Received: 15 September 2023 | Revised: 17 June 2024 | Accepted: 23 July 2024

DOI: 10.1002/qj.4834

Quarterly Journal of the Royal Meteorological Society RMetS

RESEARCH ARTICLE

A neural network to retrieve cloud cover from all-sky cameras: A case of study over Antarctica

Daniel González-Fernández[1] | Roberto Román[1] | Juan Carlos Antuña-Sánchez[1,2] | Victoria E. Cachorro[1] | Gustavo Copes[3] | Sara Herrero-Anta[1] | Celia Herrero del Barrio[1] | África Barreto[1,4] | Ramiro González[1] | Ramón Ramos[4] | Patricia Martín[1] | David Mateos[1] | Carlos Toledano[1] | Abel Calle[1] | Ángel de Frutos[1]

[1]Group of Atmospheric Optics (GOA-UVa), Universidad de Valladolid, Valladolid, Spain

[2]GRASP-SAS, Villeneuve-d'Ascq, France

[3]Servicio Meteorológico Nacional, Buenos Aires, Argentina

[4]Izaña Atmospheric Research Center, Meteorological State Agency of Spain (AEMet), Tenerife, Spain

Correspondence
Daniel González-Fernández, Group of Atmospheric Optics (GOA-UVa), Universidad de Valladolid, 47011, Valladolid, Spain.
Email: daniel@goa.uva.es

Funding information
Ministerio de Ciencia e Innovación, Grant/Award Number: PID2021-127588OB-I00; MCIN/AEI/10.13039/501100011033 and European Union, "NextGenerationEU"/PRTR, Grant/Award Number: TED2021-131211B-I00; Junta de Castilla y León, Grant/Award Number: VA227P20

Abstract

We present a new model based on a convolutional neural network (CNN) to predict daytime cloud cover (CC) from sky images captured by all-sky cameras, which is called CNN-CC. A total of 49,016 daytime sky images, recorded at different Spanish locations (Valladolid, La Palma, and Izaña) from two different all-sky camera types, are manually classified into different CC (oktas) values by trained researchers. Subsequently, the images are randomly split into a training set and a test set to validate the model. The CC values predicted by the CNN-CC model are compared with the observations made by trained people on the test set, which serve as reference. The predicted CC values closely match the reference values within ±1 oktas in 99% of the cloud-free and overcast cases. Moreover, this percentage is above 93% for the rest of partially cloudy cases. The mean bias error (MBE) and standard deviation (SD) of the differences between the predicted and reference CC values are calculated, resulting in MBE = 0.007 oktas and SD = 0.674 oktas. The MBE and SD are also represented for different intervals of measured aerosol optical depth and Ångström exponent values, revealing that the performance of the CNN-CC model does not depend on aerosol load or size. Once the model is validated, the CC obtained from a set of images captured every 5 min, from January 2018 to March 2022, at the Antarctic station of Marambio (Argentina) is compared against direct field observations of CC (not from images) taken at this location, which is not used in the training process. As a result, the model slightly underestimates the observations with an MBE of −0.3 oktas. The retrieved data are analyzed in detail. The monthly and annual CC values are calculated. Overcast conditions are the most frequent, accounting for 46.5% of all observations throughout the year, rising to 64.5% in January. The annual mean CC value at this location is 5.5 oktas, with a standard deviation of approximately 3.1 oktas. A similar analysis is conducted, separating data by hours, but no significant diurnal cycles are observed except for some isolated months.

KEYWORDS

AI, all-sky camera, Antarctic, cloud cover, convolutional neural network, image identification

1 | INTRODUCTION

Clouds form when the water vapor condenses into liquid water droplets, ice crystals, or both that can become larger particles of rain or snow (Forster *et al.*, 2021). The study of clouds and their properties is very important for the understanding of climate change, as they affect the climate in different ways. They play a fundamental role in the Earth–atmosphere energy budget (Boucher *et al.*, 2013; Forster *et al.*, 2021). Clouds contain updrafts that can carry energy, moisture, momentum, trace gases, and aerosol particles with the air from near the surface to great heights (Boucher *et al.*, 2013). The precipitation that reaches the Earth's surface shows a net warming effect of the air, despite the fact that some condensed water evaporates again (Boucher *et al.*, 2013). Also, lower clouds tend to reflect solar radiation, producing a cooling effect, whereas higher clouds tend to absorb the solar energy and create a warming effect (Forster *et al.*, 2021). On average, clouds reflect more radiation than they absorb, which results in a net cooling effect (Boucher *et al.*, 2013; Forster *et al.*, 2021). Some studies show that Antarctic clouds play a big role in the climate system, both directly at the south latitudes and indirectly globally, highlighting the need for further research on them (Lachlan-Cope, 2010, and references cited therein). However, the study of clouds in polar regions, especially in Antarctica, is normally limited to a few locations, and there is a lack of information due to the isolation and the extreme climate of these areas.

The influence of clouds in the climate is related to their properties, so it is essential to understand them in order to have a better comprehension of cloud impacts. Some important properties of clouds are the cloud cover (CC), cloud optical depth, the thermodynamic phase, and some microphysical properties such as the effective radius of the droplets, the droplets' size distribution, their number, the volume concentration, or the water content (Peris, 2021). The CC, which is the topic of this article, is an important weather indicator that represents the fraction of covered sky by all the visible clouds according to the World Meteorological Organization (WMO, 2017). Clouds cover around two-thirds of our planet's surface, ranging from 60% to 90% from 60°S to more southerly latitudes (Warren *et al.*, 2015).

The CC can be determined in a high variety of ways. One method is through meteorological observers. These field observations are hemispheric quasi-instantaneous but subjective because of the differences between human observers. They are dependent on the visible horizon and are more reliable during the daytime. They also have a limited time resolution since they are manually carried out (WMO, 2012). Satellites can also be used to determine cloud properties (e.g., Arking & Childs, 1985; Rossow & Schiffer, 1999; Zhao & Di Girolamo, 2006). Polar satellites offer a great spatial resolution as they can take data around the whole Earth, but their main disadvantage is their low time resolution. Ground-based instruments, like lidar and ceilometers, which use a pulse laser emitted vertically into the atmosphere to measure backscattered signals, are used to measure CC and cloud base height (e.g., Costa-Surós *et al.*, 2013; Martucci *et al.*, 2010; Poyer, 2008). Other ground-based instruments, like radars, radiometers, pyranometers, or sunphotometers, are frequently utilized to retrieve cloud properties (Clothiaux *et al.*, 1995; Kollias *et al.*, 2007; Mateos *et al.*, 2014; Orsini *et al.*, 2002; Vasaras *et al.*, 2001). These instruments generally present a higher time resolution but a worse spatial resolution—(for further information, see Tapakis & Charalambides, 2013, and references cited therein). Other remarkable instruments used to detect and characterize cloud properties (like CC) are the all-sky cameras.

The all-sky cameras capture hemispheric pictures of the whole sky, typically in red, green, and blue channels. These sky images allow extraction of information about the sky conditions. There are a high variety of kinds of all-sky cameras: with/without narrow filters, with/without shadowband (or shadowball) to block the Sun, looking to the zenith with a fisheye lens or to nadir directly to a mirror located in the ground, and so on; but, in general, the main advantages of all these cameras are that they can take a high number of images in a short period of time, they are cheaper than the other kinds of instruments mentioned, they are automated, and they are relatively easy to install and maintain as they usually do not require user interaction.

There are different methods to determine the CC from the images captured by all-sky cameras. Several algorithms are based on the comparison between the registered signal at two different channels, with the most typical being the red/blue ratio (RBR; e.g., Calbó & Sabburg, 2008; Johnson *et al.*, 1989; Kreuter *et al.*, 2009; Shields *et al.*, 1998; Silva & Souza-Echer, 2016); in this case, a threshold is established for this ratio to distinguish between cloud-free (ratio below the threshold) and cloudy sky pixels (ratio above the threshold). Other algorithms are based on the difference between red and blue channels instead of the ratio (Kazantzidis *et al.*, 2012). However, these criteria present difficulties in detecting certain types of clouds, such as thin cirrus clouds that may fall below the threshold. Kim *et al.* (2016) introduced a dynamic threshold for the RBR, which depended on some statistical measures of the green/blue ratio such as average, skewness, and frequency distribution, among others. Other challenges could arise from a huge concentration of aerosols, like some strong desert dust episodes, or from the circumsolar area, where a great number of pixels could be erroneously identified

as clouds (Cazorla *et al.*, 2008; Huo & Lu, 2009; Román *et al.*, 2017). Among the proposed methods to solve these issues, Cazorla *et al.* (2015) applied an adaptive threshold depending on the blue signal and the distance of the pixel to the solar position for a better performance close to circumsolar area; Huo and Lu (2009) used the fast Fourier transform to the blue/red ratio images to identify homogeneous skies (totally cloud free or overcast) to apply an RBR threshold or inhomogeneous skies to apply an adaptive threshold based on the relative solar position; Ghonima *et al.* (2012) compared clear-sky simulated images with the cloudy sky measured images, which may exhibit several changes in the sky radiances due to aerosols; Román *et al.* (2017) considered the symmetry regarding the sky principal plane to identify cloud-free points even under heavy dust conditions (also Huo and Lu (2009)) and used filters with different sizes to detect high-altitude clouds like cirrus. There are also algorithms based on pixel segmentation to detect cloudy pixels in the sky pictures (Liu *et al.*, 2015). Some algorithms also explore the synergy between the sky images with other instruments like radiometers or ceilometers (Martínez-Chico *et al.*, 2011; Román *et al.*, 2017; Wacker *et al.*, 2015). For more information about the methods mentioned, see Román *et al.* (2017) and the references cited therein.

Recently, the machine- and deep-learning techniques to train neural networks have emerged, enabling applications that involve huge problems of estimation, detection, classification, and prediction; hence, these techniques can be applied to our problem on how to detect CC in a sky image. In fact, some studies have used these techniques before to produce accurate cloud masks, such as Cazorla *et al.* (2008) and Linfoot and Allis (2008). In particular, convolutional neural networks (CNNs; Fukushima, 1980; Krizhevsky *et al.*, 2012) have been shown to achieve high accuracy and efficient results (e.g., Onishi & Sugiyama, 2017; Masuda *et al.*, 2019; Xie *et al.*, 2020, and references cited therein). The recent advances in these CNN algorithms have made it easier to adapt them to tackle the challenge of CC detection. The way these algorithms operate, mimicking the function of a human brain, appears especially effective for identifying CC in episodes with high concentrations of large aerosol particles, like desert dust. Traditional CC algorithms often misinterpret desert dust as clouds in these scenarios, which motivates the need for developing a new model for CC prediction utilizing these innovative CNN techniques. Furthermore, CNN algorithms eliminate the need to set arbitrary thresholds, which usually vary depending on the camera's characteristics as well as on the image capture configuration and image post-processing settings. In this framework, the main purpose of this work consists in developing and validating a model, based on a CNN, capable of accurately and precisely determining the CC to apply it to any daytime image from any all-sky camera. The major difference of this model compared with most of those mentioned is that it does not need to calculate which pixels see clouds or not, but rather it will focus entirely on the value of the CC regardless of the clouds' position, being the first model of this type developed for hemispherical sky images. An additional goal is to apply this developed model to a set of images captured by an all-sky camera in Antarctica (Marambio) in order to quantify and analyze the CC and its behavior in this location.

This article is structured as follows: Section 2 presents the instrumentation, sites, and the data used. The model developed to obtain the CC is introduced in Section 3, as well as the comparison of the predicted data with independent CC values. Next, the analysis and results about the CC on the Marambio station are presented in Section 4. Finally, the main conclusions are summarized in Section 5.

2 | INSTRUMENTATION, SITES AND DATA

2.1 | All-sky cameras

Two models of all-sky cameras are used in this work. The main one is the OMEA-3C model (Alcor System), which consists of a Sony IMX178 CMOS sensor coupled to a fisheye lens of 180° × 180° of field of view, both encapsuled in a weatherproof case. This camera incorporates a BK7 glass dome on top, an external temperature and humidity sensor, and an internal heating system to avoid condensation and water droplets over the dome. The sensor works with a Bayer RGGB mosaic with an infrared filter (Antuña-Sánchez, 2022). This sensor takes pictures of size 3096 × 2080 pixels with 6.44 megapixels and 14-bits resolution. This all-sky camera is controlled by a computer to collect the data with different specifications; in our case, the GOA-OMEA Capture application, developed by the Group of Atmospheric Optics of the University of Valladolid (GOA-UVa), is used to this end. The OMEA-3C cameras used in this work were configured to capture, every 5 min during the daytime (except in La Palma; see Section 2.2), a sequence of consecutive raw sky images with different exposure times, which allows one to cover a higher dynamic range. The images of each sequence are combined to obtain a high dynamic range (HDR) image, which is converted to 8 bits with a resolution of 2000 × 2000 pixels covering the whole sky (Antuña-Sánchez *et al.*, 2022).

The other all-sky camera model used in this work is the SONA202-NF (Sieltec Canarias S.L.). This model is formed by a Sony IMX249 CMOS sensor with a fisheye lens of 185°

field of view, both enclosed in a weatherproof case with glass dome (Román *et al.*, 2022). This sensor has a Bayer mosaic filter with an RGGB pattern and a tri-band filter that significantly reduce the color-channel overlap and provides a narrower spectral response of the color channels (Antuña-Sánchez *et al.*, 2021). SONA202-NF takes images of 1172 × 1158 pixels, with 2.35 megapixels and 10-bits resolution. This camera operates with a computer via a web interface, which allows the configuration of the capture options for the pictures. This camera was configured for daytime in a similar way to the OMEA-3C: capturing every 5 min a sequence of raw images with different exposure times; an 8-bits HDR image is also obtained for each sequence (Antuña-Sánchez *et al.*, 2022).

2.2 | Sites

The sky images used in this work were captured at four sites. The principal place is the Valladolid station (41.66°N, 4.71°W, 705 m a.s.l.), managed by the GOA-UVa and located on the rooftop of the Science Faculty of the University of Valladolid (Valladolid, Spain). Valladolid is a medium-size city, located in the north-center of Spain, with approximately 300,000 inhabitants, and 400,000 in the metropolitan area. It is an urban city surrounded by rural areas and has a climate classified as clean continental, with hot summers, cold winters, and occasional Saharan dust episodes (Bennouna *et al.*, 2013; Cachorro *et al.*, 2016; Román *et al.*, 2014). At this station there are various all-sky cameras installed and also other instrumentation, such as sunphotometers, a ceilometer, and various radiometers. A SONA202-NF was installed on July 11, 2018, and was working until May 29, 2022, at this site (Román *et al.*, 2022). An OMEA-3C camera was also operating since July 16, 2020, until September 27, 2021.

This camera was replaced by another OMEA-3C in Valladolid, because it was sent to the Cumbre Vieja natural park on La Palma (Canary Islands, Spain) as a consequence of the Tajogaite volcano eruption (Bedoya-Velásquez *et al.*, 2022; Salgueiro *et al.*, 2023). This camera installation was part of an unprecedented deployment of instrumentation for the monitoring of the event which was coordinated by the Meteorological State Agency of Spain (AEMET) (Córdoba-Jabonero *et al.*, 2023; García *et al.*, 2022; Milford *et al.*, 2023; Sicard *et al.*, 2022). The camera was installed in the Fuencaliente station of La Palma (Spain; 28.48°N, 17.84°W, 630 m a.s.l.) from October 6, 2021, to January 24, 2022. This camera was monitoring the clouds and aerosol emission of the eruption. In this case the camera was configured to take daytime HDR images every 2 min for a higher time resolution. La Palma is an island with an estimated population of 85,000 inhabitants, with a subtropical climate that maintain smooth temperatures all the year.

After the monitoring of the eruption, on February 4, 2022, the same camera was installed at the meteorological observatory of Izaña (Izaña Atmospheric Research Center; 28.30°N, 16.49°W, 2,400 m a.s.l.), also managed by AEMet. This high-altitude station is located in Tenerife (Canary Islands, Spain), near the Teide peak. The Izaña observatory is usually above a strong temperature inversion layer and consequently free of local anthropogenic influence with pristine and stable conditions (Barreto *et al.*, 2022; Cuevas *et al.*, 2022; Román *et al.*, 2020; Toledano *et al.*, 2018).

GOA-UVa not only manages the Valladolid station but also has instrumentation (such as sunphotometers and all-sky cameras) in other locations, including polar regions. One of these polar stations is Marambio (Argentina) in Antarctica (64.24°S, 56.52°W, 200 m a.s.l.). One OMEA-3C camera is installed in Marambio by GOA-UVa in collaboration with the National Meteorological Service of Argentina since January 26, 2018. Unfortunately, the images from this station were not captured with the GOA-OMEA Capture software until August 2020; hence, the images recorded at this station before this date have not been used in this work. Figure 1 shows a map where the four stations mentioned are marked to facilitate their location.

2.3 | Data

Some of the HDR sky images recorded in daytime, with both all-sky camera models (OMEA-3C and SONA202-NF) at the three stations mentioned (Valladolid, La Palma, and Izaña), have been classified according to the CC (in oktas) observed in each image. This task was carried out by three different human observers who were previously trained following the World Meteorological Organization guidelines (WMO, 2012). Although the observation and coding criteria should remain consistent across observers, the CC classification of the images could differ between observers since different practices are often applied in this task, with differences within ±1 oktas expected. The classified images for Valladolid are from January to June 2019 for the SONA202-NF camera and for some days of August and December 2020 for OMEA-3C. In the case of La Palma station, the classified images go from October 2021 to January 2022, and from February to March 2022 for Izaña. As a result, a total of 49,016 images have been classified by their CC. The number of images classified for each station, camera, and CC is given in Table 1, La Palma station being the one with the largest set of classified images. The most frequent

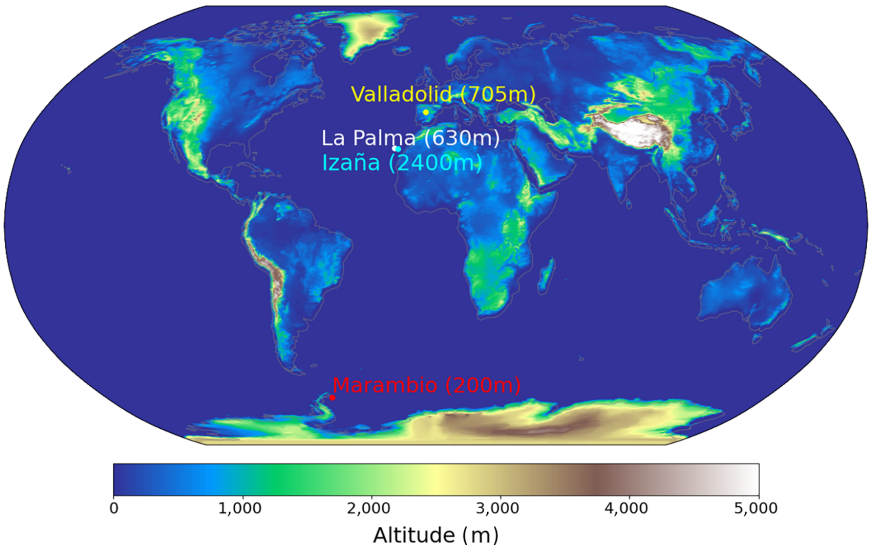

FIGURE 1 World map showing the location and the altitude of the four sites where the all-sky images were captured in this work. The color map shows the altitude in meters, in ranges of [−180,180] degrees of longitude and [−89, 89] degrees of latitude, with a resolution of 0.25° in latitude and longitude, obtained from http://research.jisao.washington.edu/data_sets/elevation/. [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 1 Number of sky images classified by human observers according to their cloud cover for different sites and cloud conditions.

Site (camera)	0 oktas	1 oktas	2 oktas	3 oktas	4 oktas	5 oktas	6 oktas	7 oktas	8 oktas	Total
Valladolid										
(SONA202-NF)	2283	1527	912	810	799	751	862	1195	3136	12257
(OMEA-3C)	503	245	95	175	91	111	140	363	382	2105
La Palma	4667	4634	2781	1963	1609	2235	2590	4119	6931	31529
Izaña	1480	326	184	97	64	112	63	106	675	3107
Total	8933	6732	3972	3045	2563	3209	3655	5783	11124	49016

conditions are cloud-free and overcast conditions, with Izaña being the site with the highest proportion of cloud-free images.

This work also uses independent CC field observations visually measured by a trained weather observer at the Marambio station. These field observations were recorded every hour, and the CC data series is available for all the period analyzed here with all sky images from January 1, 2018, to March 10, 2023. These CC observations are typically made every hour, and approximately ±10 around the o'clock hour with zero minutes. In addition to these hourly CC field observations in oktas, these field observations also contain the cloud codification for different sky types (low clouds, mid clouds, and high clouds), based on the synop key of the *WMO International Atlas of Clouds* (WMO, 2019). The annual turnover of personnel is added. The personnel responsible for these observations is replaced every year (typical of Antarctic stations), usually in October for the Marambio station. Although the observation and coding criteria should remain consistent across observers, different practices are often applied due to usage and custom, such as changes in the observation or coding time and data transmission.

3 | METHODS

The proposed model to estimate the CC is described in this section. The accuracy and precision of this model is validated against a test set of images reserved for this task. In addition, the CC estimated by the model is also compared with independent CC field observations carried out in the Marambio station.

3.1 | CNN architecture and training

The model chosen to estimate CC from the daytime sky images was a CNN, which has been named CNN-CC. The architecture of the proposed CNN is shown in Figure 2, which has been designed using Keras (Chollet, 2015), a Python application programming interface of the machine-learning platform TensorFlow (Abadi *et al.*, 2015). The CNN has been configured to receive as input HDR sky images with a size of 128 × 128 × 3 pixels, where the third dimension (size = 3) represents the image color (red, green, and blue). Therefore, all images need to be resized to these dimensions before running the model.

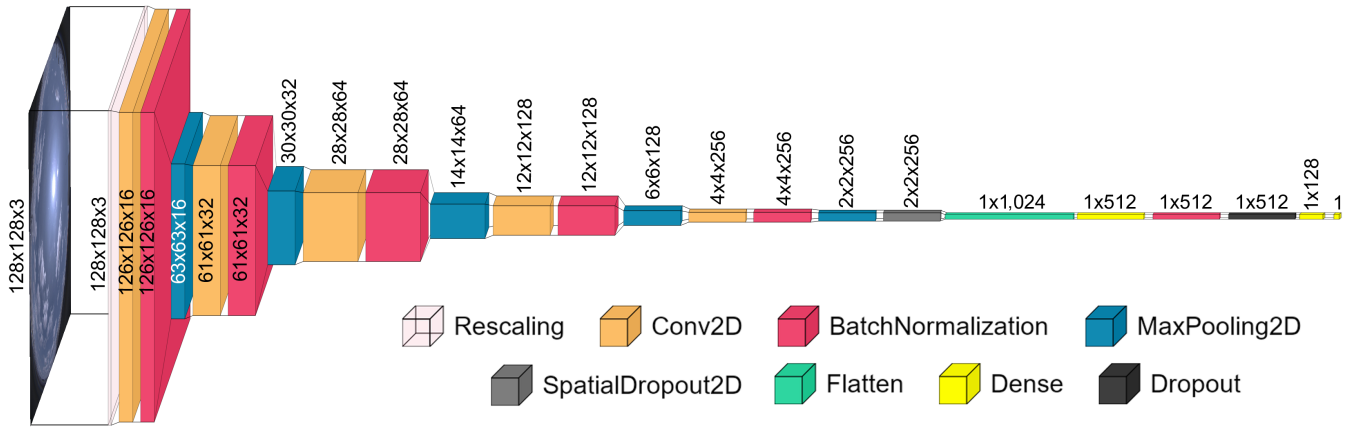

FIGURE 2 Scheme of the architecture of the proposed convolutional neuronal network. This scheme has been drawn using the VisualKeras tool (Gavrikov, 2020). [Colour figure can be viewed at wileyonlinelibrary.com]

The first layer of the CNN is a rescaling layer, which divides the 8-bits signal of each pixel of the input image by 255 in order to rescale the signal to the 0–1 interval. The next three layers are as follows: (1) a two-dimensional convolutional layer (Conv2D) with 16 filters and a kernel size of 3×3, where the rectified linear unit (ReLU) function is used as activation function and the convolution is not zero-padded, which implies a 2 pixels reduction in width and height after the convolution; (2) a normalization layer (BatchNormalization), which applies a transformation that maintains the mean output close to zero and the output standard deviation close to one (Ioffe & Szegedy, 2015); and (3) a max-pooling layer (MaxPooling2D), which downsamples the input along its spatial dimensions by taking the maximum value over a window of 2×2 pixels. These three layers appear consecutively connected four times more, doubling the number of filters of the convolution layer for each time. This process reduces the height and width of the filtered sky images in order to extract feature maps of the image.

After the last max-pooling layer, which provides 256 images with a 2×2 size, a two-dimensional dropout layer (SpatialDropout2D) is added, but it is only applied during the CNN training, dropping randomly the 20% of the entire two-dimensional feature maps (Tompson *et al.*, 2015). Then, once the feature maps are obtained, the output of the CNN at this point is resized to one dimension, in this case to 1,024 values, with another layer (*Flatten*) to start a fully connected network. The next layer is a regular densely connected layer (Dense), which provides an output of 512 values using ReLU as the activation function. This output is also normalized by a BatchNormalization layer, which is directly connected with another dropout layer (Dropout) that drops randomly the 20% of the values that reach the layer during the training. Finally, the CNN has two densely connected layers: the first one with 128 nodes and a ReLU activation function; the second and final layer with a linear activation function and only one output node, which provides a float number representing the CC of the sky image used as input. This final result is rounded to an integer between 0 and 8 oktas in order to have a proper value similar to that used in the classification. This CNN has 985,217 trainable parameters and 2016 non-trainable parameters.

This CNN has been trained with part of the classified sky images shown in Table 1. The other part of images has been used to test the CNN performance. Owing to the high difference between the number of available images between the different stations, the images have been separated into two new sets: Valladolid images (both cameras) and Canary Islands images (La Palma and Izaña). For each of the nine CC values, the minimum number of available images of both new images sets has been calculated, and 80% of this minimum value is the number of images that have been randomly selected from each of the new two sets (Valladolid and Canary Islands) to be part of the training set. This process has been done considering that at least the 20% of images of each site (or camera) must be part of the test set instead of the training set. The images not added in the training set have been included in the test set for validation. This split guarantees a diversity of images with different CCs, cameras, and sites in the validation with the test images set. Finally, the test set has been formed by 26,112 daytime sky images and the training set by 22,886 images divided into 9756 (Valladolid; SONA-202NF), 1662 (Valladolid; OMEA-3C), 10,454 (La Palma), and 1014 (Izaña) images. These data point out that the training set comprises a similar number of images from Valladolid and the Canary Islands, but also includes a high proportion of SONA-202NF images, with 43% of the training images being from this camera model.

The training set has also been randomly split for the CNN training, with 90% of the images used directly for

the CNN training and the other 10% to evaluate the performance of the CNN in each epoch during the training (the validation set during the training). The images of the training set have been randomly modified during the training to achieve more variety of images in a similar way as in data augmentation techniques; these modifications were image flipping (RandomFlip), rotation (RandomRotation), zoom (RandomZoom) and translation (RandomTranslation). This also helps to avoid temporal correlation between the images used in training and the images in the test set.

The loss function chosen in the training was the mean-square error (mse). This function has been used instead of others, like the categorical cross-entropy, since we consider the model must penalize more the cases when the predicted CC values are farther from the labeled reference than the cases when predicted and reference values are closer even when they are not the same. For example, if the CNN predicts 7 oktas in one image of the training set and the original labeled value of this image is 2 oktas, it must be more penalized than if the predicted value for the same reference is 3 oktas, since this last situation is closer to the reality. The training has been done using batches, each one formed by 32 training images (batch size). The optimizer used in the training was the Adam algorithm (adaptive moment estimation; Kingma & Ba, 2014) with an initial learning rate of 0.001. This learning rate was dynamically reduced (divided by 10) when the loss metric of the validation set (the reserved 10% of the images) was not improved in five consecutive epochs; this was computed with Keras using the ReduceLROnPlateau callback. The training was configured to stop when the loss value of the validation set was not improved after 10 consecutive epochs (early_stopping callback), trying to avoid the overfitting. With this configuration the CNN has been trained, reaching a total of 46 epochs before stopping. The version of the trained CNN chosen is the one corresponding to epoch 36, which presented the lowest loss value, mse = 0.554 oktas^2, in the validation set.

3.2 | Validation

3.2.1 | Test set

The set of images reserved for test (not used for training) has been used to evaluate the performance of the CNN model developed for CC prediction, and to quantify its accuracy and precision. The model was run for each one of the daytime images of this set to calculate the CC predicted by the model. These predicted CC values were compared with the reference ones given by the classified images.

Figure 3 shows the number of cases that the model predicts a CC value (0–8 oktas) for the different reference CC values as a confusion matrix using all available test data. In this figure, the percentage of cases for each one of the reference values is also written and shown as a color map. The majority of matching cases are concentrated along the diagonal, suggesting a strong correlation between the predicted CC values and the reference ones. A total of 68.4% of the predicted data exactly match the reference values. The same percentages but using human

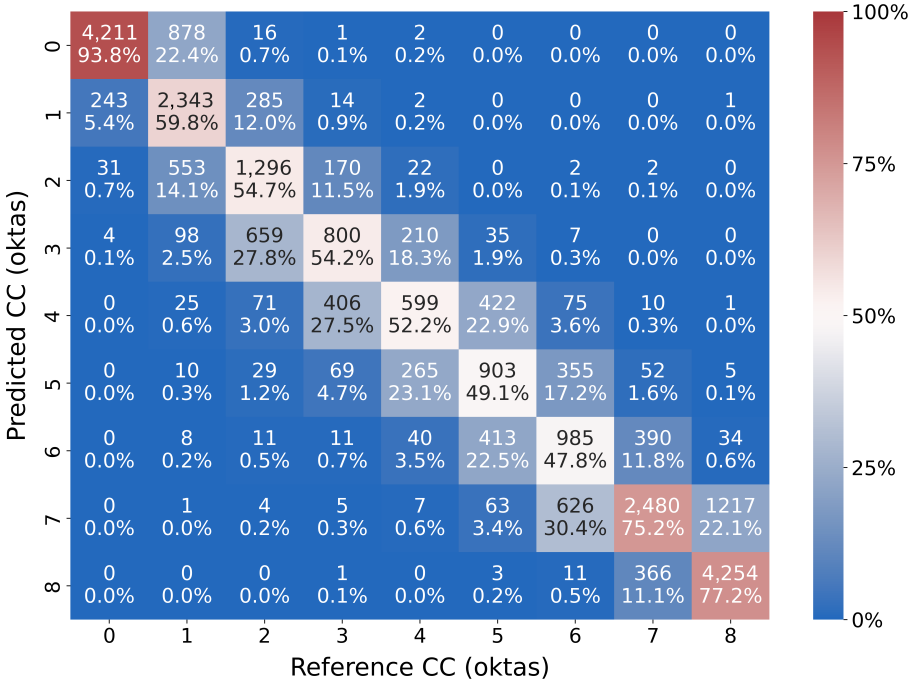

FIGURE 3 Confusion matrix with the number of cases matching for different predicted and reference cloud cover (CC) values. The reference values are the CC of classified images by trained staff in the selected set of test images. Color map represents the percentage of data considering the total as the sum of the number of data with the same reference CC value. This percentage is also written in each cell. [Colour figure can be viewed at wileyonlinelibrary.com]

field observations as reference were as follows: 41% and 49% at Payerne and Jungfraujoch (Switzerland) respectively with the algorithm of Wacker *et al.* (2015); about 60% in Thessaloniki (Greece) by Kazantzidis *et al.* (2012); about 33% in Innsbruck (Austria) as calculated by Kreuter *et al.* (2009); about 34% with the method of Cazorla *et al.* (2008) in Granada (Spain); and 45.5% in Gangneung (Korea) with the algorithm of Kim *et al.* (2016), who compared CC values in tenths instead of oktas.

In general, the extreme values of 0 and 8 oktas show the best agreement in the proposed model, as Figure 3 shows. This was also observed by Kim *et al.* (2016) and Xie *et al.* (2020), among others. The model predicts 94% of the cases assigned to the CC of 0 oktas; this percentage is 77% for overcast conditions. The partially cloudy cases present worse behavior, with the success rate being 48%, 49%, 52%, and 55% for 6 oktas, 5 oktas, 4 oktas, and 3 oktas respectively. The percentage of predicted data that fits the reference values within ±1 oktas is 97.0%, whereas other researchers, when comparing human observations with their algorithms, reported values at other locations of 60% (Granada; Cazorla *et al.*, 2008), 75.3% (Anhui, China; Xie *et al.*, 2020), 70% and 78% (Payerne and Jungfraujoch; Wacker *et al.*, 2015), 83% (Thessaloniki; Kazantzidis *et al.*, 2012), and 73% (Innsbruck; Kreuter *et al.*, 2009), and Huo and Lu (2012) reported values for within ±1 tenths of 79.9%, 66.7%, and 44.8% for Shouxian, Beijing, and Yangjiang (China) respectively. In the case of the percentage of predicted data fitting the reference values within ±2 oktas, our model presents a value of 99.4%, whereas other researchers obtained lower values when they used human field observations as reference: about 77% (Granada; Cazorla *et al.*, 2008), 90.9% (Anhui; Xie *et al.*, 2020), 84% and 89% (Payerne and Jungfraujoch; Wacker *et al.*, 2015), 94% (Thessaloniki; Kazantzidis *et al.*, 2012), and 85.3%, 76.1%, and 65.2% (Shouxian, Beijing, and Yangjiang; Huo & Lu, 2012), with this last case being for within ±2 tenths instead of oktas. The success rate within ±1 oktas in the proposed model is above 93.2% (98.1% within ±2 oktas) for all the different CC classes and above 99.2% (99.9% within ±2 oktas) for totally cloud-free and overcast conditions. On the other hand, Xie *et al.* (2020) obtained in Anhui that 50.0%, 97.0%, and 98.1% of the differences between the CC retrieved by the sky camera and the human observations were within ±1 oktas for 4 oktas, 0 oktas, and 8 oktas respectively; these percentages rose to 75.0%, 100%, and 100% when they considered within ±2 oktas.

In order to check the influence of the camera or site in the performance of the model, Figure 4 shows a similar analysis as in Figure 3 but separating the data for each different site and camera. The largest density of data corresponds to the diagonal of the confusion matrices in the four graphs, pointing out the good correlation between the predicted and reference CC values for all cameras and sites. It is better appreciated in La Palma, where the number of available data is the highest and then the results are more significant. Only in Izaña and for a CC of 6 oktas does more predicted data appear as another CC (5 oktas), but the number of data in these conditions (11 and 12) is very low to be representative. The model predicts the cloud-free conditions of 0 oktas with a success rate above 90% for all cases, being even above 99% for Izaña. Izaña also presents a high success rate of 95% for the overcast conditions; however, this value varies more with the site/camera, being 76% and 80% for La Palma and Valladolid (SONA202-NF) respectively. In the case of the OMEA-3C camera in Valladolid, the success rate is only 55% for 8 oktas, where the model considers these cases as 7 oktas in 40% of the cases; however, the number of available data is low and then this result is less representative.

The percentage of data within ±1 oktas is above 95% for La Palma for all CC conditions, whereas for the rest of the sites and cameras the same result (>95% within ±1 oktas) corresponds only for 0, 1, 7, and 8 oktas. The SONA202-NF camera shows the lowest percentage of data within ±1 oktas for 2, 3, and 4 oktas, where this value is close to 80%. This percentage is the lowest for 5 oktas in Valladolid (OMEA-3C) with a value of 83%; and for the 6 oktas case the lowest value is for Izaña with also a 83%. In general, the model fits worse the measurements for partially cloudy conditions. In general, the results are better for La Palma station, which could be partially caused by the highest amount of images at this station and so the results being more representative. The fact that the CNN was trained with more images from La Palma compared with Izaña or Valladolid (OMEA-3C) does not account for its better performance at La Palma; the number of images from La Palma and Valladolid (SONA202-NF) used in the training are similar, but the results for these two cameras differ.

Finally, the CC differences between the predicted and reference values have been calculated for the test set. The mean, also known as the mean bias error (MBE), and the standard deviation (SD) of these differences have also been calculated in order to quantify the accuracy and the precision of the model respectively. Table 2 shows these statistical estimators for the different sites and cameras and for the different CC reference values. The model overestimates (MBE > 0) the CC for scenarios with 2 oktas, showing MBE values around 0.2–0.3 oktas for all sites. This overestimation also occurs for 3 oktas except for the OMEA-3C camera in Valladolid. In general, the MBE values are within ±0.3 oktas for all CC conditions and sites, except for 8 oktas with the OMEA-3C camera in Valladolid (MBE = −0.5 oktas). These results indicate a good accuracy for all sites and cameras, with the MBE for all

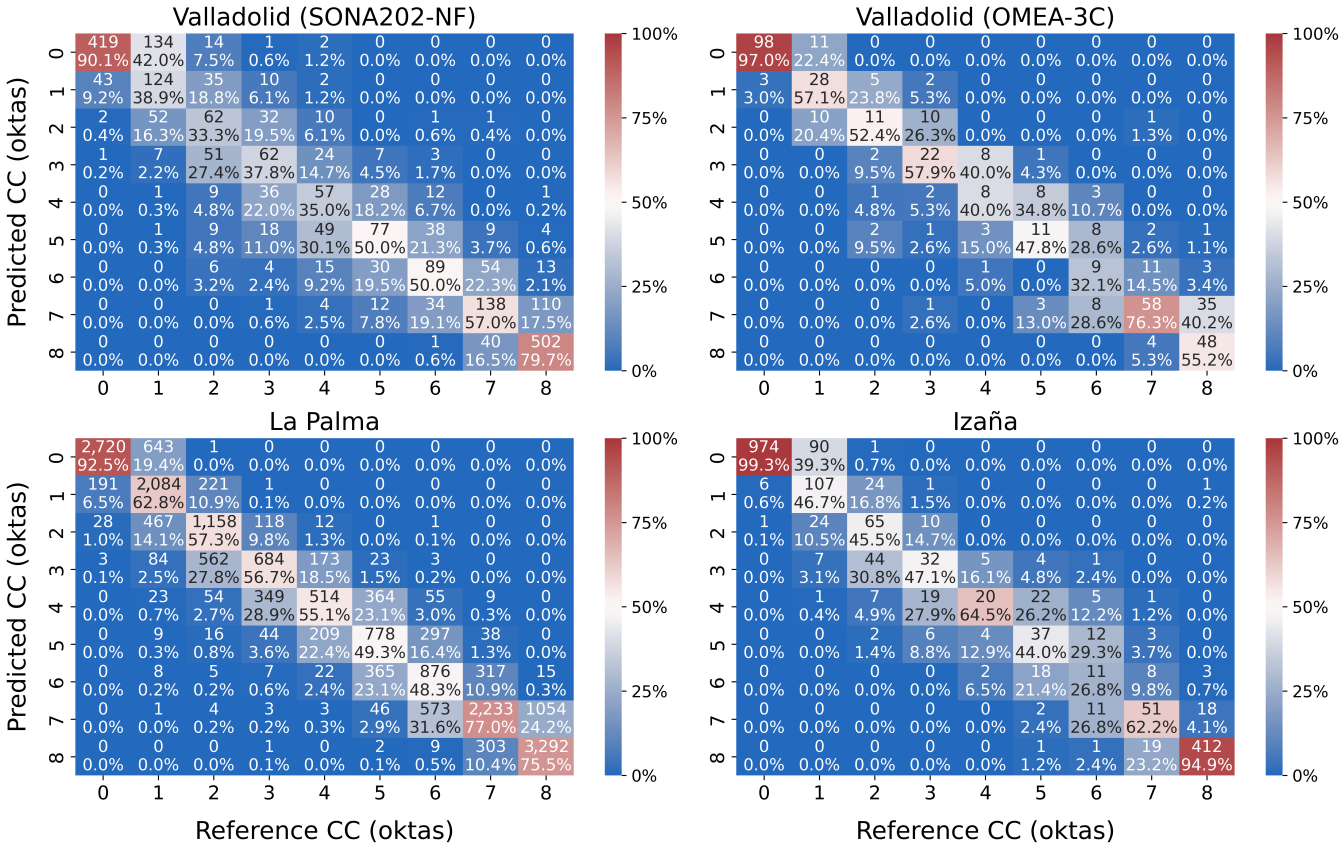

FIGURE 4 Confusion matrix with the number of cases matching for different predicted and reference cloud cover (CC) values. The reference values are the CC of classified images by the trained staff in the selected set of test images. These matrices are represented for different sites and two different cameras in the case of Valladolid. Color map represents the percentage of data considering the total as the sum of the number of data with the same reference CC value. This percentage is also written in each cell. [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 2 Mean bias error (MBE, oktas) and standard deviation (SD, oktas) of the differences between the cloud cover predicted by the convolutional neural network model and the reference values (visually measured by the trained staff) using the test set of images, for different sites and cloud cover values.

Site (camera)	Statistic	0 oktas	1 oktas	2 oktas	3 oktas	4 oktas	5 oktas	6 oktas	7 oktas	8 oktas	All
Valladolid (SONA202-NF)	MBE	0.11	−0.20	0.31	0.20	0.20	0.08	−0.22	−0.15	−0.24	−0.04
	SD	0.34	0.84	1.35	1.20	1.24	0.93	0.94	0.79	0.53	0.86
Valladolid (OMEA-3C)	MBE	0.03	−0.02	0.24	−0.16	−0.15	−0.18	−0.21	−0.21	−0.51	−0.17
	SD	0.17	0.65	1.15	1.01	0.85	1.01	0.98	0.77	0.62	0.76
La Palma	MBE	0.09	0.04	0.27	0.29	0.07	0.03	0.094	−0.04	−0.25	0.02
	SD	0.32	0.78	0.77	0.75	0.76	0.80	0.80	0.54	0.44	0.66
Izaña	MBE	0.01	−0.21	0.27	0.28	0.10	−0.06	−0.29	0.02	−0.07	−0.01
	SD	0.10	0.78	0.87	0.87	0.73	0.93	1.13	0.77	0.42	0.54
All	MBE	0.07	0.01	0.27	0.27	0.09	0.03	0.06	−0.05	−0.24	0.01
	SD	0.29	0.79	0.84	0.83	0.85	0.83	0.83	0.58	0.46	0.67

conditions ranging from −0.16 to +0.02 oktas, being practically zero when the total amount of data are considered. Huo and Lu (2012) obtained larger negative MBE values of −0.26 oktas in Shouxian, −0.10 oktas in Beijing, and −1.37 oktas in Yangjiang using human observations as reference. (Román *et al.*, 2017) compared three CC methods with human observations at two sites: in Granada, the MBE values were approximately 0.0 oktas, 0.6 oktas, and −0.1 oktas for their proposed sky camera algorithm, the RBR method, and ceilometer retrievals respectively; similarly, in Valladolid, the corresponding MBE values were −0.4 oktas, −0.4 oktas, and 0.2 oktas.

Regarding the SD, the lowest values appear for the extreme conditions of 0 oktas and 8 oktas, with SD values between 0.10 and 0.34 oktas and between 0.42 and 0.62 oktas respectively. In general, the SD is higher for partially cloudy scenarios, with the highest values for the SONA202-NF camera in Valladolid, especially from 2 to 4 oktas (SD between 1.20 and 1.35 oktas). In the rest of the cases the SD is usually below 1 oktas, the best results being for La Palma station where the SD is below 0.8 oktas for all CC conditions. The SD considering all CC cases ranges from 0.54 oktas (Izaña) to 0.85 oktas (Valladolid SONA202-NF). Considering all the data, the SD is approximately 0.67 oktas. Therefore, the overall precision of the model can be regarded as 1.34 oktas (equivalent to 2SD, or 2σ), indicating that roughly 95% of the differences between model predictions and the reference values fall within a range of ±1.34 oktas. Huo and Lu obtained higher SD values of 2.02 oktas in Shouxian, 2.32 oktas in Beijing, and 1.74 oktas in Yangjiang using human observations as reference. (Román *et al.*, 2017) compared three CC methods against human observations, obtaining SD values in Granada about 1.8 oktas, 2.3 oktas, and 1.7 oktas for their proposed sky camera algorithm, the RBR method, and the ceilometer retrievals respectively; these SD data were 1.4 oktas, 1.5 oktas, and 1.5 oktas for Valladolid.

As previously mentioned, conventional CC models often misinterpret CCs in conditions of high aerosol loads, particularly when coarse particles are present. To evaluate the performance of our proposed model under diverse aerosol loads and particle sizes, Figure 5 illustrates the MBE and SD values across different intervals of aerosol optical depth (AOD) at 440 nm and the Ångström exponent. The Ångström exponent quantifies the spectral dependence of AOD on wavelength and it is indicative of the aerosol particle size distribution, with lower values suggesting the dominance of coarse particles and higher values indicating finer particles. Both AOD and the Ångström exponent have been obtained from the Aerosol Robotic Network (AERONET; Giles *et al.*, 2019; Holben *et al.*, 1998) as level 2.0 version 3 AERONET products. Specifically, the Ångström exponent has been calculated

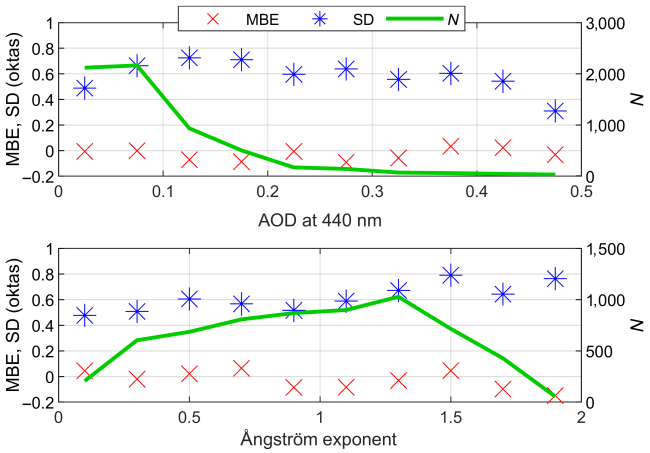

FIGURE 5 Mean bias error (MBE; red cross) and standard deviation (SD; blue asterisk) for different intervals of Aerosol Robotic Network (AERONET) aerosol optical depth (AOD) at 440 nm (upper panel) and for intervals of Ångström exponent calculated with the AERONET AOD at 440, 500, 675, and 870 nm (bottom panel). The number of data *N* related in each bin is represented by a green solid line and the right *y*-axis. [Colour figure can be viewed at wileyonlinelibrary.com]

using AERONET AOD measurements at 440, 500, 675, and 870 nm. The CC data of the test dataset without an AOD measurement in a ±5 min window have been discarded. Figure 5 also shows the number of available data per bin (*N*); most of the cases are associated with low aerosol loads (AOD at 440 nm below 0.1). Regarding AOD, the MBE values are consistently close to zero across all bins, with a range from −0.09 to +0.03 oktas. The SD varies from 0.31 to 0.72 oktas, exhibiting no apparent dependency on AOD and typically hovering around 0.6 oktas. For the Ångström exponent, the data distribution across intervals is more uniform. The MBE fluctuates slightly, from −0.15 to +0.06 oktas, without showing a clear trend with the Ångström exponent. The SD is marginally lower for intervals representing coarse particles, ranging between 0.48 and 0.79 oktas. Collectively, these findings suggest that the goodness of the CNN-CC model is not influenced by the aerosol load or particle size, maintaining robust performance even in scenarios characterized by high aerosol concentrations and coarse particles, such as desert dust.

3.2.2 | Observer measurements

In addition to the comparison with the classified images, the performance of the CNN-CC model has also been evaluated against independent field observations recorded by meteorological observers at the Marambio station, a location not used in the training of the CNN-CC. First, CC daytime data from the Marambio camera were

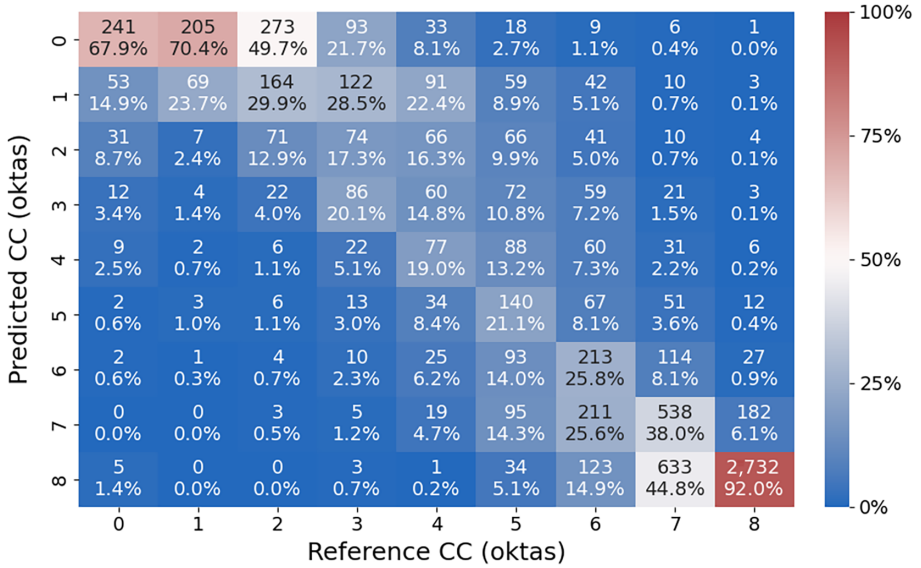

FIGURE 6 Confusion matrix between the cloud cover (CC) in oktas of the prediction model and the field observations for the Marambio data. [Colour figure can be viewed at wileyonlinelibrary.com]

retrieved using the trained CNN-CC. Subsequently, these CC retrievals and the CC observations made by meteorological observers were temporally aligned. In Marambio, CC human observations are typically conducted within a ±10 min window around each hourly measurement. Therefore, the sky image captured within this ±10 min interval, whose CC value most closely matches the observation, has been selected for comparison with this human observation.

Figure 6 displays the confusion matrix of the comparison between the CC obtained by both methods: CNN with HDR images (predicted) and visual records from weather observers (reference). The biggest success rate appears for overcast and cloud-free conditions, with 92.0% and 67.9% respectively. The rest of the matrix diagonal shows values between 19% and 38%, except for 2 oktas, which has a percentage of coincidence of 12.9%. Some 52.7% of the predicted CC fits perfectly with the human observations, which is a higher value than that reported by Wacker *et al.* (2015), Kreuter *et al.* (2009), and Cazorla *et al.* (2008) but lower than that obtained by Kazantzidis *et al.* (2012). The results improve comparing all the values that fit within ±1 oktas: the 8 oktas case carries on showing the best result (~98%), followed by 1 okta (~97%), 7 oktas (~91%), and 0 oktas (~83%). For partially cloudy conditions between 2 and 6 oktas, the percentage of predicted CC values fitting the reference within ±1 oktas is between 42% and 60%. Some 78.4% of the predicted CC fits the human observations within ±1 oktas and 90.7% within ±2 oktas; both values are higher than those obtained by Wacker *et al.* (2015), Kreuter *et al.* (2009), Cazorla *et al.* (2008), and Xie *et al.* (2020) (similar in this last case for within ±2 oktas), but lower than those reported by Kazantzidis *et al.* (2012), who obtained values of 83% and 94% respectively.

The ΔCC differences between the CC predicted by the model and the CC measured by field weather observers, which is used as the reference, have also been calculated. Figure 7 shows the distribution of these differences for the different cloud conditions. A higher frequency of ΔCC values equal to zero can be observed in the cloud-free and overcast conditions. Moreover, these distributions are wider for partially cloudy conditions, as expected. To quantify these distributions, the MBE and SD of these ΔCC have been calculated and are shown in Table 3.

The results in Table 3 indicate that the predicted data tend to underestimate the field observations, except for 7 oktas and obviously for 0 oktas. The MBE is within ±1 oktas for all values, except for the 2–4 oktas interval (MBE ~ −1.1 oktas), ranging from −0.6 to +0.7 oktas. The accuracy between both methods improves as the sky becomes more covered, with the MBE showing absolute values below 0.6 oktas for sky conditions between 5 and 8 oktas. For all cases the MBE corresponds to −0.3 oktas, pointing to a slight underestimation of the model to the field observations, which was not appreciated in the comparison with the test set. This MBE is similar to the values obtained by Román *et al.* (2017) in Valladolid comparing all-sky camera methods against human observations. Regarding the SD, all values are within 2.1 oktas, and even below 1 oktas for the cloud conditions of 1 and 8 oktas. Partially cloudy conditions, with CC values between 3 and 6 oktas, show the highest SD values, as in the comparison with the test set; this indicates a worse precision of the model for these conditions. The SD is 1.4 oktas if all available data are used, which means an overall precision (2σ) of about 2.8 oktas, approximately double that in the comparison with the test set. This SD value of 1.4 oktas is still below that reported by Román *et al.* (2017) in Valladolid and Granada with three different methods (one method

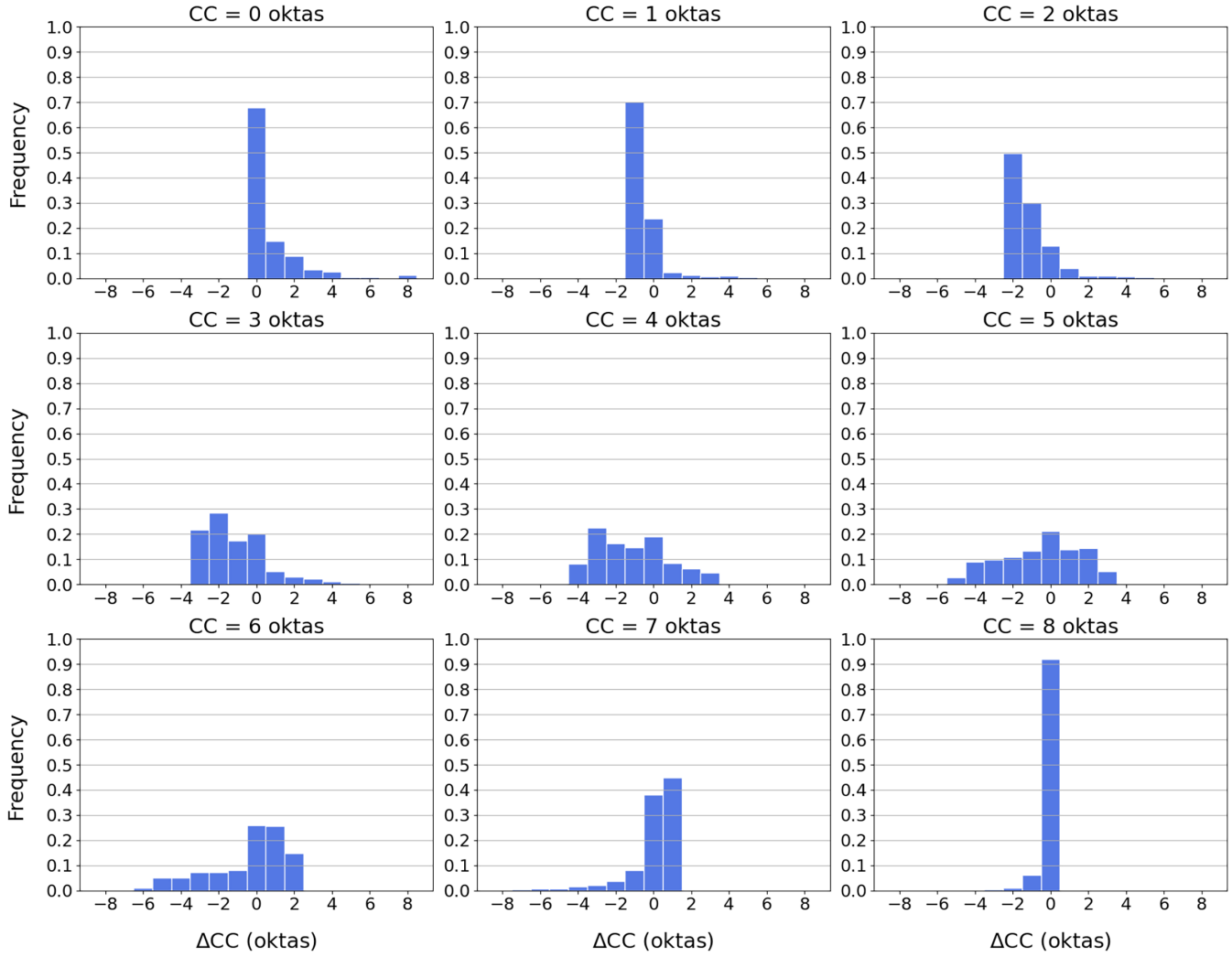

FIGURE 7 Distribution of the cloud cover (CC) differences (ΔCC) between predicted by the CNN-CC model and the field observations (reference) at Marambio for the nine CC reference values. [Colour figure can be viewed at wileyonlinelibrary.com]

presented a similar SD value at Valladolid) and those obtained by Huo and Lu (2012) in Shouxian, Beijing, and Yangjiang (see Section 3.2.1).

The performance of the CNN-CC model could be affected by the solar zenith angle (SZA), since higher SZA values decrease sky brightness, which could make it more difficult to identify CC. In order to analyze the performance of the model with respect to the SZA, the same MBE and SD values, along with other statistical estimators of the ΔCC distributions such as the 5th (P5), 10th (P10), 25th (Q1), 50th (the median), 75th (Q3), 90th (P90), and 95th (P95) percentiles, have been calculated for 1° SZA bins. These results are shown in Figure 8 as a function of SZA. In general, the MBE ranges between −0.5 and 0 oktas, pointing out the underestimation about −0.3 oktas observed in the previous results. The median is always 0 oktas, like Q3 in most of the SZA bins. Q1 varies between −1 and 0 oktas for low SZA values, but it is equal to −1 oktas in most of the higher SZA intervals. The SD values generally range from 1 to 2 oktas. For SZA data above 80° the MBE shows the lowest values and the SD the highest ones, which is in agreement with the low P5 values (about −4 oktas) reached under these conditions. These results indeed point out that the performance of CNN-CC is slightly worse for high SZA values, but these SZA conditions are the most frequent at this location, as the number of available data reveals.

Table 3 also shows the MBE and SD values calculated using data under SZA values below 80°. As expected, the accuracy and the precision improve, with SD values about 1.3 oktas if all CC conditions are considered. A confusion matrix similar to Figure 6 is included in Supporting Information Figure S1 but considering only data with SZA below 80°. In this case, 53.2% of the predicted CC fits perfectly with the human observations, and the biggest success rate appears also for overcast and cloud-free conditions, with 92.4% and 73.8% respectively.

TABLE 3 Mean bias error (MBE) and standard deviation (SD) of the differences between the cloud cover predicted by the model (convolutional neural network) and the direct field observed values at Marambio station for different sky-cloud conditions.

Cloud cover (oktas)	MBE (oktas)	SD (oktas)	$MBE_{SZA<80}$ (oktas)	$SD_{SZA<80}$ (oktas)
0	0.70	1.40	0.56	1.38
1	−0.57	0.88	−0.60	0.81
2	−1.14	1.19	−1.11	1.19
3	−1.13	1.65	−1.04	1.51
4	−1.11	1.92	−0.81	1.91
5	−0.56	2.10	−0.45	2.01
6	−0.41	2.04	−0.27	1.95
7	0.06	1.32	0.16	1.15
8	−0.12	0.54	−0.12	0.57
All	−0.31	1.41	−0.27	1.34

Note: $MBE_{SZA<80}$ and $SD_{SZA<80}$ are the MBE and SD values calculated using only data under solar zenith angles (SZA) below 80°.

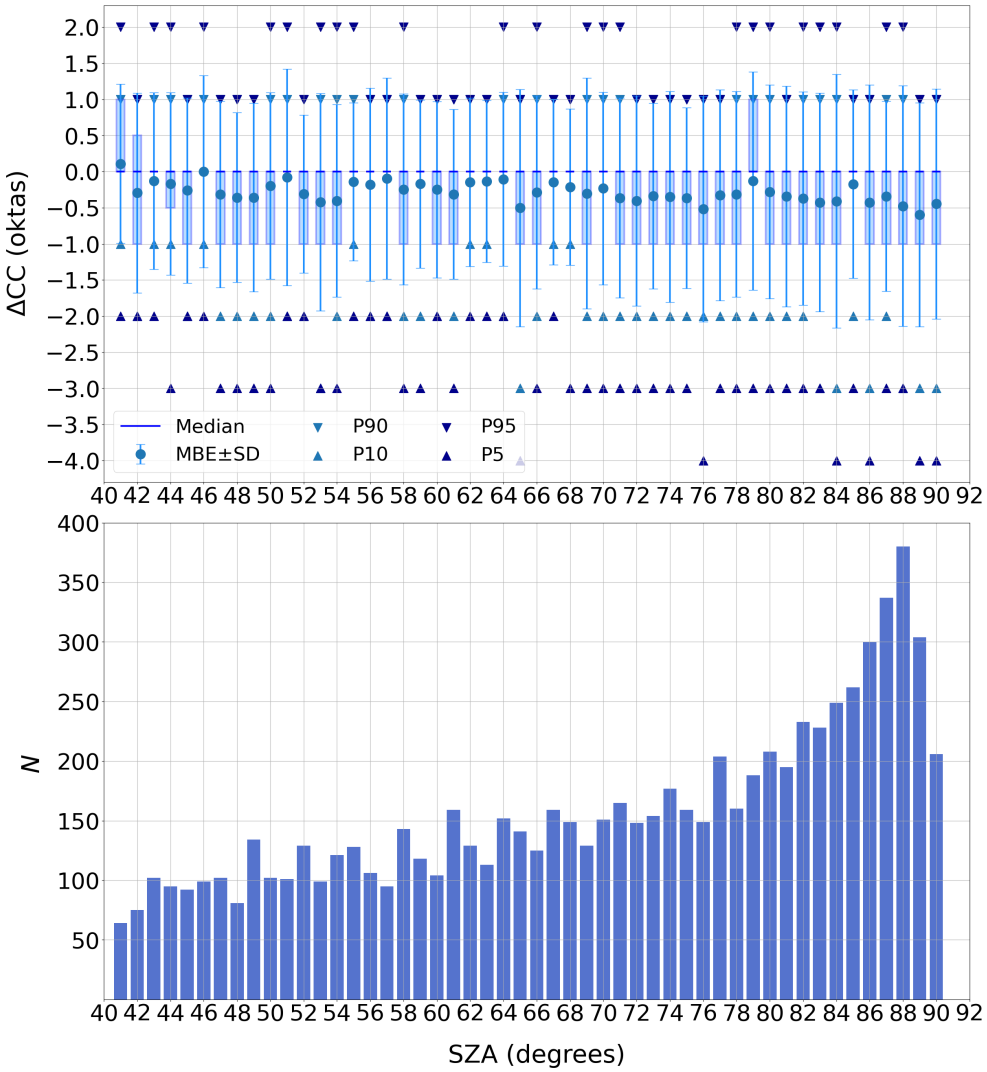

FIGURE 8 Upper panel: Mean bias error (MBE), median, standard deviation (SD), and the 5th (P5), 10th (P10), 25th (Q1), 75th (Q3), 90th (P90), and 95th (P95) percentiles of the cloud cover (CC) differences (ΔCC) between the predicted and the reference (field observations) values for ±0.5° solar zenith angle (SZA) bins. Error bar indicates the MBE ±SD. The lower and upper limits of the box represent Q1 and Q3 respectively. Lower panel: Number of available data *N* for each SZA bin. [Colour figure can be viewed at wileyonlinelibrary.com]

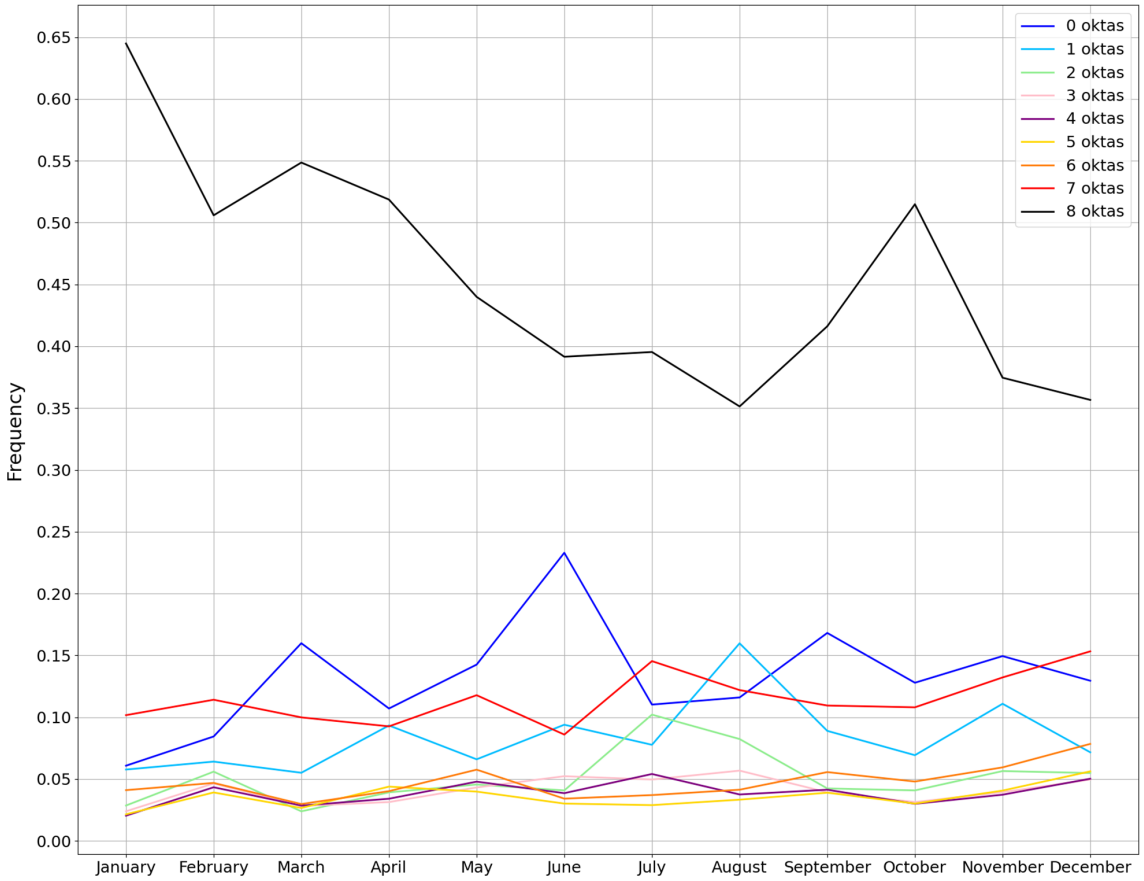

FIGURE 9 Cloud cover (CC) frequency at Marambio for each month using the CC data retrieved by the CNN-CC model. [Colour figure can be viewed at wileyonlinelibrary.com]

4 | RESULTS

Once the CNN model for CC retrieval was developed and evaluated, it was applied to the available sky images of Marambio to analyze the CC at this location. The images used were obtained with a 5-min frequency. Once the CC data were obtained, the frequency of the different CC values was calculated for each month. These results are shown in Figure 9. The most frequent sky condition is overcast (8 oktas) for all months. In fact, 46.5% of the available data correspond to 8 oktas (30,357). These overcast conditions are most common during the first part of the year, from January to April, with a frequency above 50%; these conditions are reduced to approximately 35%–42% from June to December, except in October. A decreasing trend in the frequency of overcast conditions is observed from January to August, which is just the opposite for the cloud-free (0 oktas) conditions. The frequency of the cloud-free cases is 12.3% for all data (8,045 data); this frequency ranges between 11% and 17% for all months except June (23.3%), when it is the highest, and January (6.1%) and February (8.4%), when this frequency is the lowest. August presents the highest amount of cases with 1, 2, and 3 oktas (30%), whereas December shows the most cases with 5, 6, and 7 oktas (29%). All these data mentioned can be found in the detailed information provided in Supporting Information Figure S2, including the annual values. For these annual values, the partially cloudy cases with 4, 5, and 6 oktas are, for each case, about the 3.8% of the total cases.

The monthly averaged CC and its deviation have been calculated and shown for each month in Table 4. June presents the lowest averaged CC, close to 4.6 oktas; January shows the highest average, as expected, due to the highest frequency of overcast cases, with a mean value of 6.5 oktas. The lowest CC variability appears in summer months, especially in January, with an SD value of 2.6 oktas; this is because these months present the most cases of overcast conditions with few cases of cloud-free conditions. On the other hand, June presents the highest variability, with an SD value of 3.4 oktas, which is caused by a high frequency of cloud-free cases with 0 oktas (23.3%) but also a still high frequency of the opposite conditions with 8 oktas (39.1%). In general, along the year, the CC at Marambio presents a mean value of 5.5 oktas with a variability of 3.1 oktas. These results are in concordance with the observed in Figure 9. Overall, the analysis indicates the

TABLE 4 Monthly and yearly mean and standard deviation (SD) of the cloud cover for each month at Marambio using the cloud cover data retrieved with the sky images.

Month	Mean (oktas)	SD (oktas)
January	6.49	2.61
February	5.81	2.88
March	5.70	3.18
April	5.66	3.05
May	5.37	3.40
June	4.58	3.42
July	5.19	3.05
August	4.72	3.17
September	5.08	3.23
October	5.68	3.08
November	4.97	3.19
December	5.21	3.00
Year	5.49	3.08

prevalence of high CC at Marambio, with the sky being mainly overcast or partially cloudy for the majority of the time throughout the year.

In order to study how the daytime CC varies along the day and to look for any diurnal pattern, Figure 10 presents a similar analysis to that in Figure 9 but classifying by hours instead of months. The data corresponding to this graph are shown in detail in Supporting Information Figure S3. Figure 10 shows these results using all the data available (all months together). Data from 0300 and 0400 UTC do not appear in Figure 10 because the sun does not reach SZA values below 90° in this time period along the year. The overcast conditions generally are less frequent between 0500 and 0800 UTC and from 1300 to 1800 UTC; on the other hand, the frequency of overcast conditions is higher, above 50%, at the beginning and the end of the day, but also between 0900 and 1200 UTC. Totally cloud-free conditions are slightly less frequent in central hours. The cases with 1 okta are more frequent in the 1500–1900 UTC interval. No clear hourly pattern can be appreciated, likely due to this graph being done with the data of all months, and the monthly patterns could be masked when all data are used together. To solve that, similar graphs to that in Supporting Information Figure S3 have been calculated but exclusively for each month in Supporting Information Figures S4 to S15.

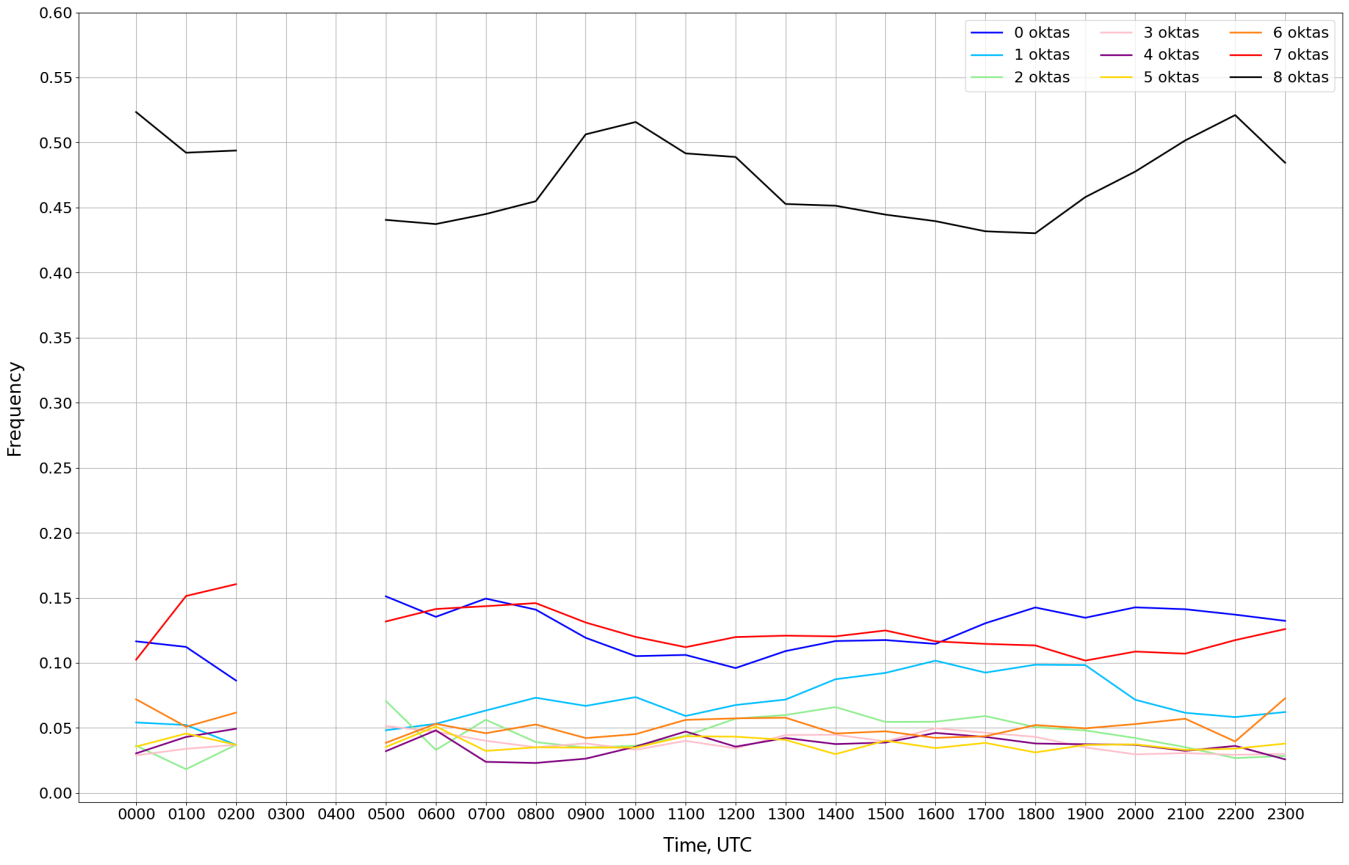

FIGURE 10 Cloud cover (CC) frequency at Marambio for each hour using the cloud cover data retrieved by the CNN-CC model. [Colour figure can be viewed at wileyonlinelibrary.com]

Regarding the most notable observed patterns, the 0 oktas conditions are more frequent in the afternoon than in the morning in the summer months (December, January, and February). The cloud-free cases in October are less frequent around midday. In the case of overcast conditions, a clear pattern can be observed in September, where these cases decrease from sunrise to sunset. November and December show less 8 oktas values around midday than close to sunrise and sunset. The rest of the conditions are less frequent, and then it is more difficult to find any pattern.

5 | CONCLUSIONS

In this work, 49,016 daytime sky images captured by all-sky cameras installed in Valladolid (Spain), La Palma (Spain), and Izaña (Tenerife, Spain) have been classified by their CC in oktas. A CNN model, named CNN-CC, has been developed and trained with part of these images to predict the CC.

The CNN-CC model has been compared against data not used in the training. The best results are for totally cloud-free (0 oktas) and overcast (8 oktas) conditions, with a percentage of coincidence between the predicted CC and the reference of 94% and 77% respectively. The predicted values within ±1 oktas match the reference values with a frequency about 99% for cloud-free and overcast conditions and above 93% for the rest of the CC conditions. The mean and SD of the differences between the predicted and the reference CC values are 0.01 oktas and 0.67 oktas respectively, pointing out a good accuracy but also a 1σ precision below 1 oktas (2σ precision of 1.34 oktas), which is within the assumed uncertainty. In general, the dependence of the performance of the model on the camera type and the location is weak; hence, it could be applied to images from other all-sky cameras.

The model for CC prediction showed worse performance when it was evaluated against independent field CC observations carried out by human observers. In this case the model tends to slightly underestimate the CC, with an MBE value about −0.3 oktas, whereas the 1σ precision is given by an SD value of 1.4 oktas (2σ precision of 2.8 oktas). The precision of the CNN-CC model improves to 1.3 oktas if it is only used for solar zenith angles below 80°.

The CNN-CC model has been applied to a set of daytime sky images recorded at the Antarctic station of Marambio (Argentina) during 3 years. The results obtained reveal that the sky condition at this location is usually overcast (8 oktas) through the year (46.5% of the time), January being the month with the highest frequency (64.5%) of overcast cases. On the other hand, totally cloud-free cases (0 oktas) appear only 12.3% in the year, but this percentage increases to 23.3% if only June is considered. No clear hourly patterns appear in the hourly evolution of CC, except in a slight way for a few months.

In general, we can conclude that the model developed for the prediction of CC provides realistic results that fit well with independent measurements. The performance of the model is in general slightly better than reported in various studies for other CC methods based on all-sky cameras. In addition, this model is capable of correctly estimating CC under the presence of high aerosol load, like in desert dust events, which usually turn the sky whiter and most of the models fail, considering these skies as cloudy. Although it has not been shown, this model has been applied to sky images from other sky cameras installed in different places. A qualitative analysis from these images points out that the predicted CC fits well with the real conditions even under high desert dust loads; on the other hand, the model seems to underestimate the CC under the presence of high thin clouds like cirrus. We want to encourage the reader and the scientific community to use the model developed in this work, which is freely available at the GOA-UVa web page (https://goa.uva.es) and in Zenodo (https://zenodo.org/records/10817551). This model can be validated or even retrained using images from other stations, or simply used to directly retrieve the CC at their respective stations. As a future prospect, a set of nighttime images will be classified to train the CC model not only for predicting CC during the daytime but also during the nighttime.

ACKNOWLEDGEMENTS

The research has been supported by the Ministerio de Ciencia e Innovación (MICINN), with Grant no. PID2021-127588OB-I00, and the Junta of Castilla y León (JCyL) with Grant no. VA227P20. This work is part of the project TED2021-131211B-I00 funded by MCIN/AEI/10.13039/501100011033 and the European Union, "NextGenerationEU"/PRTR. Special thanks are extended to Bernardo Martínez-Celda, Sergio Alegre, and Daniel García-Calzón for classifying part of the images. Sieltec Canarias is also acknowledged for lending us the SONA202-NF all-sky camera. We thank Emilio Cuevas for his effort in establishing and maintaining Izaña AERONET sites. We want to thank Rogelio Carracedo and José Luis Martín-Marcos (GOA-UVa staff) for the maintenance of the all-sky cameras at Valladolid, and also AEMet and the Meteorological Service of Argentina staff in charge of the maintenance and cleaning of the all-sky cameras at Izaña and Marambio. We want to thank the National Meteorological Service of Argentina for the support provided, especially to Ricardo Sánchez for providing the Marambio data and assisting us with the logistics of sending it to Marambio.

CONFLICT OF INTEREST STATEMENT

There are no competing interests.

DATA AVAILABILITY STATEMENT

The CNN-CC model presented in this article is available at the GOA-UVa web page (https://goa.uva.es) and in Zenodo (https://zenodo.org/records/10817551).

ORCID

Daniel González-Fernández https://orcid.org/0000-0002-3302-8080
Roberto Román https://orcid.org/0000-0003-4889-1781
Juan Carlos Antuña-Sánchez https://orcid.org/0000-0002-6786-671X
Victoria E. Cachorro https://orcid.org/0000-0002-4627-9444
Sara Herrero-Anta https://orcid.org/0000-0003-4246-1836
Celia Herrero del Barrio https://orcid.org/0009-0001-9508-3886
África Barreto https://orcid.org/0000-0001-5806-585X
Ramiro González https://orcid.org/0000-0003-0017-5591
David Mateos https://orcid.org/0000-0001-5540-4721
Carlos Toledano https://orcid.org/0000-0002-6890-6648
Abel Calle https://orcid.org/0000-0003-4161-7798
Ángel de Frutos https://orcid.org/0000-0001-5748-5078

REFERENCES

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C. et al. (2015) TensorFlow: large-scale machine learning on heterogeneous systems. Available at: https://www.tensorflow.org/

Antuña-Sánchez, J.C. (2022) Configuración y Metodología para el Uso de Cámaras de Todo Cielo en la Obtención de Parámetros Atmosféricos. Ph.D. thesis. PhD School of the University of Valladolid.

Antuña-Sánchez, J.C., Román, R., Bosch, J., Toledano, C., Mateos, D., González, R. et al. (2022) ORION software tool for the geometrical calibration of all-sky cameras. *PLoS One*, 17, e0265959. Available from: https://doi.org/10.1371/journal.pone.0265959

Antuña-Sánchez, J.C., Román, R., Cachorro, V.E., Toledano, C., López, C., González, R. et al. (2021) Relative sky radiance from multi-exposure all-sky camera images. *Atmospheric Measurement Techniques*, 14, 2201–2217. Available from: https://doi.org/10.5194/amt-14-2201-2021

Arking, A. & Childs, J. (1985) Retrieval of cloud cover parameters from multispectral satellite images. *Journal of Climate & Applied Meteorology*, 24, 322–333. Available from: https://doi.org/10.1175/1520-0450(1985)024¡0322:ROCCPF¿2.0.CO;2

Barreto, A., García, R.D., Guirado-Fuentes, C., Cuevas, E., Almansa, A.F., Milford, C. et al. (2022) Aerosol characterisation in the subtropical eastern North Atlantic region using long-term AERONET measurements. *Atmospheric Chemistry and Physics*, 22, 11105–11124. Available from: https://doi.org/10.5194/acp-22-11105-2022

Bedoya-Velásquez, A.E., Hoyos-Restrepo, M., Barreto, A., García, R.D., Romero-Campos, P.M., García, O. et al. (2022) Estimation of the mass concentration of volcanic ash using ceilometers: study of fresh and transported plumes from La Palma volcano. *Remote Sensing*, 14, 5680. Available from: https://doi.org/10.3390/rs14225680

Bennouna, Y., Cachorro, V., Torres, B., Toledano, C., Berjón, A., de Frutos, A. et al. (2013) Atmospheric turbidity determined by the annual cycle of the aerosol optical depth over north-center Spain from ground (AERONET) and satellite (MODIS). *Atmospheric Environment*, 67, 352–364. Available from: 10.1016/j.atmosenv.2012.10.065

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P. et al. (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*. Cambridge, UK: Cambridge University Press, pp. 571–657.

Cachorro, V.E., Burgos, M.A., Mateos, D., Toledano, C., Bennouna, Y., Torres, B. et al. (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. *Atmospheric Chemistry and Physics*, 16, 8227–8248. Available from: https://doi.org/10.5194/acp-16-8227-2016

Calbó, J. & Sabburg, J. (2008) Feature extraction from whole-sky ground-based images for cloud-type recognition. *Journal of Atmospheric and Oceanic Technology*, 25, 3–14. Available from: https://doi.org/10.1175/2007JTECHA959.1

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. Available from: https://doi.org/10.1016/j.solener.2015.02.006

Cazorla, A., Olmo, F. & Alados-Arboledas, L. (2008) Using a sky imager for aerosol characterization. *Atmospheric Environment*, 42, 2739–2745. Available from: https://doi.org/10.1016/j.atmosenv.2007.06.016

Chollet, F. (2015) Keras. Available at: https://keras.io

Clothiaux, E., Miller, M., Albrecht, B., Ackerman, T., Verlinde, J., Babb, D. et al. (1995) An evaluation of a 94-GHz radar for remote sensing of cloud properties. *Journal of Atmospheric and Oceanic Technology*, 12, 201–229. Available from: https://doi.org/10.1175/1520-0426(1995)012<0201:AEOAGR>2.0.CO;2

Córdoba-Jabonero, C., Sicard, M., Barreto, Á., Toledano, C., López-Cayuela, M.A., Gil-Díaz, C. et al. (2023) Fresh volcanic aerosols injected in the atmosphere during the volcano eruptive activity at the Cumbre Vieja area (La Palma, Canary Islands): temporal evolution and vertical impact. *Atmospheric Environment*, 300, 119667. Available from: https://doi.org/10.1016/j.atmosenv.2023.119667

Costa-Surós, M., Calbó, J., González, J. & Martin-Vide, J. (2013) Behavior of cloud base height from ceilometer measurements. *Atmospheric Research*, 127, 64–76.

Cuevas, E., Milford, C., Barreto Velasco, A., Bustos Seguela, J.J.d., García Rodríguez, O.E., García Cabrera, R.D. et al. (2022) *Izaña Atmospheric Research Center*. WMO/GAW. Activity Report 2019-2020. Available at: https://izana.aemet.es/wp-content/docs/Izana·Report·2019·2020.pdf

Forster, P., Storelvmo, T., Armour, K., Collins, W., Dufresne, J.-L., Frame, D. et al. (2021) The Earth's energy budget, climate feedbacks and climate sensitivity. In: *Climate change 2021: the physical science basis. Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change*. Cambridge, UK: Cambridge University Press, pp. 923–1054. Available from: https://doi.org/10.1017/9781009157896.009

Fukushima, K. (1980) Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, 36, 193–202.

García, O., Suárez, D., Cuevas, E., Ramos, R., Barreto, A., Hernández, M. et al. (2022) *La erupción volcánica de La Palma y el papel de la Agencia Estatal de Meteorología*. Revista Tiempo y Clima, 5. Available at: https://pub.ame-web.org/index.php/TyC/article/view/2516

Gavrikov, P. (2020) Visualkeras. Available at: https://github.com/paulgavrikov/visualkeras

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. *Atmospheric Measurement Techniques*, 5, 2881–2892. Available from: https://doi.org/10.5194/amt-5-2881-2012

Giles, D.M., Sinyuk, A., Sorokin, M.G., Schafer, J.S., Smirnov, A., Slutsker, I. et al. (2019) Advancements in the aerosol robotic network (AERONET) version 3 database—automated near-real-time quality control algorithm with improved cloud screening for sun photometer aerosol optical depth (AOD) measurements. *Atmospheric Measurement Techniques*, 12, 169–209. Available from: https://doi.org/10.5194/amt-12-169-2019

Holben, B.N., Eck, T.F., Slutsker, I., Tanré, D., Buis, J.P., Setzer, A. et al. (1998) AERONET—a federated instrument network and data archive for aerosol characterization. *Remote Sensing of Environment*, 66, 1–16.

Huo, J. & Lu, D. (2009) Cloud determination of all-sky images under low-visibility conditions. *Journal of Atmospheric and Oceanic Technology*, 26, 2172–2181. Available from: https://doi.org/10.1175/2009JTECHA1324.1

Huo, J. & Lu, D. (2012) Comparison of cloud cover from all-sky imager and meteorological observer. *Journal of Atmospheric and Oceanic Technology*, 29, 1093–1101. Available from: https://doi.org/10.1175/JTECH-D-11-00006.1

Ioffe, S. & Szegedy, C. (2015) *Batch normalization: accelerating deep network training by reducing internal covariate shift*. arXiv preprint arXiv:1502.03167.

Johnson, R., Hering, W. & Shields, J. (1989) *Automated Visibility & Cloud Cover Measurements with a solid state imaging system*. University of California, San Diego, Scripps Institution of Oceanography, Marine Physical Laboratory. Technical Report: SIO Ref. 89-7, GL-TR-89-0061, pp. 128.

Kazantzidis, A., Tzoumanikas, P., Bais, A., Fotopoulos, S. & Economou, G. (2012) Cloud detection and classification with the use of whole-sky ground-based images. *Atmospheric Research*, 113, 80–88. Available from: https://doi.org/10.1016/j.atmosres.2012.05.005

Kim, B.-Y., Jee, J.-B., Zo, I.-S. & Lee, K.-T. (2016) Cloud cover retrieved from skyviewer: a validation with human observations. *Asia-Pacific Journal of Atmospheric Sciences*, 52, 1–10.

Kingma, D.P. & Ba, J. (2014) Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980.

Kollias, P., Clothiaux, E., Miller, M., Albrecht, B., Stephens, G. & Ackerman, T. (2007) Millimeter-wavelength radars: new frontier in atmospheric cloud and precipitation research. *Bulletin of the American Meteorological Society*, 88, 1608–1624. Available from: https://doi.org/10.1175/BAMS-88-10-1608

Kreuter, A., Zangerl, M., Schwarzmann, M. & Blumthaler, M. (2009) All-sky imaging: a simple, versatile system for atmospheric research. *Applied Optics*, 48, 1091–1097. Available from: https://doi.org/10.1364/AO.48.001091

Krizhevsky, A., Sutskever, I. & Hinton, G.E. (2012) ImageNet classification with deep convolutional neural networks. In: Pereira, F., Burges, C.J.C., Bottou, L. & Weinberger, K.Q. (Eds.) *Proceedings of the 25th International Conference on Neural Information Processing Systems*. Red Hook, NY. Available from:: Curran Associates, Inc, pp. 1097–1105. Available from: http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf

Lachlan-Cope, T. (2010) Antarctic clouds. *Polar Research*, 29, 150–158. Available from: https://doi.org/10.1111/j.1751-8369.2010.00148.x

Linfoot, A. & Allis, R. (2008) A cloud detection algorithm applied to a whole sky imager instrument using neural networks. In: *88th Annual Meeting*. Boston, MA: American Meteorological Society, pp. 1–13.

Liu, S., Zhang, L., Zhang, Z., Wang, C. & Xiao, B. (2015) Automatic cloud detection for all-sky images using superpixel segmentation. *IEEE Geoscience and Remote Sensing Letters*, 12, 354–358. Available from: https://doi.org/10.1109/LGRS.2014.2341291

Martínez-Chico, M., Batlles, F. & Bosch, J. (2011) Cloud classification in a mediterranean location using radiation data and sky images. *Energy*, 36, 4055–4062. Available from: https://doi.org/10.1016/j.energy.2011.04.043

Martucci, G., Milroy, C. & O'Dowd, C. (2010) Detection of cloud-base height using Jenoptik CHM15K and Vaisala CL31 ceilometers. *Journal of Atmospheric and Oceanic Technology*, 27, 305–318. Available from: https://doi.org/10.1175/2009JTECHA1326.1

Masuda, R., Iwabuchi, H., Schmidt, K., Damiani, A. & Kudo, R. (2019) Retrieval of cloud optical thickness from sky-view camera images using a deep convolutional neural network based on three-dimensional radiative transfer. *Remote Sensing*, 11, 1962. Available from: https://doi.org/10.3390/rs11171962

Mateos, D., Pace, G., Meloni, D., Bilbao, J., di Sarra, A., de Miguel, A. et al. (2014) Observed influence of liquid cloud microphysical properties on ultraviolet surface radiation. *Journal of Geophysical Research: Atmospheres*, 119, 2429–2440. Available from: 10.1002/2013JD020309

Milford, C., Torres, C., Vilches, J., Gossman, A.-K., Weis, F., Suárez-Molina, D. et al. (2023) Impact of the 2021 La Palma volcanic eruption on air quality: insights from a multidisciplinary approach. *Science of The Total Environment*, 869, 161652. Available from: https://doi.org/10.1016/j.scitotenv.2023.161652

Onishi, R. & Sugiyama, D. (2017) Deep convolutional neural network for cloud coverage estimation from snapshot camera images. *Scientific Online Letters on the Atmosphere*, 13, 235–239. Available from: https://doi.org/10.2151/sola.2017-043

Orsini, A., Tomas, C., Calzolari, F., Nardino, M., Cacciari, A. & Georgiadis, T. (2002) Cloud cover classification through simultaneous ground-based measurements of solar and infrared radiation. *Atmospheric Research*, 61, 251–275. Available from: https://doi.org/10.1016/S0169-8095(02)00003-0

Peris, C. (2021) Optical properties of cloudy atmospheres through radiative transfer and remote sensing: from 1D to 3D approach. Ph.D. thesis. Ph.D. program of Remote Sensing of the University of Valencia.

Poyer, A. (2008) Evaluation of an MPL cloud detection algorithm as a reference for ceilometer testing within the ASOS PI program. In: *88th Annual Meeting*. Boston, MA: American Meteorological Society, pp. 1–6.

Román, R., Antuña Sánchez, J.C., Cachorro, V.E., Toledano, C., Torres, B., Mateos, D. et al. (2022) Retrieval of aerosol properties using relative radiance measurements from an all-sky camera. *Atmospheric Measurement Techniques*, 15, 407–433. Available from: https://doi.org/10.5194/amt-15-407-2022

Román, R., Bilbao, J. & de Miguel, A. (2014) 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. *Atmospheric Environment*, 89, 556–569. Available from: https://doi.org/10.1016/j.atmosenv.2014.02.027

Román, R., Cazorla, A., Toledano, C., Olmo, F., Cachorro, V., de Frutos, A. et al. (2017) Cloud cover detection combining high dynamic range sky images and ceilometer measurements. *Atmospheric Research*, 196, 224–236. Available from: https://doi.org/10.1016/j.atmosres.2017.06.006

Román, R., González, R., Toledano, C., Barreto, A., Pérez-Ramírez, D., Benavent-Oltra, J.A. et al. (2020) Correction of a lunar-irradiance model for aerosol optical depth retrieval and comparison with a star photometer. *Atmospheric Measurement Techniques*, 13, 6293–6310. Available from: https://doi.org/10.5194/amt-13-6293-2020

Rossow, W. & Schiffer, R. (1999) Advances in understanding clouds from ISCCP. *Bulletin of the American Meteorological Society*, 80, 2261. Available from: https://doi.org/10.1175/1520-0477(1999)080<2261:AIUCFI>2.0.CO;2

Salgueiro, V., Guerrero-Rascado, J., Costa, M., Román, R., Cazorla, A., Serrano, A. et al. (2023) Characterization of Tajogaite volcanic plumes detected over the Iberian Peninsula from a set of satellite and ground-based remote sensing instrumentation. *Remote Sensing of Environment*, 295, 113684. Available from: https://doi.org/10.1016/j.rse.2023.113684

Shields, J., Johnson, R., Karr, M. & Wertz, J. (1998) Automated day/night whole sky imagers for field assessment of cloud cover distributions and radiance distributions. In: *Tenth Symposium on Meteorological Observations and Instrumentation*. Boston, MA: American Meteorological Society.

Sicard, M., Córdoba-Jabonero, C., Barreto, A., Welton, E.J., Gil-Díaz, C., Carvajal-Pérez, C.V. et al. (2022) Volcanic eruption of Cumbre Vieja, La Palma, Spain: a first insight to the particulate matter injected in the troposphere. *Remote Sensing*, 14, 2470. Available from: https://doi.org/10.3390/rs14102470

Silva, A.A. & Souza-Echer, M. (2016) Ground-based observations of clouds through both an automatic imager and human observation. *Meteorological Applications*, 23, 150–157. Available from: https://doi.org/10.1002/met.1542

Tapakis, R. & Charalambides, A. (2013) Equipment and methodologies for cloud detection and classification: a review. *Solar Energy*, 95, 392–430. Available from: https://doi.org/10.1016/j.solener.2012.11.015

Toledano, C., González, R., Fuertes, D., Cuevas, E., Eck, T.F., Kazadzis, S. et al. (2018) Assessment of sun photometer Langley calibration at the high-elevation sites Mauna Loa and Izaña. *Atmospheric Chemistry and Physics*, 18, 14555–14567.

Tompson, J., Goroshin, R., Jain, A., LeCun, Y. & Bregler, C. (2015) Efficient object localization using convolutional networks. arXiv preprint arXiv:1411.4280.

Vasaras, A., Bais, A., Feister, U. & Zerefos, C. (2001) Comparison of two methods for cloud flagging of spectral UV measurements. *Atmospheric Research*, 57, 31–42. Available from: https://doi.org/10.1016/S0169-8095(00)00070-3

Wacker, S., Gröbner, J., Zysset, C., Diener, L., Tzoumanikas, P., Kazantzidis, A. et al. (2015) Cloud observations in Switzerland using hemispherical sky cameras. *Journal of Geophysical Research: Atmospheres*, 120, 695–707. Available from: https://doi.org/10.1002/2014JD022643

Warren, S., Eastman, R. & Hahn, C. (2015) CLOUDS AND FOG—climatology. In: *Encyclopedia of atmospheric sciences*, 2nd edition. Cambridge, MA: Academic Press, pp. 161–169. Available from: https://doi.org/10.1016/B978-0-12-382225-3.00113-4

WMO. (2012) Recomended methods for evaluating cloud and related parameters. WWRP/WGNE Joint Group on Forecast Verification Research (JWGFVR), World Weather Research Programme (WWRP) 2012-1.

WMO. (2017) International Cloud Atlas: manual on the observation of clouds and other meteors.

WMO. (2019) Manual on Codes - International Codes, Volume I.1, Annex II to the WMO Technical Regulations: part A- Alphanumeric Codes.

Xie, W., Liu, D., Yang, M., Chen, S., Wang, B., Wang, Z. et al. (2020) SegCloud: a novel cloud image segmentation model using a deep convolutional neural network for ground-based all-sky-view camera observation. *Atmospheric Measurement Techniques*, 13, 1953–1961. Available from: https://doi.org/10.5194/amt-13-1953-2020

Zhao, G. & Di Girolamo, L. (2006) Cloud fraction errors for trade wind cumuli from EOSTerra instruments. *Geophysical Research Letters*, 33, L20802. Available from: https://doi.org/10.1029/2006GL027088

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: González-Fernández, D., Román, R., Antuña-Sánchez, J.C., Cachorro, V.E., Copes, G., Herrero-Anta, S. *et al.* (2024) A neural network to retrieve cloud cover from all-sky cameras: A case of study over Antarctica. *Quarterly Journal of the Royal Meteorological Society*, 1–19. Available from: https://doi.org/10.1002/qj.4834