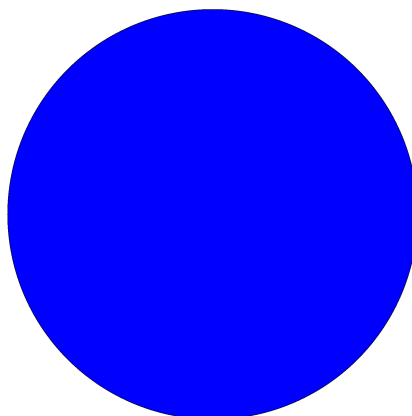
Atmospheric
Measurement
Techniques
Discussions

Open Access

EGU

Figure 11. Result of the sky cam image processing.

and, consequently, the final processed image. An error in one of the criteria presented in the tables would mean an error in the

185   cloud detection, and therefore the image would not have a valid processing and would be identified as wrong.

4   Results

In this section, we present the results of the cloud detection algorithm. In order to analyze the behavior of the software developed under different sky conditions, this section presents several pictures from various sky scenarios. A total of 850 images were taken from 2013 to 2019 at different times (from sunrise to sunset). The images were processed with the analytical objective

190   of obtaining an accurate identification of clear sky and clouds. Therefore, this section is divided into subsections as follow.

4.1   Sky images processed under all sky conditions

To analyze the quality of the developed model, several images have been processed and studied. In general, the processed image should represent the most important clouds appearing in the original image when visually inspected. Important clouds are those that can be identified clearly (not only by a few pixels). To see several examples, figures will represent the image

195   processing procedures carried out for the algorithm presented in the previous section. Figure 12 shows two examples chosen randomly, in which a clear and cloudless sky appear.

In the case of cloudless skies, the sun can vary its form and size depending on the solar altitude angle. The algorithm contemplates the solar altitude angle to identify clouds based on the variability in pixel intensity according to the sun position. Specifically, the sun is the key intensity point of the pixel value and the solar position determines a pathway for performing the

200   cloud recognition. Despite this, for the two cloudless days represented in the image, the sky is free of clouds, as represented by images with the letter H. As one can observe, for each original image (marked with the letter A), a sequence of images appears showing the steps the algorithm takes to finally obtain the image identifying the sky and cloudy pixels. In these cases, no cloud was detected and, therefore, the final images are completely blue.
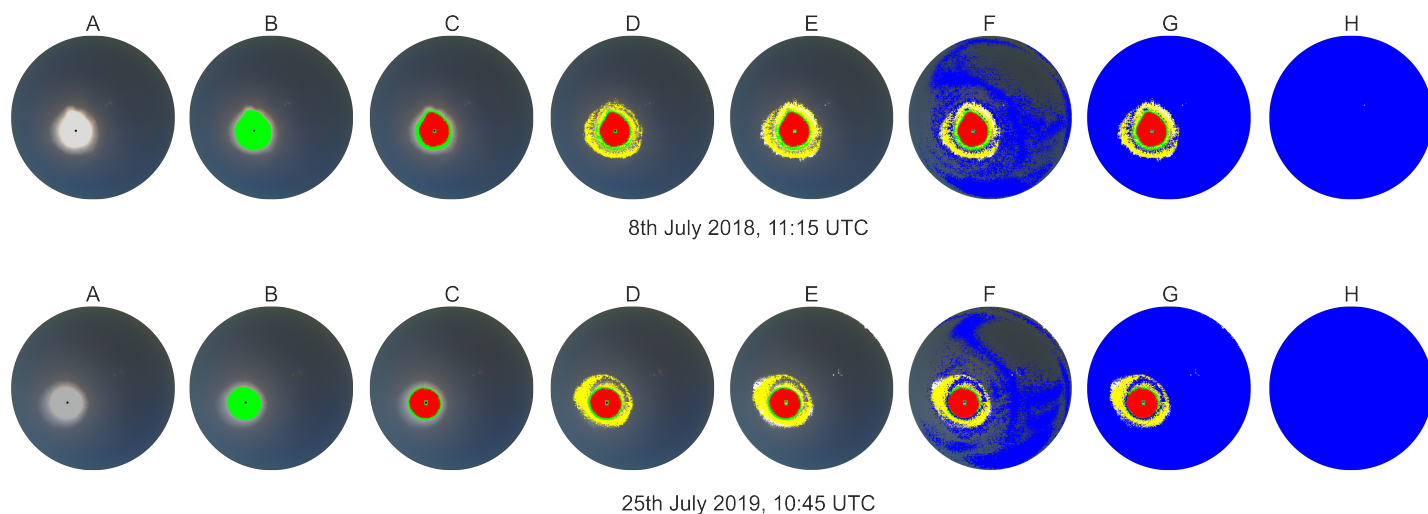
Figure 12. Result of the sky cam image processing under cloudless sky conditions.

When the sky is not completely free of clouds, one can have a sky that is either partially or completely covered. In the first
205 case, when observing a sky camera image, varying portions of the sky and cloud may appear. Therefore, it is important to
determine the boundaries between the cloud and the sky as effectively as possible. Figure 13 shows two different partially-
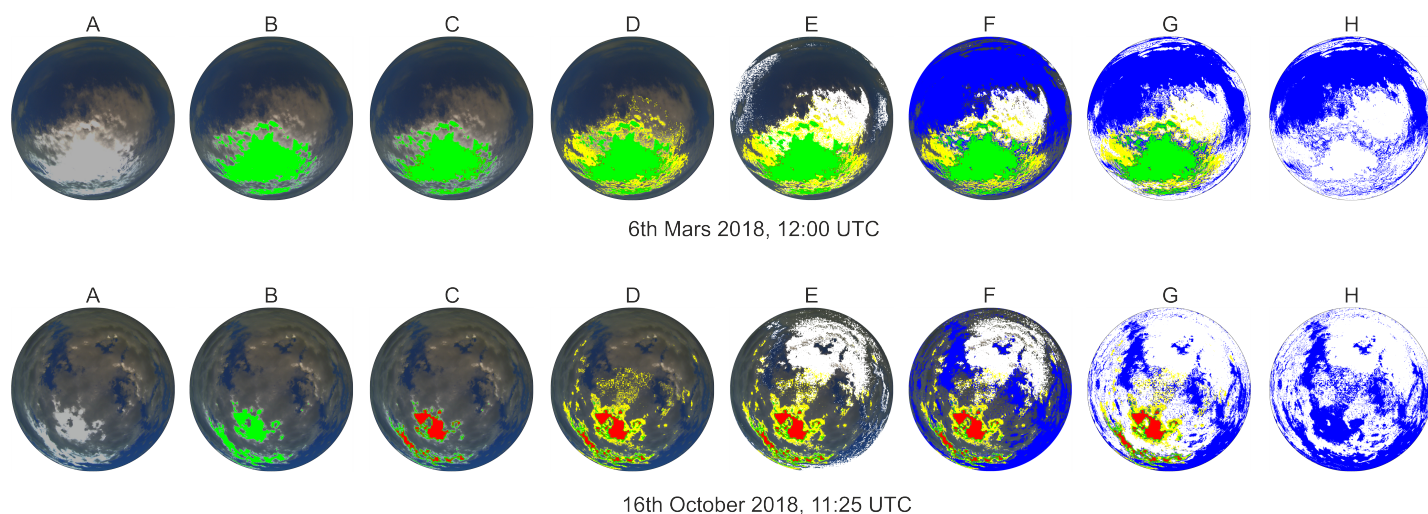cloudy situations.

Figure 13. Result of the sky cam image processing under partially-cloudy sky conditions.

Two scenes have been represented for cloud identification, with several differences between them. In the first sequence of
images (the top images), one can appreciate how most of the clouds are around the solar area. The algorithm recognizes, in
210 images B and C, large green and yellow areas. Subsequently, clouds are detected in the adjoining areas (image E) before finally

classifying those green and yellow pixels as a cloud. In the other sequence of images (the bottom line), the area detected in green turns red where the sun appears. This is because the algorithm's established criteria have not been met for identifying the pixels as a cloud; therefore, they are marked in red. Following this, the clouds are optimally detected in image E and blue sky is detected in image F. Finally, the image is resolved, classifying the pixels in red, green and yellow as cloudless. It is curious

215 how the clouds have generated significant brightness in the solar area, mistakenly classifying the solar area pixels as clouds.

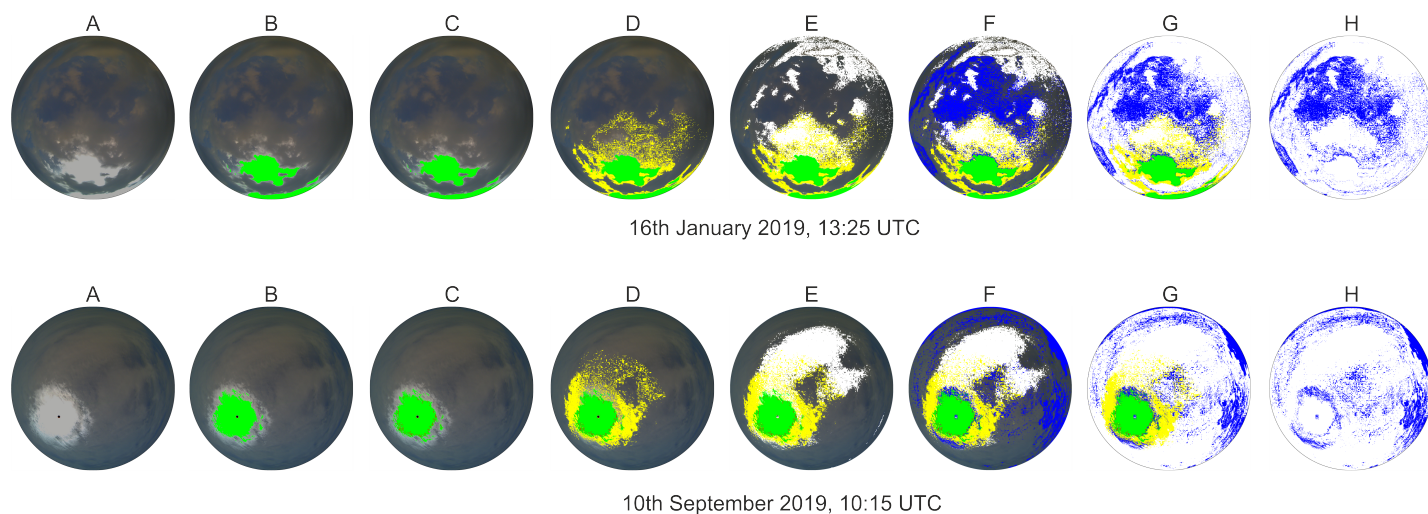Figure 14 shows two cases in which virtually the entire image is covered by clouds.

16th January 2019, 13:25 UTC

10th September 2019, 10:15 UTC

Figure 14. Result of the sky cam image processing under overcast sky conditions.

The top sequence of images shows a day when there was a lot of cloudiness and only small portions of sky. As the algorithm is executed, it is interesting to see how no red zone has been detected (attributing the solar area as cloudless) because the clouds in this case have a profile that is perfectly identified by the set of sentences presented in the previous tables. Here, the breaking

220 clouds have been correctly classified in image F. To conclude, image H shows the result of the process with the identification and classification of all the processed pixels being virtually identical to the original image. The bottom sequence of images shows another day with more clouds in the image, and as one can observe, again no red pixels have appeared. Following the steps described in the previous sections, the image processing very precisely determines the areas of blue sky and clouds (image H).

225 **4.2 Statistical results and comparison with TSI-880**

In order to make a statistical evaluation of the developed model's efficacy, we used a model that is already established and published [Alonso et al., 2014b, a,]; this works with images from a sky camera with a rotational shadow band ($TSI - 880$ model) installed on the CIESOL building, providing a hemispheric view of the sky (fish-eye vision). In short, the TSI-880 camera model is based on a sky classification using direct, diffuse and global radiation data. The sky is classified as clear,

Atmospheric
Measurement
Techniques
Discussions

Open Access

EGU

230 covered or partially covered. For each sky type, digital image processing is performed and an image is obtained with the cloud

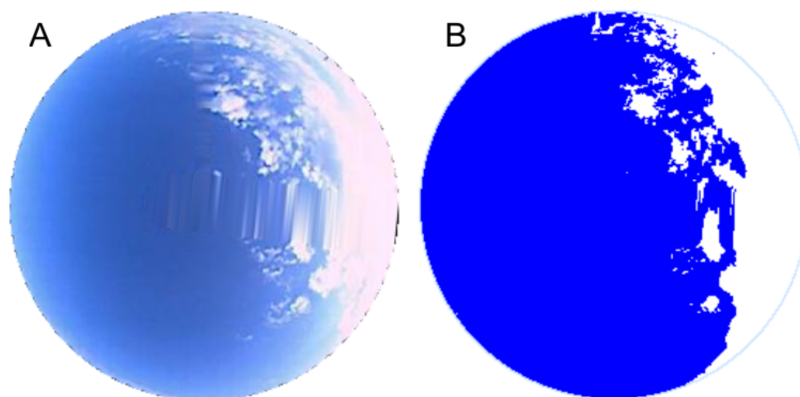identification. An example of a partially cloudy day is presented in Figure 15 at a time shortly after sunrise.

Figure 15. Image from TSI-880 on 31st March 2012, 7:15 UTC, where A is the original and B is the processed image.

Looking at the TSI-880 camera image, it is significant how the pixel intensities are very different from those represented by

the Mobotix camera. The blue sky generally appears lighter because of the camera's optics. For this reason, it is not possible

to extrapolate the development of the TSI-880 camera to the Mobotix Q24 camera.

235 Taking this model as a reference, a total of 845 images were taken between the months of August and December 2013 (by

that time, images from both cameras were available). To make the comparison, a routine was established that took an image

approximately once an hour (whenever possible) and to process the images of both cameras. In total, 419 images corresponded

to clear skies, 202 to partially covered skies and 224 to overcast skies. A probability function was used to determine the model's

success rate in percentage values. It was considered a success if the processed image represented the original image; that is, it

240 differentiated the areas of clear sky and clouds correctly. Equation (1) shows the probability function (PF):

$$PF(\%) = 100 \times \frac{Successes}{Total\ cases}, \tag{1}$$

In Equation 1, two variables are defined: Successes and Total cases. The first one is based on obtaining a processed image

with practically perfect cloud identification. Almost perfect means that the processed image adequately represents what appears

in the original image (either with the TSI-880 camera or with the Mobotix camera). Thus, a hit will be a final processed image

245 that is like the original. The total number of cases will be the total number of images analysed. The evaluation is done visually,

since there is no tool capable of detecting the difference between a raw and a processed image (that is the importance and value

of having a cloud detection algorithm). In this sense, the reference is always the original image. Any processed image should

resemble the original image.

Once the function was defined, Figure 16 shows the image processing efficacy by comparing the Motobotix Q24 camera to

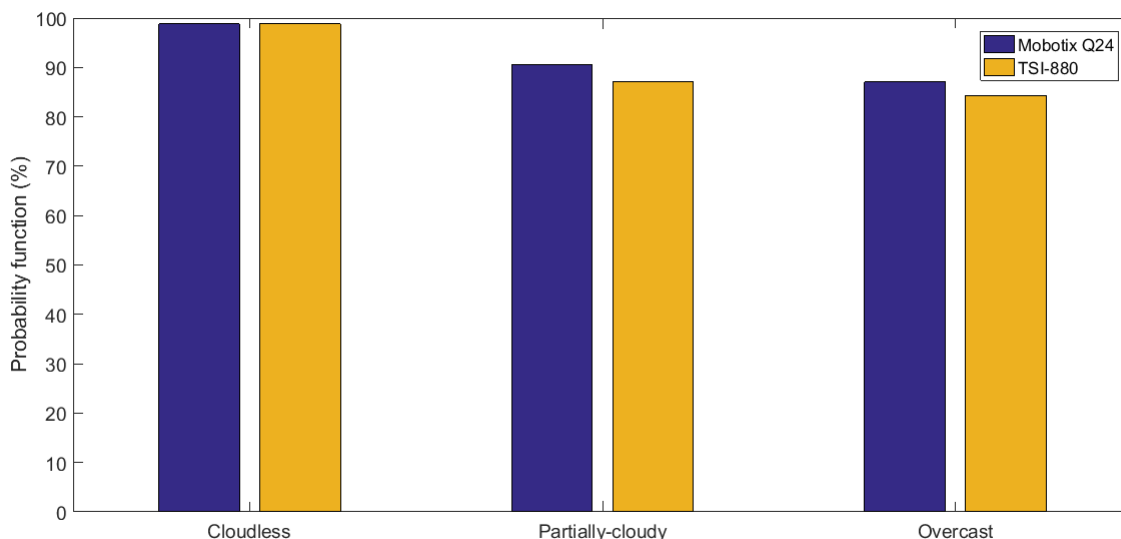250 the TSI-880 camera in terms of the sky classification.

16

Figure 16. Graphic success representation of image processing for the cloud detection by using different sky cams and algorithms, according to sky conditions.

The results have been divided into three basic groups: one representing the probabilistic results for clear skies, another for partially covered skies and the third for covered skies. In each group, a bar represents the success rate in (%). As can be seen, all the presented results are above 80%. The best results were obtained for clear skies, where the two cameras had the same success rate, at 98.8%. For partially covered skies, the Q24 camera provided better results than the TSI-880, with a success rate of 90.6%, while the TSI-880 success rate was 87.1%. In the case of overcast skies, the Mobotix camera again had a higher success rate, at 87.1%, while the TSI-880 camera had a value of 84.4%. In general, we can say that the cameras had a very similar success rate despite the slight differences found on days with clouds. Figure 17 shows a comparative graph of the overall hits in cloud detection.

As can be seen from the graph, the two cameras had very similar values overall; the cloud detection image processing for the Mobotix camera had a success rate of 93.7% while the TSI-880 camera had a value of 92.3%.

In addition, it should be noted that the TSI-880 camera requires a high level of maintenance to ensure the optimal quality of the images taken; this is due to its special design in which the glass is a rotating dome that must be cleaned periodically, taking care not to scratch it. Moreover, the glass is rotated by a motor that needs to be checked regularly in order to operate properly. In contrast, the Mobotix Q24 camera has dimensions similar to a surveillance camera, with only a small glass panel protecting the lens; this means that the maintenance requirements are reduced significantly. For our work, we can state that it was not necessary to clean the glass for several months. The device produced sharp, appropriate images allowing the algorithm to correctly identify the clouds present. Consequently, this article demonstrates that a new algorithm has been developed which
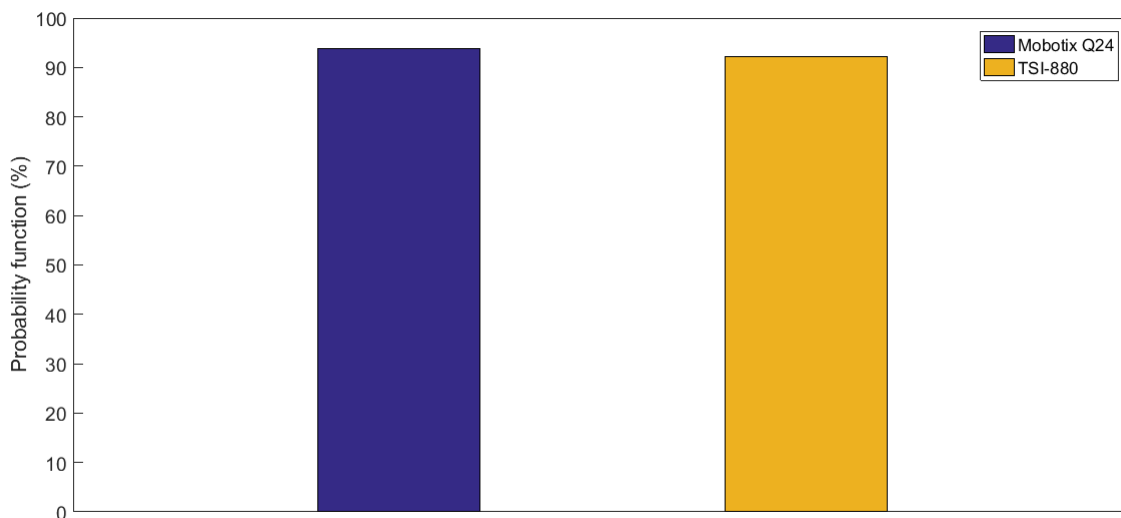
Atmospheric
Measurement
Techniques
Open Access

Discussions

EGU

Figure 17. Graphic success representation of image processing for the cloud detection by using different sky cams and algorithms, for all cases.

is capable of offering the same performance as the TSI-880 camera without needing radiation measurements to perform the digital image processing and requiring only minimal maintenance to acquire quality images for the cloud detection process.

5   Conclusions

This work presents a model to detect the cloudiness present in images in real time. It uses a low-cost Mobotix Q24 sky camera which only requires the digital levels of the image.

To detect clouds in the camera images, different areas of the image are differentiated. First, pixels are tagged in the solar area and the surrounding areas, assigning them with red, green or yellow colors. Subsequently, the algorithm detects cloudy pixels in the rest of the image and then clear sky pixels. Finally, the tagged pixels, such as cloud or sky, are classified, obtaining a final image that resembles the original.

The cloud detection system developed has been compared to a published and referenced system that is also based on digital image levels but uses a TSI-880 camera. In general, the results are very similar for both models. Under all sky conditions, the system developed with the Mobotix Q24 camera presented a higher success rate (93%) than the TSI-880 camera (around 92%). Under clear sky conditions, the processing of both cameras gave the same result (a 98% success rate). Under partially covered skies, the Mobotix camera performed better with a success rate higher than 90% (the TSI-880 success rate was 87%). Under overcast skies, the Mobotix camera had a success rate of 87% while the TSI camera's success rate was 3% less.

Atmospheric
Measurement
Techniques
Discussions
Open Access
EGU

One of the main advantages of the new system is that there is no need for direct, diffuse and global radiation data to perform the image processing (as is the case with the TSI-880); this greatly reduces costs as it makes cloud detection possible using sky cameras. Another major advantage is the minimal maintenance required to clean the camera, meaning the system is almost autonomous and can automatically obtain high quality images in which the clouds can be defined optimally.

With this method, a new system is presented which combines the digital channels from a very low-cost sky camera. It can be installed in the control panel of any solar plant or airport, of whatever type. The system represents a new development in predicting cloud cover and solar radiation over the short term.

Moreover, this new development opens the possibility to extrapolate the algorithm to another cameras. This tasks probably is not easy and direct, since each camera has its own optics, which makes it more difficult to adapt a custom-made algorithm for a camera with a determined optics. However, a novel idea has been appeared to adapt this new system to other cameras (with very slight modifications), to see if it is possible to obtain hits of the same range as for the Mobotix camera. The modifications will be necessary because each lens will have its own properties according to saturation levels, exposure, etc. Perhaps, it is possible to assume a correlation between the intensity of pixel channels for different technologies.

Atmospheric
Measurement
Techniques
Discussions

References

300   Alonso, J. and Batlles, F.: Short and medium-term cloudiness forecasting using remote sensing techniques and sky camera imagery, Energy, 73, 890–897, 2014.

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

Alonso, J., Batlles, F., Villarroel, C., Ayala, R., and Burgaleta, J.: Determination of the sun area in sky camera images using radiometric data,
305   Energy Conversion and Management, 78, 24–31, 2014b.

Alonso-Montesinos, J. and Batlles, F.: The use of a sky camera for solar radiation estimation based on digital image processing, Energy, 90, 377–386, 2015.

Alonso-Montesinos, J., Batlles, F., and Portillo, C.: Solar irradiance forecasting at one-minute intervals for different sky conditions using sky camera images, Energy Conversion and Management, 105, 1166–1177, 2015.

310   Alonso-Montesinos, J., Barbero, J., Polo, J., López, G., Ballestrín, J., and Batlles, F.: Impact of a Saharan dust intrusion over southern Spain on DNI estimation with sky cameras, Atmospheric Environment, 170, 279–289, 2017.

Chapman, A. J., McLellan, B. C., and Tezuka, T.: Prioritizing mitigation efforts considering co-benefits, equity and energy justice: Fossil fuel to renewable energy transition pathways, Applied Energy, 219, 187 – 198, 2018.

Escrig, H., Batlles, F., Alonso, J., Baena, F., Bosch, J., Salbidegoitia, I., and Burgaleta, J.: Cloud detection, classification and motion estima-
315   tion using geostationary satellite imagery for cloud cover forecast, Energy, 55, 853–859, 2013.

Fa, T., Xie, W., Wang, Y., and Xia, Y.: Development of an all-sky imaging system for cloud cover assessment, Applied Optics, 58, 5516–5524, 2019.

Francis, A., Sidiropoulos, P., and Muller, J.-P.: CloudFCN: Accurate and robust cloud detection for satellite imagery with deep learning, Remote Sensing, 11, 2019.

320   Gacal, G., Antioquia, C., and Lagrosas, N.: Trends of night-time hourly cloud-cover values over Manila Observatory: ground-based remote-sensing observations using a digital camera for 13 months, International Journal of Remote Sensing, 39, 7628–7642, 2018.

Hadizadeh, M., Rahnama, M., and Hesari, B.: Verification of two machine learning approaches for cloud masking based on reflectance of channel IR3.9 using Meteosat Second Generation over Middle East maritime, International Journal of Remote Sensing, 40, 8899–8913, 2019.

325   Li, H., Edwards, D., Hosseini, M., and Costin, G.: A review on renewable energy transition in Australia: An updated depiction, Journal of Cleaner Production, 242, 2020.

Luiz, E., Martins, F., Costa, R., and Pereira, E.: Comparison of methodologies for cloud cover estimation in Brazil - A case study, Energy for Sustainable Development, 43, 15–22, 2018.

Paszkuta, M., Zapadka, T., and Krężel, A.: Assessment of cloudiness for use in environmental marine research, International Journal of
330   Remote Sensing, 40, 9439–9459, 2019.

Qiu, S., Zhu, Z., and He, B.: Fmask 4.0: Improved cloud and cloud shadow detection in Landsats 4–8 and Sentinel-2 imagery, Remote Sensing of Environment, 231, 2019.

Román, R., Cazorla, A., Toledano, C., Olmo, F., Cachorro, V., de Frutos, A., and Alados-Arboledas, L.: Cloud cover detection combining high dynamic range sky images and ceilometer measurements, Atmospheric Research, 196, 224–236, 2017a.

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

Werkmeister, A., Lockhoff, M., Schrempf, M., Tohsing, K., Liley, B., and Seckmeyer, G.: Comparing satellite- to ground-based automated and manual cloud coverage observations - A case study, Atmospheric Measurement Techniques, 8, 2001–2015, 2015.

340 Yang, J., Min, Q., Lu, W., Yao, W., Ma, Y., Du, J., Lu, T., and Liu, G.: An automated cloud detection method based on the green channel of total-sky visible images, Atmospheric Measurement Techniques, 8, 4671–4679, 2015.