A statistical approach for rainfall confidence estimation using MSG-SEVIRI observations

Elisabetta Ricciardelli*, Filomena Romano*, Nico Cimini*, Frank Silvio Marzano°, Vincenzo Cuomo*

*Institute of Methodologies for Environmental Analysis - National Research Council (IMAA-CNR), Italy
°CETEMPS, DIESAP, University of Rome La Sapienza, Italy

Abstract

The knowledge of the rain rate and other meteorological parameters such as the cloud cover, the instability indices and the atmospheric motion wind vectors are essential for nowcasting applications. The high temporal and spatial resolution of the Meteosat Second Generation – Spinning Enhanced Visible and Infrared Imager (MSG-SEVIRI) observations are very useful in order to study precipitation events characterized by short duration and high temporal variability as well as the temporal evolution of long duration events. This work proposes a supervised classification algorithm to detect rainy clouds and estimate rainfall confidence. The spectral and textural features of infrared and visible MSG-SEVIRI images are conveniently selected and used as inputs in a supervised classification algorithm that classifies the rainfall confidence in three classes: non-rainy, low/moderate rain, heavy to very heavy rain. The algorithm is applied to the MSG-SEVIRI pixels previously classified as cloudy through a classification algorithm using physical, statistical and temporal approaches. The rainfall rates used in the training set are derived from an algorithm based on AMSU-B measurements, the Precipitation Estimation at MicroWave frequencies (PEMW). The rainfall confidence results, obtained for some case studies analysed in the Mediterranean area, were compared with the Eumetsat Multi-sensor Precipitation Estimate (MPE) as well as with active instrument rainfall estimates.

INTRODUCTION

Over the past decades several rain rate retrieval methods based on passive imager observations have been developed. Low Earth Orbiting (LEO) satellite MicroWave (MW) observations provide more direct and performing techniques for the retrieval of precipitation when compared with the IR/VIS techniques. The geostationary satellite IR/VIS techniques perform better over areas where rainfall is governed by deep convection than in areas where precipitation originates from the stratiform systems. However, the low spatial and temporal resolutions of MW observations make them unsuitable for monitoring extreme events and small-scale events characterized by a high temporal variability. Operational network of weather radars provides rain rate measurements characterized by a high temporal and spatial resolution, but leaves large areas uncovered where information on the occurrence and intensity of rainfall are missing. Rain rate estimates from passive imagers on geostationary satellite may bridge LEO satellite MW and weather radar gaps. This study suggests the statistical algorithm RACIV, RAinfall Confidence estimation from Infrared and Visible observations, that uses the Meteosat Second Generation – Spinning Enhanced Visible and InfraRed Imager (MSG-SEVIRI) Infrared/Visible observations to obtain rain rate confidences. The advantage of using MSG-SEVIRI VIS/IR radiation comes from the possibility to get measurements with a high spatial and temporal resolution. These characteristics are very important both for the continuous monitoring of extreme events and the study of events characterized by short duration, high temporal variability, and size slightly larger than the MSG-SEVIRI spatial resolution. RACIV operates in the Mediterranean area and approximately between 35 and 45 degrees of latitude north, and 10 and 20 degrees of longitude east, but it could operate in principle in any area.

The first section of this paper provides a description of the RACIV algorithm, the second presents the comparisons for some study cases between the estimates by RACIV, the Eumetsat Multisensor Precipitation Estimation (MPE) (Heinemann et al., 2002 - Eumetsat proceeding), and radar rain rate measurements.

THE RAIN RATE CONFIDENCE ESTIMATION FROM INFRARED AND VISIBLE OBSERVATION (RACIV) TECHNIQUE

The classification algorithm used in this study is the non-parametric classifier, K-Nearest Neighbour (K-NN). Non-parametric classifiers do not assume any *a priori* known parametric form to determine the probabilities, but they estimate these directly from the design samples. They implement the decision rule locally and the probabilities need to be estimated for each sample offered to the classifier. The K-NN classifier is very simple to understand and easy to implement. It assumes that the pixels close to each other in the feature space are likely to belong to the same class. For this reason the K-NN can be very effective if an analysis of the neighbour is useful as an explanation. To assign the MSG-SEVIRI pixel to a rainy/non-rainy class our K-NN classifier uses textural and spectral features estimated in boxes 3×3. Textural and spectral features characterizing each pixel are extracted from infrared and visible MSG-SEVIRI image data at 3.9 µm, 6.3 µm, 7.2 µm, 8.7 µm, 10.8 µm, 12 µm, 0.6 µm, 0.8 µm, 1.6 µm.

The spectral and textural features used by the classification algorithm are listed in Table 1.

Features
Radiance
Mean grey level Standard deviation of grey level Maximum grey level Minimum grey level Maximum/Minimum grey level ratio
Texture
Edge strength per unit area RG maximum within 4 directions of Mean, Contrast, Angular Second Moment, Entropy Mean of 4 directions of Mean, Contrast, Angular Second Moment, Entropy Edge strength per unit area of RG Maximum within 4 directions Mean, Contrast, Angular Second Moment, Entropy Mean of 4 directions Mean, Contrast, Angular Second Moment, Entropy

Table 1: **Spectral and Textural features used as inputs in the K-NN classifier**

The initial feature vector has 23x9 during day-time (23x6 during nigh-time) components, but not all of these components contribute to the rainfall classification. The features are chosen by the Fisher criterion selection for each training dataset. The training dataset to be used is chosen on the basis of solar zenith angle and land cover type as well as brightness temperature range and contrast range calculated at 10.8 µm and 6.2 µm wavelengths. RACIV is applied to the pixels previously classified as cloudy by the MACSP (cloud MAsk Coupling of Statistical and Physical methods) algorithm (Ricciardelli et al., 2008), and precisely to those pixels classified as low/middle level and high thick clouds. Figures 1 and 2 show the RACIV scheme and the RACIV statistical algorithm scheme, respectively.

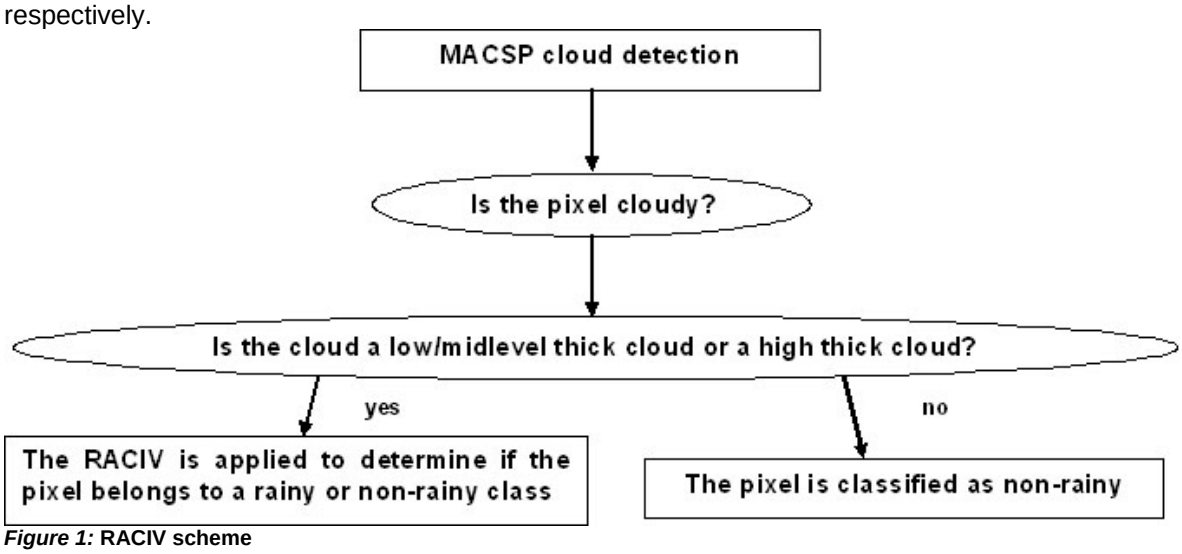

Figure 1: **RACIV scheme**

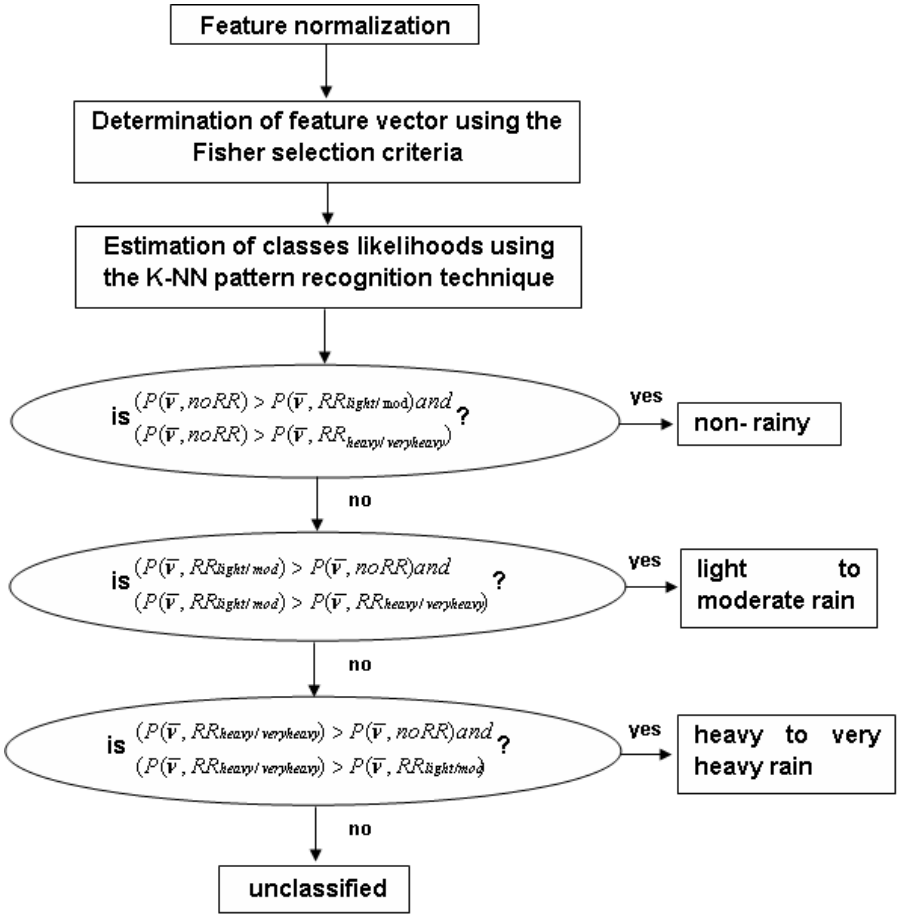

***Figure 2*: RACIV Statistical Algorithm scheme**

Training procedure

The classifier has been trained on a set of MSG-SEVIRI images collected during day- and night-time with colocated rain rate values inferred from the AMSU-B algorithm for Precipitation Estimation at MicroWave frequencies (PEMW) (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 of precipitation is 75% and 90% for rain rates greater than 1mm/h and 5mm/h,respectively (Di Tomaso et al., 2009).
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 following classes:

1) non-rainy (rain rate=0)
2) light to moderate rain (0<rain rate<4mm/h)
3) heavy to very heavy rain (rain rate>4mm/h).

These classes are chosen so as to have a sufficient number of pixels for each class to perform a significant statistical analysis. 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 2. The training-set accuracy, bias, probability of detection (POD), the Heidke Skill Score (HSS), and the False Alarm Ratio (FAR) (Erbert, 2008) were calculated for each sample included in the training set comparing against the Italian operational weather radar network rainfall values. Only the samples whose accuracy, POD, HSS, FAR and bias satisfied a relationship with threshold values – predefined for each class – are selected for the training.

Training phase – Dates for NOAA satellite overpasses
2009 September 29th, 2009 October 01th, 2009 October 02th,2010 March 04th , 2010 March 05th, 2010 March 09th, 2010 March 10th, 2010 April 26th, 2010 April 28th, 2010 May 05th, 2010 June 20th, 2010 June 21th, 2010 June 22th, 2010 June 23th, 2010 July 03th, 2010 August 04th

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

VALIDATION AND COMPARISON RESULTS

RACIV has been validated on the basis of rain rate values measured by the weather radar network operated by the Italian Civil Potection Department (DPC). Comparisons with the Eumetsat-MPE products and radar rain rate values estimated from satellite have also been considered. Only pixels completely clear or fully covered by radar rain pixels are considered in the validation stage.

The Eumetsat-MPE algorithm uses the high temporal and spatial resolution MSG/SEVIRI data in order to obtain instantaneous rain rates every 15 minutes. The Eumetsat-MPE algorithm retrieves rain rates from MSG-SEVIRI brightness temperatures on the basis of look-up tables derived from a statistical matching between MSG-SEVIRI brightness temperatures and rain rates from SSM/I passive microwave data. When considering the MPE vs. radar rain rate and the RACIV vs. radar rain rate comparisons, it must be taken into account that the MPE algorithm is limited to convective rain (Heinemann et al., 2002 - Eumetsat proceeding).

The RACIV results have been verified in a number of case studies over the Mediterranean area and are listed in Table 3. The validation method is based on a dichotomous statistical assessment.

Case	Date	Radar Measurement time (GMT)	Satellite overpass time (GMT)
1	2009 September 29th	13:00	13:00
2	2009 September 29th	17:00	17:00
3	2010 July 06th	11:30	11:30
4	2010 August 04th	14:15	14:15

Table 3: **List of the case studies**

Statistical score	Case 1 (RACIV)	Case 1 (MPE)	Case 2 (RACIV)	Case 2 (MPE)	Case 3 (RACIV)	Case 3 (MPE)	Case 4 (RACIV)	Case 4 (MPE)
Accuracy	0.99	0.99	0.99	0.99	0.91	0.91	0.98	0.98
Bias score	1.64	0.00	1.13	0.00	0.81	0.82	1.25	0.73
POD	0.64	0.00	0.28	0.00	0.29	0.29	0.38	0.07
HSS	0.49	0.00	0.75	0.00	0.27	0.27	0.33	0.07
FAR	0.60	-----	0.26	------	0.65	0.64	0.69	0.90

Table 4: **Dichotomous Statistical results (RACIV/Radar and MPE/Radar) for the case studies listed in Table 3.**

Figures 3, 4, 5 and 6 show examples of RACIV results and corresponding MPE and radar rain rate results for the case studies listed in Table 3.

For the rainfall events corresponding to cases 1, 2 and 4 (Figures 3, 4 and 6, statistical scores are shown in Table 6) RACIV algorithm overestimates rain rates in these areas (bias > 1). On the contrary, for the same cases, the MPE algorithm underestimates rain rates (bias is 0.00 for cases 1 and 2 and 0.73 for case 4). Figure 5 shows case 3 in which RACIV and MPE agree in identifying rainfall areas as confirmed by the statistics in Table 6.

The contingency values for RACIV and MPE dichotomous statistical assessment are reported in tables 5 and 6, respectively. Table 7 shows the dichotomous statistics for all the pixels considered for validation, both for RACIV and MPE rain rate results. The accuracy score shows that a large fraction (97%) of the pixels are correctly identified as rainy or non-rainy by RACIV and MPE algorithms.

However, it is recognized that this result is heavily influenced by the high occurrence of non-rainy pixels. The bias score (1.50 for RACIV and 1.29 for MPE) indicates that both RACIV and MPE have a general tendency to overestimate rainy pixels. POD shows that 71% of the rainy area is correctly detected by RACIV, while MPE detect 56% of the rainy area correctly. FAR reports that 52% and 56% of the pixels detected as rainy by RACIV and MPE, respectively, are false alarms. The statistical scores for "light to moderate" and "heavy to very heavy" classes are calculated both for RACIV and MPE algorithms. The RACIV/Radar and MPE/Radar accuracy score is the same for all classes. The bias score for "light to moderate" and "heavy to very heavy" classes indicates that both RACIV and MPE have a general tendency to overestimate rainy pixels in the two classes, and this tendency is higher for RACIV. POD shows that 87% and 50% of rainy pixels are correctly detected by RACIV and MPE in the "light to moderate" class, respectively. RACIV and MPE correctly detects 76% and 24% of rainy pixels in the "heavy to very heavy" class, respectively. FAR is about 80% both for RACIV and MPE in the "heavy to very heavy" class, while it is 71% for MPE and 59% for RACIV in the "light to moderate" class.

Weather Radar rain detection				
RaCIV rain detection		**yes**	**no**	**marginal total**
	yes	6,078	6,802	12,880
	no	2,482	340,981	343,463
	marginal total	8,560	347,783	356,343

Table 5: **Contingency table for the dichotomous statistical assessment of the RACIV algorithm for all the pixels used for validation.**

Weather Radar rain detection				
MPE rain detection		**yes**	**no**	**marginal total**
	yes	3,688	4,838	8,526
	no	2,873	344,944	347,817
	marginal total	6,561	349,782	356,343

Table 6**: Contingency table for the dichotomous statistical assessment of the MPE algorithm for all the pixels used for validation.**

Statistical scores	**All pixels (RACIV)**	**All pixels (MPE)**	**Light to moderate (RACIV)**	**Light to moderate (MPE)**	**Heavy to Very Heavy (RACIV)**	**Heavy to Very Heavy (MPE)**
Accuracy	0.97	0.97	0.97	0.98	0.99	0.99
Bias	1.50	1.29	2.12	2.59	2.20	1.30
POD	0.71	0.56	0.87	0.50	0.76	0.24
HSS	0.55	0.47	0.54	0.36	0.31	0.21
FAR	0.52	0.56	0.59	0.71	0.80	0.81

Table 7 **Dichotomous Statistics results (RACIV/Radar and MPE/Radar) for all the pixels used for validation.**

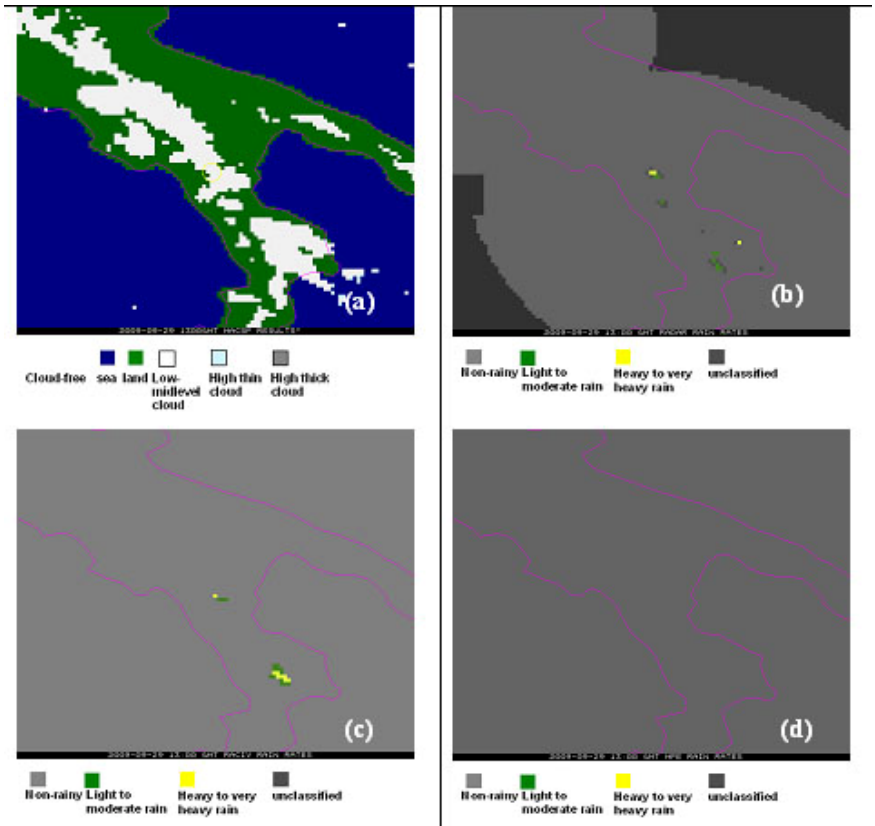

***Figure 3*: 2009 September 29th at 13:00GMT (a) MACSP results; (b) rain rate radar results collocated in MSG-SEVIRI grid; (c) RACIV rain rate results; (d) MPE rain rate results.**

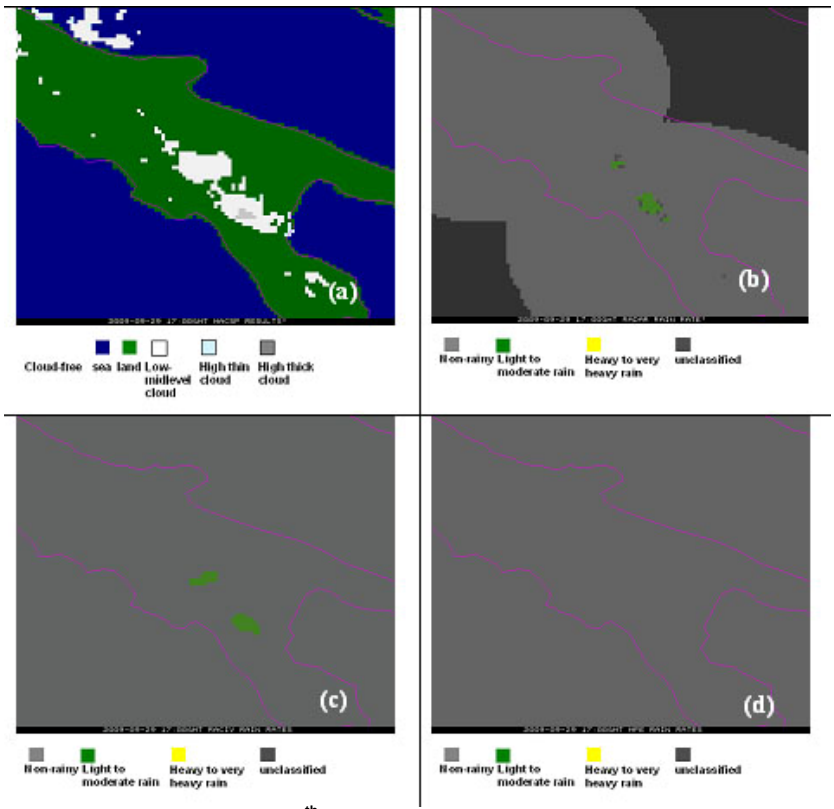

***Figure 4*: 2009 September 29th at 17:00GMT (a) MACSP results; (b) rain rate radar results collocated in MSG-SEVIRI grid; (c) RACIV rain rate results; (d) MPE rain rate results**

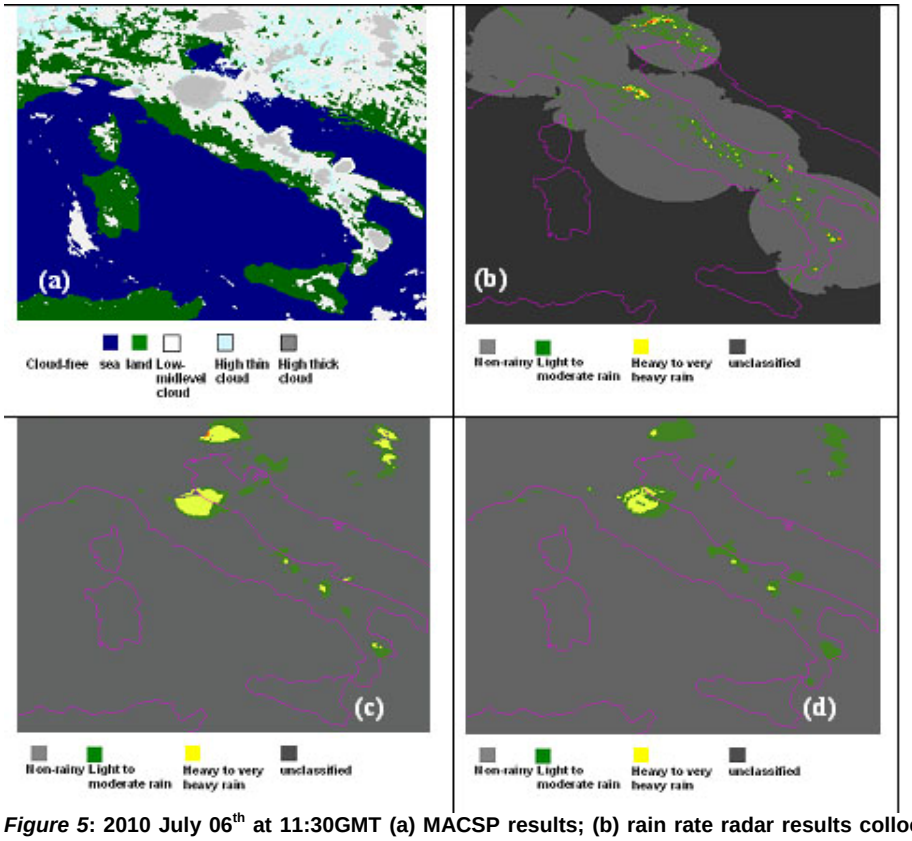

***Figure 5*: 2010 July 06th at 11:30GMT (a) MACSP results; (b) rain rate radar results collocated in MSG-SEVIRI grid; (c) RACIV rain rate results; (d) MPE rain rate results.**

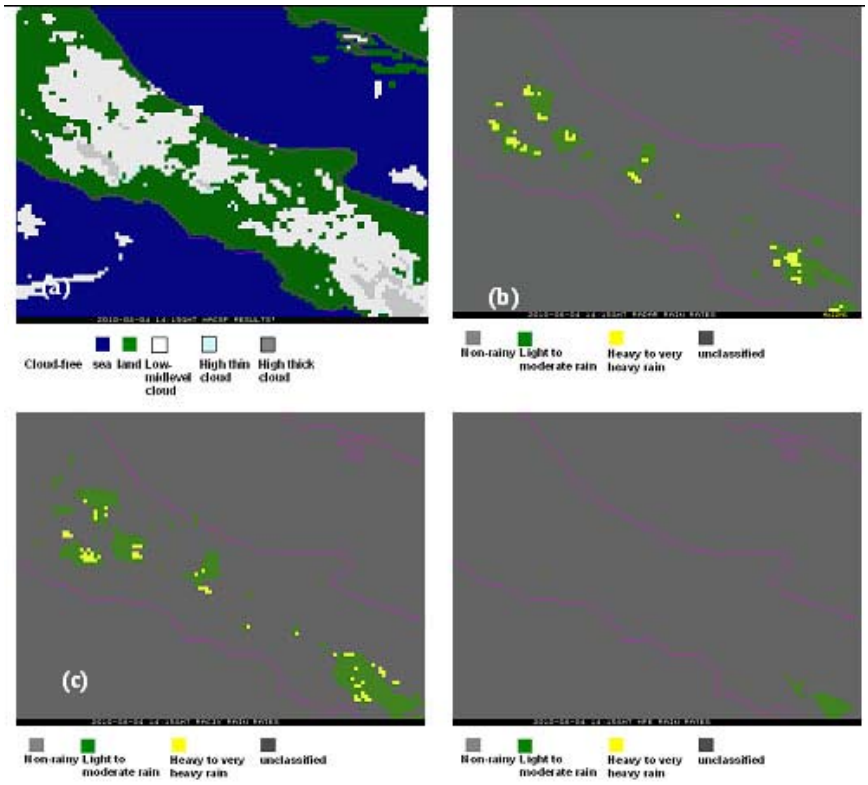

***Figure 6*: 2010 August 04th at 14:15GMT (a) MACSP results; (b) rain rate radar results collocated in MSG-SEVIRI grid; (c) RACIV rain rate results; (d) MPE rain rate results.**

CONCLUSIONS

This study proposes the statistical algorithm, RAinfall Confidence estimation from Infrared and Visible observations (RACIV), which uses the MSG-SEVIRI Infrared/Visible measurements to obtain rain rate confidences. The advantage of using MSG-SEVIRI VIS/IR observations comes from high spatial and temporal resolution. These characteristics are very important for both the continuous monitoring of extreme events and the study of events characterized by short duration, high temporal variability, and small size (of the order of the MSG-SEVIRI spatial resolution). RACIV detects rainy clouds and it classifies the rainy pixels as “light to moderate” or “heavy to very heavy” rainy. It has been trained on the AMSU-B PEMW rain rates and validated on the basis of the rain rates observation from the Italian DPC operational weather radar network. Currently RACIV works in the Mediterranean area, approximately between 35 and 45 degrees of latitude north, and 10 and 20 degrees of longitude east, but in principle it could operate in any area.
The accuracy score shows that large fractions (97%) of the pixels are correctly identified as rainy or non-rainy by RACIV. However, it is recognized that this result is heavily influenced by the high occurrence of non-rainy pixels. The bias score (1.50 for RACIV and 1.29 for MPE) indicates that both RACIV and MPE have a general tendency to overestimate rainy pixels. POD shows that 71% of the rainy area is correctly detected by RACIV, while MPE detect 56% of the rainy area correctly. FAR reports that 52% and 56% of the pixels detected as rainy by RACIV and MPE, respectively, are false alarms. The bias score for “light to moderate” and “heavy to very heavy” classes indicates that both RACIV and MPE have a general tendency to overestimate rainy pixels in the two classes, and this tendency is higher for RACIV. POD shows that 87% and 50% of rainy pixels are correctly detected by RACIV and MPE in the “light to moderate” class, respectively. RACIV and MPE correctly detect 76% and 24% of rainy pixels in the “heavy to very heavy” class, respectively. FAR is about 80% both for RACIV and MPE in the “heavy to very heavy ” class, while it is 71% and 59% in the “light to moderate” class for RACIV and MPE, respectively.

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.
The MPE rain rate products have been taken from the Eumetsat Archive.

REFERENCES

Heineman, T., Lattanzio, A., Roveda, F., The Eumetsat Multi-sensor precipitation estimate (MPE), Eumetsat Proceeding

Di Tomaso, E., Romano, F., Cuomo, V., 2009:Rainfall estimation from satellite passive microwave observations in the range 89GHz to 190GHz, *Journal of Geophysical Research*, **114**, D18203, doi:10.1029/2009JD011746.

Ebert, E, 2008: Forecast verification –Issues, Methods and FAQ, 5 Dec. (Available at http://www.bom.gov.au/bmrc/wefor/staff/eee/verif/verif_web_page.html).

Ricciardelli, E., Romano, F., Cuomo, V., 2008: Physical and statistical approaches for cloud identification using Meteosat Second Generation – Spinning Enhanced Visibile and Infrared Imager, *Remote Sensing of Environment*, **112**, 2741-2760.