

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 #1

Received and published: 16 December 2013

This paper presents a new technique (RainCEIV) to classify cloudy scenarios in terms of rain categories by exploiting the MSG-SEVIRI spectral channels. The final purpose is to provide an operational tool for continuous rainfall event monitoring (convective and stratiform), which takes advantage of the high spatial and temporal resolution of geostationary VIS and IR data in spectral and textural tests. The algorithm is composed by two modules, a cloud classification algorithm to identify clear and cloudy pixels (taking into account different cloud categories), and a second module for the delineation of the raining areas according to three rainfall intensity classes. The training processes

C6759

of the two modules are presented together with the validation results for selected case studies.

General comments In my opinion the manuscript needs a deep revision to improve the description of the algorithm, which sometimes is not so sharp at the expense of the correct comprehension of the text. In particular section 3.2 and sub-sections should be improved because they represent the core of this work and I have some specific requests and/or suggestions with respect to this part.

Authors should better emphasise the novelty and main strengths of their methodology with respect to similar products.

Also the Conclusions section is in my opinion incomplete because it simply summarizes the results from the validation but it does not provide any perspectives about the future work. From the validation some abilities of the algorithm in discriminating raining from non-raining pixels are apparent with a tendency to the overestimation of precipitating areas, but there are problems with the precipitation class attribution, especially with C2 class. I think that the authors should include in the conclusions how you will proceed to improve the performances of your algorithm.

Moreover I suggest to the authors a general revision of English.

Specific comments 1. Page 13675 lines 5-13 The blended technique by Turk et al. (1999) was also implemented among the precipitation products of the Satellite Application Facility on Support to Operational Hydrology and Water Management (H-SAF) (Mugnai et al., NHES, 13, 1959-1981, 2013).

2. Page 13676 lines 23-25 Some information about MSG satellites is wrong. MSG-1 was launched in August 2002 and MSG-4 is planned for launch in 2015. I do not understand the sentence “MSG-2 was designated as the first satellite on 11 April 2007.” Now the prime operational geostationary satellite is MSG-3 since January 2013, while MSG-1 data are available since January 2004.

C6760

3. Page 13679 lines 6-25 “The training dataset used in the previous version of MACSP has been updated in order to get a better distinction of the cloudy classes.” I think that it is better at least to include a reference to Table 5 of Ricciardelli et al. (2008) to have an idea of the previous version of the training data set, and then some further details are needed about this new version of the training data set. I understand that the C_MACSP module derives from a previous work (Ricciardelli et al., 2008), but nevertheless I think that a short description of the methodology and in particular of the used spectral features are necessary. In this paragraph is presented also the validation of the C_MACSP module but without comments about the related statistical scores. These scores are shown in Table 1, which was never cited in the text.

4. Page 13680 lines 12-16 This comment concerns the rainfall intensity classes. In my opinion the non-rainy class should range from 0 to 0.1 or 0.5 mm h⁻¹ because estimates of so light rainfall intensities (< 0.1 or 0.5 mm h⁻¹) can be very unreliable and it could be safer to include them in the non-rainy class. Could you, please, comment on this?

5. Page 13680 line 18 “. . . determines the mean value $d_{min}(x,C_i)$ ” and also the eq. (1). I think that d_{min} should be replaced by d_{mean} .

6. Page 13681 line 21 “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).” This is true not only for stratiform clouds but for all precipitating clouds. Thus considering spectral channels connected with cloud microphysical properties allows to identify raining clouds also in presence of “warm” clouds, when tests based only on IR brightness temperatures are not successful.

7. Page 13682 line 15 I do not understand when the Fisher criterion (eq. 6) is really applied in the K-NNM module to reduce the number of elements in the feature vectors, because in section 3.2.2 it seems to me that you do not use this criterion, when you

C6761

describe the methodology to determine the dimension d of the feature vectors. Improve the description of this part and all sub-section 3.2.2. (especially the procedure to determine the best values of d and k).

8. Page 13685 line 1-13 “The final bootstrap training set contains the bootstrap samples obtained for $r = N_j/4, N_j/5, N_j/10, N_j/2 - 8, N_j/2 - 6, N_j/2 - 4, N_j/2 - 2$ ”. You try 7 values of the r parameter in the construction of bootstrap samples, which is the final value of r ? “The statistical scores obtained by classifying the bootstrap samples...” I did not understand which data were used as reference data set in the validation of the K-NNM results obtained for the bootstrap data set. Specify this point in the text.

9. Page 13685 line 15 The Table 6 caption is not sufficient to explain the Table content; in particular the features are absolutely cryptic.

10. Page 13685 line 16 The title of section 4 (Validation and comparisons results) suggests that, in addition to the validation results against DPC radar rain rates, the authors present comparisons between their results and other similar products from other methodologies. But I do not see these comparisons, so I think the title should be modified by removing “comparisons”.

11. Page 13687 lines 14-20 About the case study II you stated: “The RainCEIV is able to detect rainy samples with a POD of 85 %.” But there is still a remarkable over-estimation (BIAS=1.91) of the precipitating area, and moreover the statistical scores get worse when you try the rainfall class attribution with increasing FAR and Bias values and decreasing POD and HSS. So, please, add some further comments. “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 coast of Sicily and on Southern Calabria.” The statistical score values reported in this sentence do not agree with the values in Table 10 for the case study III (FAR=0.26 and POD=0.59 for C1,C2, FAR=0.27 and POD=0.59 for C1, and FAR=0.93 and POD=0.03 for C2).

C6762

In this case the algorithm underestimate the precipitating areas, and in particular for the C2 class it seems that all precipitating pixel identified by the algorithm are actually non-precipitating (FAR=0.93), and almost all true precipitating pixels are missed (POD=0.03). Thus I think that it is not possible to state that the algorithm is able to identify regions characterized by heavy precipitation, at least for this case study.

12. Page 13688 lines 3-6 “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).” I think that it is not possible to draw this kind of conclusions on the basis of the results obtained for the case study I, statistical scores are not so good. Perhaps you could analyse other case studies of this type and consider the average behaviour of the algorithm. A single case study can penalise the algorithm performances.

Technical corrections 1. Page 13674 lines 16 and 21 “Mamoudou and Gruber (2001)” The correct citation is: Ba and Gruber (2001). Please, correct also the reference in the bibliography.

2. Page 13676 line 4 “-20° W and 20° E”. Replace with “20° W and 20° E”.

3. Page 13676 line 21 Pay attention to the name of algorithm modules. From the Introduction the name of the cloud classifier module is C_MACSP, not MACSP.

4. Page 13678 line 2 Replace DCP with DPC.

5. Page 13679 line 5 I think that the Table 2 cited in this sentence is not the correct one. Table 2 contains the AMSU-B overpasses used to build the training data set of the K-NNM module; I expected a table with the MSG-SEVIRI features, which actually are displayed in Table 6.

6. Page 13862 line 6 “. . . largest variance across the design set. . .” Is this the training data set? Replace design set with training data set.

7. Page 13682 line 13 Replace K-NN with K-NNM

C6763

8. Page 13683 line 25 AMSU-B observations used for the K-NNM training data set are displayed in Table 2, not in Table 3.

9. Page 13684 line 13 The reference Efron (1979) was not included in the bibliography.

10. Page 13684 line 21 and eq.7 I do not understand the mathematical notation used for the r nearest neighbour vectors used in the bootstrap data set construction. In my opinion $y_{rj}, y(y=1, r)$ should be replaced with $y_{kj}, z(z=1, \dots, r)$. by_{kj} (line 25) should be corrected, moreover specify the range of the index i .

13. Page 13686 line 7 “The Bias score higher for C2 ...” Replace with “The higher Bias score...”.

14. Page 13686 lines 24-25 “The statistical scores calculated for each case are listed in Table 11 (for all classes), Table 12 (for C1 class), and Table 13 (for C2 class).” In the manuscript there is only Table 10, which summarizes the results for the three case studies, so correct the sentence accordingly.

15. Page 13687 line 4 The Bias value (1.67) is not correct according to Table 10, which reports a Bias value of 1.64.

16. Page 13687 line 11 Replace “...larger temporal and spatial distribution” with “...larger temporal and spatial extent”.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 10, 13671, 2013.

C6764