AUTOMATIC OBSERVATION OF CLOUDINESS: ANALYSIS OF ALL-SKY IMAGES

Yézer Gonzalez, César López and Emilio Cuevas
Sieltec Canarias S.L. / Izana Atmospheric Research Center (CIAI),
Santa Cruz de Tenerife, Canary Islands, Spain
Tel. +34-922356013, Fax +34-901706711, e-mail: yezer.gonzalez@sieltec.es

ABSTRACT

The influence of clouds on Earth´s radiation balance is a key component of the atmospheric system. Therefore, the determination of cloud cover is an important activity traditionally performed by human observations and broadcasted by meteorological services observers in Synop and Metar reports. However, human observations are subject to large errors of estimation, and their temporal resolution is very poor.

We have developed an automatic system of cloud observation (SONA) that provides cloud cover percentage using neural networks in all-sky images. Our system consists on a 640x480 pixel resolution CCD camera with an infrared filter. All-sky images are stored in time periods from 1 to 5 minutes and real time processed to obtain cloud cover percentage during daytime. Clouds are also well detected during nighttime.

The cloud detection algorithms are based on nonlinear multilayer neural networks for image pattern recognition. They are subjected to training processes with a back-propagation method. SONA is an inexpensive, all-weather and robust system to obtain cloud information in a wide range of sky conditions.

New improvements are being developed to obtain cloud flow, cloud-base height estimation by triangulation, and an index of atmospheric dust presence. SONA products might be customized in order to satisfy requirements of diverse applications such as cloud reporting in airports, cloud and radiation nowcasting combined with satellite images for solar power plants, sky watching in astronomy observatories, public dissemination of sky conditions on beaches and resorts, etc.

1. INTRODUCTION

Clouds are an important atmospheric factor, they are part of the hydrological cycle and also have an important role in the solar radiation balance. As soon as they build from water vapor (which is the major atmospheric greenhouse gas), absorption, scattering and reflection of incoming radiation affect the weather and can reduce the solar radiation up to 99%. It is important the change in the amount of UV radiation reaching the Earth´s surface due to type and position of the clouds [1]. They also are capable to absorb Earth infrared radiation.

Cloudiness or cloud fraction is a meteorological measurement traditionally made by human observers. The observer looks at the sky and makes estimation the cloud cover in okta. These measurements are expensive and are not accurate because of human subjectivity. The same atmospheric situation could be measured by different observes with not similar results, depending on their mood, perception or fatigue.

We have developed SONA (Automatic Cloud Observation System), this system provides with enough resolution in space and time, hemispherical images of sky, which are processed and analyzed by our algorithms, generating an objective and accurate value of cloudiness percentage.

Cloud coverage is the first step, later than detection, in some conditions the system can predict short-term motion cloud flow, cloud base height and type of cloud.

2. INSTRUMENTATION

SONA comprises a computer and All-Sky imager whit the following hardware. Camera with a CCD sensor with Bayer filter with resolution 640x480 pixels, 8 bit, color response from 400 to 700nm and monochrome response from 400 to 1000nm. This camera is inside in very durable aluminum housing. CCD looks at the external environment through a borosilicate dome. The system also has a shadow band for protecting the sensor from direct sunlight. The control electronics is safe against lightning. The operating temperature is from -10º to +50ºC due to its refrigeration system.

Fig.1 Left: Image of SONA, we can see aluminum housing, dome and shadow band. Right: Sky image took by SONA.

Camera features:
- Rate: Fast frame rates (up to 70 fps).
- Adjustable JPEG compressed still-images or live MJPEG streaming video.
- Transfer of images via FTP, RTP or HTTP.
- Camera control via HTTP, XML-RPC, Telnet.
- Automatic sliding IR cut filter.

3. CLOUDS DETECTION PROCEDURE.

Nature is complex. Researchers have been trying to solve some biological, ecological and environmental problems developing models, but the main problem in modeling real problems is the lack of understanding of their underlying mechanisms because of complex and nonlinear interactions among various aspects of the problem. They are many reasons for a complex behaviour as randomness, heterogeneity, multiple causes and effects and noise. Cloud cover characterization in sky images is a complex problem, our system has many variables related with the sky as luminosity, cloud distribution, different kind of clouds, RGB pixel values, time of day, season, dust, pixel position relative to sun and horizon, etc. and other variables linked to hardware as iris aperture, exposure time, contrast, color gain, geometrical distortion and noise. The behaviour of all these variables makes our system a complex system.

The nature of these mentioned complex systems requires a system approach where the most essential features of a complex problem with multiple interactions are modeled so that the system behaviour can be predicted. Neural networks are flexible, adaptive learning systems that find patterns in the data of a nonlinear system.

Biological neuron consists of three main components (Fig. 2A): dendrites are channels for input signals, cell body accumulates and processes input weighted signals, and an axon which transmits the output signal to other neurons connected to it. The human brain receives information from the external environment via the senses, this information is procesed by brain cells o neurons, which form a large number of parallel networks. Each neuron can send signals to anywhere from a hundred to thousand neurons through neurotransmitters. The connections between neurons hold memory, this memory or adquired knowledge is stored as the connections strengths between neurons. These capacities and properties can be emulated by computing procedures.

Neuron model (Fig. 2B) or perceptron is a computing process in which signals are received, summed and processed in the cell body for producing an output. A set of perceptrons forms a neural network and the type of neural network is determinated by the distribution of connections. We use a multilayer neural networks or multilayer perceptron networks (MLP) for image pattern recognition (Fig. 2C), MLP is the most popular and widely used nonlinear network for solving many practical problems in applied sciences for classification [2].

MLP can be trained as a classifier of the following way: An input pattern (x_1, x_2, x_3, ..., x_n) is transmitted through input connections (one input connection for one input data) whose inputs weights (I_{11}, I_{12}, I_{13}, ..., I_{1n}, ..., I_{p1}, I_{p2}, I_{p3}, …, I_{pn}) are initially set to random values, that inputs are summed and processing in every neuron of the first layer, the outputs of the first layer are sent to the second layer of neurons with its initial random weights (a_{11}, a_{12}, a_{13}, ..., a_{1n}, ..., a_{p1}, a_{p2}, a_{p3}, ..., a_{pn}), in this way the information is propagating by the whole network. When the process arrives to the last neuron layer, the final outputs are generated and compared with the given target output to determine error (E) for this pattern. Inputs and patterns are presented iteratively and the weights are adjusted until the minimum possible square error (MSE) when the entire training dataset have passed through the network. MSE is calculated using a backpropagation method [3].

Training dataset must be launched in the network many times for better results.

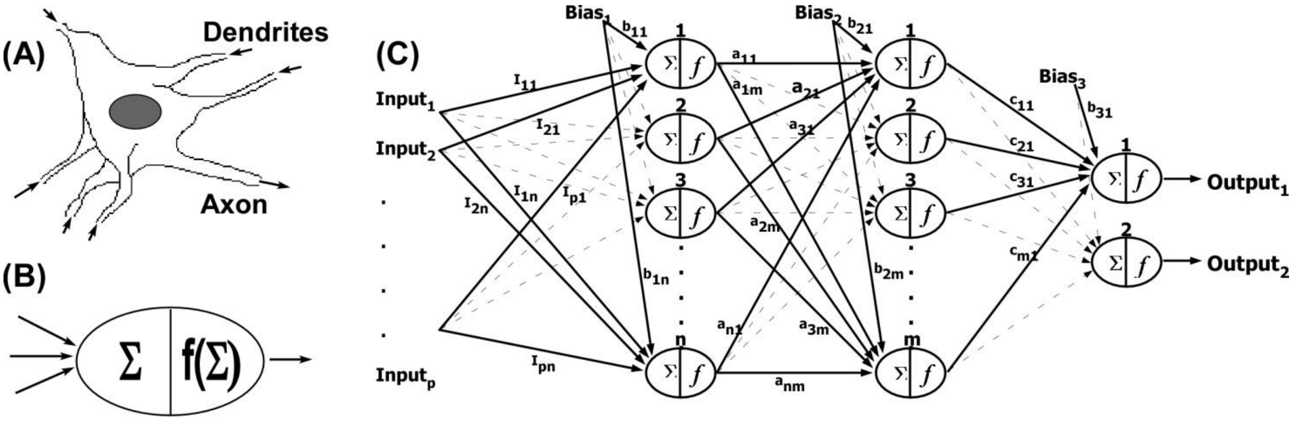

Fig.2 (A) Biological neuron. (B) Computer neuron model or perceptron. (C) Multilayer neural network with an input layer, two hidden neuron layers and a neuron output layer.

Before the cloud cover detection is necessary a previous process. First step is the choice a set of representative images from sky, which are the input for training algorithm (Fig. 3A). The algoritm places a virtual mask in the image over the shadow band and the external part of the image, that way we only have pixels belonging to sky (Fig. 3B). We select the target, cloud o clear pixels (Fig. 3C), this information is saved with all pixel input parameters.

In our system, the inputs parameters are related to every single pixel in the sky image. The input parameters are:

- Red, green, blue values (0-255).
- Red, green, blue variance values. This value is calculated in the 8-neighborhood of the pixel.
- Distance from Sun to pixel (pixels).
- Distance from zenih to pixel (pixels).
- Distance from Sun to zenith (pixels).

Configuring and training the multilayer neural network (Fig. 3D) reaching the minimum possible square error (discussed above) is the las step. The algorithm saves the configuration of the trained neural network in “memory parameters”, whose are stored in a text file.

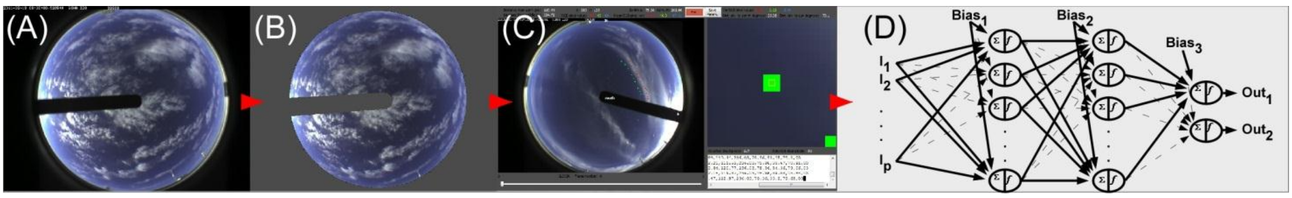

Fig.3 Training process steps. (A) Original image. (B) Masked image. (C) Program for selection cloud and clear pixel parameters. (D) Training multilayer neural network for pattern recognition.

Algorithm of cloud cover detection in sky images starts getting a sky image (Fig. 4A), later the position of the shadow band is detected placing a similar mask over it (Fig. 4B), this is important because we eliminate highlights and border effects in the image, moreover, there are less pixels and computing time is lower. The algorithm analyses every single pixel and its 8-neighborhood of the image calculating its parameters (Fig. 4C). These paremeters feed the trained neural network (Fig. 4D), that way, neural network determines whether the pixel is cloud or clear. Finally, the algorithm generates an output image with detected cloudy pixels (Fig. 4E) and a percentage value of cloudiness.

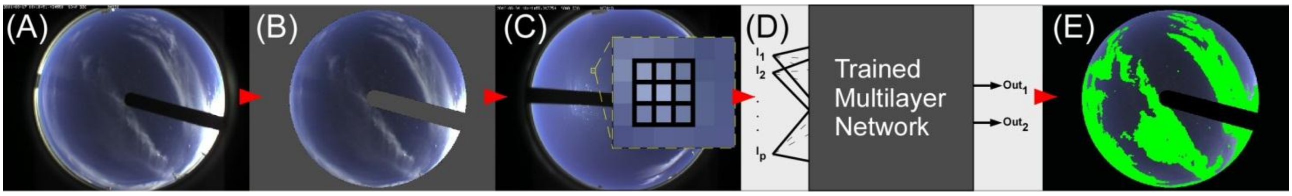

Fig.4 Cloud detection process steps. (A) Original image. (B) Masked image. (C) Algorithm analyses sky pixels and extracts input parameters. (D) Neural network generates the type of pixel. (E) Output image with cloud pixels marked as green.

In the images, the same areas close to the zenith, or close to the horizon, correspond to different solid angles. In other words, equal number of pixels in the image corresponding to different solid angle [4]. Correction is applied to every single pixel in the percentage calculation of cloudiness.

Due to forward scattering of visible light by aerosols and haze, our algorithms overestimate significantly sky cover, due to whitening near the sun and the horizon (Fig. 5A, 5D). These whitening pixels have the same characteristics of cloud pixels. For that reason, we need an extra information to feed our neural network, in order to distinguish this type of pixels.

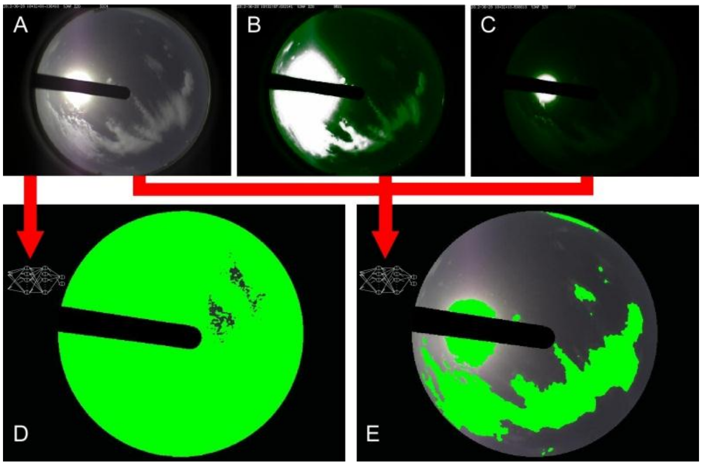

Fig.5 Single and multiple image analysis. (A) Original image. (B) Image with green gain x10. (C) Image with gamma factor x0.1. (D) Single image analysis result. Cloud cover is overestimated. (E) Multiple image analysis result.

One solution for this problem is the following strategy. Taking three different images at the same time: original image (Fig. 5A), image with green gain x10 (Fig. 5B), and image with gamma factor x0.1 (Fig. 5C). CCD sensor uses twice as many green elements as red or blue to mimic the physiology of the human eye. Retina has more rod cells than cone cells and rod cells are most sensitive to green light, that is why the modification of green gain to get additional information. Decreasing factor gamma gives data about more brilliant objects in the sky. Passing the input parameters of every single pixels of these three images through a new multilayer neural network, the algorithm generates a better cloud cover estimation (Fig. 5E).

The Sun circle in the image is a persistent problem. One characteristic of this problem is that cloud amount closes to Sun varies slowly in comparison with clouds in other areas in the image. Thus, low variability through time could be used to discriminate clouds, sampling images in a short period of time [5]. Other feature for this purpose is the “roundness” of this area at all solar elevations, looking at accuracy of a fitting function of this rounded/elliptical shape. If the Sun circle is not due to clouds, this effect would be related with an aerosol or dust load.

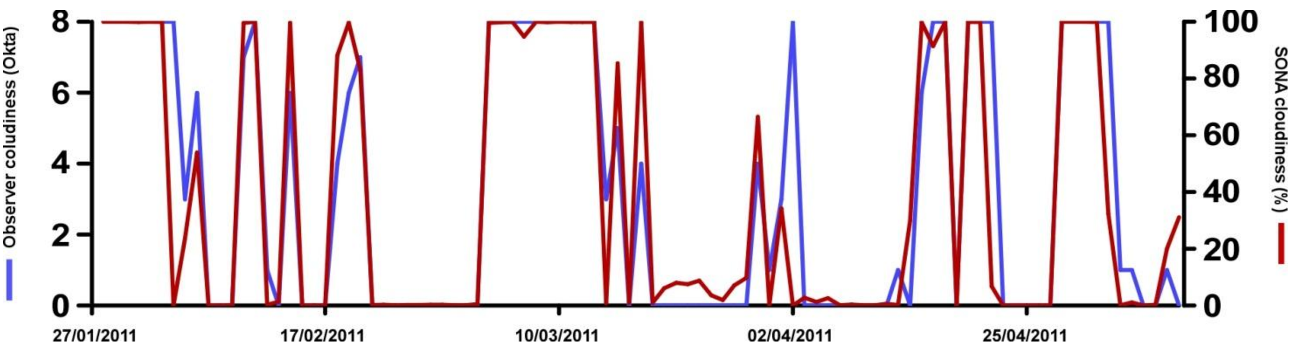

Fig.6 Cloudiness measured by observers (blue) by SONA (red). It exists a correlation between both data. Data from observes only can take eight different values (Okta), what makes the measures are less accurate.

Results from SONA are being largely tested by meteorological observers, we concluded with a satisfactory system behavior.

5. WORK IN PROGRESS

Motion flow.

Acquiring images at different instants in time makes cloud motion detection possible. Optical flow reflects the image changes due to motion during a time interval, and the optical flow field is the velocity field that represents the three-dimensional motion of objects points across a two-dimensional image. Optical flow provides tools to determinate motion, relative distance of objects, etc.

Nowcasting.

Knowledge of motion flow allows short term forecasting or nowcasting of cloud cover. Evaluation of motion field generates a probabilistic value for cloudiness in areas close to Sun in a near future.

Images from SONA combined with satellite images gives us a more accurate nowcasting. Satellite images are NWC SAF products, and allows to have information about cloud cover, cloud motion, and type of cloud, etc. Interpolating these images with SONA images (Fig. 7) is possible make nowcasting of clouds in a near-medium future. Mixing this forecast with radiative transfer models, would give a nowcasting of radiation, which is valuable information for performance of solar power plants.

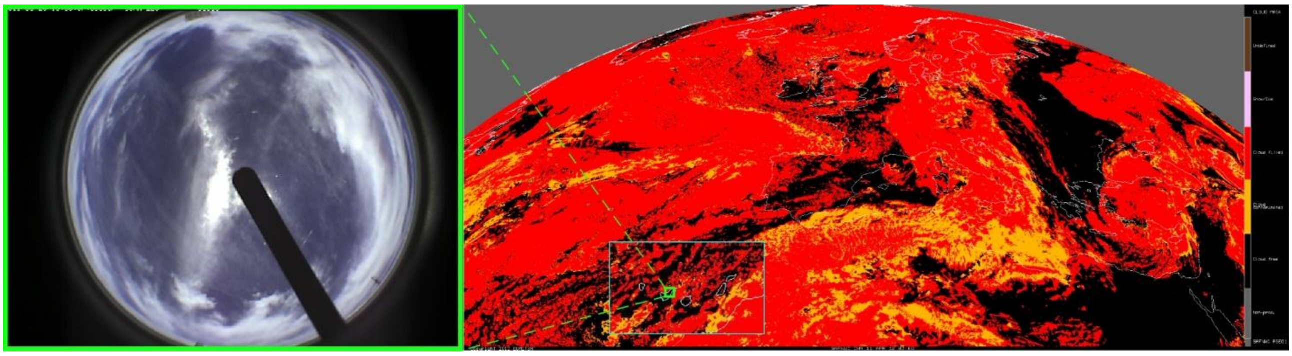

Fig.7 Nowcasting of clouds and radiation, using both images from SONA and satellite images.

Cloud-base height estimation.

Cloud-base height estimation by triangulation is an important factor in determining type of clouds and the infrared radiative properties of them. Using a SONA as primary system and other SONA or a standard camera as secondary system is necessary to calculate clouds height (Fig. 8). Interpolating images (find points or pixels which are projections from the same point on a cloud) took by both systems and detecting isolated clouds or part of clouds with “stable” features in images, make possible to detect their heights.

The main problem is to spatially match up cloud fields from two widely separated cameras, is necessary to use temporal flow fields from each camera separately.

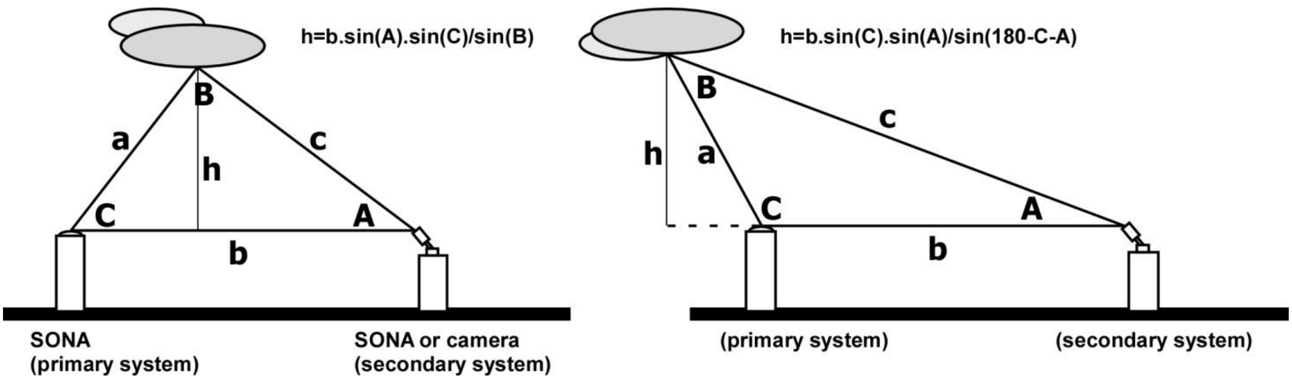

Fig.8 Cloud-base height estimation by triangulation. radiative transfer models of radiation

6. ACKNOWLEDGEMENTS

SONA Project and developed algorithms have received a subsidy in 2008 from Consejería de Hacienda and Gobierno de Canarias, with FEDER funds.

Acknowledgements to Sistemas Básicos del CIAI for their cooperation: Ramón Ramos, Rubén del Campo, Cándida Hernández, Virgilio Carreño, Fernando de Ory, Enrique Reyes, Antonio Cruz, Néstor Castro and Rocío López.

7. REFERENCES

1. Calbó, J., D. Pagès, and J.-A. González, Empirical studies of cloud effects on UV radiation: A review. Rev. Geophys., 2005. 43(2): p. RG2002.
2. Bishop, C.M., Neural Networks for Pattern Recognition. 1996: Oxford University Press, USA.
3. Samarasinghe S., Neural networks for applied sciences and engineering: from fundamentals to complex pattern recognition. 2007: Auerbach.
4. Long, C.N., Sabburg, J. M., Calbó, J., Pagés, D., Retieving cloud characteristics from ground-based daytime color all-sky images. J. Atmosf. Oceanic Technol., 2006. 23: p. 633-652.
5. Long, C.N., Correcting for Circumsolar and Near-Horizon Errors in Sky Cover Retrievals from Sky Images. Journal Name: Open Atmospheric Science Journal, 4:45-52; Journal Volume: 4, 2010: p. Medium: X.