Performances of GPM satellite precipitation over the two major Mediterranean islands

DOMENICO CARACCIOLO[1,2*], ANTONIO FRANCIPANE[3], FRANCESCO VIOLA[2], LEONARDO VALERIO NOTO[3], ROBERTO DEIDDA[2]

[1] Regional Environmental Protection Agency of Sardinia (ARPAS), viale Ciusa 6, Cagliari, Italy.

[2] Dipartimento di Ingegneria Civile, Ambientale e Architettura (DICAAR), Università di Cagliari, via Marengo 2, Cagliari, Italy.

[3] Dipartimento di Ingegneria Civile, Ambientale, Aereospaziale, e dei Materiali (DICAM), Università di Palermo, viale delle Scienze, Palermo, Italy.

[*] Corresponding author email: dcaracciolo@arpa.sardegna.it; phone: +39 070 4042660

DOI: https://doi.org/10.1016/j.atmosres.2018.06.010

Abstract

This study aims to assess the reliability of satellite-precipitation products from the Global Precipitation Measurements (*GPM*) mission in regions with complex landscape morphology. Our analysis is carried out in the European mid-latitude area, namely on the two major islands of Mediterranean Sea, i.e. Sardinia and Sicily (Italy). Both islands experience precipitation originating from the interaction of steep orography on the coasts with winds carrying humid air masses from the Mediterranean Sea. The *GPM* post real-time *IMERG* (Integrated Multi-satellitE Retrievals from Global Precipitation Measurement) "Final" run product at 0.1° spatial resolution and half-hour temporal resolution have been selected for the two-year 2015-2016 period. Evaluation and comparison of the selected product, with reference to raingauge network data, are performed at hourly and daily time scales using statistical and graphical tools. The influences of morphology and land-sea coastal area transition on the reliability of the *GPM* product have been analyzed.

Confirming previous studies, results showed that *GPM* satellite data slightly overestimate rainfall over the study areas, but they are well correlated with the interpolated raingauge data. Metrics based on occurrences above a given threshold and on total volume above the same threshold were applied and revealed better performances for the latter ones. Applying the same metrics we show how *GPM* performances improve as the temporal aggregation increases. Several drawbacks were detected in the coastal areas, which were characterized by worse performances than internal areas. Statistics are generally very similar for the two considered case studies (i.e., Sardinia and Sicily) except for correlation between topography and accuracy of *GPM* products, which was slightly higher for Sardinia.

Keywords: satellite, *GPM*, IMERG, rainfall, Mediterranean islands.

1. Introduction

Reliable and accurate precipitation measurement or estimation is crucial for water resource management, and disaster monitoring (Hou et al., 2014). However, obtaining accurate high-resolution precipitation fields is still a challenging task for scientists and practitioners, especially in the poorer regions of the world. Satellite sensing provides valuable global and regional precipitation estimates (Gourley et al., 2010; Seyyedi et al., 2015), although the biases and errors of satellite-precipitation estimates need further analysis and research.

In recent years, a large number of quasi-global satellite precipitation products with various temporal and spatial resolutions have been developed and released to the public (Prakash et al., 2018; Sun et al., 2018), such as the Tropical Rainfall Measurement Mission (*TRMM*) Multi-satellite Precipitation Analysis (*TMPA*) (Huffman et al., 2007; Maggioni et al., 2016), the Climate Prediction Center MORPHing technique (*CMORPH*) analysis (Joyce et al., 2004), the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (*PERSIANN*) (Hsu et al., 1997; Sorooshian et al., 2000), and the Global Satellite Mapping of Precipitation (*GSMap*) (Kubota et al., 2007). These free and open access products have been validated through several regional (Ciabatta et al., 2017; Lo Conti et al., 2014; Sohn et al., 2010; Xue et al., 2013; Yong et al., 2012) and global (Long et al., 2015) studies.

The most recent international mission is the Global Precipitation Measurement (*GPM*) mission (Huffman et al., 2014; Huffman et al., 2017a; NASA, 2017; Skofronick-Jackson et al., 2016) which

collects data from an international constellation of satellites, including the Core Observatory satellite and approximately ten partner satellites. The main purpose of the *GPM* mission is to establish the structure and magnitude of variations in precipitation in order to understand better the water and energy cycle (Hou et al., 2014). The measurements of the *GPM* Core Observatory satellite serve as a reference standard to combine precipitation measurements from all satellites that fly within a particular constellation. Additionally, the *GPM* precipitation measurements can be combined with other data to improve accuracy and reliability. As the successor of *TRMM* satellite, the *GPM* Core Observatory was deployed on February 28, 2014 by a joint effort of *NASA* and the Japan Aerospace Exploration Agency (*JAXA*). The *GPM* Core Observatory carries a dual-frequency precipitation radar, *DPR* (with the Ku-band at 13.6 GHz and Ka-band at 35.5 GHz), and a conical-scanning multichannel *GPM* Microwave Imager (frequency channels ranging between 10 and 183 GHz). *GPM* extends the sensor package compared to *TRMM* instruments, and, therefore, the *GPM* sensors can detect light and solid precipitation more accurately than *TRMM* sensors. The Integrated Multi-satellitE Retrievals from Global Precipitation Measurement (*IMERG*) provides three kinds of products, including the near real-time ‘‘Early” and ‘‘Late” run products, and the post real-time ‘‘Final” run product, the latter being the research level product (Huffman et al., 2015). The ‘‘Final” run product is calibrated by the Global Precipitation Climatology Centre (*GPCC*) monitoring data, whose source is the Global Telecommunications System (*GTS*) that collects data from about 7000 stations around the world (Schneider et al., 2014). The time delay between satellite observation and data release (i.e., latency) is equal to 6 hours, 18 hours, and 2.5 months for the “Early”, “Late”, and “Final” *IMERG* products, respectively. While “Early” and “Late” data are available starting from 1 of April 2015 and 7 of March

2015, respectively, the "Final" run data are available from the 12 of March 2014. Furthermore, the *IMERG* product is intended to intercalibrate, merge, and interpolate all microwave estimates of the *GPM* constellation, infrared estimates, gauge observations, and other data from potential sensors at 0.1° spatial resolution and half-hour temporal resolutions (Huffman et al., 2014). Therefore, the *GPM* products have a higher spatial and temporal resolution than *TRMM* products (whose resolutions were 0.25° in space and 3 hours in time).

Several studies examined whether the rainfall products from the *GPM* mission are able to accurately estimate global precipitation (Asong et al., 2017; Behrangi and Wen, 2017; Chen and Li, 2016; Guo et al., 2016; Houze Jr et al., 2017; Li et al., 2017; Liu, 2016; Ning et al., 2017; Ning et al., 2016; Prakash et al., 2016a; Prakash et al., 2016b; Sahlu et al., 2016; Sanò et al., 2016; Sharifi et al., 2016; Speirs et al., 2017; Tang et al., 2016b). Hereafter, a brief review of the most recent and comprehensive studies from global to local scale is proposed with the aim of providing the order of magnitude of discrepancies between observations and satellite products and highlighting the reasons for such problems.

At global scale, Libertino et al. (2016) analysed the ability of *GPM* products to evaluate intense rainfall events. They compared the date of occurrence of the most severe daily rainfall events recorded each year by a global raingauge network with the ones estimated by *GPM*. The match rate between the two was found to approach 60%, indicating significant consistency between the two data sources. Behrangi and Wen (2017) analysed and quantified the errors at global scale resulting from temporal and spatial sampling of precipitation events using the V04 version of the *IMERG* product. Relative mean square error was calculated between the degraded (temporally and spatially) and original *IMERG*

products. The temporal and spatial degradation was performed by producing temporal degradation at three-hour (T3), six-hour (T6), while keeping the original spatial resolution and spatial degradation at 0.5° (S5), and 1° (S10) with original temporal resolution. The results show generally larger errors over land than ocean, especially over mountainous regions. The relative error of T6 is almost 20% larger than T3 over tropical land, but is smaller in higher latitudes. Over land, relative error of T6 is larger than S5 across all latitudes, while T6 has larger relative error than S10 poleward of 20°S-20°N. Similarly, the relative error of T3 exceeds S5 poleward of 20°S-20°N, but does not exceed S10, except in very high latitudes. Similar results are also obtained over ocean, but the error ratios are generally less sensitive to seasonal changes; moreover the results show that the spatial and temporal relative errors are weakly correlated.

At local scale, Tang et al. (2016a) evaluated the post-real time Final *IMERG* product over Mainland China from April to December 2014, at the hourly timescale, against ground-based observations interpolated with inverse distance weighting and spline methods. The product was evaluated at gridded, regional, and national scales. *IMERG* performed well at the mid- and high-latitudes, as well as in relatively dry climate regions, and it reproduced the probability density function of 3-hour and daily rainfall with good accuracy in the low ranges. Khodadoust Siuki et al. (2017) compared the half-hourly *IMERG* data with the 3-hourly raingauge data in northwest Iran for the period March-December 2014, using different evaluation indices for validation purposes. Results showed that correlation between *IMERG* and raingauge rainfall data is higher than that of *TRMM* and raingauge data, while bias confirmed that *IMERG* underestimated rainfall over the study area.

Two important aspects that were deemed crucial in determining the performances of the *GPM* product in reproducing observed rainfall, although not fully explored, regard the effects of morphology and the land-sea coastal area transition. One of the few works dealing with these topics was carried out by Kazamias et al. (2017), which evaluated *GPM* daily precipitation products over Greece and found reasonable agreement against raingauge observations, with the exception of coastal areas in which low correlations were observed. The *GPM* daily precipitation product overestimated rainfall, especially in complex terrain areas with high annual precipitation. In particular, rainfall estimates in western Greece had a strong positive bias. Another important contribution is from Kim et al. (2017), which assessed precipitation products from *GPM* using gauge-based precipitation data from Far-East Asia during the pre-monsoon and monsoon seasons. In both mountainous and coastal regions, the *GPM* showed uncertainties attributable to the presence of orographic convection and land-ocean transition.

Given these premises, the interest in testing *GPM* products in complex domains and in investigating the role of morphology in driving performances is clear. Following this research question, we selected the two major islands of the Mediterranean Sea, i.e. Sardinia and Sicily, to test satellite-precipitation *GPM-IMERG* products against data provided by dense raingauges over the two islands. Indeed, the particular combination of geographic position, climate, shape and morphology of both islands offers an interesting opportunity for the validation of satellite-precipitation data in the European mid-latitude area and in complex domains. The two islands are characterized by a complex morphology and by small spatial scale and long sea-land transition borders. Moreover, they experience precipitation originating from the interaction of steep orography on the coasts with winds carrying humid air masses from the Mediterranean Sea. The *GPM-IMERG* post real-time “Final” run product (version V04

145 released in spring 2017) at 0.1° spatial resolution and half-hour temporal resolution has been selected
146 for the two-year period 2015-2016. Evaluation and comparison of the selected product are performed
147 with reference to data provided by the raingauge network of the two islands. Both *GPM* and raingauge
148 data have been aggregated at hourly and greater time scales, then the raingauge data have been spatially
149 interpolated and resampled at the *GPM* grid resolution. In order to obtain general information about the
150 performances of estimates related to the entire two islands, features of rainfall spatial distribution and
151 the influence of the aggregation time scale have been investigated using statistical and graphical tools.
152 The influences of landscape morphology and land-sea costal area transition on the correct estimation of
153 rainfall through the *GPM* product have been analysed.

154 This work is organized as follows: description of the case study and details on raingauges and *GPM*
155 datasets are provided in the sections 2.1, 2.2 and 2.3, respectively, while the evaluation indices used to
156 compare the *GPM* and the interpolated raingauge grids are described in section 3. Results are
157 summarized in section 4: the agreements of rainfall spatial distribution at annual, hourly and daily time
158 scales is discussed in sections 4.1 and 4.2, and the influence of the orography on *GPM* performances is
159 investigated in section 4.3; an analysis on the overall performances at different time scale aggregations
160 is illustrated in section 4.4. Conclusions are drawn in section 5.

161

162 **2. Datasets**

163 **2.1. Description of the study area**

164 The study investigates the performances of the *GPM-IMERG* "Final" product in reproducing
165 precipitation fields observed on the two largest islands of the Mediterranean Sea, Sardinia and Sicily

(Italy). Sardinia has an area of about 24,000 km^2 and it is located between the 38.8°N and 41.4°N latitudes, about 400 km western the Italian peninsula, while the area of Sicily is about 25,700 km^2 and it is located between the 36°N and 39°N latitudes (see Figure 1). The waterfront length of the two islands is similar, namely it is 1,849 km for Sardinia and 1,639 km for Sicily. Both islands are characterized by complex orography as shown in Figure 1.

In Sardinia the elevation ranges between 0 and 1,820 m a.s.l.. Almost all the island is quite flat, except for a long mountain range, called Sardinian-Corse Mountain System, which is located in the eastern part of the island, running from north to south, and a smaller isolated mountain range located in the southwest. The climate of Sardinia shows a typical Mediterranean behaviour with strong seasonality characterized by hot and dry summers (June-August), rainfall mainly concentrated during autumn and winter, and a significant interannual variability often characterized by dry multi-year periods, leading to prolonged droughts. The mean annual precipitation ranges between 500 mm in lower elevation areas and 1,160 mm in the highest mountains (with a mean equal to about 650 mm) (Caracciolo et al., 2017b; Mascaro et al., 2013).

The morphology of Sicily is characterized by a mountain range along the longitudinal direction on the northern side and the Etna volcano on the eastern side, and the elevation ranges between 0 and 3,323 m a.s.l.. The mean annual precipitation over Sicily is about 715 mm with the largest rainfall amount falling during the winter months (Viola et al., 2017). The July-August months are usually characterized by little or no rainfall (Caracciolo et al., 2017a). A marked spatial variability of precipitation is observed, ranging from an average of 400 mm in the south-eastern region to an average

of 1,300 mm in the north-eastern region (Cannarozzo et al., 2006; Di Piazza et al., 2011) which is also the region characterized by the most intense rainfall at the hourly scale (Forestieri et al., 2018).

2.2. Reference raingauges networks

Sardinia's raingauge dataset is provided by the Regional Environmental Protection Agency (*ARPAS*) service that collects meteo-climatic information and provides a quality-controlled dataset. The dataset is comprised of 105 tipping-bucket raingauges, with an average density equal to about 228 km^2/gauge, as shown in Figure 1. Data are retrieved with high-temporal resolution (1 minute).

Sicily's raingauge dataset is provided by the Osservatorio delle Acque - Agenzia Regionale per i Rifiuti e le Acque (*OA-ARRA*) (*http://www.osservatorioacque.it/*). The dataset includes 195 tipping-bucket raingauges with a rather homogeneous spatial distribution, and an average density equal to about 130 km^2/gauge (see again Figure 1). Data are retrieved with hourly resolution.

For both islands the data collected in the period 2015-2016 (i.e. two years) were aggregated at hourly and longer time scales and then used as reference to evaluate the performance of the *GPM* satellite data. For sake of notation, hereafter, observed precipitation will be referred to as P_{obs}, while *GPM* estimated precipitation as P_{est}.

A direct quantitative comparison between satellite-product and raingauge data (i.e., a grid-to-point evaluation) can be problematic mainly because the first product provides an estimation of mean precipitation in each grid cell, while the latter relies on point observation. In order to overcome this issue, the Thiessen polygon method was adopted to interpolate the raingauge records over a regular grid structure. In order to reproduce the mentioned method, a high-resolution (0.01°, i.e., approximately

equal to 1 km) grid covering the islands was defined, and the rainfall value in each high-resolution grid cell was assumed equal to the value observed in the nearest raingauge. The resulting rainfall field was then resampled at the same resolution of the satellite product grid (i.e., 0.1º resolution) by simply averaging the values of the high-resolution grid.

2.3. *GPM* satellite data

In this study, we used precipitation estimates from the version V04 of the *IMERG* "Final" product, released in spring 2017 (*ftp://arthurhou.pps.eosdis.nasa.gov/gpmallversions/V04/*) (Huffman et al., 2017a; Huffman et al., 2017b). *IMERG* provides timely observation of precipitation at 0.1° resolution every 30 minutes. Overlapping raingauge locations in Sardinia and Sicily with the *GPM* grid, we found that in Sardinia about 40% of *GPM* pixels contain at least one raingauge, while in Sicily this number rises to about 60%.

As anticipated in the Introduction, the "Final" product is calibrated by the *GPCC* monitoring product whose data source is from the *GTS* with about 7000 stations around the world (Schneider et al., 2014). It is important to underline that there are only 14 *GPCC* stations in Sardinia, and 24 in Sicily and these stations are scarcely representative of the elevation distribution. Since the few stations of the global raingauge network do not allow for an optimal calibration of the product (Lo Conti et al., 2012; Lo Conti et al., 2014), the comparison of the "Final" *GPM* product with the data measured by the two regional raingauge networks, described in section 2.2, is meaningful.

3. Evaluation indices

The following set of indices has been chosen to describe satellite-precipitation performances in reproducing interpolated maps, in turn obtained from the raingauge network datasets. Performance metrics have been classified as continuous (section 3.1), categorical (section 3.2), and volumetric (section 3.3) indices, which take into account precipitation values and precipitation and volumetric occurrences, respectively.

3.1. Continuous evaluation indices

The continuous indices used in this work are: the correlation coefficient (*CC*), the Standardized Root Mean Square Error ($S-RMSE$), and the Standardized Mean Bias Error ($S-MBE$). Moreover, the normalized Taylor diagram (Taylor, 2001) has been used to graphically analyze the results.

The Correlation Coefficient (*CC*) is equal to:

$$CC = \frac{\mathrm{cov}(P_{est}, P_{obs})}{\sigma(P_{est}) \cdot \sigma(P_{obs})} \quad (1)$$

where P_{est} and P_{obs} are the precipitation estimation provided by the satellite product for a single pixel and the resampled precipitation value provided by raingauges, respectively, while cov(*X*,*Y*) is the empirical covariance between *X* and *Y* variables, and $\sigma(X)$ is the empirical standard deviation of *X*.

The Standardized Root Mean Square Error, $S-RMSE$, can be written as:

$$S-RMSE = \sqrt{\frac{\sum_{i=1}^{n}(P_{obs}^{(i)} - P_{est}^{(i)})^2}{n}} \Big/ \left(\frac{\sum_{i=1}^{n} P_{obs}^{(i)}}{n} \right) \quad (2)$$

where *n* is the number of observations.

The Standardized Mean Bias Error, $S-MBE$, can be defined as:

$$S-MBE=\frac{\sum_{i=1}^{n}(P_{obs}^{(i)}-P_{est}^{(i)})}{\sum_{i=1}^{n}P_{obs}^{(i)}} \quad (3)$$

Besides numerical indices, we decided to use also the normalized Taylor diagram (Taylor, 2001), that is based on the geometrical relationship between correlation coefficient, standard deviation and centered mean square error. This diagram has been used to summarize error statistical performances with reference to spatial averaged time series.

3.2. Categorical indices

The categorical indices used in this work are: the Probability of Detection (*POD*), the False Alarm Ratio (*FAR*), the Miss Index (*MISS*), and the Critical Success Index (*CSI*) (Wilks, 2011). We used a slightly modified version of these indices, which take into account a precipitation depth threshold.

The Probability of Detection, *POD*, sometimes referred to as Hit Score, has been estimated as:

$$POD=\frac{\sum_{i=1}^{n}I(P_{est}^{(i)}>t\,|P_{obs}^{(i)}>t)}{\sum_{i=1}^{n}I(P_{obs}^{(i)}>t)} \quad (4)$$

where t is a threshold on precipitation depth above which the precipitation occurrences are considered for *POD* computation, and $I(a|b)$ is a function counting the number of occurrences when conditions a and b are both satisfied. For t=0 equation (4) provides the common and widely used *POD* estimate. A higher threshold can be used to evaluate the performances in reproducing the occurrences of the higher precipitation quantiles. In this study, the threshold value for categorical indices is set equal to the 5th

and 50th percentiles of non-zero rainfall records in each pixel, considering only rainfall values higher than 1 mm. *POD* provides the relative number of rainfall occurrences above the threshold *t* that are correctly detected by the considered estimation product. *POD* is equal to 1 if the analysed dataset is able to represent all occurrences and 0 if no occurrences are detected. The categorical Miss Index (*MISS*) is equal to *1-POD*.

The False Alarm Ratio, *FAR*, can be defined as follows:

$$FAR = \frac{\sum_{i=1}^{n} I(P_{est}^{(i)} > t \,|P_{obs}^{(i)} < t)}{\sum_{i=1}^{n} I(P_{est}^{(i)} > t)} \quad (5)$$

FAR indicates the relative number of rainfall occurrences above a threshold *t* detected by the satellite product when the reference dataset is lower than *t*. *FAR* is equal to 0 if estimates do not reproduce any false occurrence and 1 if all estimated occurrences do not correspond to observed threshold exceedances.

The Critical Success Index (*CSI*), also known as the Threat Score, combines different aspects of the *POD* and *FAR*, describing the overall skill of the estimation relative to reference observation:

$$CSI = \frac{\sum_{i=1}^{n} I(P_{est}^{(i)} > t \;\big|P_{obs}^{(i)} > t)}{\sum_{i=1}^{n} I(P_{obs} > t) + \sum_{i=1}^{n} I(P_{est}^{(i)} > t \;\big|P_{obs}^{(i)} < t)} \quad (6)$$

The *CSI* ranges from 0 to 1, where 0 indicates no skill and 1 indicates perfect skill.

3.3. Volumetric indices

The volumetric measures, defined by AghaKouchak and Mehran (2013), provide additional information that cannot be achieved from the above categorical metrics. For most climate variables one may need to go beyond the *POD* and estimate the volume of the variable of interest detected correctly. For this reason, the Volumetric Hit Index (*VPOD*) is defined as the volume of correctly detected estimations relative to the volume of the correctly detected estimations and missed observations:

$$VPOD = \frac{\sum_{i=1}^{n}\left[P_{est}^{(i)}\middle|(P_{est}^{(i)} > t \ \& \ P_{obs}^{(i)} > t)\right]}{\sum_{i=1}^{n}\left[P_{est}^{(i)}\middle|(P_{est}^{(i)} > t \ \& \ P_{obs}^{(i)} > t)\right] + \sum_{i=1}^{n}\left[P_{obs}^{(i)}\middle|(P_{est}^{(i)} \le t \ \& \ P_{obs}^{(i)} > t)\right]} \quad (7)$$

The *VPOD* ranges from 0 to 1, with 1 being the perfect score.

The Volumetric False Alarm Ratio (*VFAR*) can be expressed as the volume of false P_{est} above the threshold t relative to the sum of rainfall volume on all estimated occurrences:

$$VFAR = \frac{\sum_{i=1}^{n}\left[P_{est}^{(i)}\middle|(P_{est}^{(i)} > t \ \& \ P_{obs}^{(i)} \le t)\right]}{\sum_{i=1}^{n}\left[P_{est}^{(i)}\middle|P_{est}^{(i)} > t\right]} \quad (8)$$

The *VFAR* ranges from 0 to 1, with 0 being the perfect score.

The fraction of the volume of missed P_{est} relative to P_{obs} can be expressed using the Volumetric Miss Index (*VMISS*):

$$VMISS = \frac{\sum_{i=1}^{n}\left[P_{obs}^{(i)}\middle|(P_{est}^{(i)} \le t \ \& \ P_{obs}^{(i)} > t)\right]}{\sum_{i=1}^{n}\left[P_{est}^{(i)}\middle|(P_{est}^{(i)} > t \ \& \ P_{obs}^{(i)} > t)\right] + \sum_{i=1}^{n}\left[P_{obs}^{(i)}\middle|(P_{est}^{(i)} \le t \ \& \ P_{obs}^{(i)} > t)\right]} \quad (9)$$

The *VMISS* ranges from 0 to 1, with 0 being the perfect score.

Finally, following the original *CSI* concept, the Volumetric Critical Success Index (*VCSI*) is defined as an overall measure of volumetric performance**:**

$$VCSI = \frac{\sum_{i=1}^{n}\left[P_{est}^{(i)}\middle|(P_{est}^{(i)} > t \ \& \ P_{obs}^{(i)} > t)\right]}{\sum_{i=1}^{n}\left[P_{est}^{(i)}\middle|(P_{est}^{(i)} > t \ \& \ P_{obs}^{(i)} > t)\right] + \sum_{i=1}^{n}\left[P_{obs}^{(i)}\middle|(P_{est}^{(i)} \le t \ \& \ P_{obs}^{(i)} > t)\right] + \left[P_{est}^{(i)}\middle|(P_{est}^{(i)} > t \ \& \ P_{obs}^{(i)} \le t)\right]} \quad (10)$$

The *VCSI* ranges from 0 to 1, with 1 being the perfect score.

4. Results

In this section we show first the results of a preliminary analysis aimed at the evaluation of the capability of *GPM* products in reproducing the long-term spatial pattern and the temporal evolution of cumulated mean areal precipitation (*MAP*) over the whole islands of Sardinia and Sicily (section 4.1). The analysis is then aimed to evaluate the reproduction of spatial patterns at hourly and daily scale using a wide range of statistical indices (section 4.2), to reveal morphology effects (section 4.3) and finally to investigate the spatially averaged regional performances at different time scale aggregations (section 4.4).

4.1. Preliminary analysis

A preliminary analysis aimed to compare, over the two-year period 2015-2016, the spatial distribution of precipitation retrieved by the *GPM* satellite product (Figure 2a,b) with that obtained by spatially interpolating the raingauge data (Figure 2c,d) for Sardinia and Sicily. For Sardinia the *GPM* (raingauge interpolated) map in Figure 2a (2c) shows values ranging between 526 (424) mm/yr and 769

(1017) mm/yr, with a spatial mean of 689 (616) mm/yr, for Sicily the *GPM* (rainfall interpolated) map in Figure 2b (2d) shows values ranging between 562 (323) mm/yr and 1009 (1596) mm/yr, with a spatial mean value equal to 766 (726) mm/yr. It can be observed that, for both Sardinia and Sicily, while the precipitation range provided by gauge data is greater than that estimated by *GPM*, the mean spatial value of the *GPM* map is greater than that derived by observations in accordance with the results obtained by Kazamias et al. (2017) in the Mediterranean area of Greece.

Another preliminary analysis aimed to evaluate the capability of the *GPM* product to reproduce the temporal evolution of observed *MAP* over the whole of Sardinia and Sicily. The cumulative *MAP* from 2015 and 2016 obtained by spatially averaging the rainfall values of the *GPM* and gauge-interpolated pixels covering Sardinia and Sicily is shown in Figure 3a and 3b, respectively. We can observe that, for both Sardinia and Sicily, the two time series in each subplot of Figure 3 (i.e., *GPM* and interpolated raingauge values) evolve in time with a very similar behaviour. While the *GPM* time series slightly overestimates the precipitation measured by the raingauges on both islands, *GPM* cumulated values are closer to observations in Sicily than in Sardinia. A first general consideration is that the *GPM* measures are able to correctly detect the rainfall timing over the two years on both the islands, with a good agreement in the total cumulated rainfall during the analysed period.

4.2. Spatial analysis

In order to obtain a quantitative comparison of satellite product performances, the continuous evaluation indices, described in section 3.1, have been calculated for each grid cell at hourly and daily scale for Sardinia and Sicily.

Figure 4 shows the spatial distribution of the correlation coefficient (*CC*, defined by Eq. 1), the standardized root mean square error (*S-RMSE*, defined by Eq. 2), and the standardized mean bias error (*S-MBE*, defined by Eq. 3), obtained by comparing hourly time series in each grid cell for Sardinia (Panels a, c, e) and Sicily (Panels b, d, f). Table 1 reports minimum, mean and maximum values of each index at hourly and daily time scale for both islands. Looking in detail at Figure 4a we can observe the largest values of the *CC* mainly in the northern part of Sardinia, while for Sicily (Figure 4b) the largest values of the *CC* are located in the southwestern part. The lowest values of *S-RMSE* are in the eastern part of Sardinia, while the largest values are in the north-west (Figure 4c). In Sicily, *S-RMSE* ranges from lowest values in the north-east and north-west to largest values in the south-east and north-west (Figure 4d). The *S-MBE* reaches the best performance (green pixels in Figure 4e,f) in the centre of the two islands. This could be due to an issue arising from coastal treatments because retrieval algorithms can suffer from some weaknesses due to different radiative properties of hydrometeors over the land and the ocean respectively (Lo Conti et al., 2014). From Table 1 it can be observed that *GPM* performances at hourly time scale are slightly better in Sardinia than in Sicily.

At daily time scale, in general, the performances increase, as demonstrated by the largest values of *CC* (Table 1), which are equal to 0.82 in the centre of Sardinia (Figure 5a), while for Sicily (Figure 5b) the largest values of *CC* are equal to 0.85 in the south-centre of the island. Looking at the spatial distribution of *S-RMSE* in Sardinia (Figure 5c) the lowest value can be observed in the central part, while the largest values can be found in the northwestern part. The spatial distribution in Sicily (Figure 5d) shows maximum values in the south-east and north (as at the hourly time scale) and minimum values in the western-central part of the island. The *S-RMSE* values are comparable to those obtained

by Xu et al. (2017) and Li et al. (2017). *S-MBE* is identical to the hourly case because of its mathematical definition and thus it is not shown in Figure 5. At daily time scale, the *GPM* performances in reproducing observations are slightly better in Sicily than in Sardinia (Table 1). Comparing these results with that obtained by Lo Conti et al. (2014) in Sicily using the *PERSIANN* and *TMPA* data, an increase of the performances of the new satellite products can be highlighted.

Results arising from the analysis of categorical and volumetric indices described in section 3.2 and 3.3, respectively, have been summarized in Table 1 and depicted in Figures 6-9 which display sample *POD, VPOD, FAR, VFAR, MISS, VMISS, CSI*, and *VCSI* values for Sardinia and Sicily, for hourly (Figures 6-7) and daily (Figures 8-9) time scales, respectively. Threshold values adopted for categorical and volumetric indices are set in each grid cell equal to 5^{th} and 50^{th} percentiles of the empirical distribution of the resampled observed non-zero precipitation in the selected grid cell, considering only rainfall values higher than 1 mm. For the sake of simplicity, we will use the notation t_α to refer to thresholds t corresponding to α^{th} percentile. *POD*, *FAR*, *MISS* and *CSI* maps allow the comparison of product capability of reproducing precipitation occurrences for each location at different probability levels, while the volumetric indices (i.e., *VPOD, VFAR, VMISS, VCSI*) provide additional information beyond the classical 2x2 contingency table. An overall survey of Figures 6-9 reveals a systematic worsening of *GPM* performances as the threshold rises from t_5 (left subplots) to t_{50} (right subplots), i.e. *POD, VPOD, CSI, VCSI* indices decrease, while *FAR, VFAR, MISS, VMISS* indices increase.

In Figure 6, the *POD* and *VPOD* values of the hourly data in Sardinia decrease when the thresholds t increase for α equal to 5% and 50%. The *POD* values range between 0.83 (0.52) and 1 (0.87) using thresholds t_5 (t_{50}), (Table 1). The *VPOD* values indicate that *GPM* detects between 87% (57%) and

100% (95%) of the volume of observed precipitation exceeding the 5^{th} (50^{th}) percentile, showing a good performance of *GPM* satellite data in detecting the volume of observed precipitation in almost all of the island of Sardinia. Using thresholds t_5 (t_{50}), the *FAR* and *VFAR* values over Sardinia show that the false precipitation is below 10% (51%) with highest values clusterized in the northeastern and southwestern parts. Similarly, the *MISS* index values are lower than 0.17 (0.48) using thresholds t_5 (t_{50}), showing a good capability of the *GPM* to detect a large fraction of precipitation. Based on *VMISS*, for t_5 and t_{50}, the fraction of precipitation amount that *GPM* does not detect is very close to zero. Considering the *CSI* index, the overall performance is between 0.78 (0.37) and 1 (0.68) for thresholds t_5 (t_{50}), whereas taking into account the volume of precipitation the *VCSI* indicates a higher performance score. It is worth noting that the volumetric indices (i.e., *VPOD, VFAR, VMISS, VCSI*) often provide better results than those obtained by the categorical indices (i.e., *POD, FAR, MISS, CSI*) highlighting that, in terms of total volume, the agreement between *GPM* and measured values is better than a simple counting of threshold exceeding.

Looking at results for hourly data in Sicily in Figure 7 and Table 1, we found similar indices values to those obtained for Sardinia (Figure 6), although with different spatial patterns. The *POD* values range between 0.73 (0.30) and 1 (0.91), while *VPOD* values range between 0.80 (0.39) and 1 (0.96) for thresholds t_5 (t_{50}) showing a good performance of *GPM* data in detecting volume of observed precipitation on almost all the island. The *FAR* values are between 0 (0.19) and 0.12 (0.59), while *VFAR* values range between 0 (0.13) and 0.15 (0.67) for thresholds t_5 (t_{50}) with lowest values in the southwestern part of Sicily. The *MISS* values are lower than 0.27 (0.67) using thresholds t_5 (t_{50}), showing good *GPM* performances. Focusing attention on *VMISS* and using thresholds t_5 and t_{50}, the

fraction of precipitation that *GPM* does not detect is relatively small. Moving to the last two indices, the *CSI* values are between 0.72 (0.26) and 1 (0.72) for thresholds t_5 (t_{50}), whereas the *VCSI* indicates a higher performance score between 0.80 (0.24) and 1 (0.77) for thresholds t_5 (t_{50}), testifying good overall performances of *GPM*. These results are in accordance with those obtained in Mediterranean areas and in Sicily by Panegrossi et al. (2016).

The same performance metrics were computed also for rainfall fields aggregated at daily scales. Results are shown in Figures 8 and 9 for Sardinia and Sicily, respectively. Comparing these maps with those in previous Figures 6 and 7, a general improvement can be observed, as expected, together with a very similar pattern (see also Table 1).

In order to investigate better the existence of possible spatial patterns of categorical and volumetric indices for daily and hourly precipitation, we calculated the probabilistic metric *Global Moran's I* (Moran, 1950). The *Moran's I*, which ranges from −1 to +1, provides information about the spatial autocorrelation of a feature on the basis of feature locations and feature values simultaneously. It is able to evaluate whether the pattern expressed by a feature is clustered (positive values), dispersed (negative values), or random (values close to zero).

Results of *Moran's I* obtained for all of the categorical and volumetric indices previously calculated for thresholds t_5 and t_{50} are provided in Table 2. A statistical test with 1% significance revealed that only in four cases (relative to Sardinia) the null hypothesis of random fields cannot be rejected (italic bold font in Table 2), while for all the other cases the spatial distribution can be considered clusterized. No cases of dispersed pattern were found.

4.3. Elevation effects and coastal influence

In order to analyse landscape morphology effects on the accuracy of *GPM* products, we investigated the dependence of *S-RMSE* at hourly time scale on the DEM elevations of grid cells at 0.1° resolution. With this aim, the values of *S-RMSE* have been plotted versus DEM elevations in Figure 10a, b. The dependence was analysed by linear regressions characterized by a coefficient of determination R^2 equal to 0.33 for Sardinia (Figure 10a) and 0.13 for Sicily (Figure 10b) and negative slopes, suggesting an improvement in *GPM* accuracy with elevation which is in disagreement with the results found by Xu et al. (2017) (i.e., improvement of *GPM* accuracy reducing the elevation). This disagreement could be related to the fact that our results may be influenced by the unfair distribution of the elevations; in fact there are few high-elevation pixels where we found higher accuracy, and many low-elevation pixels where performances are highly variable, due also to the proximity of coastal line.

To exclude that results can be influenced by the interpolation methods, the analysis was repeated by calculating the *S-RMSE* between the *GPM* pixel values and the hourly rainfall values observed at the gauge inside the *GPM* pixel. Results are shown in Figure 10c,d where we can observe a slightly stronger dependence: while for Sicily R^2 is low and equal to 0.16 (Figure 10d), a slightly higher correlation has been found in Sardinia, where the R^2 is equal to 0.48 (Figure 10c). Nevertheless, this kind of analysis would require more data to assess a robust estimate of the relation between *GPM* accuracy and terrain morphology.

As mentioned in the introduction, some authors highlighted several drawbacks of *GPM* products in coastal regions, because the *GPM* algorithm uses a combined process for ocean and land that contains problematic ambiguity in discerning between rain and no-rain (Wolff and Fisher, 2009). This issue has

been investigated also with our case studies. Specifically, an analysis has been carried out to verify if *GPM* performances are statistically different in coastal or internal pixels. This kind of analysis has been performed focusing on the *CC*, *S-RMSE* and *S-MBE* at hourly time scale. Namely, we have divided all the considered pixels into two subsets, one including only the internal pixels and the other considering only the coastal ones, and we calculated the values of *CC*, *S-RMSE* and *S-MBE* at hourly time scale for each subset. Results are reported in Table 3 where we can observe, in general, better performances (i.e., higher *CC* and smaller *S-RMSE* and *S-MBE*) when only the internal pixels are considered, while the worst results have been detected in the coastline pixels, in accordance with the result obtained by Kazamias et al. (2017) and Kim et al. (2017). Finally, in order to verify if performances are statistically different between coastal and internal pixels, we performed the two-sample *t*-test comparing the two vectors (i.e., internal and coastal) and testing if they come from independent normal distributions with equal means and equal but unknown variances. The test has been rejected with significance level equal to 0.05 for all the three considered indices and for both the regions, confirming the worst performances detected along the coastal line.

4.4. Analysis of spatially averaged precipitation values

To further investigate how *GPM* product performances vary with different temporal aggregation scales, *CC*, *S-RMSE*, *POD*, and *FAR* indices have been computed considering precipitation depths at 1, 3, 6, 12 hours, and from 1 to 60 days. Results of such analysis have been summarized in Figure 11, where spatially averaged indices for each time aggregation scale are plotted for Sardinia (Figure 11a, c, e, g) and Sicily (Figure 11b, d, f, h). Figure 11 describes an improvement of performances as time

aggregation increases, showing about the same behaviour for both islands and also confirming similar results obtained by Lo Conti et al. (2014), Sohn et al. (2010) and Behrangi and Wen (2017). In Figure 11a,b the black solid line represents the regional mean of the *CC* obtained at pixel scale, while the black dashed line represents the *CC* between regional aggregated *MAP* time series. The *CC* values increase as the time resolution decreases, while the *S-RMSE* decreases as the time-aggregation interval increases (Figure 11c,d). *POD* (Figure 11e,f) and *FAR* (Figure 11g,h) values have been calculated considering thresholds t_5 and t_{50} for each pixel: *POD* (*FAR*) increases (decreases) as the time-aggregation scale increases. Almost all the plotted indices improve in the first time scales (i.e., from hourly to 5-day) and then tend to become constant for time aggregation larger than 10 days. The *S-MBE* levels do not change with the temporal aggregation because it does not affect the ratio between mean bias and mean precipitation and, for this reason, the *S-MBE* plot has not been shown.

The normalized Taylor diagram has been used finally to compare the time series of *MAP* values as obtained from observation and from *GPM* outputs, both spatially averaged over the whole of the islands. The comparison has been carried out at hourly and daily time scale for both Sardinia and Sicily and it is displayed in Figure 12. The normalized Taylor diagram summarizes the distance between testing and reference series in terms of normalized standard deviation, correlation coefficient (the same of Figure 11a,b, dashed black line), and normalized *RMSD* (root mean square difference). The black star refers to the observed value for which the normalized standard deviation is equal to 1, the radial distance from the black star quantifies the centered *RSMD* normalized by the standard deviation of observed *MAP*, the azimuth and the radial distance from the origin quantify *CC* and normalized standard deviation, respectively. The observed standard deviations are 0.285 and 4.421 mm (0.285 and

4.427 mm) for hourly and daily *MAP* time series in Sardinia (Sicily), respectively. Hourly data (circles in Figure 12) characterized by a normalized standard deviation close to 1.6 for both Sardinia and Sicily, have a correlation coefficient of about 0.68 and normalized *RMSD* lower than 1.3. Looking at daily data (squares in Figure 12), the normalized standard deviation is about 1.4 and the correlation coefficient improves up to 0.85, while the normalized *RMSD* is lower than 0.8. Interestingly, hourly and daily statistics are very similar for the two considered cases (i.e., Sardinia and Sicily). This suggests that the influence of temporal aggregation on *GPM* performances is greater than the effects of geographic context.

5. Conclusions

In this study, *GPM-IMERG* post real-time "Final" product data (version V04 released in spring 2017) were evaluated against measured rainfall over the two largest islands of the Mediterranean Sea (Sardinia and Sicily) during a two-year window (2015-2016) at hourly and longer time scales.

For an accurate evaluation, a pool of statistical indices, including continuous, categorical, and volumetric indices, were applied to evaluate the performances of satellite products. The results of this study showed that correlation among *GPM* satellite data and interpolated gauge rainfall reaches values of *CC* equal to 0.8 at daily scale. The values of bias confirmed that *GPM* satellite data slightly overestimates rainfall over the study areas. The volumetric indices (i.e., *VPOD, VFAR, VMISS, VCSI*) often provided better results than categorical ones (i.e., *POD, FAR, MISS, CSI*), meaning that the agreement between *GPM* and measures is better in terms of total volume above a given threshold rather than in terms of occurrences above the same threshold.

Interestingly, hourly and daily statistics are very similar for the two considered cases (i.e., Sardinia and Sicily). This suggests that *GPM* satellite data performances are driven by temporal aggregation more than the geographic context. This behaviour has been investigated better by analysing how *GPM* performances improve as the temporal aggregation increases and it was found that a stable performance is reached at approximately 10 days. The *GPM* products showed some weaknesses in reproducing observed rainfall in proximity of land-sea transition: indeed the worst results for all the considered indices have been obtained for the coastline pixels in both islands. Finally, we searched for a relation between accuracy of *GPM* products and landscape morphology. This kind of connection has not been found in Sicily, while in Sardinia a weak correlation indicates increasing accuracy with elevation.

In conclusion, although we found promising results indicating the potential of *GPM* estimates for hydrological applications, there is still room for improvement under technological and scientifical perspectives. While the first involves the long-term continuous instrumentation update and innovation, scientific efforts can certainly contribute to improve retrieval algorithms to account for coastal and morphological effects, even including bias-correction procedures.

Acknowledgments

This work was partially supported by the Sardinian Region with law 7/2007, funding call 2013, project CRP 78576: “Anthropogenic and climatic impacts on the hydrological cycle at basin and hillslope scale” (*Impatti antropogenici e climatici sul ciclo idrologico a scala di bacino e di versante*).

This work was partially supported by the *Fondazione di Sardegna*, funding call 2017, project: Impacts of climate change on water resources and floods, CUP: F71I17000270002.

525

References

AghaKouchak, A., Mehran, A., Extended contingency table: Performance metrics for satellite observations and climate model simulations, *Water Resources Research* **49**(2013), pp. 7144-7149.

Asong, Z., Razavi, S., Wheater, H., Wong, J., Evaluation of Integrated Multisatellite Retrievals for GPM (IMERG) over Southern Canada against Ground Precipitation Observations: A Preliminary Assessment, *Journal of Hydrometeorology* **18**(2017), pp. 1033-1050.

Behrangi, A., Wen, Y., On the Spatial and Temporal Sampling Errors of Remotely Sensed Precipitation Products, *Remote Sensing* **9**(2017), p. 1127.

Cannarozzo, M., Noto, L., Viola, F., Spatial distribution of rainfall trends in Sicily (1921–2000), *Physics and Chemistry of the Earth, Parts A/B/C* **31**(2006), pp. 1201-1211.

Caracciolo, D., Arnone, E., Conti, F.L., Noto, L.V., Exploiting historical rainfall and landslide data in a spatial database for the derivation of critical rainfall thresholds, *Environmental Earth Sciences* **76**(2017a), p. 222.

Caracciolo, D., Deidda, R., Viola, F., Analytical estimation of annual runoff distribution in ungauged seasonally dry basins based on a first order Taylor expansion of the Fu's equation, *Advances in Water Resources* **109**(2017b), pp. 320-332.

Chen, F., Li, X., Evaluation of IMERG and TRMM 3B43 monthly precipitation products over mainland China, *Remote Sensing* **8**(2016), p. 472.

Ciabatta, L. *et al.*, Daily precipitation estimation through different microwave sensors: Verification study over Italy, *Journal of Hydrology*(2017).

Di Piazza, A., Conti, F.L., Noto, L., Viola, F., La Loggia, G., Comparative analysis of different techniques for spatial interpolation of rainfall data to create a serially complete monthly time series of precipitation for Sicily, Italy, *International Journal of Applied Earth Observation and Geoinformation* **13**(2011), pp. 396-408.

Forestieri, A., Lo Conti, F., Blenkinsop, S., Fowler, H.J., Noto, L.V., Regional frequency analysis of extreme rainfall in Sicily (Italy), *International Journal of Climatology*(2018).

Gourley, J.J., Hong, Y., Flamig, Z.L., Li, L., Wang, J., Intercomparison of rainfall estimates from radar, satellite, gauge, and combinations for a season of record rainfall, *Journal of Applied Meteorology and Climatology* **49**(2010), pp. 437-452.

Guo, H. *et al.*, Early assessment of integrated multi-satellite retrievals for global precipitation measurement over China, *Atmospheric Research* **176**(2016), pp. 121-133.

Hou, A.Y. *et al.*, The global precipitation measurement mission, *Bulletin of the American Meteorological Society* **95**(2014), pp. 701-722.

Houze Jr, R.A. *et al.*, The Olympic Mountains Experiment (OLYMPEX), *Bulletin of the American Meteorological Society* **98** (2017), pp. 2167-2188.

Hsu, K.-l., Gao, X., Sorooshian, S., Gupta, H.V., Precipitation estimation from remotely sensed information using artificial neural networks, *Journal of Applied Meteorology* **36**(1997), pp. 1176-1190.

Huffman, G. *et al.*, GPM Integrated Multi-Satellite Retrievals for GPM (IMERG) Algorithm Theoretical Basis Document (ATBD) Version 4.4. PPS, NASA/GSFC, 30 pp. (2014).

Huffman, G.J., Bolvin, D.T., Braithwaite, D., Hsu, K., Joyce, R., et al., Algorithm Theoretical Basis Document (ATBD) Version 4.6 NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (IMERG), *NASA 1-25.< https://pmm.nasa.gov/index.php?q=data-access/downloads/gpm>*(2017a).

Huffman, G.J., Bolvin, D.T., Nelkin, E.J., Integrated Multi-satellitE Retrievals for GPM (IMERG) technical documentation, *NASA/GSFC Code* **612**(2015), p. 47.

Huffman, G.J., Bolvin, D.T., Nelkin, E.J., Stocher, E.F., V04 IMERG Final Run Release Notes, *NASA 1-2.< https://pmm.nasa.gov/data-access/downloads/gpm/ V04 IMERG Final Run Release Notes >*(2017b).

Huffman, G.J. *et al.*, The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales, *Journal of Hydrometeorology* **8**(2007), pp. 38-55.

Joyce, R.J., Janowiak, J.E., Arkin, P.A., Xie, P., CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution, *Journal of Hydrometeorology* **5**(2004), pp. 487-503.

Kazamias, A., Sapountzis, M., Lagouvardos, K., Evaluation and intercomparison of GPM-IMERG and TRMM 3B42 daily precipitation products over Greece. *Fifth International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2017)*, International Society for Optics and Photonics (2017), p. 1044413.

Khodadoust Siuki, S., Saghafian, B., Moazami, S., Comprehensive evaluation of 3-hourly TRMM and half-hourly GPM-IMERG satellite precipitation products, *International Journal of Remote Sensing* **38**(2017), pp. 558-571.

Kim, K., Park, J., Baik, J., Choi, M., Evaluation of topographical and seasonal feature using GPM IMERG and TRMM 3B42 over Far-East Asia, *Atmospheric Research* **187**(2017), pp. 95-105.

Kubota, T. *et al.*, Global precipitation map using satellite-borne microwave radiometers by the GSMaP project: Production and validation, *IEEE Transactions on Geoscience and Remote Sensing* **45**(2007), pp. 2259-2275.

Li, N. *et al.*, Statistical assessment and hydrological utility of the latest multi-satellite precipitation analysis IMERG in Ganjiang River basin, *Atmospheric Research* **183**(2017), pp. 212-223.

Libertino, A., Sharma, A., Lakshmi, V., Claps, P., A global assessment of the timing of extreme rainfall from TRMM and GPM for improving hydrologic design, *Environmental Research Letters* **11**(2016), p. 054003.

Liu, Z., Comparison of integrated multisatellite retrievals for GPM (IMERG) and TRMM multisatellite precipitation analysis (TMPA) monthly precipitation products: initial results, *Journal of Hydrometeorology* **17**(2016), pp. 777-790.

Lo Conti, F., Caracciolo, D., Pumo, D., Viola, F., Noto, L., Analisi della affidabilità dei database pluviometrici globali a scala locale. Applicazione al territorio siciliano, *XXXIII Convegno Nazionale di Idraulica e Costruzioni Idrauliche*(2012), pp. 1-11.

Lo Conti, F., Hsu, K.-L., Noto, L.V., Sorooshian, S., Evaluation and comparison of satellite precipitation estimates with reference to a local area in the Mediterranean Sea, *Atmospheric Research* **138**(2014), pp. 189-204.

Long, D., Longuevergne, L., Scanlon, B.R., Global analysis of approaches for deriving total water storage changes from GRACE satellites, *Water Resources Research* **51**(2015), pp. 2574-2594.
Maggioni, V., Meyers, P.C., Robinson, M.D., A review of merged high-resolution satellite precipitation product accuracy during the Tropical Rainfall Measuring Mission (TRMM) era, *Journal of Hydrometeorology* **17**(2016), pp. 1101-1117.
Mascaro, G., Deidda, R., Hellies, M., On the nature of rainfall intermittency as revealed by different metrics and sampling approaches, *Hydrology and Earth System Sciences* **17**(2013), p. 355.
Moran, P.A., Notes on continuous stochastic phenomena, *Biometrika* **37**(1950), pp. 17-23.
NASA, Algorithm Theoretical Basis Document GLOBAL PRECIPITATION MEASUREMENT (GPM) 714 MISSION, < *http://rain.atmos.colostate.edu/ATBD/ATBD_GPM_June1_2017.pdf*>(2017).
Ning, S. *et al.*, Error Analysis and Evaluation of the Latest GSMap and IMERG Precipitation Products over Eastern China, *Advances in Meteorology* **2017**(2017).
Ning, S., Wang, J., Jin, J., Ishidaira, H., Assessment of the Latest GPM-Era High-Resolution Satellite Precipitation Products by Comparison with Observation Gauge Data over the Chinese Mainland, *Water* **8**(2016), p. 481.
Panegrossi, G. *et al.*, Use of the GPM constellation for monitoring heavy precipitation events over the Mediterranean region, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **9**(2016), pp. 2733-2753.
Prakash, S. *et al.*, A preliminary assessment of GPM-based multi-satellite precipitation estimates over a monsoon dominated region, *Journal of Hydrology*(2016a).
Prakash, S., Mitra, A.K., Gairola, R.M., Norouzi, H., Pai, D.S., Status of High-Resolution Multisatellite Precipitation Products Across India. *Remote Sensing of Aerosols, Clouds, and Precipitation*, Elsevier (2018), pp. 301-314.
Prakash, S., Mitra, A.K., Pai, D., AghaKouchak, A., From TRMM to GPM: How well can heavy rainfall be detected from space?, *Advances in Water Resources* **88**(2016b), pp. 1-7.
Sahlu, D., Nikolopoulos, E.I., Moges, S.A., Anagnostou, E.N., Hailu, D., First evaluation of the Day-1 IMERG over the upper Blue Nile Basin, *Journal of Hydrometeorology* **17**(2016), pp. 2875-2882.
Sanò, P. *et al.*, The new Passive microwave Neural network Precipitation Retrieval (PNPR) algorithm for the cross-track scanning ATMS radiometer: description and verification study over Europe and Africa using GPM and TRMM spaceborne radars, *Atmospheric Measurement Techniques* **9**(2016), p. 5441.
Schneider, U. *et al.*, GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle, *Theoretical and Applied Climatology* **115**(2014), pp. 15-40.
Seyyedi, H., Anagnostou, E.N., Beighley, E., McCollum, J., Hydrologic evaluation of satellite and reanalysis precipitation datasets over a mid-latitude basin, *Atmospheric Research* **164**(2015), pp. 37-48.
Sharifi, E., Steinacker, R., Saghafian, B., Assessment of GPM-IMERG and other precipitation products against gauge data under different topographic and climatic conditions in Iran: preliminary results, *Remote Sensing* **8**(2016), p. 135.

Skofronick-Jackson, G. *et al.*, The global precipitation measurement (GPM) mission for science and society, *Bulletin of the American Meteorological Society*(2016).

Sohn, B., Han, H.-J., Seo, E.-K., Validation of satellite-based high-resolution rainfall products over the Korean Peninsula using data from a dense rain gauge network, *Journal of Applied Meteorology and Climatology* **49**(2010), pp. 701-714.

Sorooshian, S. *et al.*, Evaluation of PERSIANN system satellite–based estimates of tropical rainfall, *Bulletin of the American Meteorological Society* **81**(2000), pp. 2035-2046.

Speirs, P., Gabella, M., Berne, A., A Comparison between the GPM Dual-Frequency Precipitation Radar and Ground-Based Radar Precipitation Rate Estimates in the Swiss Alps and Plateau, *Journal of Hydrometeorology* **18**(2017), pp. 1247-1269.

Sun, Q. *et al.*, A Review of Global Precipitation Data Sets: Data Sources, Estimation, and Intercomparisons, *Reviews of Geophysics*(2018).

Tang, G., Ma, Y., Long, D., Zhong, L., Hong, Y., Evaluation of GPM Day-1 IMERG and TMPA Version-7 legacy products over Mainland China at multiple spatiotemporal scales, *Journal of Hydrology* **533**(2016a), pp. 152-167.

Tang, G. *et al.*, Statistical and hydrological comparisons between TRMM and GPM level-3 products over a midlatitude basin: Is day-1 IMERG a good successor for TMPA 3B42V7?, *Journal of Hydrometeorology* **17**(2016b), pp. 121-137.

Taylor, K.E., Summarizing multiple aspects of model performance in a single diagram, *Journal of Geophysical Research: Atmospheres* **106**(2001), pp. 7183-7192.

Viola, F., Caracciolo, D., Forestieri, A., Pumo, D., Noto, L., Annual runoff assessment in arid and semiarid Mediterranean watersheds under the Budyko's framework, *Hydrological Processes* **31**(2017), pp. 1876-1888.

Wilks, D.S., *Statistical methods in the atmospheric sciences*. Academic press (2011).

Xu, R. *et al.*, Ground validation of GPM IMERG and TRMM 3B42V7 rainfall products over southern Tibetan Plateau based on a high-density rain gauge network, *Journal of Geophysical Research: Atmospheres* **122**(2017), pp. 910-924.

Xue, X. *et al.*, Statistical and hydrological evaluation of TRMM-based Multi-satellite Precipitation Analysis over the Wangchu Basin of Bhutan: Are the latest satellite precipitation products 3B42V7 ready for use in ungauged basins?, *Journal of Hydrology* **499**(2013), pp. 91-99.

Yong, B. *et al.*, Assessment of evolving TRMM-based multisatellite real-time precipitation estimation methods and their impacts on hydrologic prediction in a high latitude basin, *Journal of Geophysical Research: Atmospheres* **117**(2012).

686 **Tables**

Indices	Hourly						Daily					
	Sardinia			Sicily			Sardinia			Sicily		
	min	mean	max	min	mean	max	min	mean	max	min	mean	max
CC	0.17	0.32	0.54	0.02	0.26	0.44	0.45	0.67	0.81	0.11	0.64	0.85
S-RMSE	6.46	11.80	16.88	7.29	12.07	38.83	2.09	3.44	4.91	2.15	3.39	9.30
S-MBE	-0.70	-0.17	0.34	-1.21	-0.12	0.48	-0.70	-0.17	0.34	-1.21	-0.12	0.48
POD α=5%	0.83	0.97	1	0.73	0.96	1	0.86	0.96	1	0.85	0.95	1
POD α=50%	0.52	0.69	0.87	0.30	0.65	0.91	0.46	0.67	0.89	0.45	0.70	0.95
VPOD α=5%	0.87	0.98	1	0.80	0.98	1	0.94	0.98	1	0.89	0.98	1
VPOD α=50%	0.57	0.79	0.95	0.39	0.75	0.96	0.60	0.80	0.96	0.60	0.83	0.98
FAR α=5%	0	0.03	0.10	0	0.03	0.12	0	0.02	0.06	0	0.02	0.08
FAR α=50%	0.17	0.35	0.51	0.19	0.36	0.59	0.11	0.23	0.37	0.11	0.25	0.45
VFAR α=5%	0	0.02	0.10	0	0.02	0.15	0	0.01	0.05	0	0.01	0.08
VFAR α=50%	0.11	0.30	0.51	0.13	0.75	0.67	0.05	0.15	0.30	0.05	0.16	0.44
MISS α=5%	0	0.03	0.17	0	0.04	0.27	0	0.04	0.14	0	0.05	0.15
MISS α=50%	0.12	0.31	0.48	0.09	0.35	0.67	0.11	0.33	0.54	0.05	0.30	0.55
VMISS α=5%	0	0.02	0.13	0	0.03	0.20	0	0.02	0.06	0	0.02	0.11
VMISS α=50%	0.05	0.21	0.43	0.034	0.25	0.61	0.04	0.20	0.40	0.02	0.17	0.41
CSI α=5%	0.78	0.94	1	0.72	0.94	1	0.84	0.94	1	0.83	0.93	1
CSI α=50%	0.37	0.51	0.68	0.26	0.48	0.72	0.42	0.56	0.73	0.39	0.56	0.73
VCSI α=5%	0.85	0.96	1	0.80	0.95	1	0.91	0.97	1	0.88	0.97	1
VCSI α=50%	0.42	0.59	0.77	0.24	0.57	0.77	0.55	0.70	0.86	0.47	0.71	0.85

687

688 **Table 1.** Spatial minimum (min), mean and maximum (max) values of each index at hourly and daily
689 time scale for both islands.

690

Moran's Index	Hourly				Daily			
	Sardinia		Sicily		Sardinia		Sicily	
α	5%	50%	5%	50%	5%	50%	5%	50%
POD	0.090	0.354	0.091	0.347	***0.046***	0.677	0.254	0.552
VPOD	0.102	0.402	0.168	0.389	0.116	0.605	0.244	0.582
FAR	***0.052***	0.432	0.115	0.297	0.173	0.418	0.265	0.351
VFAR	0.101	0.489	0.102	0.342	0.121	0.561	0.182	0.300
MISS	0.090	0.354	0.092	0.347	***0.046***	0.677	0.254	0.552
VMISS	0.102	0.402	0.168	0.389	0.116	0.605	0.244	0.582
CSI	***0.071***	0.365	0.150	0.340	0.104	0.532	0.134	0.489
VCSI	0.093	0.490	0.149	0.409	0.150	0.475	0.171	0.506

Table 2. Values of *Moran's I* for categorical and volumetric indices at hourly and daily time scales for Sardinia and Sicily. α is the percentile of the threshold value, t_α, above which the precipitation occurrences are considered. Number format indicates whether the spatial distribution of the feature is to be considered random (italic bold), or clustered (serif).

Indices	Sardinia		Sicily	
	Coastal	**Internal**	**Coastal**	**Internal**
CC	0.29	0.35	0.22	0.27
S-RMSE	12.60	10.83	13.46	11.77
S-MBE	-0.20	-0.13	-0.17	-0.11

Table 3. *GPM* performance in coastal and internal pixels. The values of mean *CC*, *S-RMSE* and *S-MBE* refer to hourly time scale.

Figures

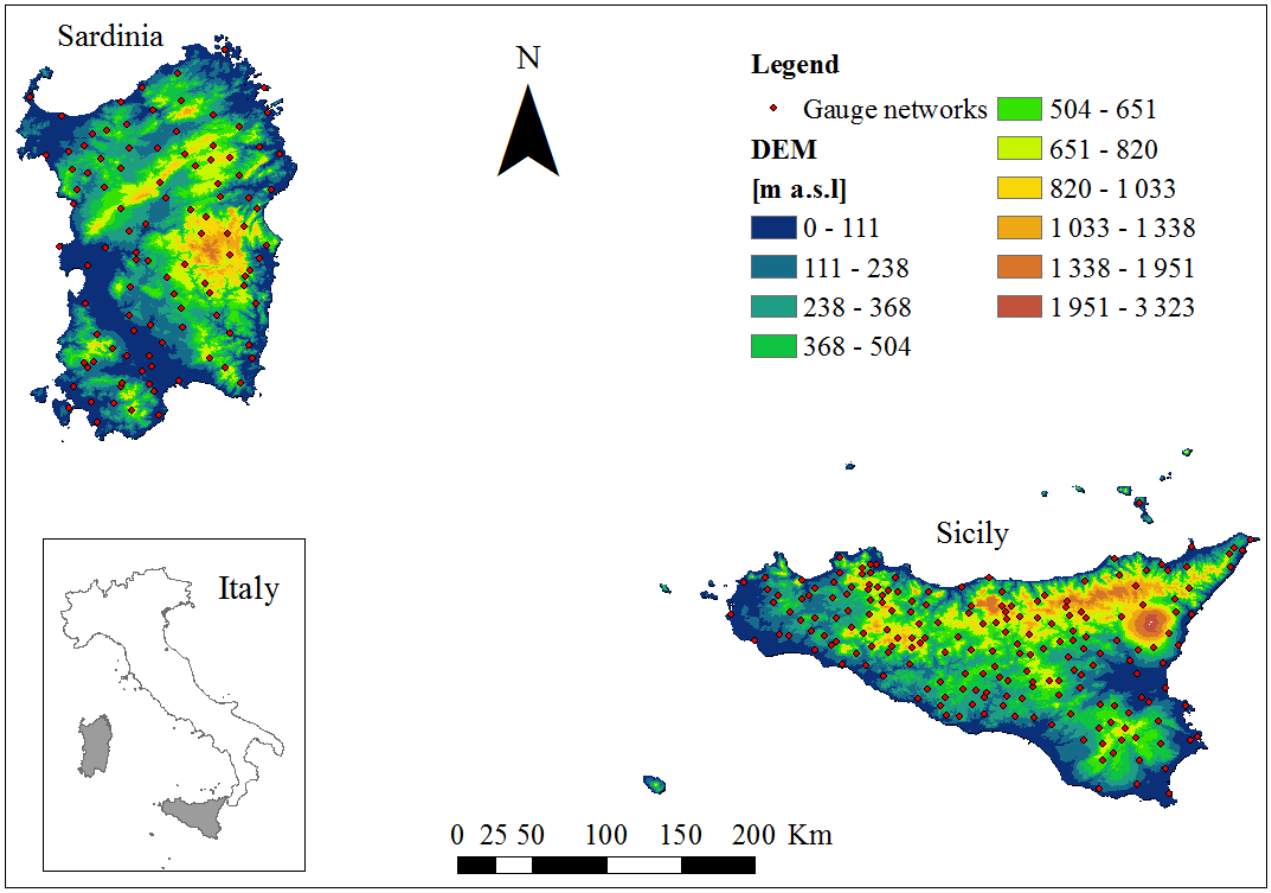
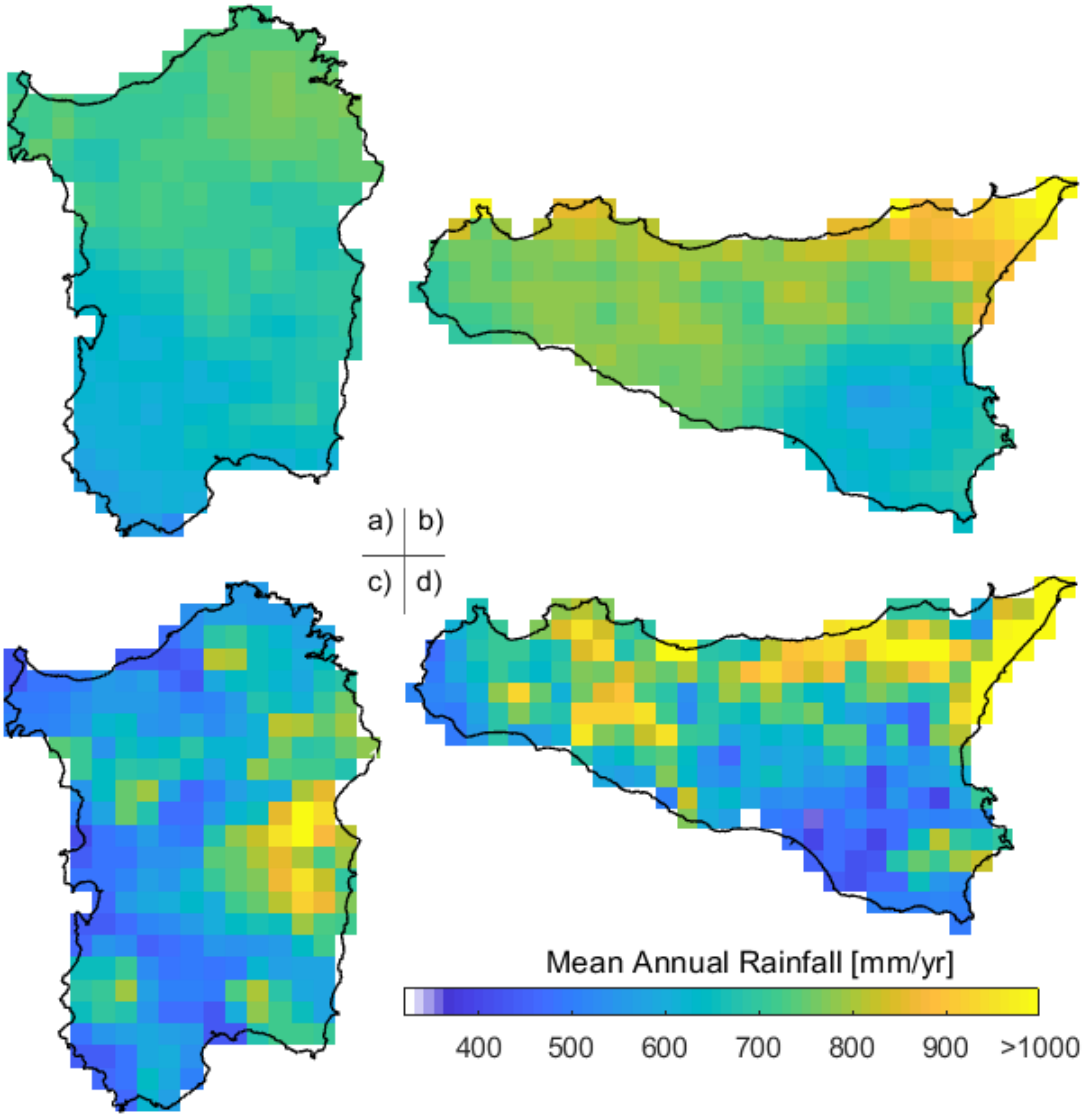


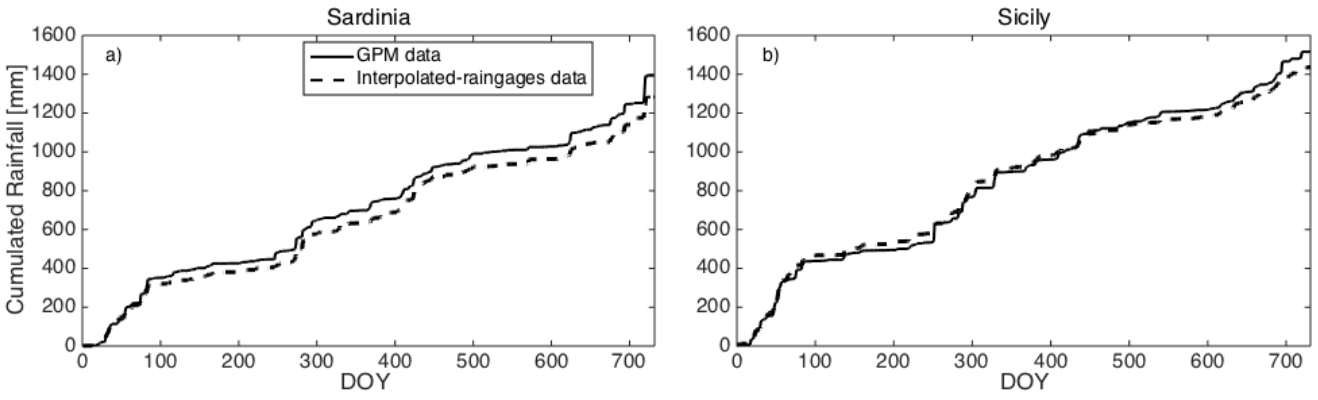
Figure 1. Digital elevation model of the study areas (i.e., Sicily and Sardinia islands) and raingauge locations (red dots).



706

707 **Figure 2.** Maps of the mean annual rainfall (2015 and 2016 period) measured with *GPM* (a,b) or
708 obtained spatially interpolating the raingauge data (c,d) for Sardinia (a,c) and Sicily (b,d).

709

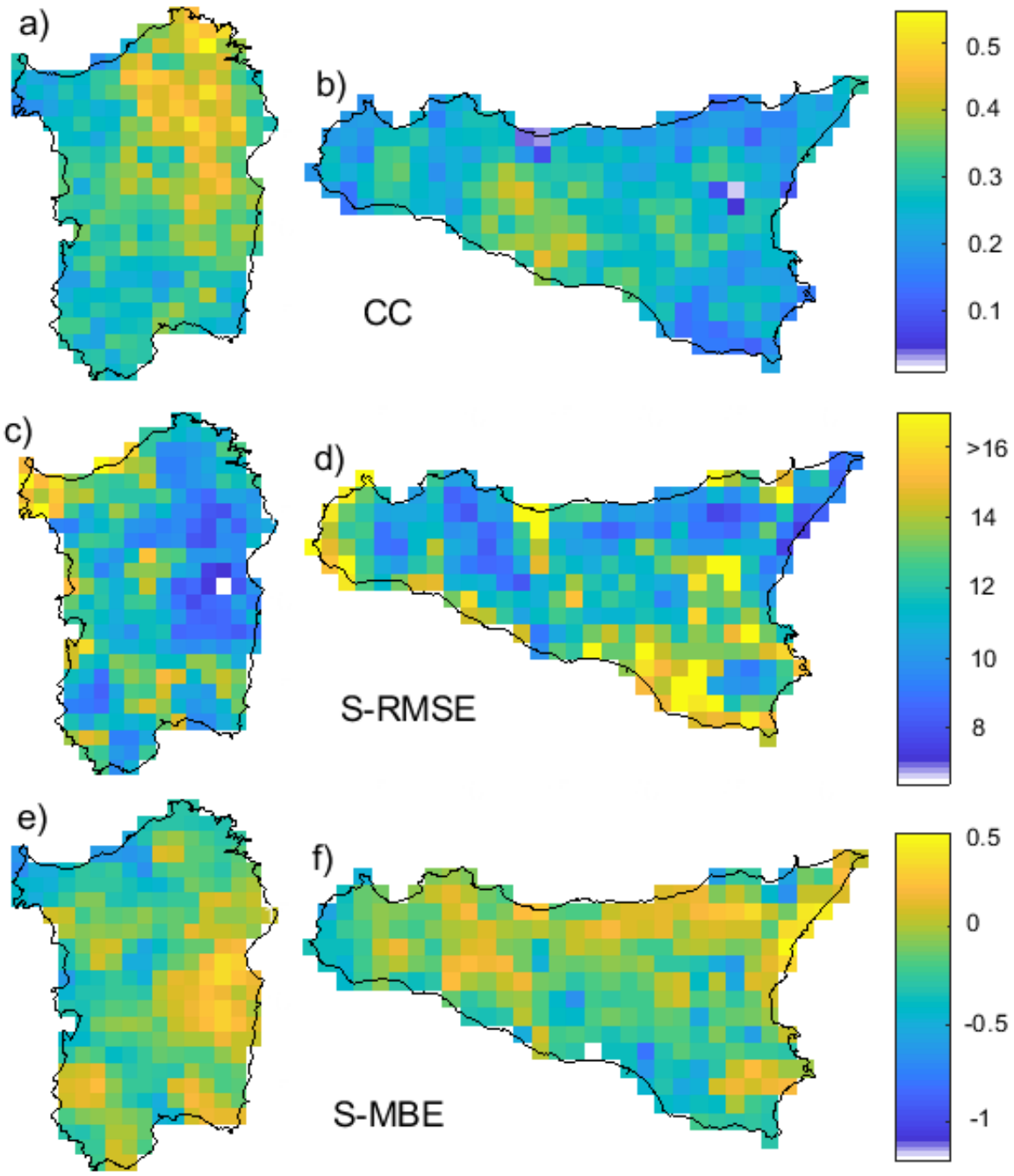


710

711 **Figure 3.** Comparison of cumulative depths of *MAP* from *GPM* (continuous line) and raingauges
712 (dashed line) over the whole of Sardinia (a) and Sicily (b) for the two-year period 2015-2016.

713

714



715

716 **Figure 4.** Comparison of hourly *GPM* satellite precipitation and interpolated-raingauge data for
717 Sardinia (a,c,e) and Sicily (b,d,f) by continuous indices: *CC*, correlation coefficient (a,b); *S-RMSE*,
718 standardized root mean square error (c,d); *S-MBE*, standardized mean bias error (e,f).

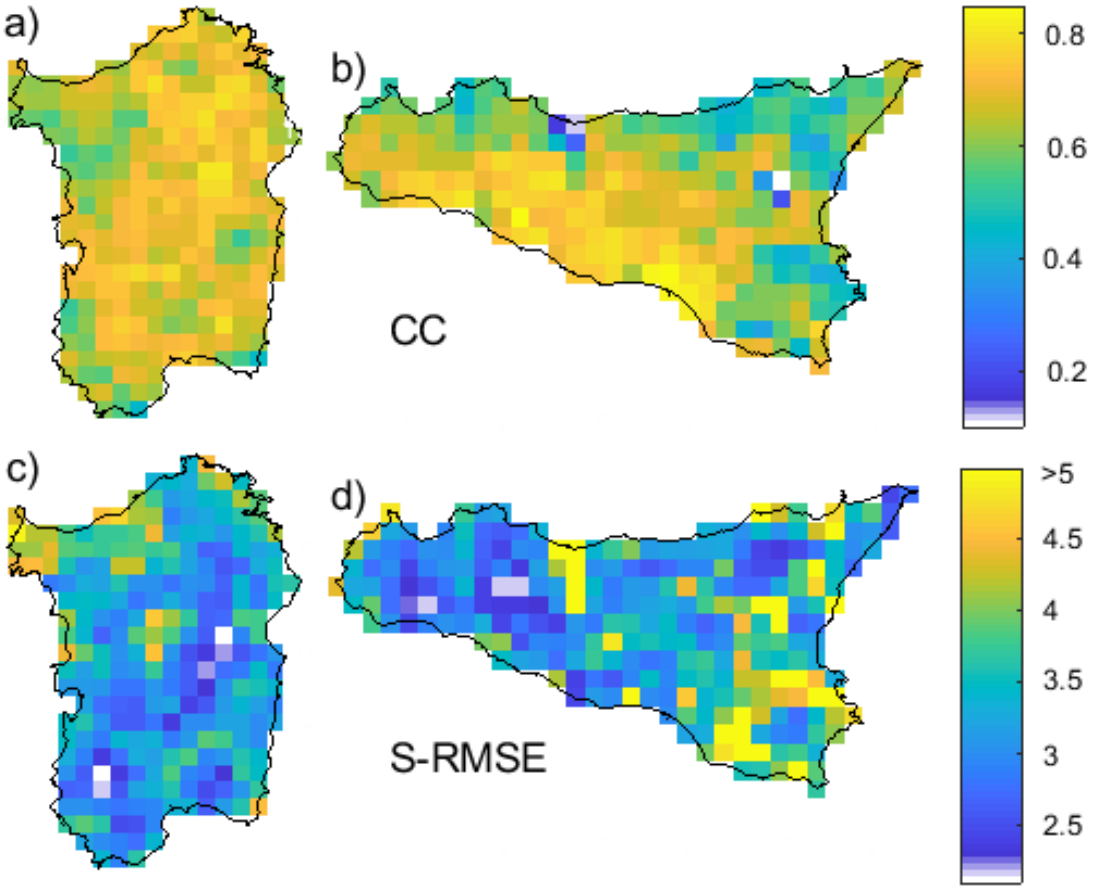


Figure 5. Same as Figure 4, but for daily time scale. *S-MBE* is identical to the hourly case because of its mathematical definition and thus it is not represented.

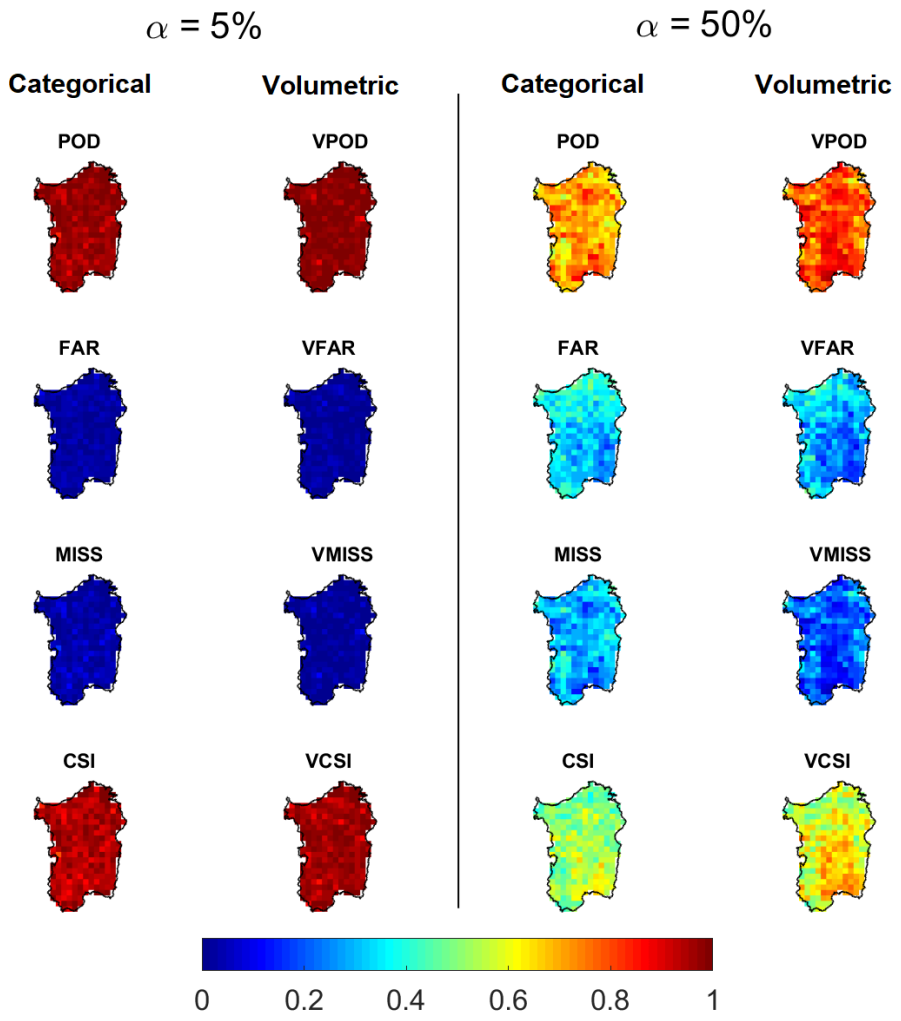


Figure 6. Comparison of hourly *GPM* satellite precipitation and interpolated-raingauge data for Sardinia by categorical and volumetric indices: *POD*, *VPOD*, *FAR*, *VFAR*, *MISS*, *VMISS*, *CSI*, and *VCSI* (see text for acronym definitions). Results using thresholds t_α with α equal to 5% and 50% are reported in the left and right plots, respectively.

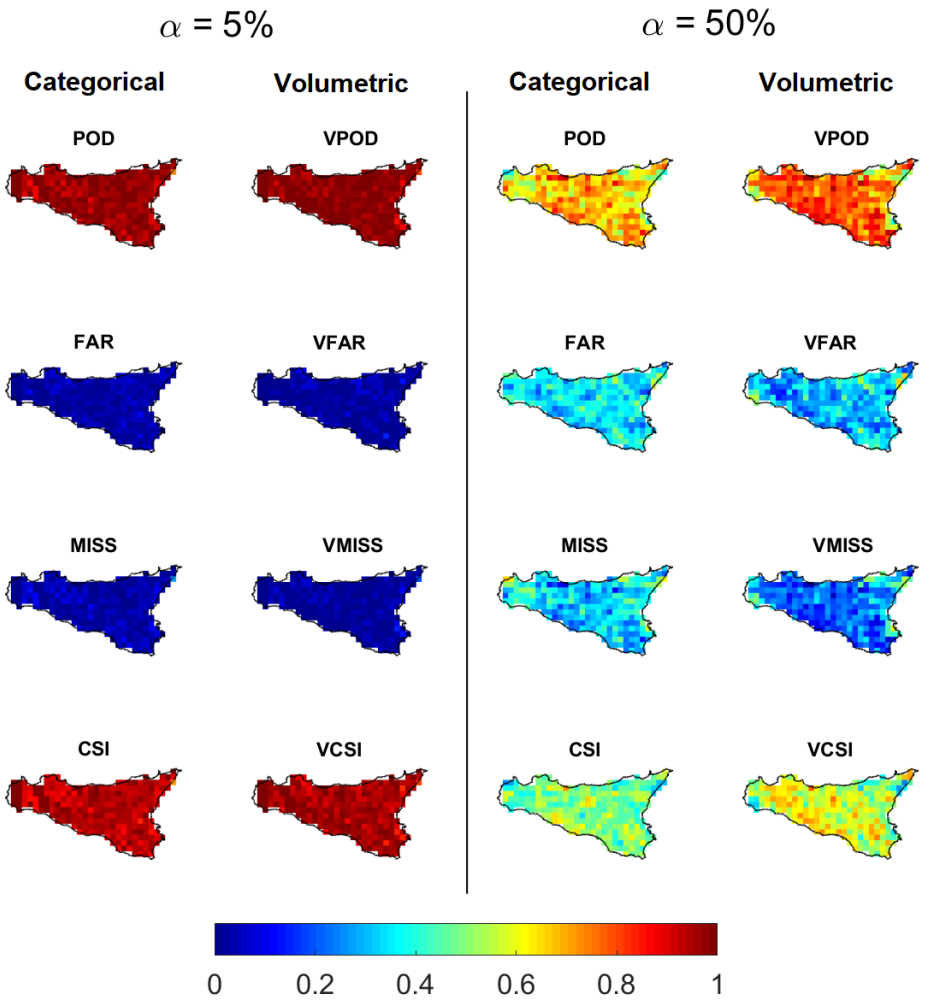


Figure 7. Same as Figure 8, but for hourly precipitation in Sicily.

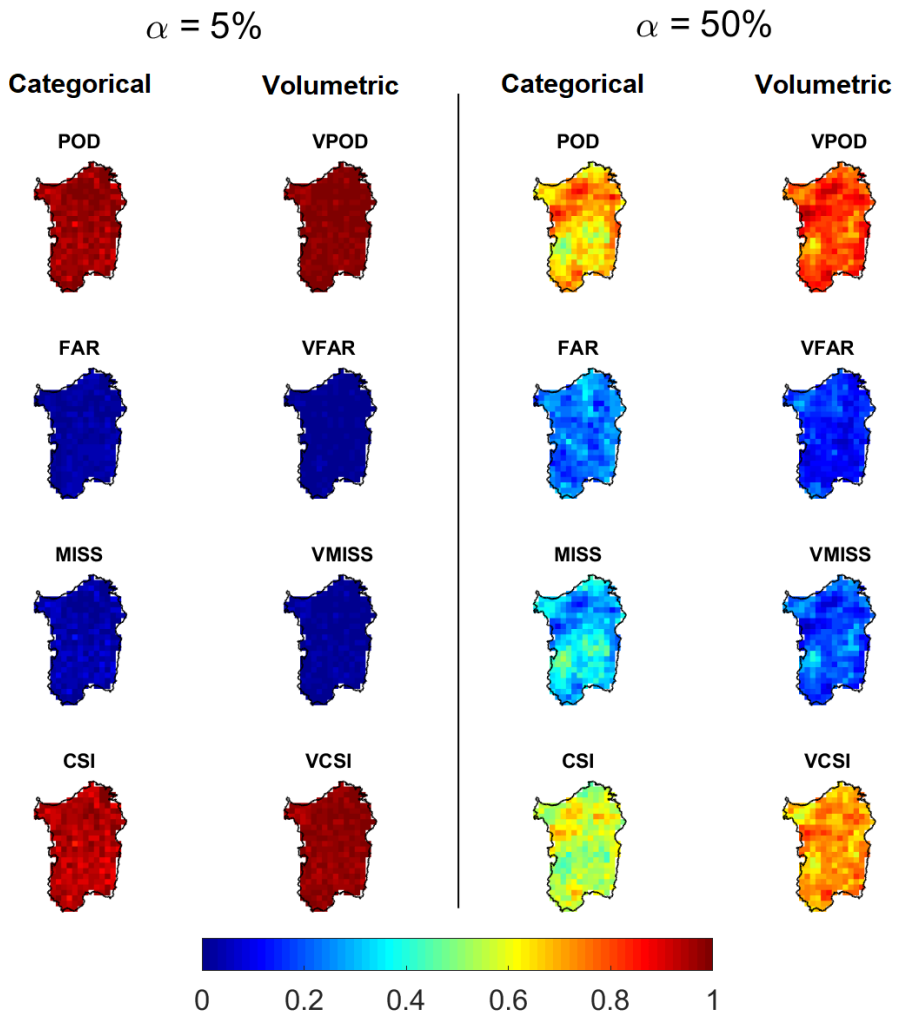


Figure 8. Same as Figure 8, but for daily precipitation in Sardinia.

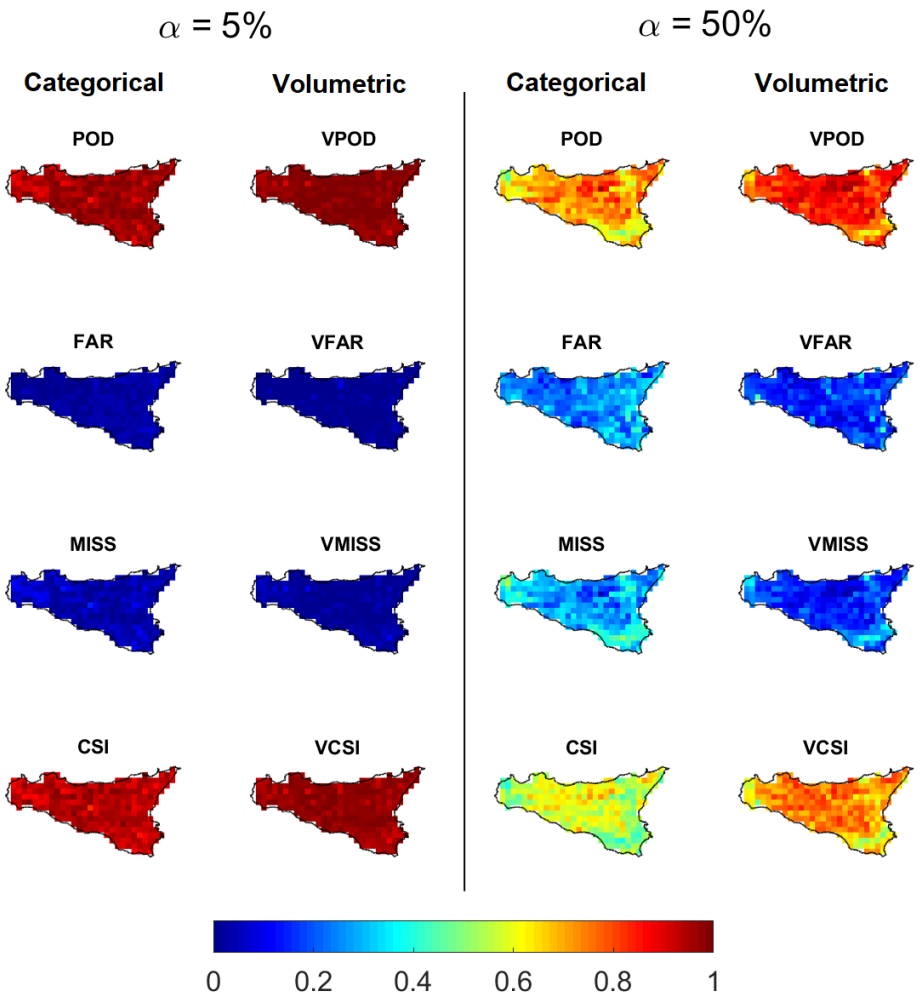
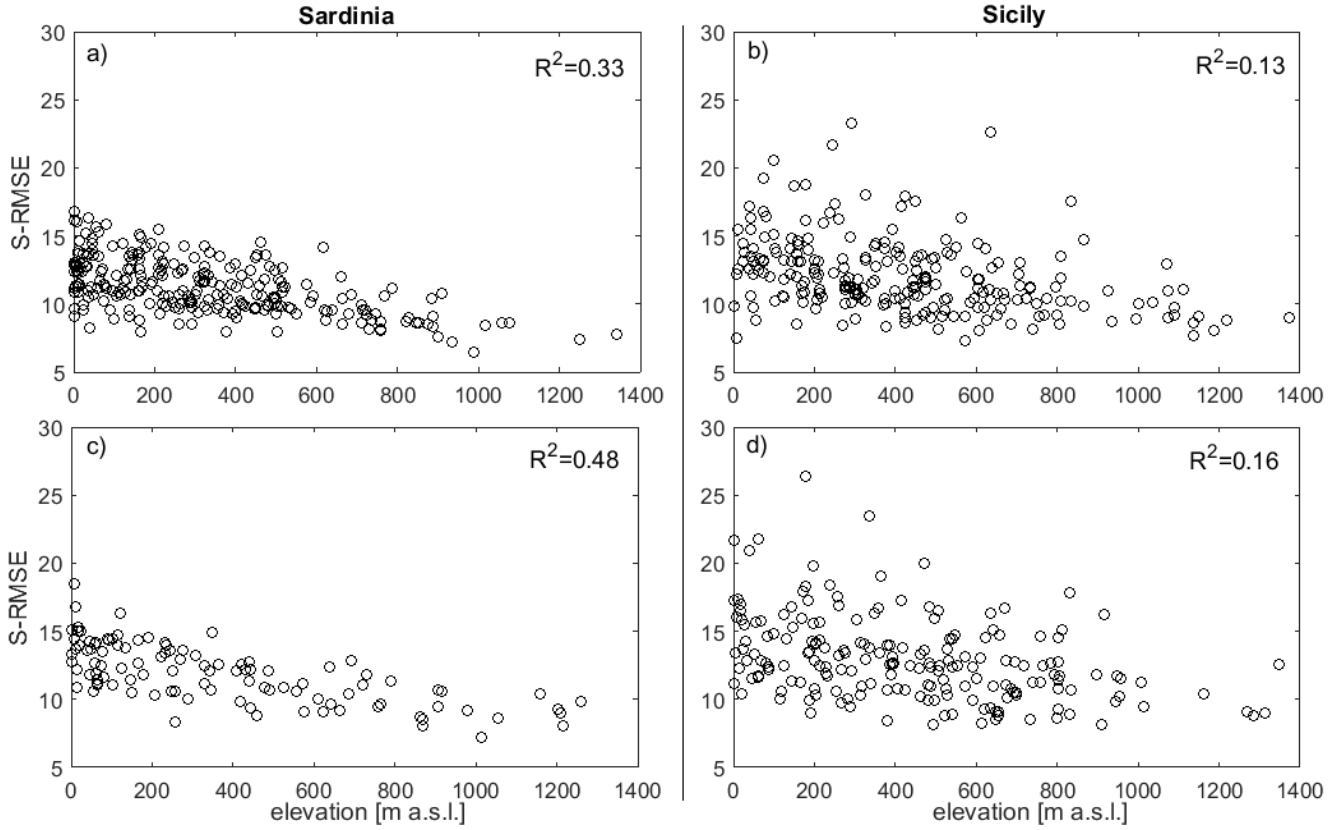
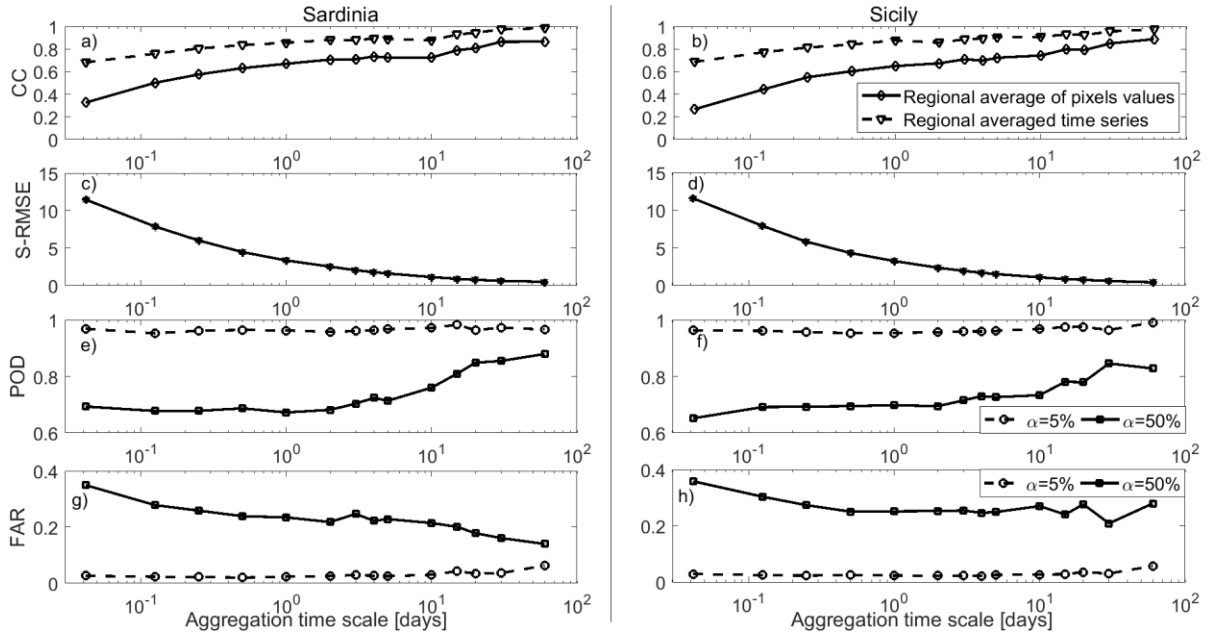


Figure 9. Same as Figure 8, but for daily precipitation in Sicily.



738

739 **Figure 10.** Relation between elevation and S-RMSE at hourly time scale for each pixel (a,b) and for
740 each raingauge (c,d), for Sardinia (a,c) and Sicily (b,d).



741

742 **Figure 11.** *GPM* product performances for different aggregation time scales (i.e., 1, 3, 6, 12 hours, 1, 2,
743 3, 4, 5, 10, 15, 20, 30, 60 days) in Sardinia (a,c,e,g) and Sicily (b,d,f,h): *CC*, correlation coefficient
744 (a,b); *S-RMSE*, standardized root mean square error (c,d); *POD*, probability of detection (e,f); *FAR*,
745 false alarm ratio (g,h). Last two indices are computed for thresholds t_α with α equal to 5% and 50%. In
746 subplots a) and b) the black solid line represents the regional mean of the correlation coefficient
747 obtained at pixel scale, while the black dashed line represents the correlation coefficient between
748 regional aggregated time series.

749

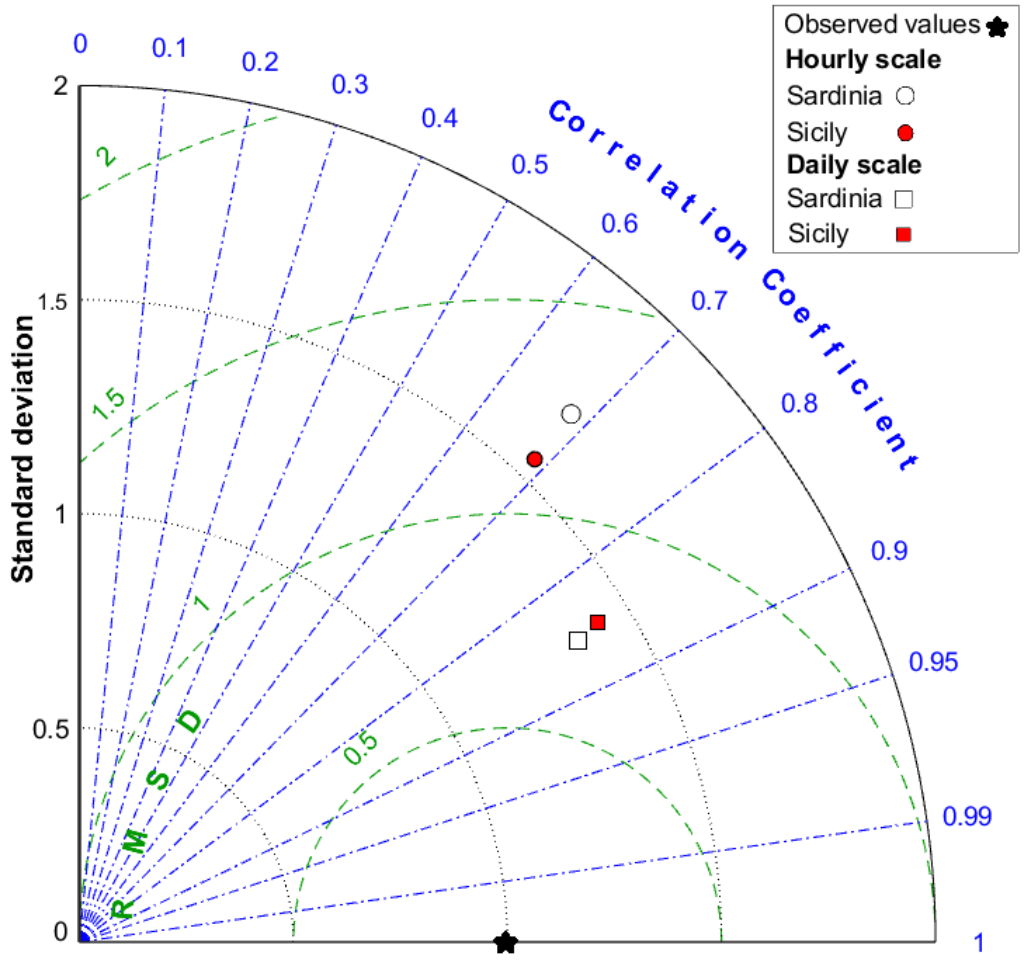


Figure 12. Normalized Taylor diagram obtained from *MAP* time series at hourly (circles) and daily (squares) scales in Sicily and Sardinia. The black star refers to the observed value where the normalized standard deviation is equal to 1, the radial distance from the black star quantifies the centered *RSMD* normalized by the standard deviation of observed *MAP*, the azimuth and the radial distance from the origin quantify instead *CC* and normalized standard deviation, respectively.