Interactive comment on “A statistical approach for rain class evaluation using Meteosat Second Generation-Spinning Enhanced Visible and InfraRed Imager observations” by E. Ricciardelli et al.

Anonymous Referee #3

The paper “A statistical approach for rain class evaluation using Meteosat Second Generation-Spinning Enhanced Visible and InfraRed Imager observations” by Ricciardelli et al., proposes a statistical technique to infer precipitation classes from SEVIRI radiances and radiance spatial and temporal features. The calibration of the technique is carried out by using AMSU derived estimates and it is validated against radar rain fields. The subject of the paper is of some interest for this journal, but is poorly written, with a number of serious weaknesses that I do not believe could be addressed through a standard major revision. I suggest to reject the paper for a number of reasons: I listed below the most relevant ones (page numbers refer to the discussion paper, from 1 to 36).

The aim of the paper seems to provide a tool to benefit short term hydrology and long term climate studies (lines 1-3 on page 3): the author should explain the usefulness of a technique that gives as output only two precipitation levels.

Author Comment (A.C.):
We admit that in the abstract we described the utility of the rainfall measurements generically thus missing the actual RainCEIV purpose that is the continuous monitoring of the precipitating events and the gathering of information on the cloud cover and classification.
The abstract and the introduction as well as each section of the paper are improved in order to explain the utility of the RainCEIV technique more in-depth. In particular the abstract now reads:

“This study exploits the Meteosat Second Generation (MSG)–Spinning Enhanced Visible and Infrared Imager (SEVIRI) observations to evaluate the rain class at high spatial and temporal resolutions and, to this aim, proposes the Rain Class Evaluation from Infrared and Visible observation (RainCEIV) technique. RainCEIV is composed of two modules: a cloud classification algorithm which characterizes and individuates the cloudy pixels, and a supervised classifier that delineates the rainy areas according to the three rainfall intensity classes, the *non-rainy* (rain rate value<0.5 mm×h^{-1}) class, the *light-to-moderate rain* class (0.5 mm×h^{-1}≤rain rate value<4 mm×h$^{-1)}$, and the *heavy-to-very-heavy rain* class (rain rate value≥4 mm×h^{-1}). The second module considers in input the spectral and textural features of the infrared and visible SEVIRI observations for the cloudy pixels detected by the first module. It also uses the temporal differences of the brightness temperatures related to the SEVIRI water vapour channels indicative of the atmospheric instability strongly related to the occurrence of rainfall events.
The rainfall rates used in the training phase are obtained through the Precipitation Estimation at Microwave frequencies, PEMW (an algorithm for rain rate retrievals based on Atmospheric Microwave Sounder Unit (AMSU)-B observations). RainCEIV provides a continuous monitoring both of the cloud coverage and rainfall events without using real–time ancillary data. Its principal aim is that of supplying preliminary qualitative information on the rainy areas within the Mediterranean basin where there is no radar network coverage. The results of RainCEIV have been validated against radar-derived rainfall measurements by the Italian Operational Weather Radar Network.”

The abstract will be updated by introducing the statistical scores obtained for the enlarged validation dataset, as will be explained in the discussion at the end of this document.

In the introduction, there is no need to mention early works on satellite precipitation estimation in the ‘80s and ‘90s, since they used very different approaches and instruments.
On the other side, many works on SEVIRI data use for precipitation are missing (the mentioned Kidd and Levizzani reports on them).

A.C.:
The introduction has been updated following your suggestion and including some works missed in the previous version of the paper.

The correct reference for Mamoudou and Gruber is Ba and Gruber (page 4 and reference list).
A.C.: Thanks for the correction.

Section 2. The history and launch schedule of Meteosat spacecrafts are not necessary for the aim of this paper. Please, add a reference for the Italian radar network, and report on the quality of the data used. Since the radar data are used here to validate satellite product, it is mandatory a more detailed description of the radar network and its reliability.

A.C.:
Agreed. We have removed the history and launch schedule of Meteosat spacecrafts from the manuscript. References and more details concerning the Italian radar network are now provided in Section 2. The following text has been added to provide information on data quality:
Procedures for mitigating ground clutter, anomalous propagation, beam blockage effects are applied (Vulpiani et al., 2008a). The sri product is derived applying a reflectivity-rainfall (Z-R) relationship to the Lowest Beam Map (LBM), i.e. the reflectivity values at the lowest level of the corrected radar volumes. The sri product used here represents the best estimate from the radar network available for the period under analysis, and it has been already used to validate satellite rainfall estimates (Cimini et al., 2013), including EUMETSAT H-SAF products (Puca et al., 2013). Procedures to improve the quality of the sri product, including attenuation compensation, polarimetric rainfall inversion techniques, and adaptive algorithms to retrieve mean vertical profiles of reflectivity have been recently developed at DPC (Vulpiani et al., 2012; Rinollo et al., 2013).

Section 3. Section 3.1 roughly describes the cloud classification algorithm. Is table 1 related to this section? How is accuracy defined for cloud classes? Are clear sky pixels included in the accuracy calculation?

A.C.:
Yes, Table 1 is related to this section and lists the accuracy scores (defined as the ratio between the number of the test samples classified correctly and the total number of the test samples examined) for cloud and clear classes. In order to explain more in-depth how C_MACSP works, section 3.1 “**3.1- Cloud classification algorithm description**” has been modified in the revised version.

What are the outliers mentioned in line 13 on page 9? Are they damaged pixels, noise, or what?

A.C.:
We defined as outliers the samples that during the training phase are misclassified. (e.g. as for C_MACSP a thin cloud could be misclassified as clear, or a low/middle cloud could be misclassified as high thick cloud, as for RainCEIV heavy rain could be misclassified as moderate rainy pixel). This information is now provided in the revised version.

Only two images out of the nine used to validate the classification are during nighttime: are there enough pixels to verify correct classification of all cloud classes?

A.C.:
Initially, we decided to consider a fixed number of test samples for each cloud class and for the clear class, making no distinction between the samples acquired during night-time and those acquired during daytime. In the revised version, the accuracy shown in Table 7 (that was Table 1 in the previous version) is determined for each C_MACSP class for night-time and daytime samples, separately.

I think that the validation dataset should be much larger

A.C.:
Agreed. We followed your suggestion and the validation dataset has been enlarged so to include more night-time scenes. In addition, we are going to follow the advice of referee#2 who proposes to show the C_MACSP validation results in a sub-section of Section 4. As a consequence, section 4 "Validation results" is now divided into two sub-sections: "4.1 C_MACSP validation results" and "4.2 RainCEIV validation results

In section 3.2.1 there are a number of sentences that have to be canceled (my suggestion) or discussed with much more detail. I report here few examples, but the entire section should be rewritten or canceled. How can SEVIRI observation "individuate precipitation processes" (lines 16-17 on page 11) ? especially in convective clouds? Which processes can be individuated (coalescence, riming, breakup, melting)?

A.C.:
We apologize for the incorrect use of the English language, the term "precipitation processes" was erroneously used to mean "precipitation events". The purpose of RainCEIV is to determine a precipitation class not the precipitation process.

The radiance measured in the SEVIRI channels comes from the very top layers of the cloud. Few lines below it is said that "features related to radiances acquired at 3.9 and 1.6µm bear on the cloud drop size distribution": as a matter of fact, "cloud drop size distribution", unfortunately, cannot be derived by any feature related to SEVIRI channels.

A.C.:
The paragraph purpose was to describe the characteristics and the usefulness of the 3.9µm,1.6µm 12.0µm, 10.8µm, 0.6µm SEVIRI spectral channels to derive some cloud microphysical properties in order to make it clear that the choice of these spectral channels was made because of their connection with cloud microphysical properties so as to allow the identification of rainy clouds.
Consequently, sub-section 3.2.1 from line 11 on page 13681 to line 24 on page 13681 is rewritten as follows:

"All the spectral and textural features defined for the IR/VIS SEVIRI images acquired at 0.6 µm, 0.8 µm, 1.6 µm, 3.9 µm, 6.3 µm, 7.2 µm, 10.8 µm, and 12 µm were initially considered as components of the feature vector $\vec{x}$. Among the spectral channels listed above, some are usually utilized to infer information on microphysical properties related to the cloud top. In particular, the observations acquired at 10.8µm and 12.0µm are used to provide information on cloud top temperature and cloud optical thickness, the observations at 0.6µm are used to get information on cloud optical thickness, while the 3.9µm and 1.6µm observations are used to infer information on cloud thermodynamic phase and cloud drop-size distribution. 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). Consequently, this study, considering features derived from spectral channels connected with cloud microphysical properties, could allow a more accurate identification of rainy clouds."

The temperature of WV channels are related with tropospheric moisture content over clear sky areas, but in case of mid- and high- level clouds the contribution to the radiance measured by satellite sensor has a dominant contribution from the cloud top. How can the temperature differences mentioned on lines 3-4 on page 12 "characterize convective as well as stratiform precipitation" ?

A.C.:
As before we used the verb "characterize" wrongly. In fact, the temporal differences have been used as input for the classifier in order to associate a pixel to the class C_0, C_1, or C_2. The WV temporal differences are useful to distinguish different rainy/non-rainy classes only when used with the other components of the feature vector.
In order to clarify how the WV spectral channels have been considered for the RainCEIV purposes, sub-section 3.2.1 from line 24 on page 3681 to line 4 on page 13682 is modified as follows:

"The spectral channels centred at 6.2 μm and 7.3 μm are indicative of the water vapour (WV) content in the troposphere at levels lower than 350hPa and 500hPa, respectively. The features related to WV spectral channels when considered alone do not give useful information on the presence of a rainy cloud, on the contrary when considered with the other spectral channels features, in particular with those related to the 10.8μm channel, they are useful to individuate convective events (Mosher, 2001, 2009). Moreover, the WV temporal changes are indicative of the atmospheric instability that is a useful index in the detection of precipitating area. Because of this, the temporal differences $\Delta TB_{(6.2)15-30}$, $\Delta TB_{(6.2),15-45}$, $\Delta TB_{(6.2),30-45}$, $\Delta TB_{(7.2)15-30}$, $TB_{(7.2),15-45}$, $TB_{(7.2),30-45}$, between the brightness temperature of WV channel observations acquired 15, 30 and 45 minutes before the time of interest are exploited to get information on the WV temporal changes at different levels in the atmosphere. Obviously, the temporal change of WV brightness temperature related to a pixel does not always mean that the pixel is rainy, and as for the other features it gains usefulness in discriminating rainy/non-rainy classes when used in combination with the other features opportunely chosen, as will be described in the following sub-section."

Section 3.2.2. Probably Table 3 means Table 2 (line 25 on page 13).

Agreed. In the revised version, due to the fact that Table 1 is renamed Table 7, Table 2 (wrongly named Table 3) is renamed Table 1.

On line 26- 28 (page 13) is described the matching between SEVIRI and AMSU rain product. It seems that the rain value estimated over an area ranging between 200 km2 (at nadir) and 1000 km2 (on the edge of the swath) is assigned to a SEVIRI pixel of around 25 km2 in the considered area. This implies a number of assumptions on the rainfall spatial and temporal structure that are not usually verified in real rain.

A.C.:
We apologize for not making the collocation process clearer.
The collocation of PEMW-derived rain rate values in the SEVIRI grid is now described in Section "2- Instruments and data" approximately at line 25 on page 13677, as follows:

"The PEMW rain rate value is assigned to the SEVIRI pixel only when the latter is entirely enclosed in the corresponding AMSU-B/MHS FOV. The re-sampling of the PEMW rain rate values on the SEVIRI grid was done by considering the area of each AMSU-B/FOV calculated using the orbital parameters described in (Bennartz, 2000). The temporal matching has been carried out considering a maximum delay of 7 minutes and 30 seconds between the acquisition time of the SEVIRI pixel and the AMSU/MHS FOV."

Table 6 has to be better introduced and discussed in the text, and the caption should be rewritten accordingly.

A.C.:
Table 6 is now split into two tables: Table 5 and 6 that list the features to be used during daytime and night-time, respectively. The captions of Tables 5 and 6 have been re-written so to be clearer. A description of Tables 5 and 6 is now added at the end of sub-section 3.2.2 as follows:

"The features chosen as components of the feature vector $\vec{x}$ related to daytime and night-time acquisition are listed in Tables 5 and 6, respectively. The features used over land and over sea are the same, but in some cases they vary for different cloud classes, e.g. the max and min value of the ASM are very useful in order to determine the confidence that a low/middle cloud is precipitating, but its discriminant power is not so high as to distinguish the other rainy classes. On the contrary, the minimum and maximum values of Contrast and Mean give an useful contribution in detecting both light-to-moderate rainy class and heavy rainy class for all the cloudy class."

Section 4. A good validation practice requires that the datasets used for calibration and validation are independent. In the work reported in this paper, it seems this condition is not satisfied for all the considered cases. Comparing table 2 and table 6, for 4 out of 11 cases (29/09/09, 23/06/10, 04/08/10 and 10/10/10) the satellite overpasses used for validation are very close to the slot used for the calibration, and this should be avoided. I suggest to remove the mentioned cases from the validation, and to add more slots of the other cases.

A.C.:
Agreed. Although some training and validation samples have been acquired on the same day, the Solar Zenith Angle (SZA) ranges of the training and validation samples are different. Consequently, the cases study of 29 September 2009 at 13:00UTC and 23 June 2010 at 15:00UTC were not classified by using the training samples acquired on the same day. In fact, as explained in the revised version of section 3.2.2, the training samples have been grouped (and successively chosen by the k-NNM classifier) both on the basis of the 10.8µm SEVIRI channel brightness temperature ranges and the SZA ranges.
In detail, we agree on removing the case related to 04 October 2010 at 19:30 UTC because it is very close to the training samples related to the same day, but we would rather leave the other cases for validation purposes:

- 29 September 2009 at 13:00UTC: the training samples related to 29 September 2009 at 17:00UTC have not been used as training samples to classify the SEVIRI observations acquired on 29 September 2009 at 13:00UTC because their SZA ranges do not correspond (for the samples acquired at 13:00 UTC SZA<58º, while for the ones acquired at 17:00UTC SZA>80º);
- 23 June 2010 at 15:00UTC: the SZA ranges for the training and validation samples related to 23 June 2010 are different, in fact the samples acquired at 15:00UTC for validation are characterized by a SZA>48º, while those acquired at 12:52UTC have a SZA<35º.

Moreover, the AMSU-B/MHS passes related to 29 September 2009 at 15:16UTC, 4 August 2010 at 12:26 UTC and 14:46 UTC, 21 February at 13:10 UTC have been removed from Table 2 because they were used only to carry out the test dataset as described in sub-section 3.2.2 of the revised version. In fact, the AMSU-B/MHS passes used to build both training and test dataset were wrongly listed in Table 2 without distinction. This point was not explained in depth in the previous version.
Table 2 is now renamed Table 1 and has been modified on the basis of the above considerations and removing some inaccuracy (the correct time for the AMSU-B/MHS pass of 29 September 2009 at 15:22UTC is 17:22 UTC, the AMSU/MHS passes at 03:54UTC, 06:15UTC, 10:16UTC, 13:14 UTC and 15:17UTC were wrongly related to 04 August 2010 but actually they correspond to 04 October 2010).

In table 7 the last column title is “satellite overpass time”, but the number reported in the column are probably the nominal time of delivery of the SEVIRI slot. Since the SEVIRI starts scanning the earth from the South, the Mediterranean region is scanned few minutes before the end of the scan, at 12, 27, 42 and 57 minutes every hour. In this table should be reported the real scan time of the Mediterranean region.
A.C.:
Yes, the time reported in the column is the nominal time of the acquisition of the SEVIRI slots. Following your suggestion, it has been corrected indicating the real scan time of Mediterranean region, that ends approximately 2 minutes before the end of the scan.

The accuracy indicator is of a very limited meaning in evaluating the technique performances, since it includes the number of correct negatives, which is always very high, and can be arbitrarily increased by enlarging the considered area. See as an example table 8 and figures 2, 3 and 4.
A. C.:
Agree. The accuracy indicator is highly influenced by the number of corrected negatives, because of this the other statistical scores (HSS, POD, FAR and Bias) are considered. Moreover, in order to increase the number of light-to-moderate-rainy samples and heavy-to-very-heavy samples, we are enlarging the validation dataset by adding more daytime and night-time scenes and choosing cases study characterized by more convective events both during daytime and night-time.