

# Evaluation of a Satellite-Based Rain Estimation Algorithm Using a Network of Meteorological Stations. Preliminary Results in a Region with Complex Terrain

Stavros Kolios <sup>1,\*</sup>, Nikos Hatzianastassiou <sup>1</sup>, and Christos J. Lolis <sup>1</sup>

<sup>1</sup> Laboratory of Meteorology and Climatology, Department of Physics, University of Ioannina

\* Correspondence: stavroskoliosl@yahoo.gr

**Abstract:** The present study is a first attempt to quantitatively evaluate an existing satellite-based rain estimation algorithm using measurements from a network of ground-based meteorological stations. The study domain is the Epirus Region (the rainiest region in Greece) where the laboratory of Meteorology and Climatology of Ioannina University operates eight meteorological stations distributed across the study domain. The utilized version of the rain estimation algorithm is using the Meteosat-11 Brightness Temperature in the 10.8  $\mu\text{m}$  channel ( $\text{BT}_{10.8\mu\text{m}}$ ) to estimate the rain intensity on a 4 Km pixel basis, after discriminating the rain/non-rain pixels with a simple thresholding method. The rain recordings of the meteorological stations' network were spatiotemporally correlated with the Meteosat-11 data. These correlations led to a dataset with 1323 pairs of rain recordings and their relative rain estimations from the satellite-based algorithm. A statistical analysis of these pairs of values was conducted revealing a Mean Error (ME) of 0.22 mm/hr (14% error with respect to the mean value of the recordings). The computed Probability of False Detection (POFD), Probability of Detection (POD) and the bias score are equal to 0.22, 0.69 and 0.88, respectively. The evaluation statistics are promising for operationally using this algorithm for rain estimation on a real-time basis.

**Keywords:** rain; satellite-based estimation; Meteosat; meteorological stations; Epirus region (Greece)

---

## 1. Introduction

Rainfall as part of precipitation is a key component of the hydrological cycle, having abrupt spatial and temporal variations. An accurate knowledge of the amount of regional rainfall is essential for the welfare of society. Rainfall, especially during rainstorms, has also a direct impact on human life in terms of damages. Indeed, significant losses are caused by heavy rain events and flash floods. To improve weather and climate predictions, accurate and spatially complete rainfall records are necessary. Rain gauge data are available over land only, mainly in densely populated areas, while they cannot capture the strong spatial variations of rainfall since they provide only point measurements.

With the advent of geostationary weather satellites in the 1960s and 70's, various techniques have been developed to estimate rainfall from visible and infrared (IR) radiation upwelling from the Earth into space. The geostationary weather satellite instrumentation ensures the rapid temporal update cycle necessary to capture the growth and decay of precipitating clouds. Operational

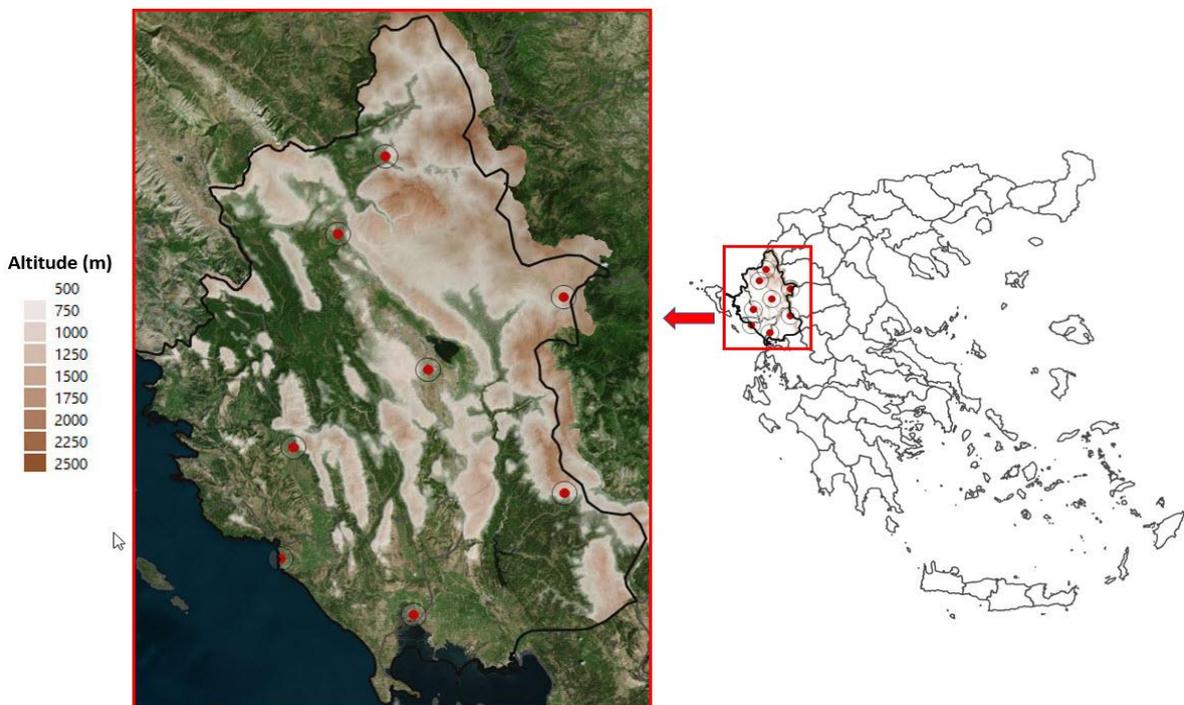
applications, however, require quantitative rainfall determination from a variety of precipitating systems, which differ both dynamically and microphysically. This fact prompts for non-unique solutions based on the physics of precipitation formation processes [1]. The operational geostationary satellites, and more specifically the Meteosat Second Generation (MSG) satellites, are highly suitable for weather monitoring over Europe and Africa due to their viewing position. Additionally, the high temporal and spatial resolution of the MSG in the visible (VIS) and infrared (IR) wavelength regions allows for the capture of growth and microstructure of precipitating clouds.

The use of satellite rainfall estimates is advantageous, because they are able to provide information for locations where ground-based observations do not exist. However, in general, the satellite estimates contain bias, since they are produced by algorithms that transform the sensors response into rainfall values. Especially during the last two decades, numerous techniques have been proposed for rain estimation using geostationary satellite data, but they differ significantly in the channels used, the implemented statistical approaches and the study domain for which they operate, thus leading to results with different accuracy [2-6].

This study provides the preliminary results of the first attempt to quantitatively evaluate an existing satellite-based rain estimation algorithm using a network of ground-based meteorological stations in an area with complex terrain and high rainfall amounts of convective nature. In Section 2, a brief description of the study domain, the data used, and the methodologies is provided. Section 3 includes the discussion of the obtained results, while conclusions are drawn in the last section.

## 2. Data and Methods

For the needs of the study, namely the validation of the satellite algorithm, measurements from eight different meteorological stations, spread across the whole Epirus Region (Greece), were used (Figure 1). The meteorological stations provide measurements for the rainfall every 30 minutes [7, 8].



**Figure 1.** The study domain (Epirus Region, Greece). The circles with red dots, refer to the locations of the meteorological stations whose rainfall measurements were used in the present study.

The rain algorithm makes use of Meteosat satellite data. More specifically, images from the satellite instrument SEVIRI (Spinning Enhanced Visible and Infrared Imager) on board the Meteosat satellite

platform available at four channels (Table 1) are used in the rain algorithm. This algorithm is part of an automated satellite-based and autonomous system [9] that using exclusively five SEVIRI channels in order to detect and nowcast mainly cloud convection. This system has been extended to provide dust estimates [10] while in the present study it was further extended to also provide rain estimations whose the accuracy is evaluated in the present study. The rain rate is estimated by the satellite algorithm every 15 minutes.

**Table 1.** Spectral characteristics of the four channels of SEVIRI instrument onboard the Meteosat satellite that are used in the satellite rain algorithm.

Channel (Band)	Spectral interval ( $\mu\text{m}$ )	Spectral center ( $\mu\text{m}$ )
5	5.35 - 7.15	6.2
6	6.85 - 7.85	7.3
9	9.8 - 11.8	10.8
10	11-13	12.0

Regarding the procedure of the rain estimation by the algorithm, firstly a cloud mask is applied on the SEVIRI multispectral data in order to characterize a specific pixel as rainy/no rainy. More specially, if for a pixel (i) the Brightness Temperature (BT) in the channel of  $10.8\mu\text{m}$  ( $BT_{10.8\mu\text{m}}$ ) is lower than 250K, (ii) the Brightness Temperature Difference (BTD) between the channels  $6.2\mu\text{m}$  and  $7.3\mu\text{m}$  ( $BTD_{6.2\mu\text{m}-7.3\mu\text{m}}$ ) is larger than -20K and (iii) the Temperature Difference (BTD) between the channels  $10.8\mu\text{m}$  and  $12.0\mu\text{m}$  ( $BTD_{10.8\mu\text{m}-12.0\mu\text{m}}$ ) is smaller than 3K then, then a rain rate (mm/hr) estimate is assigned to this pixel. Otherwise, the pixel is considered to be no rainy. It is mentioned that subsequently the rain estimate is performed by a non-linear (exponential) model having as independent variable the  $BT_{10.8\mu\text{m}}$ .

Finally, an automated procedure is followed that spatiotemporally correlates the two different datasets. More analytically, every 30-min, the total satellite-based rainfall estimations in pixel basis, were correlated with the relative (in time and space) half-hourly rainfall measurements. This procedure has been developed and applied for 20selected rainy days of the year 2019, during which significant amounts of rainfall were recorded from the network of the eight ground-based meteorological stations. As a result of this procedure, a final dataset of 1323 pairs of values in total, was created (satellite-based rain estimates and rainfall measurements from the network of the meteorological stations). In the following, two different categories of statistics were calculated for the evaluation of the rain algorithm; error statistics (Table 2) and statistical scores based on a 2x2 contingency table (Table 3).

### 3. Results

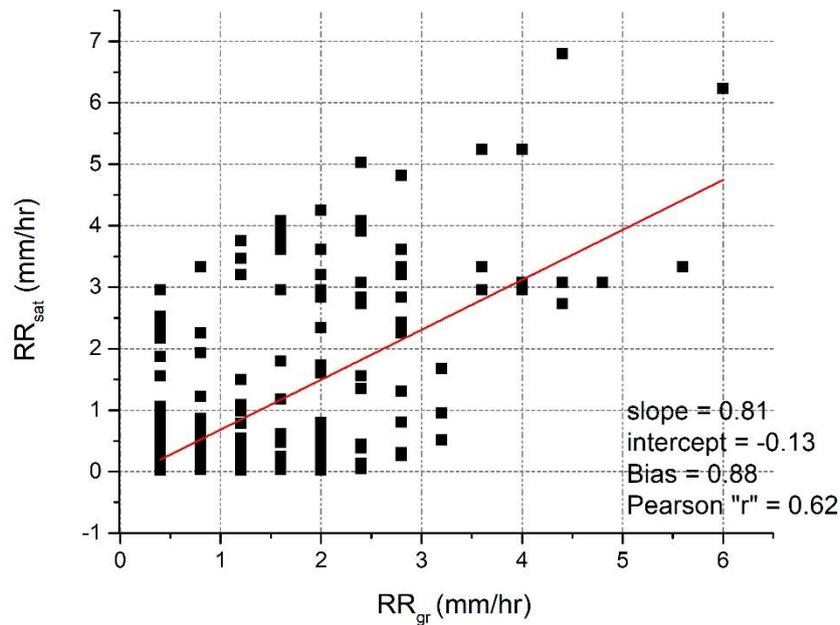
In the first part of the analysis, the statistical metrics of the Mean Error (ME), the Mean Absolute Error (MAE), the Root Mean Square Error and the Pearson Correlation Coefficient were calculated using the equations of Table 2, where the computed values are also listed. In the equations of Table 2, the term " $RR_{gr}$ " is the rain rate calculated from the meteorological stations, " $RR_{sat}$ " is the rain rate estimated from the satellite-based rain algorithm, and " $n$ " is the total number of utilized values.

**Table 2.** Analytical equations of statistical metrics used for the evaluation of the rain satellite algorithm.

Statistical Parameter	Equation	Value
MAE	$MAE = \sum_{i=1}^n ( RR_{gr_i} - RR_{sat_i} ) / n$	1.57
ME	$ME = \left( \sum_{i=1}^n (RR_{gr_i} - RR_{sat_i}) \right) / n$	0.22

$$\begin{aligned}
 \text{RMSE} \quad RMSE &= \sqrt{\left(\sum_{i=1}^n (RR_{gr_i} - RR_{sat_i})^2\right) / n} & 2.36 \\
 r_p \quad r_p &= \frac{\sum_{i=1}^n (RR_{gr_i} - \overline{RR_{gr}}) (RR_{sat_i} - \overline{RR_{sat}})}{\sqrt{\left(\sum_{i=1}^n (RR_{gr_i} - \overline{RR_{gr}})^2\right)} \sqrt{\left(\sum_{i=1}^n (RR_{sat_i} - \overline{RR_{sat}})^2\right)}} & 0.62
 \end{aligned}$$

It should be mentioned that the calculation of the statistics of Table 2 was made only for pairs with “ $RR_{gr}$ ” values larger than 0.5 mm/hr. This choice was made because it is crucial to evaluate the accuracy of the rain algorithm exclusively in cases for which notable rainfall occurred. Figure 2 displays the scatterplot between the ground-based measurements and the satellite-based estimates of rain rates, with an overall quite satisfactory correlation coefficient (0.62).



**Figure 2.** Scatterplot of rain rates of ground-based measurements ( $RR_{gr}$ , in mm/hr) versus the corresponding satellite-based rain estimates ( $RR_{sat}$ , in mm/hr) and the computed Pearson’s Correlation Coefficient.

In the second part of the evaluation, the statistical scores listed in Table 3 were calculated using the equations (1) - (3). Such statistical scores are widely used to highlight the differences between the mean of the forecasts and the mean of the observations [11]. More specifically, the POD stands for the probability of detection, FAR stands for the false alarm ratio, POFD stands for probability of false detection, and BS stands for the overall (systematic) bias. The symbols “H” (hit), “M” (miss), “CN” (correct negative), and “FA” (false alarm) are parameters that are presented in Table 3. From a practical standpoint, “H” refers to the number of correct rain estimates (measurements and estimates were found as rainy), “FA” refers to the total number of satellite estimated pixel values that were wrongly assigned with a value larger than zero (“rain”) while the relative real measurements were zero (“no rain”), “M” refers to the observed rainy values that were wrongly estimated as no-rainy, and finally, “CN” refers to the total number of paired values that both (satellite and ground) do not belong to the rainy ones (no rain values for both of the measurements and estimates). In this type of statistics, the whole dataset (1323 paired values) was used, because the nature of these statistical metrics apparently includes the cases where either the rain gauge measurement or the correlated

satellite estimate, have zero rain value. The computed Probability of False Detection (POFD), Probability of Detection (POD) and the bias (BS) score are equal to 0.22, 0.69 and 0.88, respectively, indicating a relatively good performance of the rain satellite algorithm.

**Table 3.** Contingency table with calculated statistical scores used to evaluate the ability of the satellite-based algorithm to estimate the rain intensity over the study domain.

Threshold value	Ground-based measurements		
	Yes	No	
Rain estimations	Yes	Hit (H)	False Alarm (FA)
	No	Miss (M)	Correct Negative (CN)

$$POD = \frac{H}{(H + M)} \quad (1)$$

$$POFD = \frac{FA}{(FA + CN)} \quad (3)$$

$$BS = \frac{(H + FA)}{(H + M)} \quad (4)$$

#### 4. Discussion

The first results of the evaluation of rain estimates obtained with a satellite-based algorithm, using exclusively the channel 10.8 $\mu$ m of Meteosat-11, against ground-based measurements show a satisfactory agreement and good performance of the algorithm. This is highlighted by the computed statistical metrics, revealing a ME value equal to 0.22 mm/hr (14% error with respect to the mean value of station recordings) which indicates a slight underestimation. The MAE of 1.57 mm/hr denotes a quite satisfactory accuracy, taking into account the completely different nature of the two compared datasets. Indeed, the spatial resolution of Meteosat-11, which has a spatial resolution of about 4x4 km<sup>2</sup> across the study domain, cannot capture any local spatial variations of the actual rain rate occurring in a spatial scale finer than its pixel resolution. Nevertheless, the first results are promising. The computed RMSE error equal to 2.36 mm/hr suggests that there are not significant differences between the compared pairs of values, while the correlation coefficient of 0.62 highlights a satisfactory covariance between the real measurements and their relative estimates.

These conclusions and promising results concerning the rain satellite algorithm are corroborated by the additionally computed statistical scores. Indeed, the computed Probability of False Detection (POFD) was found equal to 0.22, meaning that only 22 % of the total number of the paired values were classified wrongly (cases where the algorithm estimated rain, while the meteorological stations did not record rain). The Probability of Detection (POD) highlights a 69 % chance for a right satellite estimation (cases where the algorithm estimated rain and the meteorological stations recorded rain) while the bias score is 0.88, confirming that the estimations are very close to the observations in the majority of the examined cases.

#### 5. Conclusions

This study presents the results of a preliminary quantitative evaluation of a satellite-based rain estimation algorithm using measurements from a network of ground-based meteorological stations. The study domain is the Epirus Region (Greece), which is one of the rainiest areas of Greece, with a

complex relief, and where a network of eight meteorological stations distributed across the study area, is available. The undertaken statistical analysis led to satisfactory results, which highlight a promising performance of the satellite-based rain estimation using Meteosat data.

The automatic and reliable rain rate estimation with an algorithm using exclusively images from the Meteosat satellites, can be very useful and support both operational and research activities aiming at the real-time monitoring and nowcasting (very short-range forecasting) of rain, providing modern and useful services to final users and the wide public.

Further improvements are the objective of ongoing work, updating the existing statistical scheme for making estimates by using multispectral information and also developing sophisticated decision trees in the estimation procedure, with a main scope to optimize these encouraging preliminary results. Also, the use of a larger than the one utilized here, sample of paired values will be considered in a relative future work.

**Author Contributions:** Conceptualization, N.H. and S.K.; methodology, S.K.; software, S.K.; validation, S.K., C.L. and N.H.; formal analysis, S.K.; investigation, S.K.; resources, S.K, N.H and C.L.; data curation, S.K and C.L.; writing—original draft preparation, S.K.; writing—review and editing, N.H, S.K., and C.L.; visualization, S.K.; supervision, N.H.; project administration, N.H.; funding acquisition, N.H. All authors have read and agreed to the published version of the manuscript.”, please turn to the [CRediT taxonomy](#) for the term explanation.

**Funding:** This research was funded by European Regional Development Fund (ERDF), through Structural Funds and Cohesion Fund 2014-2020, Operational Program “Epirus”, Priority Axis “Strengthening regional competitiveness by developing innovation and ICT”, title “Development and use of digital systems, products and services by the Local Authorities” and specific title “Online platform for automatic information of the public and local authorities for possible occurrence of severe weather phenomena in Epirus Prefecture, Greece”, grant number 82740.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1. Levizzani, V.; Amorati, R.; Meneguzzo, F. A Review of Satellite-based Rainfall Estimation Methods. MUSIC – MUltiple-Sensor Precipitation Measurements, Integration, Calibration and Flood Forecasting Project, 2002. Deliverable 6.1. MUSIC – EVK1-CT-2000-00058.
2. Levizanni, V.; Schmetz, J.; Lutz, H.L.; Kerkmann, J.; Alberoni, P.P.; Cervino, M. Precipitation estimations from geostationary orbit and prospects for METEOSAT Second Generation. *Meteorol. Appl.* 2001, 8, 23-41
3. Sehad, M.; Ameer, S. 2020. A multilayer perceptron and multiclass support vector machine based high accuracy technique for daily rainfall estimation from MSG SEVIRI data. *Adv. Space Res.* 2020, 66(4), 1250-1262, doi: 10.1016/j.asr.2019.11.018
4. Lazri, M.; Labadi, K.; Brucker, J.M.; Ameer, S. 2020. Improving satellite rainfall estimation from MSG data in Northern Algeria by using a multi-classifier model based on machine learning. *J Hydrol.*, 584, doi: 10.1016/j.jhydrol.2020.124705
5. Bouaita, B.; Moussaoui, A.; Bachari, N.E.I. Rainfall estimation from MSG images using fuzzy association rules. 2020. *J Intelli. Fuz. Sys.* 2019, 37(1), 1357-1369, doi: 10.3233/JIFS-182786
6. Tapiador F.J.; Marcos C.; Sancho M.J. 2019. The Convective Rainfall Rate from Cloud Physical Properties Algorithm for Meteosat Second-Generation Satellites: Microphysical Basis and Intercomparisons using an Object-Based Method. *Rem Sens*, 11, 527; doi:10.3390/rs11050527
7. Bartzokas A.; Kotroni V.; Lagouvardos K.; Lolis C.J.; Gkikas A.; Tsirogianni M.I. 2010. Weather forecast in north-western Greece: RISKMED warnings and verification of MM5 model. *Nat Haz Earth Sys Sci* 10: 383-394.
8. Christofilakis V.; Tatsis G.; Votis C.T, Chronopoulos S.K.; Kostarakis P.; Lolis C.J.; Bartzokas A. 2018. Rainfall Measurements Due to Radio Frequency Signal Attenuation at 2 GHz. *J Signal Inform Processing* 9: 192-201.

9. Kolios S.; Hatzianastassiou N. 2019. Quantitative Aerosol Optical Depth detection during dust outbreaks from Meteosat imagery using an Artificial Neural Network model. *Rem Sens*, 11(9), 1022, doi.org/10.3390/rs11091022
10. Kolios S. 2018. Study of Mesoscale Cloud System Oscillations Capable of Producing Convective Gravity Waves. *Climate*, 6(2), 25, <https://doi.org/10.3390/cli6020025>
11. Wilks, 2019. *Statistical Methods in the Atmospheric Sciences*, 4<sup>th</sup> Edition, Elsevier Publishing, ISBN: 9780128158234



© 2020 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).