Renewable Energy 97 (2016) 155–161

Contents lists available at ScienceDirect

Renewable Energy

journal homepage: www.elsevier.com/locate/renene

The application of Bayesian network classifiers to cloud classification in satellite images

J. Alonso-Montesinos [a, b, *], M. Martínez-Durbán [c], J. del Sagrado [c], I.M. del Águila [c], F.J. Batlles [a, b]

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

ARTICLE INFO

Article history:
Received 20 January 2016
Received in revised form
15 April 2016
Accepted 17 May 2016

Keywords:
Cloud classification
Electricity generation
Remote sensing
Bayesian classifiers
Satellite images

ABSTRACT

The need to reduce the impact of traditional electricity generation necessitates an increase in the optimization of alternative systems that produce less environmental contamination. Renewables play a key role, with solar energy considered one of the most important energy supply sources. Solar power plants have to be perfectly designed to optimize electricity generation, and their placement must be as suitable as possible for the meteorological conditions. Clouds are the most mitigating factor in solar energy production and their study is decisive in locating the plant. Apart from the importance of studying clouds before building the solar plants, cloud detection is equally decisive in adapting plant operation to cloud types during solar power plant operation.

This adaptation benefits plant performance and allows electricity management to be integrated into the electricity grid. Nonetheless, the majority of cloud studies determine atmospheric parameters, which are sometimes not available. In this work, we have developed an automatic, fully-exportable cloud classification model, where Bayesian network classifiers were applied to satellite images so as to determine the presence of clouds, classifying the sky as cloudless or with high, medium and low cloud presence. There was an average success probability of 90% for all sky conditions.

1. Introduction

One of the determining factors for a country's economic and industrial development is whether it can combine energy supply with the need to meet increasingly restrictive environmental requirements [1]. Renewables, therefore, offer a significant competitive advantage as they harness their own inexhaustible supply and do not emit carbon dioxide emissions into the atmosphere. Over recent years, solar energy has become ever more relevant with an expansion in the number of operational solar plants [2]. For this reason, solar thermal and solar photovoltaic technologies play an indispensable role in electricity generation.

Accurate understanding and modelling of solar radiation is essential to evaluate renewable energy resources, the climate and hydrological models. Precise historical solar radiation data are required for the places where solar plants are to be built. However, radiation stations are not always sited in these geographical locations. In such cases, interpolation and/or extrapolation techniques are used, which are imprecise methods in many situations.

In order to improve the results for interpolation/extrapolation techniques of solar radiation data, different studies have shown that using satellite imagery improves the solar radiation estimation in places where it is not possible to obtain radiation data from solar radiation stations. Hence, geostationary satellites, such as Meteosat, provide continuous land observations, which allow us to observe changes in the satellite path to land occurring over short time intervals due to solar irradiance. This is the reason why satellite imagery is one of the best options for estimating solar radiation and evaluating its energy potential at any location.

Over recent decades, various researchers have estimated solar radiation from satellite images. Zelenka et al., 1992 [3] used empirical relationships to evaluate clouds. Following this, the authors implemented daily global radiation estimations.

The Centre Energétique et Procédés (CEP) of the École des Mines

* Corresponding author. Department of Chemistry and Physics, University of Almería, 04120 Almería, Spain.
E-mail address: joaquin.alonso@ual.es (J. Alonso-Montesinos).

http://dx.doi.org/10.1016/j.renene.2016.05.066
0960-1481/

in Paris, collaborating with other European research centres, developed a statistical model to estimate solar radiation incidence on the Earth's surface using Meteosat images. This model is widely known as Heliosat [4–6] and provides a correlation between cloud cover and the global radiation incident at a point on the Earth's surface to estimate solar radiation. Using visible MSG satellite images, Badescu and Dumitrescu 2014 [7] estimated global radiation, where measured values were tested against estimations for a region in Romania. Furthermore, simple models have been proposed for solar radiation computation using the Meteosat Cloud Fractional Coverage dataset [8] under different sky conditions (cloudless and overcast). By classifying the sky into clear, overcast or cloudy, solar global radiation was estimated, offering the best cloudy sky model—with a relative root mean square error value ranging from approximately 17 to 68% [9,10].

Bosch et al., 2010 [11] combined the Heliosat-2 model with digital terrain models (DTM) to estimate a clear sky index (Kc) along with global daily radiation at a mountainous site, taking into account the sun obstruction caused by the presence of mountains. The root-mean-square error (RMSE) values oscillated between approximately 9-15%, when the estimated radiation was compared to data from four stations.

Analyzing the impact of atmospheric constituents on the study of solar radiation, clouds are presented as the highest attenuating factor. Up until the end of last century, cloud classification was performed by human observers, identifying cloud types visually and dividing the sky into oktas [12]. In Ref. [13], the authors made a prediction of cloud cover using the number of oktas, obtaining satisfactory results. Nowadays, the cloud classification scheme, which can be found in "The International Clouds Atlas" [14], is based on a Latin Linnean system [15], which were the first studies related to cloud classification. Subsequently, many authors have studied cloud coverage using various techniques and technologies.

Martínez-Chico et al. [16] combined DNI (direct normal irradiance) and sky camera imagery to perform a cloud classification (into four groups) aiming to show radiation attenuation based on cloud type. In addition to this, a cloud identification system was developed where the sky camera images were processed according to a sky classification [17], thus solving the pixel saturation problem in the solar area [18]. These studies were concluded with a solar radiation estimation using the pixel information from the sky image [19] applying the cloud motion vectors to obtain solar irradiance forecasting over the short term [20]. Moreover, satellite images have also been used to study cloud cover and cloud classification. Probabilistic neural networks and fuzzy logic techniques were employed to determine cloud classification from AVHRR satellite images, as shown in Refs. [21] and [22], respectively. Furthermore, using Landsat-7 satellite images, the clouds were characterized with satisfactory results [23].

Meteosat Second Generation (MSG) satellite images have been widely used for the study of atmosphere composition, including the presence of clouds. Escrig et al., 2013 [24] used these images for studying cloud cover and for classifying the clouds into three different layers (high, medium and low) depending on the cloud ceiling. In this article, different multispectral tests were applied to satellite images, showing over 85% agreement between the clouds detected by a whole-sky camera and the clouds derived from multispectral tests. The importance of such cloud identification allows us to reproduce the motion of clouds observed using cloud motion vectors. This objective was presented in a study, where satellite and sky camera imagery where combined to make a cloudiness forecast for the short- and medium-term [25] with satisfactory results. The cloudiness forecast was represented using a real-time GUI designed especially for solar power plant operators, providing useful information about the cloud presence over a solar field in near time horizons [26]. Therefore, the importance of cloud classification is valued in various fields, but overall in solar power plant management. Nonetheless, most studies developed for cloud classification require atmospheric parameters that are not always available.

The aim of this work is to apply and evaluate an automatic classification model construction based on Bayesian network classifiers to develop a reliable cloud classification model from satellite imagery data; this can be integrated into solar power plant control systems to improve a plant's performance as well as the management of electricity for integration into the electricity grid. This model can be adapted to other latitudes providing a similar dataset following the methodology proposed and avoiding the necessity to have atmospheric parameters for the classification.

2. Materials and methods

2.1. Data collection

This work used satellite images for cloud classification. The testing facility was located at the Solar Energy Research Centre (CIESOL) at the University of Almería, Spain (36.8°N, 2.4°W, at sea level), which has a Mediterranean climate and a high maritime aerosol presence.

Satellite data from the years 2011–2013, and from all possible sky types, were used.

All satellite channels were collected every 15 min when the solar altitude was higher than 10° so as to avoid mistakes caused by low image brightness. Table 1 shows the twelve spectral channels and details their properties. Besides satellite channels, each data sample also collected solar altitude, along with diffuse and direct solar radiation data.

SYNOP reports provided by the State Meteorological Agency (AEMET) for the years 2011–2013 (situated at Almería airport—36.5°N, 2.21°W, at 21 m above sea level, and with a distance of about 4 km from CIESOL building) were also used in the validation process. We used a total sky camera with a rotational shadow band (namely a *TSI* 880 model) with hemispheric sky vision.

2.2. Cloud types

The atmospheric altitude range in which clouds are generally encountered varies from sea level to the tropopause, i.e. 8 km in polar regions, 13 km in medium latitudes and 18 km in the tropics. By convention, the part of the atmosphere in which clouds typically appear is vertically divided in three layers: high, medium and low. Each layer is defined by the range of levels at which certain cloud genera occur most frequently. Thus, in the different cloud layers are [14,27–30]:

- Cirrus (Ci), Cirrocumulus (Cc) and Cirrostratus (Cs) for high cloud layers. These are typically thin and white in appearance but can appear in a magnificent array of colours when the sun is low on the horizon.
- Altocumulus (Ac), Altostratus (As), and Nimbostratus (Ns) for medium cloud layers. These are composed primarily of water droplets; however, they can also be composed of ice crystals when temperatures are sufficiently low.
- Cumulus (Cu) Stratocumulus (Sc), Stratus (St), and Cumulonimbus (Cb) for low cloud layers. In this case, the clouds are mainly composed of water droplets.

The layers overlap and their limits vary with latitude. Table 2 indicates the approximate limit heights [14,28,31].