Publication Year	2019
Acceptance in OA	2021-02-02T11:57:31Z
Title	Measurements of High Energy Cosmic Rays and Cloud presence: A method to estimate Cloud Coverage in Space and Ground-Based Infrared Images
Authors	ANZALONE, ANNA, BRUNO, ALESSANDRO, Isgrò, F.
Publisher's version (DOI)	10.1016/j.nuclphysbps.2019.07.017
Handle	http://hdl.handle.net/20.500.12386/30152
Journal	NUCLEAR AND PARTICLE PHYSICS PROCEEDINGS
Volume	306-308

Available online at www.sciencedirect.com
ScienceDirect
Nuclear Physics B Proceedings Supplement 00 (2019) 1–8
Nuclear Physics B Proceedings Supplement

Measurements of High Energy Cosmic Rays and Cloud presence: A method to estimate Cloud Coverage in Space and Ground-Based Infrared Images

A. Anzalone[a], A. Bruno[a], F. Isgrò[b]

[a] *Istituto di Astrofisica Spaziale e Fisica Cosmica IASF-PA/INAF, Palermo via U.La Malfa 153, 90146 Palermo, Italy*

[b] *Dipartimento di Ingegneria Elettrica e Tecnologie dell'Informazione, Università di Napoli Federico II, via Claudio 21, 80125 Napoli, Italy*

Abstract

Several projects and already-operative observatories aimed at detecting High Energy Cosmic Rays (HECR) are/will be equipped with instruments to monitor the atmosphere. Since cloud presence can affect the night-time indirect measurements of the HECRs and Cherenkov radiation, it is crucial to know the meteorological conditions during the observation period of the HECR detectors. Several meteorological satellites already provide useful information, however to obtain accurate reconstructions of the detected events it is more suitable using devices that operate synchronously with the main detector. To this purpose, infrared cameras that acquire images of the whole field of view are thought to support the atmosphere monitoring during observations from both space and ground. Meaningful parameters, like cloud coverage and cloud top/bottom height, can be retrieved from the analysis of those data. Multi-spectral information are typically analyzed and combined to obtain cloud masks for each image, where a cloudy/cloud-free probability flag is associated with each pixel. These algorithms normally use several spectral bands that are not always available in non-meteorological sensors. Th a different approach is presented in this paper. It only relies on the gray level values of the image pixel, and it can be applied on thermal infrared as well as visible images acquired from both space and ground. To test the method on real cloudy scenes, images from polar satellite and all-sky data archives are considered, and the results are compared to the corresponding cloudiness masks provided by the same data repositories.

Keywords: Cloud Detection, Atmosphere Monitoring, Infrared Imaging

1. Introduction

Accurate cloud observations and their property analysis are necessary to understand the climate system, to validate models and to monitor weather change. Nevertheless the role that clouds play in the Earth radiative balance it is not the only aspect that other scientists apart from the meteorologists, are interested in. Cloud presence has in fact, a great influence on the detection of high energy cosmic rays and Cherenkov radiation, from ground and from space. Future space-based missions such as the Joint Experiment Missions for Extreme Universe Space Observatory (JEM-EUSO) [1] [2], already operative observatories from ground such as the Pierre Auger Observatory [3] and future ground-based projects such as Cherenkov Telescope Array (CTA) [4], and Astrofisica con Specchi Replicanti a Tecnologia Italiana (ASTRI) [5], need to include suitable systems to monitor the

cloud coverage in the field of views of their telescopes. The night-time observations of the fluorescence light produced when an UCHR reaches the atmosphere can be affected by the presence of clouds, depending on their height and depth. This is one among the reasons why cloud coverage [6], cloud top height [7] and cloud top temperature [8] detection still represent a challenge for the scientific community. The light intensity and the temporal evolution of the detected events can be correctly reconstructed if the data acquired by the atmosphere monitoring system (such as LIDAR - Light Imaging Detecting And Ranging, infrared cameras and so on) are elaborated to provide all the necessary information. In this article we focused our attention on the extraction of cloudiness coverage from images of satellite and ground cameras.

In most common approaches multi-spectral information are typically analyzed and combined to obtain cloud masks for each image, where a cloudy/cloud-free probability flag is associated with each pixel. These algorithms normally provide information by way of multispectral test results, but all bands are not always available in non-meteorological sensors. Therefore a different approach is presented in this paper that relies only on the gray level values of the image pixel, and for this reason it can be applied on thermal infrared as well as visible images.

We aim at detecting cloudiness map taking into account the following considerations: regarding the relations between clouds and the Earth surface, clouds often show higher reflectance and lower temperature values than the underlying Earth surface; simple visible and infrared window threshold approaches are generally reliable in cloud detection but not always sufficient; some cloud types such as thin cirrus and low-level stratus at night, and small-scale cumulus are difficult to detect because of insufficient contrast with the surface radiance. Cloud edges increase difficulty since the instrument field of view is not completely cloudy or clear.

In this paper we tackle the cloudiness coverage detection with the purpose of achieving results as accurate as possible by using only one feature, the brightness temperature pixel value as an example. The proposed solution is based on three main steps:

- Brightness Temperature histogram analysis for detecting the K main modalities of the image;
- Pixel Region Clustering through K-means;
- Cloud segmentation through a simple thresholding.

The rest of the paper is organized as follows: in section 2 an overview of the state of the art methods for cloud detection is given; in section 3 each step of the proposed method is described; in section 4 the most important findings are highlighted; section 5 ends the paper with conclusions and future works.

2. State of the art

Most of state of the art methods on cloud detection can be grouped as follows: radiometric feature based; multispectral band based; multiple image based (different dates); contextual pixel and texture based; object classification based; machine learning and deep learning methods.

Satellite imagery products such as those from MODIS (Moderate Resolution Imaging Radiometer) [23] and AATSR (Advanced Along-Track Scanning Radiometer) [9] data, and from satellite missions like Sentinel-3 [10], provide cloudiness masks, extracted as a result of both radiative and atmospheric property analysis. The cloudiness masks are with different spatial resolution depending on the camera sensor, and are obtained by combining a big amount of info from several bands.

Zhu et al. [6] proposed a method for the improvement and expansion of the Function of Mask (Fmask) Algorithm. Foga et al. [11] created a new validation dataset consisting of 96 Landsat 8 scenes, representing different biomes and proportions of cloud cover. They found that CFMask, a C code based version of the Fmask algorithm, and its confidence bands have the best overall accuracy among many algorithms tested using several validation data.

A method for detecting cloud contamination in radiances measured by high-spectral-resolution infrared sounders is proposed by McNally and Watts [12]. Other than classic research methods involved to detect cloud clearing or cloud parameters, McNally and Watts proposed to detect cloud in radiance data by infrared sounders.

In [13] the authors perform cloud detection by employing visual saliency methods, based on visual attention mechanisms. The document in [14] describes all 36 channels from MODIS for multispectral approaches to cloud detection. Bulgin et al. [15] presented five cloud detection algorithms over land based on dynamic threshold or Bayesian

techniques, applicable to AATSR instrument and compared these with the standard threshold based cloud detection scheme. Bulgin et al. [16] also proposed a Bayesian image classification scheme for discriminating cloud, clear and sea-ice observations at high latitudes to improve identification of areas of clear-sky over ice-free ocean for SST retrieval.

A novel multilevel cloud detection method based on deep learning is proposed by Xie et al. [17] for remote sensing images. It consists of a simple linear iterative clustering (SLIC) method and a Convolutional Neural Network (CNN) with two branches, they are combined to predict superpixels as one of three classes including thick cloud, thin cloud, and non cloud. Yabuki et al. [18] focused their effort on developing a cloud detection method from whole-sky color images. The method is based on spectral contrast between the RGB components of color all-sky images. The work presented by Chauvin et al. [19] is about the improvement of the cloud cover estimation, which is the main step in direct normal irradiance forecasting on all-sky imagery.

3. Materials and Methods

The main task of the method is to provide a general method for cloud coverage reconstruction starting from images taken by a camera sensor. Although our approach is not strictly dependent on the chosen band, here we propose an application on the thermal band because of more interest for the above mentioned astrophysical projects. The overall scheme of the method is as follows:

- Image Histogram Analysis to detect the K main Modalities;
- Pixel Clustering (K-Means) based on Gray Level Intensity analysis;
- Threshold-based Binarization to obtain the Cloudiness Mask.

We started from the following considerations: the brightness temperature values of BT image pixels are in most cases function of the real temperature of the area captured on an infrared sensor camera; the hotter the region captured, the higher the BT pixel value is expected. Since clouds are generally cooler than most elements (all but land ice) captured on an infrared camera, it is expected to find coherent values with respect the BT intensity values. Therefore, the cloudy pixels show lower BT values (or at most equal) than the sea surface, soil, vegetation pixels. Hence, BT values change with respect to different macro-categories (high, low, middle cloud, sea, land); high cloud pixel (figure 1) and low cloudy pixel (figure 2) usually show lower BT values than the sea surface pixels (figure 3).

3.1. Histogram Analysis

As it can be seen in figures 1-3, the elements of interest for atmospheric monitoring such as sea, high, middle and low clouds are characterized by distinct gray-level intensity values. Our aim is to detect clouds from BT images by considering the gray-level distribution. For this purpose we started by analyzing the histogram of the images looking for a direct relation between the image modalities and the elements of interest of our research, that is the clouds.

As an example, in figure 4 is shown an histogram that provides a complete description of the gray-levels of a MODIS image taken on 7 January 2018 including details related to small signal fluctuations. Since we are interested to analyze the main modalities of the histogram, we decided to use a straightforward histogram sub-sampling. The idea behind the sub-sampling is to extract the most important peaks of the histogram by neglecting the smallest fluctuations. To accomplish this, a simple signal sub-sampling is sufficient to highlight only the most prominent fluctuations of the given signal. So it is intuitive to compute the number of K main modalities of an image as the number of peaks from the resulting histogram (see the number of peaks in figure 5). The size of the sub-sampling step has been experimentally determined, further details will be given in the next sections.

3.2. Pixel Clustering

In our work we adopt the k-means algorithm [20] to clustering pixels of the images. For our purpose we assume that the k parameter to be used with k-means algorithm, corresponds to the number of the main modalities described in the previous section.

K-means allows to label each pixel value within a corresponding cluster, this is the first step to visualize the structural elements of the image (low, middle and high clouds, land and sea), as it can be seen in figure 7. New pixel values are assigned within the normalized range [0,1] and they are functions of the original gray-level intensity values. Let's consider, just for instance and sake of clarity, that an image

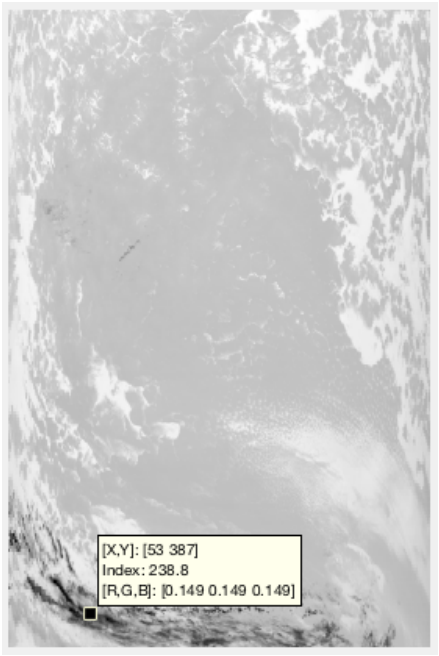

Figure 1: Bightness Temperatures from MODIS, taken using the 11-µm band on 22 March 2017 - 17:04:05 UTC

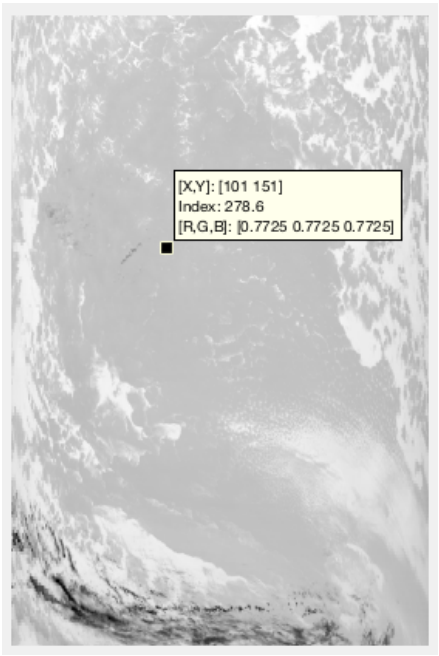

Figure 2: Brightness Temperatures from MODIS, taken using the 11-µm band on 22 March 2017 - 17:04:05 UTC

mainly contains pixels of elements such as clouds (low, middle and high), land and sea. It is observed that pixels falling within land and sea regions will

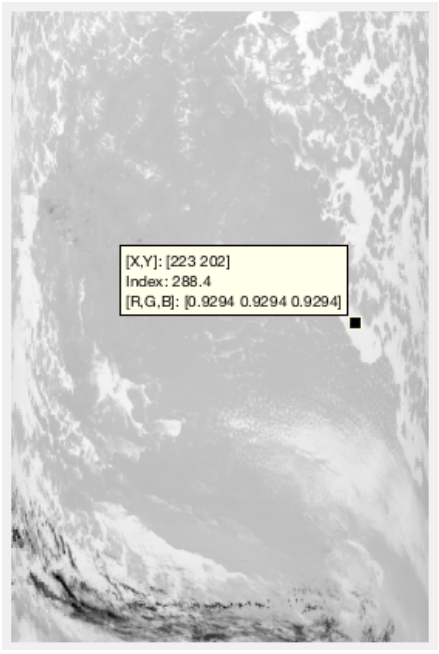

Figure 3: Brightness Temperatures from MODIS, taken using the 11-µm band on 22 March 2017 - 17:04:05 UTC

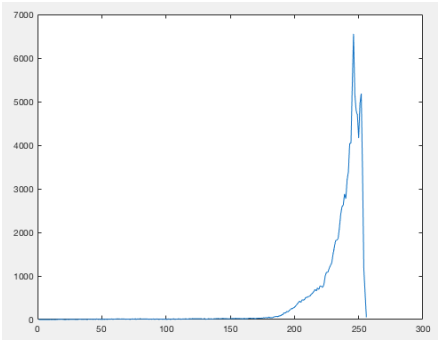

Figure 4: Histogram of the BT image from MODIS taken using the 11-µm band on 7 January 2018 -12:00 UTC

often show intensity values different from the ones falling within cloudy regions. Following this hypothesis on the correspondences between the clusters and the regions of the image, we expect that pixels laying on the same region will be labelled with the same label values. In a nutshell, we can say that we use k-means clustering as a sort of labelling step to generate a labelled image. By comparing the images in figures 6 and 7 we can see that pixel label values still maintain a coherent relation with respect to the original pixel intensity values. It means that the overall performance of k-means by using a proper K value related to the number of

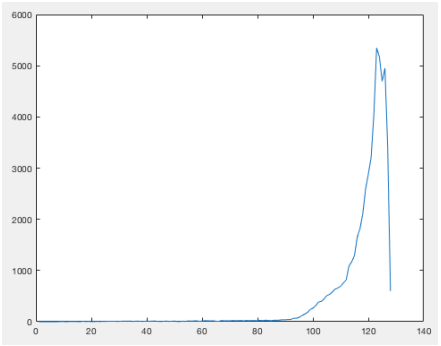

Figure 5: Histogram sub-sampling allows to detect the K main modalities of the BT Image

the main modalities is able to achieve good results in terms of image labelling.

It is necessary to mention that the performance of k-means on the images we treat depends on several factors. It comes out that the best conditions for a good clustering are when the intensity values of each region are well separated. The worst case is when pixels of regions such as sea and low cloud regions have quite similar temperature values, giving rise to a sort of misleading clustering. The aforementioned case is when we have both clouds and sea-ice, differentiating the spectral signatures of cloud and sea-ice is a well known problem in the remote sensing community. When pixels belonging to different atmospheric categories (i.e. sea and low clouds) are assigned to the same cluster it means that two different classes are mixed up together. This case is just to describe how much important is the role played by the clustering phase.

3.3. Thresholding

As mentioned in the previous section, we adopt the k-means algorithm to cluster the pixels of BTs. For a given image, the K-means algorithm returns an image labelled with values falling into the continuous range [0,1]. The values are functions of the real temperature of the regions captured by the cameras. In most of the cases we noticed a certain distance between the hottest regions (sea and land) and the coldest regions (clouds) in terms of pixel gray-level intensity. Since k-means returns clustered pixel values as a function of BTs, we set the lowest label value as threshold to be applied to obtain a binary cloud mask. The idea behind our choice is that in most cases we have at least three atmospheric elements (sea, cloud, land). Our preliminary experimental results show that clouds are

Figure 6: Input BT image taken by MODIS using the 11-µm band on 7 January 2018 - 12:00 UTC

usually cooler than seas and lands, they are clustered with the lowest labels into the range [0,1]. With the exception of land, sea and lake ice, most BT images show cloud regions having lower gray-scale values than other regions. Since clouds and sea-ice may have some similar spectral signatures we have been dealing with the building of a new classifier suited for the recognition of clouds from sea ice. If we refer to I(k) as the gray-scale value of the clustered pixels of an image, for a given number of clusters N we choose the threshold value Th as in equation 1.

$$Th(s) = Min\{I(k) : k = 1, ..., N\} \quad (1)$$

Therefore, for each image, the lowest cluster label value is set to be the threshold value. In the next section a more detailed description of the experimental results will be given in terms of precision and recall with respect to the binary masks from MODIS repositories.

4. Results and Validation

A set of data across several repositories of remote sensing images has been used to validate our proposed method. We conducted several tests on

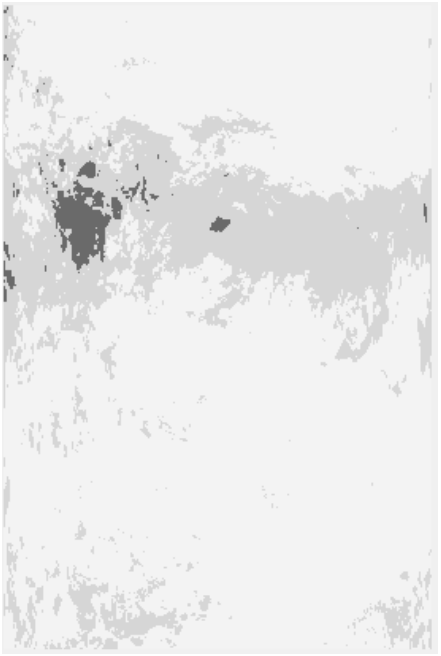

Figure 7: Clustered BT image

Figure 8: Cloudiness Mask extracted by our Method

images coming from MODIS onboard Terra and Aqua satellites [23] and ground based images from PRISMA (Prima Rete per la Sorveglianza sistematica di Meteore e Atmosfera) project [21]. Several products, such as MOD06 and MOD32 [22] of

Figure 9: MODIS Cloud Mask used as Ground-truth

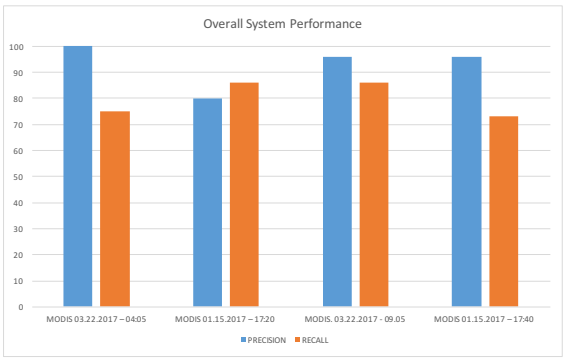

Figure 10: Overall system performance with precision and recall metrics

MODIS, provide the corresponding cloud mask for any given BT image with a good reliability since it is computed by using several features from both radiative and non-radiative processes. Precision and recall metrics ((2), (3)) give an effectiveness measure by considering the number of TP (True Positives), FP (False Positives), FN (False Negatives). We obtained Precision and Recall values by comparing the results of our method with the MODIS masks used as ground-truth.

A quite general and qualitative evaluation can be carried out by analyzing the images in figures 6-9 where we show the input BT image from MODIS

(figure 6) the clustered image (figure 7), the cloudiness mask (figure 8) and the ground-truth reference (figure 9).

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

Precision and recall values remain at substantial levels, 90% precision and %75 recall on average. It means that our system is able to detect high number of true positive (TP) but still shows a non negligible number of false negatives (FN).

Summing up we can say that a general purpose method based on low level image processing is proposed respectively through the detection of the main modalities of the image, the k-means clustering of the pixels and a simple binarization based on an adaptive threshold value as a function of the cluster label values. The proposed method has been evaluated over some images from MODIS repository, it comes out that we still need to improve the detection of the lowest clouds of the regions of interest in some phase misclassified with regions other than clouds. We need to work on boosting the difference between non cloudy pixels and the pixels falling within the lowest cloud regions (i.e. discriminating pixels belonging to regions showing quite similar spectral signatures). Since our first interest was to maintain a cloud free conservative approach in estimating the sky cloud coverage we chose the lowest clustering label value as threshold value. It is likely that this choice makes us lose some true positives from those regions showing BT values close to sea and land pixel values.

We have also been working with all-sky images, that is a particular kind of images taken from ground, we have been focusing our attention on the images within Prisma project. Preliminary studies and results of Prisma data are encouraging but we still need some preprocessing step to keep the scene clear of static objects such as antennas.

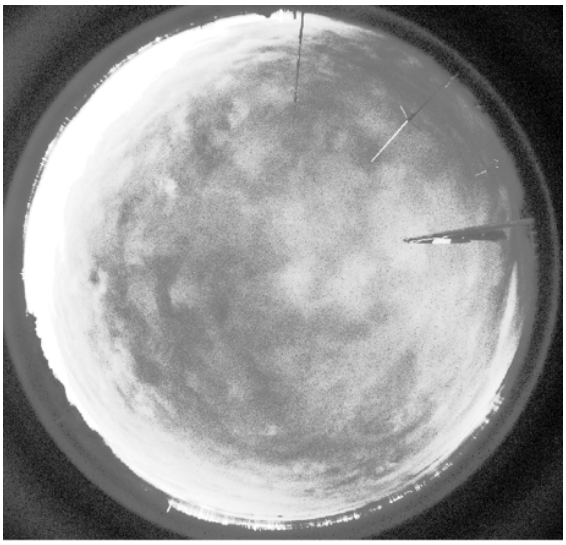

Figure 11: An all-sky infrared image taken from Prisma project

5. Conclusions

The proposed method allows to achieve good performance in terms of cloud detection accuracy on remote sensing imagery products such as those from MODIS, using only information from BT images. As shown by the quality metrics in figure 10, the method reaches good overall precision and recall. Despite a low number of false positives, there is still a non negligible number of false negatives given by the some misclassified pixels likely belonging to those regions showing quite similar spectral signatures. In future works a further experimentation will be needed to improve the overall accuracy in the cloud detection by adopting several clustering methods other than k-means like superpixel clustering and k-medoids and combining the results with a voting mechanism system.

6. Acknowledgements

The authors acknowledge the MODIS Science team for the Science Algorithms, the Processing Team for producing MODIS data, and the GES DAAC MODIS Data Support Team for making MODIS data available to the user community.

References

[1] J. H. Adams, Jr., et al.,The JEM-EUSO instrument, Experimental Astronomy, vol.40, no.1, pp.19–44, (2015).

[2] J. H. Adams, Jr., et al., The atmospheric monitoring system of the JEM-EUSO instrument, Exp. Astron., vol.40, no.1, pp.45–60, (2015).

[3] A.Aab and et al.(Pierre Auger Coll.), The Pierre Auger cosmic ray Observatory,Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, Vol.A798, pp. 172-213, (2015)

[4] CTA Cherenkov Telescope Array [Online]. Available: https://www.cta-observatory.org/about/

[5] ASTRI Project https://www.cta-observatory.org/cta-prototype-telescope-astri-achieves-first-light/
[6] Zhu, Zhe and Wang, Shixiong and Woodcock, Curtis E, Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images, Remote Sensing of Environment, vol.159, pp.269-277, Elsevier, (2005)
[7] A. Anzalone, M.E. Bertaina, S. Briz, C. Cassardo, R. Cremonini, A.J. de Castro, S. Ferrarese, F. Isgrò, F. López, I. Tabone, Methods to Retrieve the Cloud-Top Height in the Frame of the JEM-EUSO Mission, IEEE Transactions on Geoscience and Remote Sensing, n. 99, pp. 1-15, IEEE, (2018)
[8] Wei, Chiang and Yeh, Hui-Chung and Chen, Yen-Chang, Assessment on spatiotemporal relationship between rainfall and cloud top temperature from new generation weather satellite imagery,EGU General Assembly Conference Abstracts, vol. 19, pag.7155, (2017)
[9] AATSR: Global-change and surface-temperature measurements from Envisat, Llewellyn-Jones, D and Edwards, MC and Mutlow, CT and Birks, AR and Barton, IJ and Tait, H, ESA bulletin, vol.105,no.10-21, pag. 25, (2001)
[10] The global monitoring for environment and security (GMES) sentinel-3 mission, Donlon, Craig and Berruti, B and Buongiorno, A and Ferreira, M-H and Féménias, P and Frerick, J and Goryl, P and Klein, U and Laur, H and Mavrocordatos, C and others,Remote Sensing of Environment, vol.120, pp.37–57, Elsevier (2012)
[11] Foga, Steve and Scaramuzza, Pat L and Guo, Song and Zhu, Zhe and Dilley, Ronald D and Beckmann, Tim and Schmidt, Gail L and Dwyer, John L and Hughes, M Joseph and Laue, Brady, Cloud detection algorithm comparison and validation for operational Landsat data products, Remote sensing of environment, vol.194, pp. 379-390, Elsevier, (2017)
[12] McNally, AP and Watts, PD, Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography, vol. 129, n.595, pp.3411–3423, Wiley Online Library, (2003)
[13] Hu, Xiangyun and Wang, Yan and Shan, Jie, Automatic Recognition of Cloud Images by Using Visual Saliency Features, IEEE Geosci. Remote Sensing Lett., vol.12, n.8, pp.1760–1764, (2015)
[14] Discriminating clear-sky from cloud with modis algorithm theoretical basis document (mod35), Team, MODIS Cloud Mask and Ackerman, Steve and Strabala, Kathleen and Menzel, Paul and Frey, Richard and Moeller, Chris and Gumley, Liam and Baum, Bryan and Schaaf, Crystal and Riggs, George, ATBD Ref. ATBD-MOD-06, version, vol. 4, p. 115p, Citeseer, (1997)
[15] Bulgin, CE and Sembhi, H and Ghent, D and Remedios, JJ and Merchant, CJ, Cloud-clearing techniques over land for land-surface temperature retrieval from the Advanced Along-Track Scanning Radiometer, International journal of remote sensing, vol.35, number 10, pp. 3594–3615, Taylor & Francis, (2014)
[16] Bulgin, C. E., Eastwood, S., Embury, O., Merchant, C. J. and Donlon, C. (2015) Sea surface temperature climate change initiative: alternative image classification algorithms for sea-ice affected oceans. Remote Sensing of Environment, 162. pp. 396-407. ISSN 0034-4257
[17] Xie, Fengying and Shi, Mengyun and Shi, Zhenwei and Yin, Jihao and Zhao, Danpei, Multilevel cloud detection in remote sensing images based on deep learning, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol.10, n. 8, pp. 3631–3640, IEEE, (2017)
[18] Development of a cloud detection method from whole-sky color images, Yabuki, Masanori and Shiobara, Masataka and Nishinaka, Kimiko and Kuji, Makoto,Polar Science, vol.8, no.4, pp.315–326, Elsevier (2014)
[19] Chauvin, Remi and Nou, Julien and Thil, Stephane and Traore, Adama and Grieu, Stephane, Cloud detection methodology based on a sky-imaging system, Energy Procedia, vol.69, pp.1970–1980, (2015)
[20] Alsabti, Khaled and Ranka, Sanjay and Singh, Vineet, An efficient k-means clustering algorithm, (1997)
[21] http://www.prisma.inaf.it/
[22] MODIS Cloud Optical Properties: User Guide for the Collection 6 Level-2 MOD06/MYD06 Product and Associated Level-3 Datasets. Accessed: Feb. 13, 2017. [Online] Available:https://modis-atmosphere.https://modis-atmos.gsfc.nasa.gov/sites/default/files/ModAtmo/C6MOD06OPUserGuide_1.pdf
[23] MODIS (MODerate Resolution Imaging Spectroradiometer) NASA. [Online]. Available: http://modis.gsfc.nasa.gov/