Hydrol. Earth Syst. Sci. Discuss., 10, 13671–13706, 2013
www.hydrol-earth-syst-sci-discuss.net/10/13671/2013/
doi:10.5194/hessd-10-13671-2013

Hydrology and Earth System Sciences
Open Access
Discussions

A statistical approach for rain class evaluation using Meteosat Second Generation-Spinning Enhanced Visible and InfraRed Imager observations

E. Ricciardelli[1], **D. Cimini**[1,2], **F. Di Paola**[1], **F. Romano**[1], **and M. Viggiano**[1]

[1]National Research Council of Italy – Institute of Methodologies for Environmental Analysis, c. da S. Loja, 85050 Potenza, Italy
[2]CETEMPS, University of L'Aquila, via Vetoio 1, 67100 L'Aquila, Italy

Received: 8 October 2013 – Accepted: 2 November 2013 – Published: 12 November 2013

Correspondence to: E. Ricciardelli (elisabetta.ricciardelli@imaa.cnr.it)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

Precipitation measurements are essential for short term hydrological and long term climate studies. Operational networks of rain gauges and weather radars provide fairly accurate rain rate measurements, but they leave large areas uncovered. Because of this, satellite remote sensing is a useful tool for the detection and characterization of the raining areas in regions where this information remains missing. This study exploits the Meteosat Second Generation – Spinning Enhanced Visible and Infrared Imager (MSG-SEVIRI) observations to evaluate the rain class at high spatial and temporal resolutions. The Rain Class Evaluation from Infrared and Visible (RainCEIV) observations technique is proposed. The purpose of RainCEIV is to supply continuous monitoring of convective as well as of stratiform rainfall events. It applies a supervised classifier to the spectral and textural features of infrared and visible MSG-SEVIRI images to classify the cloudy pixels as *non rainy, light to moderate rain*, or *heavy to very heavy rain*. The technique considers in input also the water vapour channels brightness temperatures differences for the MSG-SEVIRI images acquired 15/30/45 min before the time of interest. The rainfall rates used in the training phase are obtained with the Precipitation Estimation at Microwave frequencies (PEMW), an algorithm for rain rate retrievals based on Atmospheric Microwave Sounder Unit (AMSU)-B observations. The results of RainCEIV have been validated against radar-derived rainfall measurements from the Italian Operational Weather Radar Network for some case studies limited to the Mediterranean area. The dichotomous assessment shows that RainCEIV is able to detect rainy areas with an accuracy of about 91 %, a Heidke skill score of 56 %, a Bias score of 1.16, and a Probability of Detection of rainy areas of 66 %.

HESSD

10, 13671–13706, 2013

1 Introduction

Precipitation measurements are essential for many applications from short term hydrological to long term climate studies, including the near-real-time monitoring of rainfall events that is crucial for timely alerting and early intervention. Rainfall estimation is accomplished by operational networks of rain gauges and weather radars, which provide fairly accurate rain rate measurements characterized by high temporal and spatial resolutions. However, rain gauges limit the measurements to certain points only, and weather radars leave large areas uncovered (as remote land regions and oceans) where information on the occurrence and intensity of rainfall remains therefore missing. Because of these limitations, scientists have increasingly turned to satellite remote sensing as a possible mean for the detection and characterization of the raining areas. Over the past decades, several rain rate retrieval methods based on passive radiometer observations from satellite have been developed. Microwave (MW) observations from Low Earth Orbiting (LEO) satellites provide more direct and performing techniques for the retrieval of precipitation compared to Infrared/Visible (IR/VIS). However, the low spatial and temporal resolutions of MW observations from LEO satellites make them unsuitable for monitoring extreme and small-scale events characterized by a high spatial and temporal variability. MW observations for geostationary platforms have been proposed (Gasiewski et al., 2003; Lambrigtsen et al., 2006; Bizzarri et al., 2007) but not launched to date. Therefore, rain rate estimates from passive IR/VIS imagers on geostationary satellite may be useful to bridge the gap between LEO MW and weather radar rainfall observations. Many techniques based on geostationary satellite IR/VIS observations have been developed in order to estimate rain rate values or confidences. A recent overview is given by Kidd and Levizzani (2011). The geostationary satellite techniques perform better over areas where rainfall originates from deep convection than in areas where it originates from the stratiform systems. In particular, Negri and Adler (1981) examined relation between cloud top temperature and rain rate by analysing Geostationary Operational Environmental Satellite (GOES) and radar

HESSD
10, 13671–13706, 2013

data associated to a sample of thunderstorms. Adler et al. (1985) proposed a Thunderstorm Index (TI) to give probability to observe heavy precipitation. Successively, Adler et al. (1988) extended their interest to stratiform precipitation (produced under the anvils of mature and decaying convective systems) from GOES satellite infrared data. Adler and Mack (1984) investigated a relation between radar based rain-rate and cloud height determined from GOES satellite data. Wu and Weinman (1985) used GOES data in order to estimate rainfall by means of a pattern recognition algorithm trained and tested on different sets of rain rate measurements obtained from NOAA operational radars. They classify rain into three classes (non rainy, light rain, heavy rain). Adler et al. (1993) were the first to successfully combine the advantages of both types of instrument by using matched MW and IR data. Vicente et al. (1998) introduced the auto-estimator in order to estimate rainfall from GOES measurements focusing on heavy precipitation. The auto-estimator differs from previous IR methods for rainfall estimation because it considers other factors in addition to IR window cloud top temperature. In particular, information about environmental moisture are used to obtain a more correct estimation of rainfall as well as for the screening of no-rainy pixels. Mamoudou and Gruber (2001) used the GOES visible (0.65 µm), near infrared (3.9 µm), water vapour (6.7 µm) and windows channels (10.7 and 12.0 µm) to estimate rainfall rate. They distinguished raining from no-raining clouds by taking into account the cloud top temperature, the effective radius of cloud particles and the temperature gradient. Moreover, in an attempt to give more reliable values of rain rates, Mamoudou and Gruber (2001) used the moisture factor correction developed by Scofield (1987) and modified by Vicente et al. (1998). Other authors have used artificial neural networks to derived precipitation estimates using satellite IR images (Hsu et al., 1997). Many authors developed techniques to determine rain rate from Meteosat data, both physical and statistical. Physical techniques consist of brightness temperature differences threshold tests or consider effective radius as well as cloud top height/temperature in order to determine rainfall rate and/or probability by the use of look-up tables. The look-up tables usually are build by considering rainfall measurements obtained by rain-gauges instru-

HESSD
10, 13671–13706, 2013

ments or radar as well as rain rate values determined by MW data. An example of IR method that uses rain rate values determined by MW observations was developed by Jobard and Desbois (1994), the RAin and Cloud Classification method (RACC), that used the SSM/I and Meteosat data in order to classify the Meteosat images into several categories of rain. Turk et al. (1999) proposed a blended geostationary-microwave technique for the retrievals of rain rate measurements. This technique has been taken as role model by several investigators (Kidd et al., 2003; Marzano et al., 2004), including Heinemann et al. (2002) who developed the Multi-Sensor Precipitation Estimate (MPE) technique operating at the European agency for the deployment of meteorological satellites (EUMETSAT). MPE product consists of the near-real-time rain rates for each Meteosat Second Generation (MSG)-Spinning Enhanced Visible and Infrared Imager (SEVIRI) images in original pixel resolution. Kuhnlein et al. (2010) also investigated the SEVIRI potential to determine rain rates, assuming a relationship between rain rates and optical thickness as well as effective radius. In particular they have established a relation between the reflectance observations acquired at 0.6 and 1.6 µm – giving information about cloud optical thickness and effective radius – and ground-based rainfall rate. Recently, Feidas and Giannakos (2012) have proposed an algorithm that works with SEVIRI observations by combining physical and statistical methods to characterize convective and stratiform precipitation areas. They calibrated the algorithm using rain rate measurements derived from a substantial number of rain gauge stations in Greece. Other techniques are based on cloud motion, exploiting IR observations to provide an estimate of cloud movement to be used for advecting the more direct MW rainfall observations (Joyce et al., 2004). Di Paola et al. (2012) proposed the Precipitation Evolving Technique (PET) for convective rain cell continuous monitoring. The PET propagates forward in space and time the last rain rate map inferred by AMSU and MHS MW observations by using SEVIRI IR brightness temperature maps. This technique is able to propagate forward in time the last available rain field for 2–3 h.

The aim of this study is to propose a technique based on a statistical classification algorithm that uses spectral and textural features of SEVIRI IR/VIS observations to

HESSD
10, 13671–13706, 2013

classify the cloudy pixels as *non rainy*, *light to moderate rain*, or *heavy to very heavy rain*. The proposed technique, the Rain Class Evaluation from Infrared and Visible observations (RainCEIV) operates in a fixed area, that is the Mediterranean basin, approximately between 35 and 45 °N, and −20° W and 20 °E. RainCEIV firstly discriminates cloudy from not-cloudy pixels, then it determines the rain class only for the pixels classified as cloudy. It deploys the *k*-Nearest Neighbour Mean classifier (*k*-NNM) with input the spectral and textural features determined from the SEVIRI VIS/IR images and the brightness temperatures differences of SEVIRI water vapour channels acquired 15, 30, and 45 min before the time of interest. The RainCEIV has been validated against the radar-derived rain rate values obtained from the Italian *Operational Weather Radar Network* observations managed by the Italian Department of Civil Protection (DPC). RainCEIV is proposed as an useful tool to achieve real-time monitoring of rainfall events, both the intense convective and the stratiform moderate ones.

Section 2 provides a description of the satellite sensors whose observations and/or products have been used for the RainCEIV implementation; Sect. 3 describes the two modules of RainCEIV algorithm (C_MACSP cloud classification algorithm and the RainCEIV *k*-NNM classifier); Sect. 4 shows the statistical scores obtained by comparing RainCEIV and radar-derived rain rate measurements.

2 Instruments and data description

The spectral and textural features of the MSG-SEVIRI images are used as input for both the MACSP cloud classification algorithm and the RainCEIV *k*-NNM classifier. SEVIRI is the main payload on board the MSG series, composed by MSG-1 (Meteosat 8), launched in February 2003, MSG-2 (Meteosat 9), launched in December 2005, MSG-3 (Meteosat 10) launched in July 2012, and future MSG-4 (Meteosat 11), planned for launch in 2014. MSG-2 was designated as the first satellite on 11 April 2007. The SEVIRI is a 50 cm-diameter aperture line-by-line scanning radiometer. It observes the Earth-atmosphere system in 11 channels at full-disk with a spatial sampling of 3 km at

HESSD
10, 13671–13706, 2013

the sub-satellite point. In addition, the High Resolution Visible (HRV) channel covers half of the full disk with a 1 km spatial sampling at the sub-satellite point. The actual instantaneous field of view is about 4.8 km at the sub-satellite point for all channels except 1.67 km for the HRV channel. The major improvements with respect to previous sensors are its enhanced spectral characteristics, its higher temporal resolution (15 min), the improved signal-to-noise ratio, and the higher precision of data storing, going from 8 bits (256 levels) on Meteosat-7 to 10 bits (1024 levels) on Meteosat-8 (Schmetz et al., 2002).

The RainCEIV *k*-NNM classifier has been trained on the rain rate product from the Precipitation Estimation at Microwave Frequencies (PEMW). PEMW was developed by Di Tomaso et al. (2009) at the Institute of Methodologies for Environmental Analysis of the National Research Council of Italy (IMAA-CNR) for inferring surface rain intensity from satellite MW LEO observations. These are provided by the Advanced Microwave Sounding Unit-B (AMSU-B) and the Microwave Humidity Sounder (MHS) on board the National Oceanic and Atmospheric Administration (NOAA) satellites and the European Polar Satellite MetOp-A, respectively. AMSU-B and MHS are cross-track, line scanning microwave radiometers measuring radiances in five channels in the frequency range from 89 to 190 GHz. The centre frequencies for the two window channels are 89, 150 GHz, while the three opaque (water vapour) channels are centred at 183 ± 1, 183 ± 3, and 183 ± 7 GHz. The AMSU-B and MHS fields of view have a circular shape (with a diameter of about 16 km) at nadir, while away from the nadir their shape become ellipsoidal (the axes length is 51 km for the cross-track direction and 25 km for the along-track direction at the maximum scanning angle) (Bennartz, 2000). The purpose of these instruments is to measure radiation from different layers of the atmosphere in order to obtain global data on humidity profiles.

The RainCEIV results have been validated on the basis of the rain rate values derived from the weather radar network operated by the Italian DPC in collaboration with regional authorities, research centres, the Air Traffic Control service (ENAV), and the Meteorological Service of the Italian Air Force (CNMCA). The radar network consists of

HESSD
10, 13671–13706, 2013

twenty microwave weather radars belonging to regional authorities (ten C-band radars), ENAV (two C-band radars) and DCP (six C-band radars and two X-band polarimetric radars). The surface rate intensity (sri, in mm h^{-1}) and other products such as the Vertical Maximum Intensity (VMI), the Constant Altitude Plan Position Indicator (CAPPI) and the one-hour accumulated surface rain total (srt, in mm), are retrieved from measured reflectivity volumes. All the products are available on a grid of 1400 km × 1400 km with a spatial resolution of circa 1 km and a temporal resolution of 15 min.

3 RainCEIV description

The RainCEIV technique consists of two modules:

I. a cloud classification algorithm that discriminates *clear* from *cloudy* pixels and further classifies the cloudy pixels;

II. a *k*-Nearest Neighbour Mean (*k*-NNM) classifier that evaluates the rain class for each pixel classified as cloudy by the first module.

3.1 Cloud classification algorithm description

The cloud Mask Coupling of Statistical and Physical methods algorithm – MACSP (Ricciardelli et al., 2008) – is used to distinguish *cloudy* from *no-cloudy* pixels. The version used for RainCEIV purposes is called C_MACSP, which stands for cloud Classification Mask Coupling of Statistical and Physical methods. The current version has been updated in order to give information about the cloud class and in particular to classify clouds belonging to the generic *high cloud* class into two additional classes (*high optically thin* and *high optically thick*). In addition, the *convective cloud* class has been added, not just for the module II but also to individuate the possible occurrence of extreme events. In details, a pixel can be classified in 5 different classes considered both over land and sea: *clear*, *low*/*middle cloud*, *high optically thin cloud*, *high optically thick cloud* and *convective cloud*. The C_MACSP screening is useful to:

HESSD
10, 13671–13706, 2013

– reduce the number of the input pixels to the RainCEIV *k*-NNM classifier by removing the pixels classified as *clear* and as *high thin cloud*;

– define the components of the feature vector in input to the RainCEIV classifier (as will be described in the following sub-section, the components chosen for each cloud class are shown in Table 2).

The training dataset used in the previous version of MACSP has been updated in order to get a better distinction of the cloudy classes. The cloud classification for the training dataset was made through a careful visual inspection of the SEVIRI images. The clear and cloudy pixels have been selected manually after observing the spectral characteristics in SEVIRI IR/VIS images as well as in the their RGB composition, a useful practice for distinguishing cloudy classes (Lensky and Rosenfeld, 2008). In order to get a reliable training dataset, the outliers have been removed by means of the Condensed Nearest Neighbour Rule (CNN) (Hart, 1968) and the cross-validation method has been applied so to refine it. Successively, the validity of the C_MACSP algorithm has been tested by applying it to an independent dataset made of nine SEVIRI images acquired on 5 May 2012 at 20:30 UTC, 19 May 2012 at 11:00 UTC, 23 July 2012 at 10:30 UTC, 5 December 2012 at 08:45 UTC, 19 September 2009 at 19:15 UTC, 6 July 2010 at 11:30 and 12:30 UTC, 4 August 2010 at 14:30 UTC. This test dataset has been built by selecting images characterized by various cloud types corresponding to the C_MACSP cloud classes. The validation has been carried out by comparing the C_MACSP classification results and the samples manually collected for each class from the above-mentioned images. The manual classification has been made through a careful observation of the SEVIRI RGB composition. The deep convective cloud classification results have been validated by considering the rain rate maps derived from the weather radar network operated by the Italian Civil Protection Department (DPC).

HESSD
10, 13671–13706, 2013

3.2 *k*-Nearest Neighbour Mean classifier description

The pattern classifier used to determine the confidence that a cloudy pixel belongs to a rainy/non-rainy class is the non-parametric supervised classifier *k*-Nearest Neighbour Mean (*k*-NNM) proposed by Viswanath and Sarma (2011). This classifier has been chosen because of its simplicity and good performance (Dasarathy, 1991, 2002; Babu and Viswanath, 2009) and because it does not assume any a priori known probabilities as in the case of the Bayes classifier, but it estimates these directly from the design samples. It implements the decision rule locally. The *k*-NNM classifier has demonstrated to perform better than the *k*-NN classifier and it is suitable for parallel implementation so to reduce classification time, as asserted by Viswanath and Sarma (2011).

Let $\boldsymbol{x}$ be the vector of features related to the pixel to be classified and C_i the *rainy/non-rainy* class with $i = 0,1,2$ defined as follows:

1. *non-rainy* (rain rate = 0) (C_0)
2. *light to moderate rain* ($0 <$ rain rate $\leq 4\,\mathrm{mm\,h^{-1}}$) ($C_1$)
3. *heavy to very heavy rain* (rain rate $> 4\,\mathrm{mm\,h^{-1}}$) (C_2).

The *k*-NNM classifier finds the k (where $k \geq 1$) nearest neighbours of $\boldsymbol{x}$ for each class C_i and determines the mean value $d_{\min}(\boldsymbol{x}, C_i)$ of their distances ($d(\boldsymbol{x}, \boldsymbol{x}_{i,j})$) for each class C_i:

$$d_{\min}(\boldsymbol{x}, C_i) = \frac{\sum_{j=1}^{k} d(\boldsymbol{x}, \boldsymbol{x}_{i,j})}{k} \quad i = 0,1,2 \tag{1}$$

where $d(\boldsymbol{x}, \boldsymbol{x}_{i,j})$ is the Euclidean distance between the vector of feature $\boldsymbol{x}$ to be classified and the j-th nearest training sample for the class C_i, $\boldsymbol{x}_{i,j}$. The pixel is labelled as the class characterized by the lowest mean distance $d_{\mathrm{mean}}(\boldsymbol{x}, C_i)$:

HESSD
10, 13671–13706, 2013

$$(d_{\text{mean}}(\boldsymbol{x}, C_0) < d_{\text{mean}}(\boldsymbol{x}, C_1)) \text{ and } (d_{\text{mean}}(\boldsymbol{x}, C_0) < d_{\text{mean}}(\boldsymbol{x}, C_2)) \rightarrow \boldsymbol{x} \in C_0 \quad (2)$$

$$(d_{\text{mean}}(\boldsymbol{x}, C_1) < d_{\text{mean}}(\boldsymbol{x}, C_0)) \text{ and } (d_{\text{mean}}(\boldsymbol{x}, C_1) < d_{\text{mean}}(\boldsymbol{x}, C_2)) \rightarrow \boldsymbol{x} \in C_1 \quad (3)$$

$$(d_{\text{mean}}(\boldsymbol{x}, C_2) < d_{\text{mean}}(\boldsymbol{x}, C_0)) \text{ and } (d_{\text{mean}}(\boldsymbol{x}, C_2) < d_{\text{mean}}(\boldsymbol{x}, C_1)) \rightarrow \boldsymbol{x} \in C_2. \quad (4)$$

Figure 1 shows the scheme of the RainCEIV technique.

3.2.1 Features selection and description

The *k*-NNM classifier uses textural and spectral features estimated in 3 × 3 -pixel boxes in order to associate each SEVIRI pixel to a rainy/no-rainy class. The textural and spectral features used in this study and their different weight in the grid element – that is a 3 × 3-pixel box where both textural and tonal features have significant values – are described in Ricciardelli et al. (2008). Textural and spectral features characterizing each pixel are extracted from IR and VIS SEVIRI images acquired at the following wavelengths: 0.6, 0.8, 1.6, 3.9, 6.3, 7.2, 10.8, and 12 µm. The SEVIRI observations provide very useful information about cloud top microphysical structure to characterize precipitation processes and then to determine rain rate confidences. In particular, features related to radiance acquired at 10.8, 12.0 and 0.6 µm are useful to individuate precipitation processes that develop in optically thick convective clouds with very cold tops. Moreover, features related to radiances acquired at 3.9 and 1.6 µm bear on the cloud drop size distribution as well as the thermodynamic phase, and thus they are useful to individuate precipitation from stratiform clouds with cloud tops warmer than the convective ones. In fact, in stratiform clouds the precipitation processes are strongly related to the microphysical structure of the cloud top and, in particular, rain rate confidence is high for cloud top with large cloud droplets or in presence of ice (Lensky and Rosenfeld, 1997). The channels centred at 6.3 and 7.2 µm are indicative of the water vapour content in the troposphere at levels lower than 350 and 500 hPa, respectively. The differences between the brightness temperatures for water vapour channels acquired 15, 30 and 45 min before the time of interest are exploited to get information

HESSD
10, 13671–13706, 2013

about the water vapour content changes at different levels in the atmosphere, so to characterize convective as well as stratiform precipitations. For this reason, the water vapour brightness temperature differences $\Delta TB_{(6.7),15-30}$, $\Delta TB_{(6.7),15-45}$, $\Delta TB_{(6.7),30-45}$, $\Delta TB_{(7.2),15-30}$, $TB_{(7.2),15-45}$, $TB_{(7.2),30-45}$, have been added in the features vector.

It is important to note that the features (x^i) characterized by the largest variance across the design set tend to dominate the Euclidean distance. Therefore, it is useful to normalize each feature by applying the following equation:

$$\widetilde{x}^i = \frac{x^i - \bar{x}^i}{\sigma^i} \tag{5}$$

where x^i is the i-th component of the features vector $\boldsymbol{x}$ to be normalized, $\widetilde{x}^i$ is the i-th component of the normalized $\widetilde{\boldsymbol{x}}$, $\bar{x}^i$ and σ^i are the mean and standard deviation for the feature x^i calculated for all the training set samples. This equation is applied also to the features vector related to the pixels to be classified.

By taking in mind that in general the k-NN classifier performance decreases with the dimension of the features vector, it is useful to reduce the numbers of the components (x^i) of $\boldsymbol{x}$. The Fisher criterion (Ebert, 1987; Parikh, 1977) has been applied in order to evaluate the discriminatory power of the individual features. Let $\bar{x}_j^i$ and σ_j^i be the mean and standard deviation of the feature x^i for the training set from class C_j, thus the Fisher distance is defined as:

$$D_{ijk} = \frac{\left|\overline{x_j^i} - \overline{x_k^i}\right|}{\left(\sigma_j^i - \sigma_k^i\right)}. \tag{6}$$

It measures the ability of the variable x^i to differentiate class C_j from class C_k. The Fisher distance has been determined for the following combinations: (C_0, C_1); (C_0, C_2); (C_1, C_2).

HESSD
10, 13671–13706, 2013

Jain and Chandrasekaran (1982) point out that the ratio between the number of the training samples for each class and the feature vector dimension d should be at least five. In practical situations, as asserted by Hamamoto et al. (1997), this ratio is not easily reached. Initially, the number of the training samples for each class was between 40 and 45 and we have decided to chose $d = 10$ to remove outliers. The definitive d and k values for the RainCEIV k-NNM classifier have been obtained as will be described in the next sub-section.

3.2.2 Training procedure

The training dataset was built by collecting a set of SEVIRI images during day- and night-time with co-located rain rate values inferred from AMSU-B observations processed with the PEMW algorithm (Di Tomaso et al., 2009), both over land and sea. PEMW exploits the observations made in both window and water vapour channels. PEMW estimates show a very good agreement with ground-based observations in the detection of rainfall and a reasonably good estimation of rain rate values. The Probability of Detection (POD) of precipitation is 75 and 90 % for rain rates greater than 1 and 5 mm h^{-1}, respectively (Di Tomaso et al., 2009). At the present, the operative version of PEMW algorithm (OPEMW) is operationally run 24/7 at IMAA-CNR. The OPEMW has been validated by Cimini et al. (2013) against radar-derived rain rate and rain gauge surface rain intensity. The analysis shows an accuracy of 98 % in identifying rainy and non-rainy areas and a Heidke skill score of 0.45 (with respect to radar-derived rain rates) and 0.42 (with respect to rain gauge rain rates). The *accuracy*, *Bias score, Probability of Detection* (POD), *False Alarm Ratio* (FAR), *Heidke Skill score* (HSS) are described in Ebert (2013). The AMSU-B observations used for building the training database are collected during the NOAA satellite passes over the Mediterranean area on the dates listed in Table 3.

During the training phase the MSG-SEVIRI pixel closest to the AMSU-B FOV classified as rainy/non-rainy by the PEMW algorithm is assigned to one of the classes C_0, C_1, C_2 mentioned above.

HESSD
10, 13671–13706, 2013

Generally, the performance of a pattern recognition classifier depends strictly on the correctness of the training samples, on the number of the training samples and on the dimension of the features vector. The training dataset reliability has been initially improved by removing the outliers by means of the CNN rule (Hart, 1968). The CNN rule has been applied after assuming d equal to 10. Successively, in order to test the performance of the classifier and to decide the best value for d and k, a set of *bootstrap* samples have been classified by varying d and k. The choice of a *bootstrap* artificial dataset, rather than a real dataset, for determining the classifier parameters, has been done by considering the need to have a reliable test dataset larger, smoother and more versatile than the real one. The bootstrap method proposed by Hamamoto et al. (1997) allows to get a larger dataset starting from a true test dataset by combining the original training samples. Consequently, the *bootstrap* training set obtained is smoother than the one presented by Efron (1979).

Let $X = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n\}$ be the training dataset including nd-dimensional features vectors and $Y = \{\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_m\}$ the test dataset build by examining the same PEMW rain rate results related to AMSU overpasses listed in Table 2 but different from the X samples. Let $N_j = 20$ be the total number of the feature vectors $\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_{N_j}$ for the class C_j. The bootstrap samples for each class are determined as follows:

the vector $\boldsymbol{y}_{k_j}$ is selected from the training set samples Y (the initial samples are chosen so that no sample is selected more than once);

the r nearest neighbours samples $\boldsymbol{y}_{r_j, y(y=1,r)}$ are found by using the Euclidean distance;

a bootstrap sample is calculated by applying the equation:

$$b\_y^i_{k_j} = \frac{1}{r} \sum_{z=1}^{r} y^i_{r_j, z} \tag{7}$$

to the components of the nearest features vector $\boldsymbol{y}_{k_j}$ so to obtain the bootstrap feature vector ${}_b\boldsymbol{y}_{k_j}$.

HESSD
10, 13671–13706, 2013

The bootstrap samples have been calculated for each class. The artificial test dataset was built by considering different values for r. The final *bootstrap* training set contains the bootstrap samples obtained for $r = \frac{N_j}{4}, \frac{N_j}{5}, \frac{N_j}{10}, \frac{N_j}{2} - 8, \frac{N_j}{2} - 6, \frac{N_j}{2} - 4, \frac{N_j}{2} - 2$. A careful screening has been done to eliminate the redundant features vectors. The *bootstrap* samples have been classified by means of the k-NNM using X as training dataset. The statistical scores obtained by classifying the *bootstrap* samples for ($k = 3$; $d = 10$, $d = 15$, $d = 20$), ($k = 5$; $d = 10$, $d = 15$, $d = 20$), ($k = 7$; $d = 10$, $d = 15$, $d = 20$) are listed in Tables 3, 4, and 5, respectively. Other combinations of d and k have been investigated obtaining results worst than the ones listed in the above mentioned tables. In particular, for $k < 3$, $d < 10$ the FAR related to moderate class is higher than 0.4 and POD is lower than 0.6, while for $k > 7$ the FAR for all the classes is higher than 0.44 and the other statistical scores are lower than the ones obtained for the other combinations.

On the basis of the statistical scores obtained in the operation of cross validation, it results that the best choice for k and d is 5 and 15, respectively. The features chosen as components of the features vector $\boldsymbol{x}$ are listed in Table 6.

4 Validation and comparisons results

RainCEIV results have been validated against the rain rate values derived from the weather radar network operated by the DPC. The co-location process of the radar-derived rain rate measurements into the SEVIRI grid consists in associating to each SEVIRI pixel the rain rate measurements obtained by averaging the radar- derived rain rate measurements included into the SEVIRI pixel. The case studies used for validation are listed in Table 7. The radar-derived measurements co-located into the SEVIRI grid have been grouped in the three RainCEIV class C_0 (*non rainy*), C_1 (*light to moderate rain*), and C_2 (*heavy to very heavy rain*) described above. The validation method is based on a dichotomous statistical assessment. Table 8 sums up the contingency values for RainCEIV dichotomous statistical assessment. Table 9 shows the dichotomous statistics for all the pixels considered for validation. The statistical scores have

HESSD
10, 13671–13706, 2013

been calculated for all the classes together and for *light to moderate* and *heavy to very heavy* classes separately. The accuracy score for RainCEIV is 91 % both for all rainy pixels and for C_1 class only. The accuracy obtained for C_2 class is 99 %.This result is heavily influenced by the number of the *correct negatives* that is higher (492, 645) than that related to C_1 class (419, 488) and to all classes (417, 344). The Bias score obtained for RainCEIV is 1.16 for all classes and 1.13 when only C_1 class is considered, indicating a slight over-estimate of rain frequency. The Bias score higher for C_2 class (2.88) testifies the general tendency that the RainCEIV algorithm has to overestimate the "*heavy to very heavy"* areas. In fact, the FAR related to C_2 class is 78 %. It must be remarked that the high FAR value is highly influenced by the *light to moderate*-samples wrongly classified as *heavy to very heavy* by the RainCEIV. When both the rainy classes are considered, the FAR lowers to 43 % (44 % when C_1 class is considered separately). The POD statistical score reveals that RainCEIV correctly detects the 66, 63 and 64 % of the rainy areas when all classes, C_1 only, and C_2 only, respectively, are considered.

The HSS has also been considered. It is a measure of the correct forecasts after eliminating the forecasts that would be correct due exclusively to random chance. The HSS values obtained for RainCEIV are 56, 55 and 33 % when all classes, C_1 only, and C_2 only, respectively, are considered.

The proposed case studies, related to 29 September 2009 (case I) at 01:00 UTC, 4 August 2010 at 14:15 UTC (case II), and 21 February 2013 at 15:00 UTC (case III) are analyzed separately and the RainCEIV results are shown in Figs. 2, 3, and 4 together with C_MACSP results and radar-derived rain rate measurements. The statistical scores calculated for each case are listed in Table 11 (for all classes), Table 12 (for C_1 class), and Table 13 (for C_2 class).

The case I has been chosen because it highlights the ability of RainCEIV to detect very small rainy areas. On 29 September 2009 approximately at 13:00 UTC a very rapid rainfall event, characterized by heavy rain, affected a small area between Basilicata and Calabria regions in Southern Italy. The accuracy score is high (99 %), because the ac-

HESSD
10, 13671–13706, 2013

curacy results is largely influenced by the high occurrence of non-rainy pixels. The POD obtained for RainCEIV shows that it correctly detects the 64 % of the rainy samples, while Bias and FAR scores reveal its tendency to overestimate rainy samples (FAR score is 61 % and Bias score is 1.67). In details, the FAR and Bias score are higher for C_2 class than for C_1 class. In spite of its inclination to overestimate heavy rainy areas, RainCEIV is able to detect correctly the rainy areas. Moreover, it is important to note that this rainfall event was very rapid and even if the temporal distance between SEVIRI and radar network acquisition is little, it is able to affect the statistical scores negatively.

The RainCEIV statistical scores related to cases II and III (Figs. 3 and 4, respectively) are better than those of the case study discussed above. This is because they are related to rainy events characterized by a larger temporal and spatial distribution. The case study II bears on a set of heavy and moderate rainfall events that affected Central and Southern Italy on 4 August 2010 at 14:15 UTC. RainCEIV accuracy score is about 99 %. The RainCEIV is able to detect rainy samples with a POD of 85 %. The case study III is related to the analysis of an extreme convective event characterized by very heavy precipitation that occurred on 21 February 2013 on the east cost of Sicily and caused a flash flood over Catania. Also in this case, RainCEIV detects as rainy pixels that are no-rainy for the radar network (FAR is 0.27), but it is able to monitor the areas characterized by very heavy precipitation as well as by moderate precipitation (POD is 0.62) both on the east cost of Sicily and on Southern Calabria.

5 Conclusions

This paper proposes the RainCEIV technique as a useful tool for the characterization of rainy/non-rainy areas. RainCEIV consists of two modules that uses geostationary observations from the SEVIRI in order to detect cloudy pixels and, successively, to associate them to a rainy/non rainy class. The RainCEIV uses both IR and VIS measurements to determine if the SEVIRI pixel belongs to the *non rainy*, *light to moderate rainy* or *heavy to very heavy rainy* class. It is well known that the IR/VIS observations

HESSD
10, 13671–13706, 2013

do not have the same potentiality of MW observations in characterizing rainy areas, but their high spatial and temporal resolution are exploited to get a continuous monitoring of stratiform and convective events. Regarding the convective events, the RainCEIV is a useful tool for the study and characterization of the rainfall events characterized by short duration, high temporal variability, and small size area (of the order of the MSG-SEVIRI spatial resolution). RainCEIV has been trained on the AMSU-B PEMW rain rates and validated on the basis of the rain rates observations from the Italian DPC operational weather radar network. The dichotomous statistical scores indicate that a good fraction (91 %) of the pixels examined are correctly identified as rainy or non-rainy by the RainCEIV. The bias score (1.16) suggests that RainCEIV has a light tendency to overestimate rainy pixels. This tendency is emphasized when the *heavy to very heavy* class is considered. In fact, the Bias score for this class is 2.88. The POD obtained for RainCEIV shows that it detects 66 % of the rainy areas correctly. FAR reports that 43 % of the pixels detected as rainy by the RainCEIV are false alarms. The Bias score for "*light to moderate*" and "*heavy to very heavy*" classes indicates that RainCEIV has a general tendency to overestimate rainy pixels in the two classes. This high value is related to tendency that RainCEIV has to classify as "*heavy to very heavy"* the "*light to moderate"* samples. Because of this also the FAR values for "*heavy to very heavy"* class is quite high, that is 78 %. However, the RainCEIV POD is 63 % for the "*light to moderate*" and 64 % for the *"heavy to very heavy"* classes. In remarking upon the comparisons results, it is important to take in mind the different spatial resolution as well as the temporal distance between radar and satellite observations. In fact, especially for rapid convective events, even if the time distance between radar and SEVIRI acquisitions is little, it could affect the statistical scores negatively.

By taking advantage of the spatial and temporal resolution of the MSG satellite, RainCEIV can be very useful for near real-time weather monitoring purpose. In this study, a first attempt to use the temporal differences between SEVIRI observations to evaluate rain rate class is presented.

HESSD
10, 13671–13706, 2013

Acknowledgements. The authors acknowledge the Italian Department of Civil Protection (DPC) for providing data from the national radar network in the framework of the CETEMPS-DPC IDRA project. The staff of HIMET s.r.l. is also acknowledged for greatly facilitating the access to the data archive. This research activity has been supported by the NextData project.

References

Adler, R. F. and Mack, R. A.: Thunderstorm cloud height-rainfall rate relations for use with satellite rainfall estimation techniques, J. Clim. Appl. Meteorl., 23, 280–296, 1984.

Adler, R. F. and Negri, A. J.: A satellite infrared technique to estimate tropical convective and stratiform rainfall, J. Appl. Meteorol., 27, 30–51, 1988.

Adler, R. F., Markus, M. J., and Fenn, D. D.: Detection of severe Midwest thunderstorms using geosynchronous satellite data, Mon. Weather Rev., 113, 769–781, 1985.

Adler, R. F., Negri, A. J., Keehn, P. R., and Hakkarinen, I. M.: Estimation of monthly rainfall over Japan and surrounding waters from a combination of low-orbit microwave and geosynchronous IR data, J. Appl. Meteorol., 32, 335–356, 1993.

Babu, V. S. and Viswanath, P.: Rough-fuzzy weighted *k*-nearest leader classifier for large data sets, Pattern Recogn., 42, 1719–1731, 2009.

Bennartz, R.: Optimal convolution of AMSU-B to AMSU-A, J. Atmos. Ocean. Tech., 17, 1215–1225, 2000.

Bizzarri, B., Gasiewski, A. J., and Staelin, D. H.: Observing rain by millimetre–submillimetre wave sounding from geostationary orbit, in: Measuring Precipitation From Space – EURAINSAT and the Future, edited by: Levizzani, V., Bauer, P., and Turk, F. J., Springer, Dordrecht, 675–692, 2007.

Cimini, D., Romano, F., Ricciardelli, E., Di Paola, F., Viggiano, M., Marzano, F. S., Colaiuda, V., Picciotti, E., Vulpiani, G., and Cuomo, V.: Validation of satellite OPEMW precipitation product with ground-based weather radar and rain gauge networks, Atmos. Meas. Tech. Discuss., 6, 4279–4312, doi:10.5194/amtd-6-4279-2013, 2013.

Dasarathy, B. V. (Ed.): Nearest Neighbor (NN) Norms: NN Pattern Classification Techniques, IEEE Computer Society Press, Los Alamitos, California, 1991.

HESSD

10, 13671–13706, 2013

Dasarathy, B. V.: Data mining tasks and methods: classification nearest-neighbor approaches, in: Handbook of Data Mining and Knowledge Discovery, edited by: Klösgen, W. and Zytkow, J. M., Oxford University Press, New York, 288–298, 2002.

Di Paola, F., Casella, D., Dietrich, S., Mugnai, A., Ricciardelli, E., Romano, F., and Sanò, P.: Combined MW-IR Precipitation Evolving Technique (PET) of convective rain fields, Nat. Hazards Earth Syst. Sci., 12, 3557–3570, doi:10.5194/nhess-12-3557-2012, 2012.

Di Tomaso, E., Romano, F., and Cuomo, V.: Rainfall estimation from satellite passive microwave observations in the range 89 GHz to 190 GHz, J. Geophys. Res., 114, D18203, doi:10.1029/2009JD011746, 2009.

Ebert, E.: A pattern recognition technique for distinguishing surface and cloud types in the polar regions, J. Clim. Appl. Meteorol., 26, 1412–1427, 1987.

Ebert, E.: Forecast Verification Issues, Methods and FAQ, available at: http://www.cawcr.gov.au/projects/verification/, last access: 25 July 2013.

Feidas, H. and Giannakos, A.: Classifying convective and stratiform rain using multispectral infrared Meteosat Second Generation satellite data, Theor. Appl. Climatol., 108, 613–630,2012.

Gasiewski, A. J., Voronovich, A., Weber, B. L., Stankov, B., Klein, M., Hill, R. J., and Bao, J. W.: Geosynchronous Microwave (GEM) Sounder/Imager Observation System Simulation, Geoscience and Remote Sensing Symposium, 2003, IGARSS'03, Proceedings, 2003 IEEE International, 2, 21–25 July 2003, 1209–1211, 2003.

Hamamoto, Y., Uchimura, S., and Tomita, S.: A bootstrap technique for nearest neighbor classifier design, IEEE T. Pattern Anal., 19, 73–79, 1997.

Hart, P. E.: The condensed nearest neighbor rule, IEEE T. Inform. Theory, 14, 515–516, 1968.

Heinemann, T., Lattanzio, A., and Roveda, F.: The Eumetsat Multi-Sensor Precipitation Estimate (MPE), in: Second International Precipitation Working group (IPWG) Meeting, September 2002, Madrid, Spain, 2002.

Hsu, K., Gao, X., Sorooshian, S., and Gupta, H. V.: Precipitation estimation from remotely sensed information using artificial neural networks, J. Appl. Meteorol., 36, 1176–1190, 1997.

Jain, A. K. and Chandrasekaran, B.: Dimensionality and sample size considerations in pattern recognition practice, in: Handbook of Statistics, 2, edited by: Krishnaiah, P. R. and Kanal, L. N., North-Holland Publishing Company, Amsterdam, 835–855, 1982.

Jobard, I. and Desbois, M.: Satellite estimation of the tropical precipitation using the Meteosat and SSM/I data, Atmos. Res., 34, 285–298, 1994.

HESSD
10, 13671–13706, 2013

Joyce, R. J., Janowiak, J. E., Arkin, P. A., and Xie, P.: CMORPH: a method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolutions, J. Hydrometeorol., 5, 487–503, 2004.
Kidd, C. and Levizzani, V.: Status of satellite precipitation retrievals, Hydrol. Earth Syst. Sci., 15, 1109–1116, doi:10.5194/hess-15-1109-2011, 2011.
Kidd, C., Kniveton, D. R., Todd, M. C., and Bellerby, T. J.: Satellite rainfall estimation using a combined passive microwave and infrared algorithm, J. Hydrometeorol., 4, 1088–1104, 2003.
Kühnlein, M., Thies, B., Nauß, T., and Bendix, J.: Rainfall-rate assignment using MSG SEVIRI data – a promising approach to spaceborne rainfall-rate retrieval for midlatitudes, J. Appl. Meteorol. Clim., 49, 1477–1495, 2011.
Lambrigtsen, B., Tanner, A., Gaier, T., Kangaslahti, P., and Brown, S.: A Microwave Sounder for GOES-R Developing the GeoSTAR Mission, in: Geoscience and Remote Sensing Symposium, 2006, IGARSS 2006, IEEE International Conference, 31 July–August 2006, 3964–3967, 2006.
Lensky, I. M. and Rosenfeld, D.: Clouds-Aerosols-Precipitation Satellite Analysis Tool (CAPSAT), Atmos. Chem. Phys., 8, 6739–6753, doi:10.5194/acp-8-6739-2008, 2008.
Lensky, M. I. and Rosenfeld, D.: Estimation of precipitation area and rain intensity based on the microphysical properties retrieved from NOAA AVHRR Data, J. Appl. Meteorol. Clim., 36, 234–242, 1997.
Mamoudou, B. and Gruber, A.: GOES multispectral rainfall algorithm (GMSRA), J. Appl. Meteorol., 40, 1500–1514, 2001.
Marzano, F. S., Palmacci, M., Cimini, D., Giuliani, G., and Turk, J. F.: Multivariate statistical integration of satellite infrared and microwave radiometric measurements for rainfall retrieval at the geostationary scale, IEEE T. Geosci. Remote, 42, 1018–1032,2004.
Negri, A. J. and Adler, R. F.: Relation of satellite based thunderstorm intensity to radar estimated rainfall, J. Appl. Meteorol., 20, 288–300, 1981.
Parikh, J.: A comparative Study of Cloud classification Techniques, Remote Sens. Environ., 6, 67–81, 1977.
Ricciardelli, E., Romano, F., and Cuomo, V.: Physical and statistical approaches for cloud identification using Meteosat Second Generation – Spinning Enhanced Visibile and Infrared Imager, Remote Sens. Environ., 112, 2741–2760, 2008.

HESSD
10, 13671–13706, 2013

Schmetz, J., Pili, P., Tjemkes, S., Just, D., Kermann, J., Rota, S., and Ratier, A.: An introduction to Meteosat Second Generation (MSG), B. Am. Meteorol. Soc., 83, 977–992, 2002.

Scofield, R. A.: The NESDIS operational convective precipitation technique, Mon. Weather Rev., 115, 1773–1792, 1987.

Turk, J. F., Rohaly, G., Hawkins, J., Smith, E. A., Marzano, F. S., Mugnai, A., and Levizzani, V.: Meteorological applications of precipitation estimation from combined SSM/I, TRMM, and geostationary satellite data, in: Microwave Radiometry and Remote Sensing of the Environment, edited by: Pampaloni, P., VSP Intern. Sci. Publisher, Utrecht, the Netherlands, 353–363, 1999.

Vicente, G. A., Menzel, W. P., and Scofield, R. A.: The operational GOES infrared rainfall estimation technique, B. Am. Meteorol. Soc., 79, 1883–1898, 1998.

Viswanath, P. and Sarma, H.: An Improvement to *k*-Nearest Neighbor Classifier, IEEE Recent Advances in Intelligent Computational Systems, 227–231, 2011.

Wu, R., Weinman, J. A., and Chin, R. T.: Determination of rainfall rates from GOES satellite images by a pattern recognition technique, J. Atmos. Ocean. Tech., 2, 314–330, 1985.

HESSD

10, 13671–13706, 2013

Table 1. Accuracy of the C_MACSP algorithm on an independent dataset.

Classes	Classification accuracy
Clear over land	94.0 %
Clear over sea	96.0 %
Low/midlle clouds over land	91.6 %
Low/middle clouds over sea	92.0 %
High thin clouds over land	83.0 %
High thin clouds over sea	85.0 %
High thick clouds over land	95.8 %
High thick clouds over sea	96.0 %
Convective clouds	94.5 %
Convective clouds	94.5 %

HESSD
10, 13671–13706, 2013

Table 2. List of the NOAA satellite overpasses for the AMSU-B PEMW rain rate results considered in the training phase.

Training phase – Dates for NOAA satellite overpasses	
Dates	HH-MM of the NOAA passes over the area of interest (UTC)
29 Sep 2009	15:16, 15:22
1 Oct 2009	04:37, 05:13, 08:30, 13:03, 15:56, 16:37, 19:18, 19:50
2 Oct 2009	01:25, 04:13, 05:00
4 Mar 2010	14:23, 16:03, 16:28, 20:05
5 Mar 2010	00:56, 01:48, 04:16, 06:24, 08:20, 11:40
26 Apr 2010	12:47, 13:20, 14:49
28 Apr 2010	12:26, 15:45
2 May 2010	15:45, 16:32, 19:44
20 Jun 2010	11:42, 11:58, 14:28
21 Jun 2010	02:00
23 Jun 2010	12:52
4 Aug 2010	10:43, 12:19, 14:46, 16:24, 18:03, 18:56, 20:38, 03:54, 06:15, 10:16, 13:14, 15:17
4 Oct 2010	17:44, 19:33
21 Feb 2013	11:20, 13:10

HESSD
10, 13671–13706, 2013

Table 3. Statistical scores related to the RainCEIV rain rate results obtained classifying the artificial test dataset ($k = 3$). The statistical scores are shown for all rainy classes (C_1, C_2), light to moderate rain (C_1), and heavy to very heavy rain (C_2).

	C_1, C_2 $k = 3, d = 10$	C_1	C_2	C_1, C_2 $k = 3, d = 15$	C_1	C_2	C_1, C_2 $k = 3, d = 20$	C_1	C_2
Accuracy	0.92	0.80	0.84	0.93	0.81	0.85	0.93	0.80	0.84
Bias	0.93	1.10	1.05	0.92	1.13	1.03	0.93	1.13	1.03
POD	0.91	0.75	0.79	0.91	0.78	0.79	0.91	0.77	0.78
HSS	0.81	0.56	0.65	0.83	0.59	0.67	0.82	0.57	0.65
FAR	0.02	0.31	0.25	0.01	0.31	0.23	0.01	0.32	0.24

HESSD
10, 13671–13706, 2013

Table 4. Statistical scores related to the RainCEIV rain rate results obtained classifying the artificial test dataset ($k = 5$). The statistical scores are shown for all rainy classes (C_1, C_2), light to moderate rain (C_1), and heavy to very heavy rain (C_2).

	C_1, C_2 $k = 5, d = 10$	C_1	C_2	C_1, C_2 $k = 5, d = 15$	C_1	C_2	C_1, C_2 $k = 5, d = 20$	C_1	C_2
Accuracy	0.94	0.86	0.92	0.95	0.87	0.92	0.94	0.84	0.90
Bias	0.94	1.05	1.05	0.95	1.05	1.05	0.96	1.05	1.03
POD	0.93	0.82	0.91	0.94	0.83	0.90	0.94	0.79	0.86
HSS	0.87	0.70	0.82	0.89	0.71	0.81	0.87	0.65	0.77
FAR	0.01	0.22	0.14	0.01	0.21	0.14	0.01	0.25	0.16

HESSD
10, 13671–13706, 2013

Table 5. Statistical scores related to the RainCEIV rain rate results in classifying the artificial test dataset ($k = 7$). The statistical scores are shown for all rainy classes (C_1, C_2), light to moderate rain (C_1), and heavy to very heavy rain (C_2).

	C_1, C_2 $k = 7, d = 10$	C_1	C_2	C_1, C_2 $k = 7, d = 15$	C_1	C_2	C_1, C_2 $k = 7, d = 20$	C_1	C_2
Accuracy	0.90	0.81	0.91	0.89	0.79	0.90	0.89	0.78	0.89
Bias	0.94	1.05	1.07	0.94	1.08	1.05	0.94	1.05	1.08
POD	0.90	0.74	0.90	0.89	0.73	0.88	0.89	0.70	0.88
HSS	0.77	0.58	0.79	0.75	0.57	0.81	0.75	0.52	0.76
FAR	0.01	0.29	0.16	0.01	0.32	0.16	0.01	0.33	0.19

HESSD
10, 13671–13706, 2013

Table 6. Summary of the features considered for use in the RainCEIV *k*-NNM classifier(A = daytime for all C-MACSP classes, **A** = only nighttime for all C_MACSP classes; ***A*** = daytime and nighttime for all C_MACSP classes; LM = daytime for low/middle cloud class; **LM** = at nighttimes for low/middle cloud class; ***LM*** = day time and nighttimes for low/middle cloud class, HT/C = daytime and nighttimes for high thick and convective cloud class).

Features	MSG-SEVIRI spectral bands							
	VIS 0.6	VIS 0.8	NIR 1.6	IR 3.9	IR 6.7	IR 7.3	IR 10.8	IR 12.0
Max Gray level			A				***A***	
Min Gray level				**A**			***A***	
Mean Gray level	A	A						
Max/Min (Gray level)							***A***	
Max (Contrast 0°, 45°, 90°, 135°)						***A***		
Max (Entropy 0°, 45°, 90°, 135°)				**A**				
Max (Mean 0°, 45°, 90°, 135°)				**A**		***A***	LM	
Max (ASM 0°, 45°, 90°, 135°)				***LM***				
Min (Contrast 0°, 45°, 90°, 135°)								
Min (Entropy 0°, 45°, 90°, 135°)							***A***	
Min (Mean 0°, 45°, 90°, 135°)								***A***
Min (ASM 0°, 45°, 90°, 135°)				**LM**				***A***
ΔTB_{15-30}					***A***			
ΔTB_{15-45}					***A***	HT/C		
ΔTB_{30-45}						HT/C		

HESSD
10, 13671–13706, 2013

Table 7. List of case studies used for validation.

Date	Radar measurement time (UTC)	Satellite overpass time (UTC)
2 May 2009	15:00	15:00
19 Sep 2009	19:15	19:15
29 Sep 2009	13:00 (case I)	13:00
8 Jan 2010	11:00, 13:00, 16:30	11:00, 13:00, 16:30
9 Mar 2010	17:00	17:00
23 Jun 2010	15:00	15:00
1 Jul 2010	16:45	16:45
6 Jul 2010	11:30, 12:30	11:30, 12:30
4 Aug 2010	12:00, 14:15 (case II)	12:00, 14:15
4 Oct 2010	19:30	19:30
21 Feb 2013	13:00, 15:00 (case III)	13:00, 15:00

HESSD
10, 13671–13706, 2013

Table 8. Contingency table for the dichotomous statistical assessment of the RainCEIV algorithm for all the pixels used for validation.

		Radar-derived rain rate results		
		yes	no	marginal total
RainCEIV results	yes	34 434	26 140	60 574
	no	17 882	417 844	435 726
	marginal total	52 316	443 984	496 300

HESSD
10, 13671–13706, 2013

Table 9. Dichotomous Statistics scores (RainCEIV vs. radar-derived rain rate measurements) for all the cases listed in Table 3. The statistical scores are shown for all rainy classes (C_1, C_2), light to moderate rain (C_1), and heavy to very heavy rain (C_2).

Statistical Scores	C_1, C_2	C_1	C_1
Accuracy	0.91	0.91	0.99
Bias	1.16	1.13	2.88
POD	0.66	0.63	0.64
HSS	0.56	0.55	0.33
FAR	0.43	0.44	0.78

HESSD
10, 13671–13706, 2013

Table 10. Dichotomous statistical scores shown for all rainy classes (C_1, C_2), light to moderate rain (C_1), and heavy to very heavy rain (C_2), for the case studies I, II and III.

Statistical score	Case I 29 Sep 2009, 13:00 UTC			Case II 4 Aug 2010, 14:15 UTC			Case III 21 Feb 2013, 15:00 UTC		
	C_1, C_2	C_1	C_2	C_1, C_2	C_1	C_2	C_1, C_2	C_1	C_2
Accuracy	0.99	0.99	0.99	0.99	0.98	0.99	0.86	0.86	0.99
Bias score	1.64	1.45	2.33	1.91	2.04	1.25	0.80	0.81	0.41
POD	0.64	0.64	0.33	0.85	0.72	0.31	0.59	0.59	0.03
HSS	0.49	0.52	0.20	0.58	0.47	0.28	0.57	0.57	0.04
FAR	0.61	0.56	0.86	0.55	0.65	0.75	0.26	0.27	0.93

HESSD
10, 13671–13706, 2013

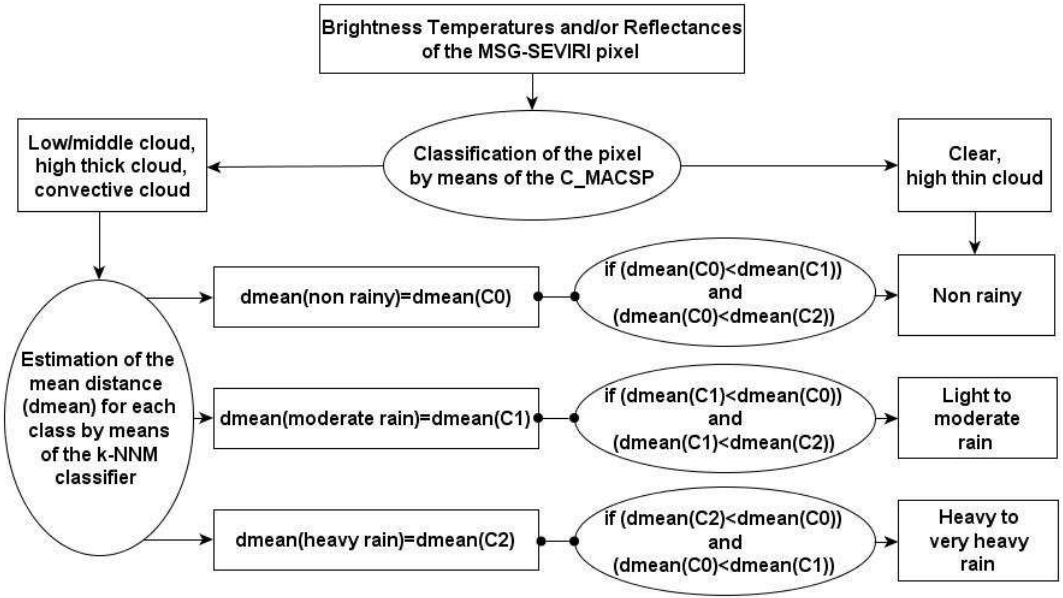

Fig. 1. Flowchart of the RainCEIV algorithm.

HESSD
10, 13671–13706, 2013

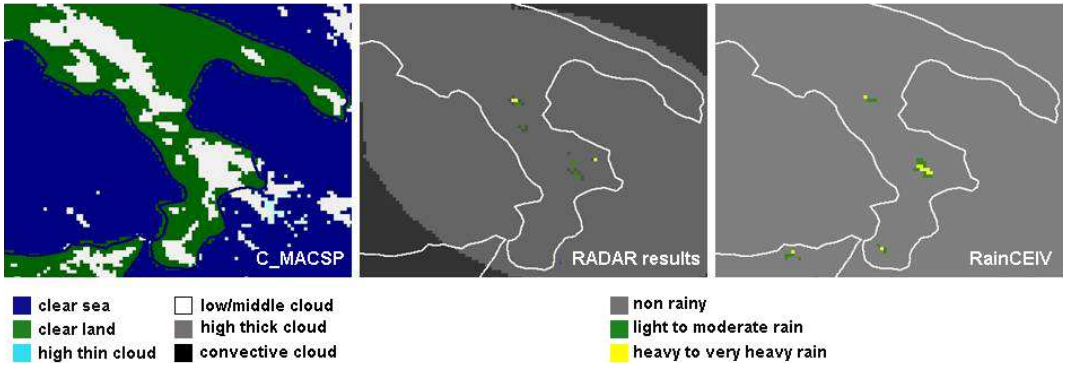

Fig. 2. 29 September 2009 at 13:00 UTC. From left to right panels: C_MACSP cloud classification results, radar-derived rain rate results, RainCEIV rain rate results.

HESSD
10, 13671–13706, 2013

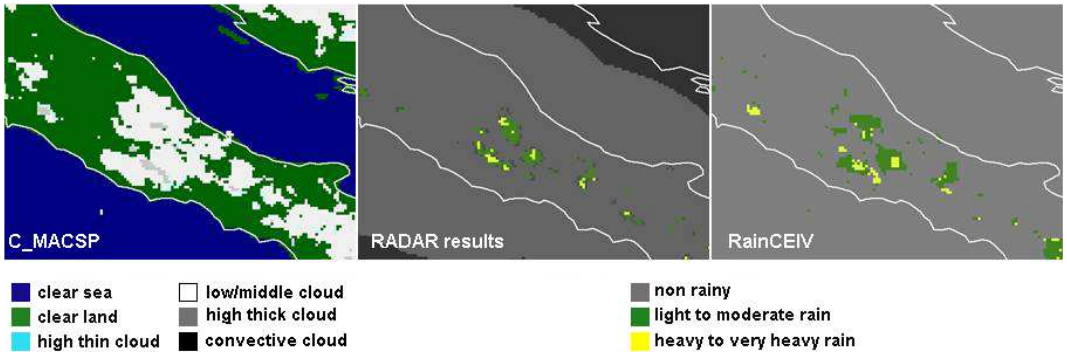

Fig. 3. 4 August 2010 at 14:15 UTC. From left to right panels: C_MACSP cloud classification results, radar-derived rain rate results, RainCEIV rain rate results.

HESSD
10, 13671–13706, 2013

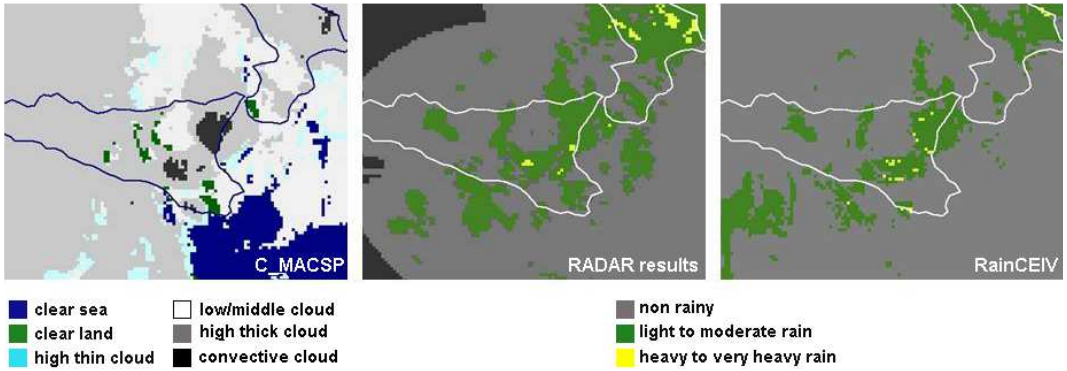

Fig. 4. 21 February 2013 at 15:00 UTC. From left to right panels: C_MACSP cloud classification results, radar-derived rain rate results, RainCEIV rain rate results.

HESSD
10, 13671–13706, 2013