SHORT TERM MOBILITY 2011
SCIENTIFIC REPORT

"Comparison of TRMM satellite rainfall estimates with rain gauge data and landslide empirical rainfall thresholds under different morphological and climatological conditions in Italy"

Mauro Rossi
Istituto di Ricerca per la Protezione Idrogeologica, Consiglio Nazionale delle Ricerche
Tel.: +39 075.5014420
Fax: +39 075.5014420
mauro.rossi@irpi.cnr.it

Hosting Institution:
NASA-GSFC (Goddard Space Flight Center) – Hydrological Sciences Laboratory, Greenbelt, Maryland 20771 (USA)

Contact reference:
Dalia Kirschbaum
dalia.b.kirschbaum@nasa.gov

Abstract

Quantitative information on rainfall is necessary to predict the possible occurrence of rainfall-induced landslides. Landslide early warning systems attempt to predict rainfall-induced landslides through the comparison of quantitative rainfall information with empirical rainfall thresholds for the possible occurrence of landslides. Most of the systems exploit rainfall measurements obtained from networks of gauges, and only a few systems use satellite rainfall estimates. All the systems exploit empirical rainfall thresholds defined using rainfall measurements obtained from networks of rain gauges. Despite the availability of quantitative satellite rainfall estimates, and their experimental use in existing warning systems, surprisingly limited research has been done to compare satellite rainfall estimates and rain gauge rainfall measurements, for the forecasting of possible landslide occurrence. In this work, we analyse the relationships between rainfall measurements obtained from a network of > 1950 rain gauges in Italy and rainfall satellite estimates for the same area obtained by the NASA Tropical Rainfall Measuring Mission (TRMM-RT and TRMM-v6), for the period 2009-2010. Coupling point rain gauge measurements and TRMM rainfall estimates at individual grid cells (0.25 latitude × 0.25 longitude), we evaluate the correlation between the gauge rainfall measurements and the satellite rainfall estimates in different morphological and climatological conditions, using linear and power-law fitting models. We use cumulative rainfall measurements/estimates for different periods, from 24 to 72 hours. We analyse and compare the distributions of the ground-based rainfall measurements and the satellite rainfall estimates using standard non-parametric and parametric statistical methods. In the Umbria region we determine rainfall events starting from rain gauge and satellite raw rainfall data series, and we compare their relative distribution to that characterizing rainfall events associated to landslide phenomena. We observe significant differences in the distributions of cumulative rainfall data, for different morphological and climatological areas in Italy. Differences are larger in mountainous areas, and collectively reveal a complex relationship between the ground-based rainfall measurements and the satellite rainfall estimates. Power law correlation models have the best fitting performance, at the expenses of large prediction intervals, particularly for large values of cumulated rainfall. An exponential distribution provides a better fit for satellite rainfall estimates, compared to rain gauge measurements. Distributions of rainfall events obtained from rain gauge and satellite raw rainfall data series are significantly different. TRMM data result generally underestimated compared to rain gauge measurements. Both rainfall events distributions result significantly lower than the distribution of rainfall events associated to landslide phenomena. These preliminary results indicate that satellite rainfall estimates are not straightforward comparable with rain gauge rainfall measures, requiring a detailed local investigation of degree of correlation. Further specific empirical rainfall thresholds have to be defined to fully exploit satellite rainfall estimates in existing early warning system.

Key words: Landslide, TRMM, Rain gauges, rainfall threshold

Index

1. Introduction

Rainfall-induced landslides are widespread phenomena that cause every year casualties and extensive damages. In Italy, rainfall is the primary trigger of landslides causing every year damaging failures. In the 60-year period 1950-2009 rainfall-triggered landslides caused more than 6300 casualties (**Salvati et al., 2010**). Predicting rainfall-induced landslide is a problem of scientific and societal interest.

Investigators have long attempted to determine the amount of precipitation necessary to trigger slope failures. Rainfall-induced landslides are triggered by the increase of water pressure into the soil. Groundwater conditions leading to a slope failure depend from rainfall infiltration, and hence from soil properties, soil moisture conditions, and rainfall quantity (**Wieczorek, 1996**). To analyze the triggering of rainfall-induced landslides two modeling approaches are possible: (i) empirical, mainly consisting in the determination of rainfall thresholds (**Caine, 1980; Innes, 1983; Guzzetti et al., 2007; 2008**), and (ii) deterministic, consisting in simplified physical-based models coupling an infiltration model to an instability model (**Wu & Sidle, 1995; Montgomery & Dietrich, 1994; Iverson, 2000**). While the first is largely used to analyze the triggering of rainfall-induced landslides at different scales from local to global, the second is mainly applied to single instability phenomena or to single basins. In both cases spatial and temporal quantitative information on rainfall is necessary to analyze the triggering mechanisms of rainfall-induced slope failures. At present these approaches only exploit rainfall measurements from rain gauge networks, even if satellite or radar quantitative rainfall estimates are available for long periods and large areas.

The empirical modeling approach have been integrated in early-warning systems to predict the possible occurrence of rainfall-induced landslides over space and time. Commonly the existing landslide early warning systems attempt to predict rainfall-induced landslides through the comparison of quantitative rainfall information with ID (Intensity/Duration) empirical rainfall thresholds for the possible occurrence of landslides. Most of these systems exploit rainfall measurements obtained from networks of gauges (**Rossi et al., 2012**), but some systems use radar or satellite rainfall estimates (**Kirschbaum et al., 2011**).

Despite the availability of quantitative satellite rainfall estimates, and their experimental use in existing warning systems, surprisingly limited research has been done to compare satellite rainfall estimates and rain gauge rainfall measurements for the analysis of the triggering mechanism and for the forecast of the rainfall-induced landslides.

This work analyses the relationships between rainfall measurements obtained from a network of > 1950 rain gauges in Italy (available through the Experience Platform of the Italian Civil Protection Department) and rainfall satellite estimates for the same area obtained by the NASA Tropical Rainfall Measuring Mission for the period 2009-2010 (TRMM v6 Algorithm 3B42, http://trmm.gsfc.nasa.gov/3b42.html). The correlation between rainfall measurements and satellite rainfall estimate and the distribution of both type of rainfall data were analysed in different morphological and climatological conditions of the Italian territory.

The correlation between rain gauge rainfall measurements and satellite rainfall estimates in different morphological and climatological conditions were evaluated considering cumulative rainfall measurements/estimates for different periods, from 3 to 72 hours. The distributions of the ground-based rainfall measurements and the satellite rainfall estimates were analysed and

compared, using standard non-parametric and parametric statistical methods. Finally rainfall events have been determined from rain gauge and satellite data exploiting an automatic procedure. Distribution of rainfall events obtained from both type of rainfall data were estimated and compared to the distribution of rainfall events associated to slope instability phenomena.

2. Data

The analysis of the relationships between rainfall ground measurements and rainfall satellite estimates was realized along the entire Italian territory comparing: (i) rainfall data from the Italian rain gauge network, and (ii) rainfall estimates obtained by the NASA Tropical Rainfall Measuring Mission. We compare rainfall data of one hydrological year from 01/09/2009 to 31/08/2010. The analyses was focused to highlight the differences between the two rainfall data type in different morphological and climatological conditions corresponding to different morphological subdivisions of the Italian territory.

2.1 Rain gauge measurements

Rainfall measurements (*R*) were obtained from the rain gauge network available through the Experience Platform of the Italian Civil Protection Department. **Figure 1** shows the location of 1950 rain gauges in the Italian territory. Rainfall data are available from 2003 to now and are updated every 6 hours.

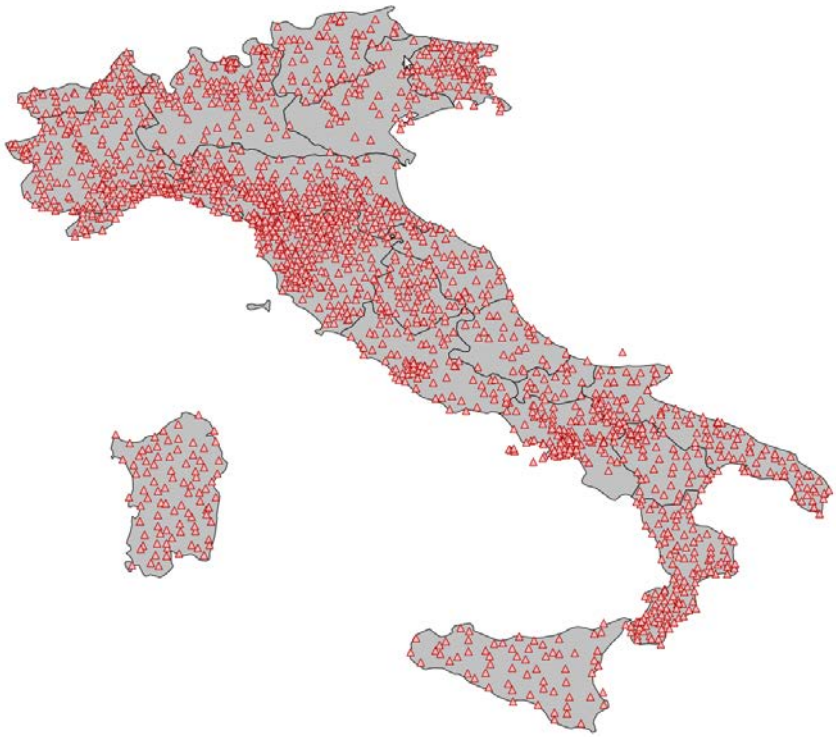

Figure 1. Location of the rain gauges along the Italian territory (available through the Experience Platform of the Italian Civil Protection Department).

Rain gauge sensors in the database are heterogeneous and rainfall data can have different temporal resolutions, mostly semi-hourly. Sensor accuracy varies from 0.2 to 1 mm. Gaps, errors and discontinuous time series can be present in the database. A specific procedure was used to evaluate the quality of rainfall data in the analysed period. The procedure estimates the reliability of rain gauge sensors comparing the relative Cumulative Annual Precipitation (*CAP*) with Mean Annual Precipitation (*MAP*) values calculated in the period 1961-1990 exploiting data available at global scale from the Global Historical Climatology Network-Monthly (GHCN-M) archive of the National Climatic Data Center (NCDC) (http://www.ncdc.noaa.gov/ghcnm/) realized by the National Oceanic and Atmospheric Administration (NOAA). Criteria to estimate rain gauge reliability and results obtained for the period from 01/09/2009 to 31/08/2010 are summarized in **Table 1**.

Table 1. Criteria for the estimation of rain gauge reliability. *CAP*: Cumulative Annual precipitation; *MAP*: Mean Annual Precipitation.

CONDITION	QUALITY INDEX	MALFUNCTIONING PROBLEM	RAIN GAUGES # (%)
No rainfall values in the period ($CAP = NULL$)	1	Rain gauge not acquired	287 (15%)
$CAP = 0$	2	Malfunctioning rain gauge	2 (0%)
$CAP < MAP - \frac{MAP}{2}$	3	Rain gauge with discontinuous values	148 (8%)
$MAP - \left(\frac{MAP}{2}\right) < CAP < MAP + (MAP \cdot 2)$	4	Rain gauge properly working	1488 (76%)
$CAP > MAP + (MAP \cdot 2)$	5	Rain gauge with excessive values	25 (1%)

2.2 TRMM satellite rainfall estimates

Satellite rainfall data (*T*) were estimated by NASA exploiting the TRMM (Tropical Rainfall Measuring Mission) data by mean of the TRMM v6 Algorithm 3B42. The purpose of the algorithm is to produce TRMM merged high quality (HQ)/infrared (IR) precipitation and root-mean-square (RMS) precipitation-error estimates. The 3B-42 estimates are produced in four stages: (i) the microwave estimates precipitation are calibrated and combined; (ii) infrared precipitation estimates are created using the calibrated microwave precipitation; (iii) the microwave and IR estimates are combined; and (iv) rescaling to monthly data is applied. Each precipitation field is best interpreted as the precipitation rate effective at the nominal observation time. These gridded estimates are on a 3-hour temporal resolution and a 0.25 × 0.25-degree spatial resolution in a global belt extending from 50 degrees South to 50 degrees North latitude. Data available for the Italian territory in the period from 01/0/1998 to 31/08/2010 were extracted. TRMM data were stored in a PostgreSQL/PostGIS database structure (www.postgresql.org) to simplify the data management. Rainfall estimates were associated spatially to the relative cell

centroid location. The location of the 638 cell centroids covering the Italian territory is shown in **Figure 2**.

Figure 2. Location of TRMM v6 3B42 cell centroids covering the Italian territory.

2.3 Morphological subdivision of Italy

The analyses of the distributions and correlations of rain gauge measures and satellite rainfall estimates were performed in different zones, to consider different morphological and climatological conditions of the Italian territory. The Italian morphological subdivision was derived by that realized by **Guzzetti and Reichenbach (1994)** exploiting morphometric data calculated from a 230m DTM obtained mosaicking the entire Italian mean elevation archive. The original subdivision classify the Italian territory in eight major topographic division (provinces) and 30 minor divisions (sections). The subdivisions were identified using a step-wise semi-quantitative approach combining: (i) an unsupervised three-class cluster analysis of altitude derivatives; (ii) an heuristic visual inspection of morphometric maps; (iii) and the interpretation of small-scale geological and structural maps. Provinces and sections reflect different morphological conditions, measured by four morphological variables derived by the altitude map: (i) altitude, (ii) slope curvature, (iii) frequency of slope reversal, (iv) and elevation-relief ratio. Highland, upland and lowland topographic types were identified. In this work we aggregated heuristically the 30 minor subdivisions (sections) in 10 new macro-subdivision, considering in addition latitude and aspect, to reflect different morphological and climatological conditions of the Italian territory. **Table 2** summarize values of the morphological parameters associated to each morpho-climatological subdivision.

Table 2. Criteria for the estimation of rain gauge reliability. *CAP*: Cumulative Annual precipitation; *MAP*: Mean Annual Precipitation.

SUBDIVISION	PARAMETER	Min Value	Max Value	Lowland (%)	Upland (%)	Highland (%)
Tyrr Central Tyrrhenian coast	Elevation (m)	s.l.	1738	55.4	40.0	4.6
	Slope (°)	0	41			
	Elevation relief ratio	0	0.98			
	Slope reversal (1/km^2)	0.14	12.07			
	Curvature (1/m)	-1.82	0.86			
Sici Southern/Western Sicily	Elevation (m)	s.l.	3340	49.3	45.6	5.1
	Slope (°)	0	43			
	Elevation relief ratio	0.01	0.91			
	Slope reversal (1/km^2)	0.14	10.07			
	Curvature (1/m)	-2.57	1.65			
Sard Sardinia	Elevation (m)	s.l.	1786	37.7	46.1	16.2
	Slope (°)	0	48			
	Elevation relief ratio	0.01	0.93			
	Slope reversal (1/km^2)	0.14	9.43			
	Curvature (1/m)	-5.25	2.71			
Popl Po plain and Alpine foothills	Elevation (m)	s.l.	842	92.9	7.1	0.0
	Slope (°)	0	38			
	Elevation relief ratio	0	0.99			
	Slope reversal (1/km^2)	0.14	9.14			
	Curvature (1/m)	-1.39	4.02			
Lang Liguria/Piedmont hills	Elevation (m)	s.l.	1287	31.3	61.9	6.8
	Slope (°)	0	36			
	Elevation relief ratio	0.06	0.88			
	Slope reversal (1/km^2)	0.14	10.79			
	Curvature (1/m)	-0.79	0.71			
ApeU Northern Apennines	Elevation (m)	s.l.	2121	3.4	60.0	36.6
	Slope (°)	0	49			
	Elevation relief ratio	0.01	0.86			
	Slope reversal (1/km^2)	0.14	10.79			
	Curvature (1/m)	-3.84	1.23			
ApeL Southern Apennines	Elevation (m)	s.l.	2267	8.7	55.4	35.9
	Slope (°)	0	48			
	Elevation relief ratio	0.03	0.98			
	Slope reversal (1/km^2)	0.14	9.64			
	Curvature (1/m)	-1.94	1.33			

ApeC Central Apennines	Elevation (m)	27	2914	1.5	63.0	35.5
	Slope (°)	0	57			
	Elevation relief ratio	0.02	0.86			
	Slope reversal (1/km^2)	0.14	10.86			
	Curvature (1/m)	-5.59	3.46			
Alps Northern alpine area	Elevation (m)	s.l.	4810	13.8	32.0	54.1
	Slope (°)	0	72			
	Elevation relief ratio	0.02	0.96			
	Slope reversal (1/km^2)	0.14	10.28			
	Curvature (1/m)	-6.65	5.66			
Adri Central Southern Adriatic coast	Elevation (m)	s.l.	1485	44.0	55.7	0.3
	Slope (°)	0	35			
	Elevation relief ratio	0	0.97			
	Slope reversal (1/km^2)	0.14	11.43			
	Curvature (1/m)	-0.64	0.68			

The map of the morpho-climatological subdivision partitioning the Italian territory and the elevation map used to derived the subdivision are shown in **Figure 3**.

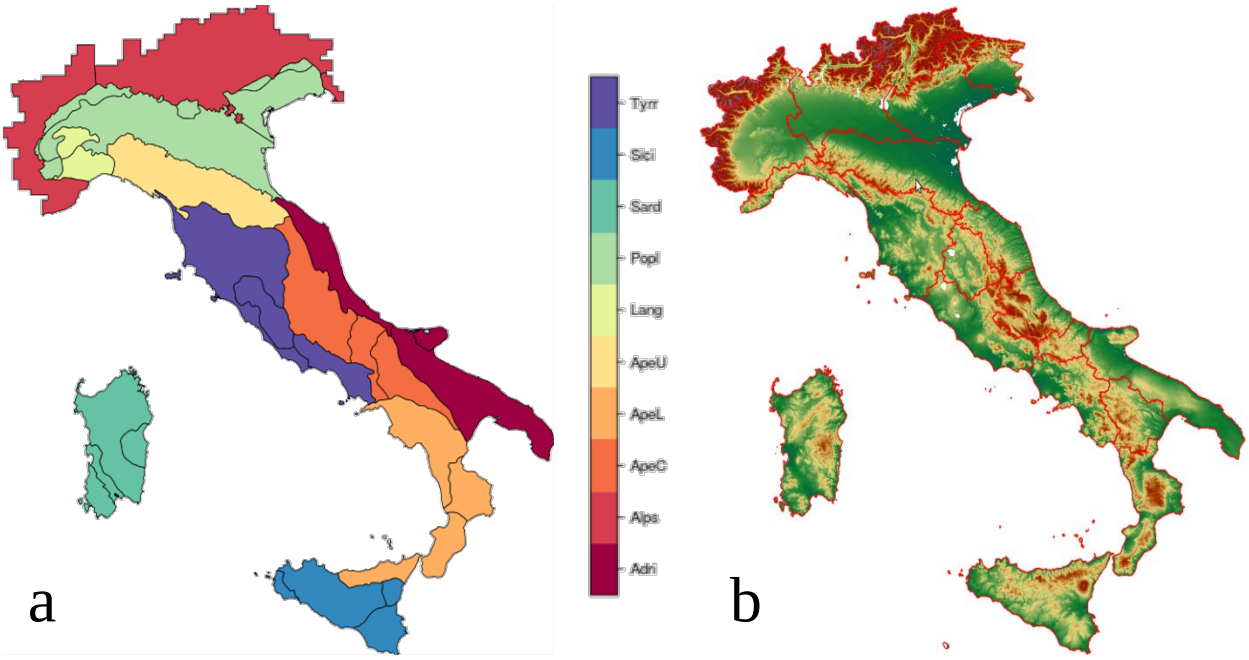

Figure 3. Morphological subdivision to the Italian territory (a) derived by Guzzetti & Reichenbach, 1994 from a 230m DTM (b) obtained mosaicking the entire Italian mean elevation archive.

3. Analysis of cumulative rainfall data

The analysis of cumulative rainfall data was performed to estimate the correlation existing between rain gauge measures and satellite rainfall estimates and to characterize their relative distributions. Preliminarily a minimum distance criteria was used to couple rain gauge stations and satellite cell centroids. A specific GRASS GIS (a free Geographic Information System software used for geospatial data management and analysis, image processing, graphics/maps production, spatial modelling, and visualization, http://grass.fbk.eu/, **Neteler and Mitasova 2008**) procedure allowed to associate to each rain gauge the closest satellite centroid. The procedure associated to each centroid one or more rain gauges with a one-to-many relationship. **Figure 4** show an example of the result of the GIS procedure. Starting from raw data series, the cumulative rainfall for centroids and for rain gauges was estimated for different rainfall duration corresponding to 3, 12, 24, 48 and 72 hours. The procedure realized in R (a free software environment for statistical computing and graphics, **R Development Core Team, 2011**) excluded incomparable rain gauge/centroid couples due to: (i) missing data, gaps and errors in rainfall series, (ii) incomparable temporal resolutions, and (iii) partial overlapping problems.

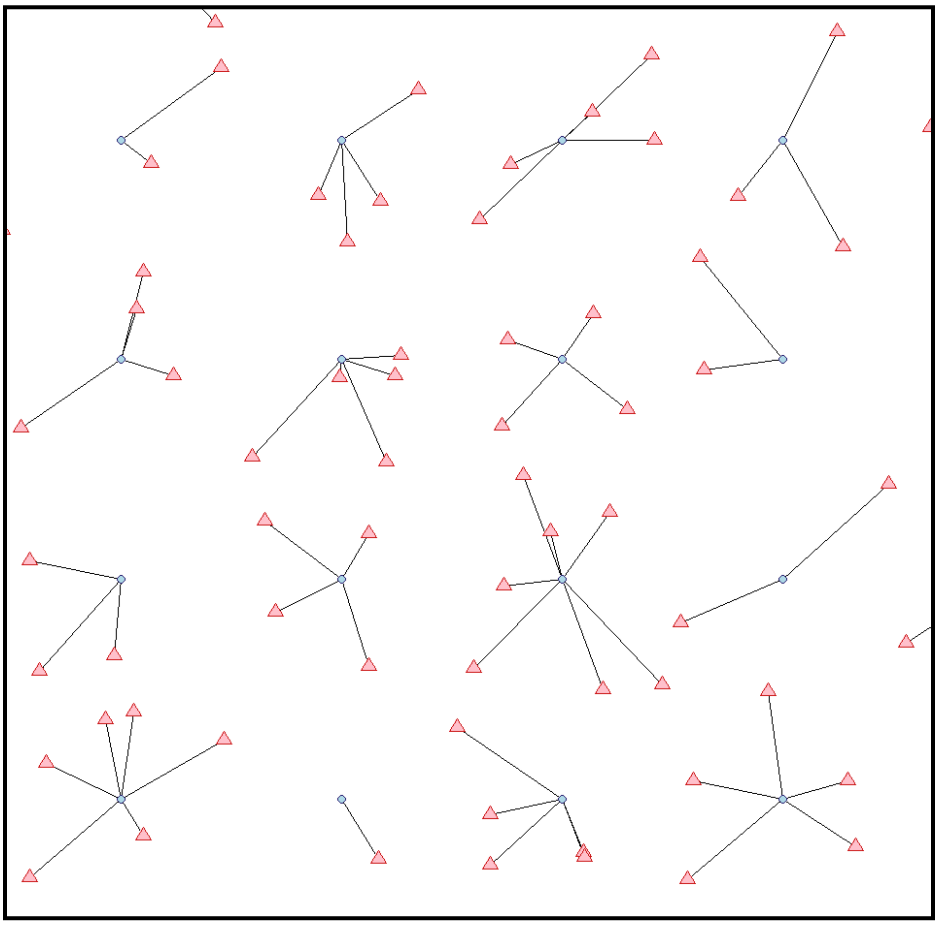

Figure 4. Minimum distance criteria used to couple, with a one-to-many relationship, to each TRMM centroid (blue circles) one or more rain gauge stations (red triangles).

3.1 Analysis of correlation

To analyse the correlation between cumulative rainfall measured by rain gauges (*R*) and estimated by satellites (*T*), we exploit linear (**Eq. 1**) and power law (**Eq. 2**) regression models.

$$R = a \times T \qquad \text{Eq. [1]}$$

$$R = T^{b} \qquad \text{Eq. [2]}$$

In the equations *a* is the linear coefficient and *b* is the power law exponent.

R and *T* couple values correspond to cumulative rainfall estimated from the analysis of raw rainfall data series for different rainfall durations. In the analyses we exploit rainfall period of 24 and 72 hours. Once cumulative rainfall values were estimated for each rain gauge/centroid couple, we aggregate those in different morpho-climatological domains corresponding to morphological zones derived from the 230m DEM analysis. Using a peak over threshold approach we then filter rainfall series taking values above different thresholds: 0, 2, 4, 6, 8, 10, 12, 14, 16, 18 and 20 mm. For each threshold we estimated linear (*a*) and power low (*b*) fitting parameters (**Figure 5**), the relative determination coefficients (R^2, measuring the goodness of fit), and confidence and prediction intervals (**Figure 6a,b**). For both linear and power law models and for each morphological subdivision we realize boxplots showing the variability of the fitting parameters (**Figure 6c,d**) and of the determination coefficients associated to the different rainfall thresholds (**Figure 6e,f**).

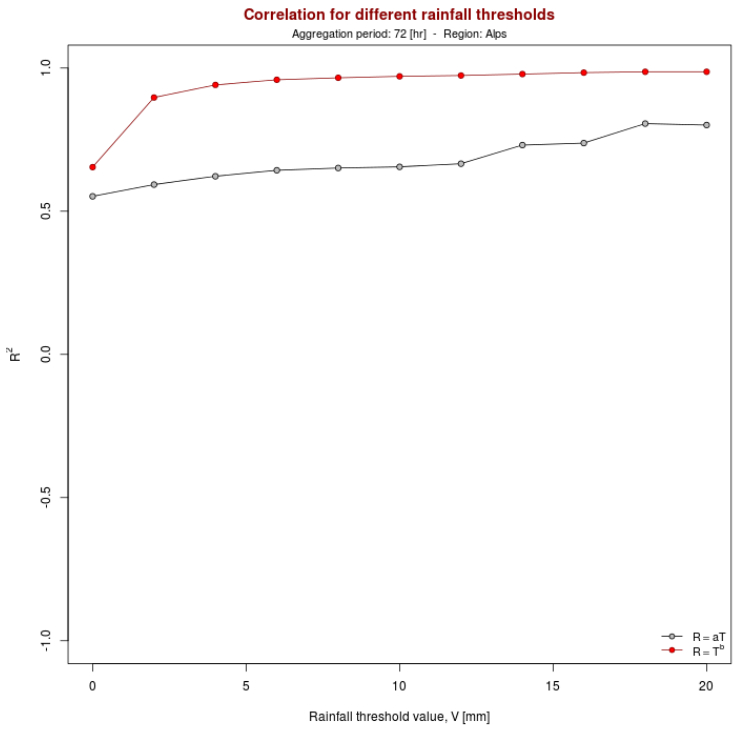

Figure 6. Determination coefficients estimated in the Alpine area for linear (grey circles) and power law fitting correlation calculated considering 72-hours cumulative rainfall above different thresholds.

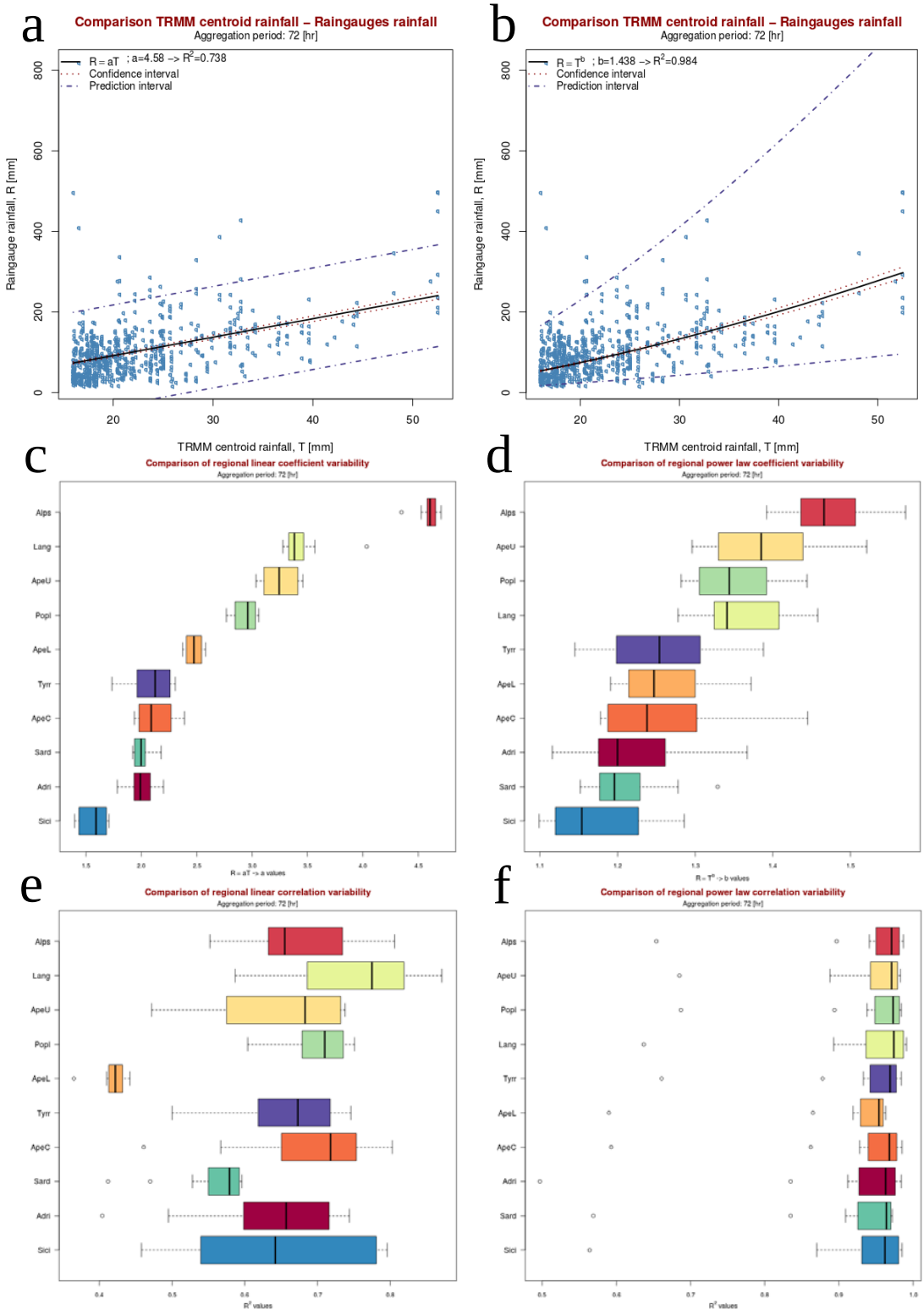

Figure 6. Linear (a) and power law (b) fitting between 72-hour TRMM rainfall estimates *T* and rain gauge measures *R*. Variability of linear (c) and power law (d) coefficients and determination coefficients (e, f) for different rainfall thresholds in the different Italian morphological subdivisions.

3.2 Analysis of cumulative rainfall distribution

We further analyzed the distribution of satellite and rain gauge cumulative rainfall. First since data appear to be heavy tailed distributed we estimated the frequency of the logarithm of the rainfall of both series (**Figure 7a,b**) using the histogram estimation method (**Venables and Ripley, 2002**). Then using the kernel density estimation method (**Parzen, 1962**) we calculate the kernel densities of the logarithm of both series using a Gaussian kernel, and we compare these with the densities obtained by the histogram estimation method (**Figure 8**). Finally using maximum likelihood estimation method (**Fisher, 1922a, 1922b; White et al., 2008**) we estimated the parameter of the exponential distribution we assumed to model cumulative rainfall data (**Figure 9**). Using a bootstrap procedure we estimate the error associated to the exponential models. For the purpose we sampled the original rainfall series 100 times (bootstrap resampling), then for each sample we estimated the exponential model parameters and the associated modeled probabilities. We then calculated the mean (μ) and the standard deviation (σ) of the modeled probabilities and we use $\mu \pm \sigma$ as error bars.

To measure the goodness of fit we first generated rainfall series exponential distributed and we realized QQplot (**Wilk and Gnanadesikan, 1968**) (insets in **Figure 9**) comparing raw and exponential modeled data. Further we performed a bootstrapped version of the two-sided Kolmogorov-Smirnov test (**Kolmogorov,1933; Smirnov, 1933**) to quantify the distance between the empirical distribution function of the rainfall data and the exponential distribution function estimated from data (used as reference distribution). The Kolmogorov-Smirnov test statistics (D and pvalue) calculated are shown in **Figure 9**.

The analyses allowed the direct comparison of the distribution of rain gauge and satellite cumulative rainfall data. All the analyses were repeated considering 24 hour cumulative rainfall.

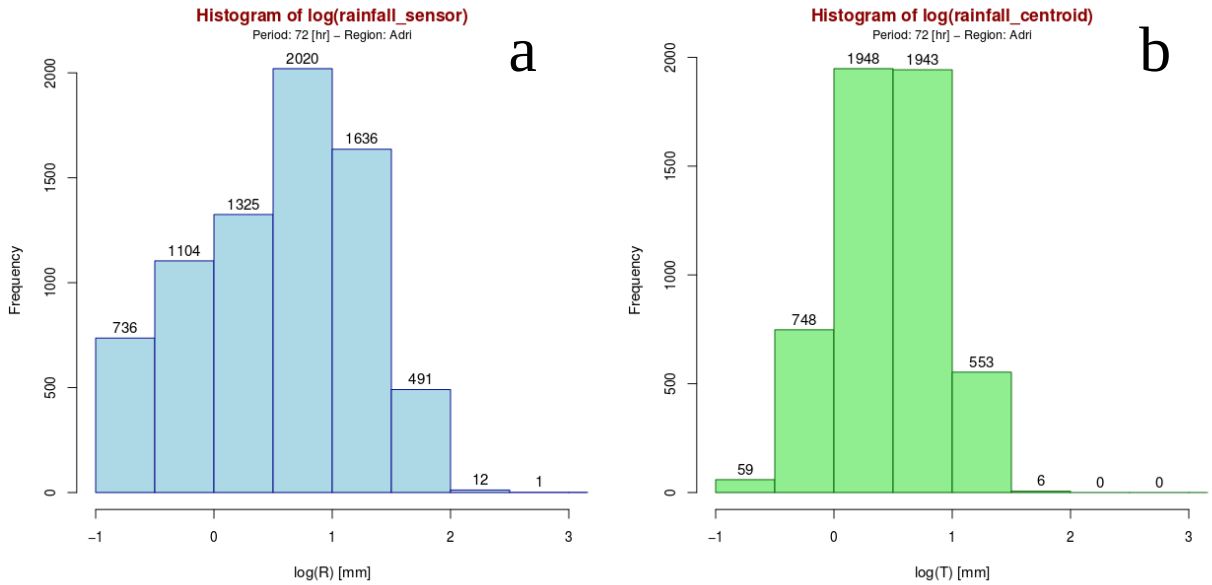

Figure 7. Frequency distribution of the logarithm of 72-hours cumulative rainfall calculated for (a) rain gauge measures *R* and for (b) TRMM rainfall estimates *T* in the Adriatic morphological subdivision.

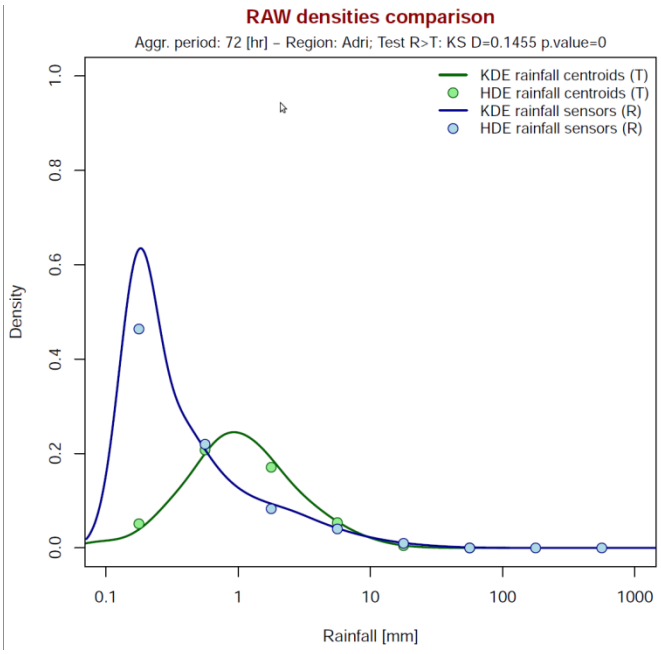

Figure 8. Probability density of the logarithm of 72-hours cumulative rainfall estimated using Histogram Density Estimation (HDE, cirlces) and Kernel Density Estimation (KDE, lines) for rain gauge measures *R* (in blue) and for TRMM rainfall estimates *T* (in green) in the Adriatic morphological subdivision.

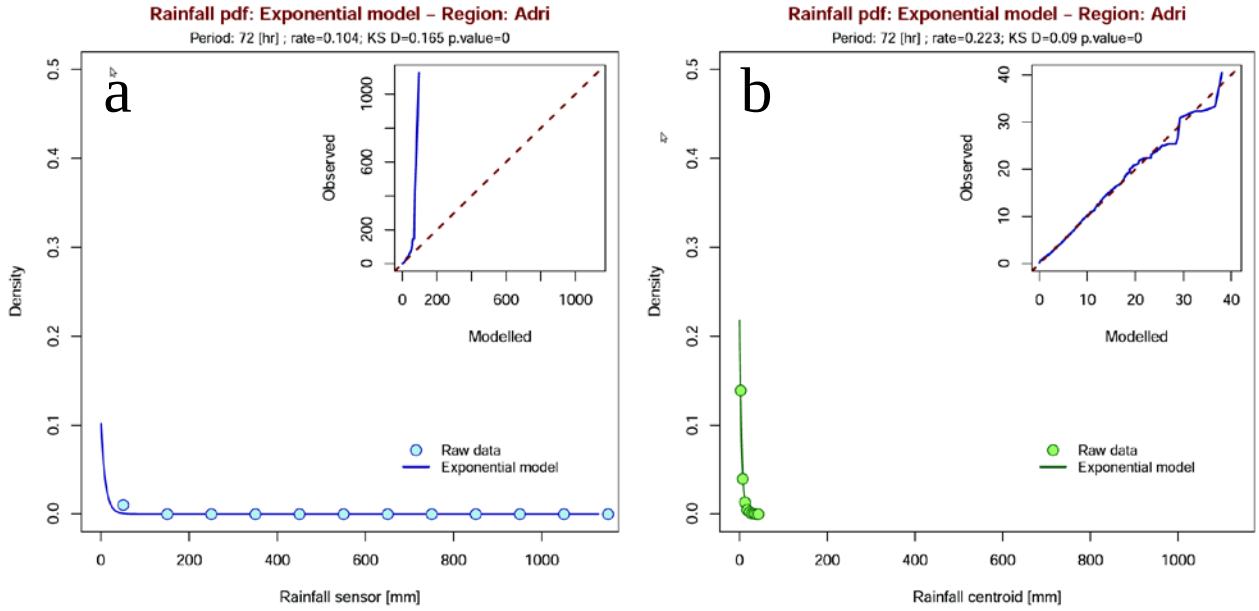

Figure 9. Exponential distribution estimated for 72-hours cumulative rainfall using Maximum Likelihood Estimation (MLE) for rain gauge measures *R* (a) and for TRMM rainfall estimates *T* (b) in the Adriatic morphological subdivision. Insets show QQplot measuring the fitting performance of the exponential model. KS D and p.value are Kolmgorov Smirnow test statistics.

4. Analysis of rainfall events

The increase in pore water pressure due to the rainfall infiltration process is the real triggering mechanism of rainfall-induced landslides. Infiltration models could play a fundamental role in landslide initiation modelling and hence in landslide prediction. Despite this, physical infiltration models can rarely be applied, given the high demand in terms of geotechnical and hydrological parameters, and often their application is limited to single slope or small basins. Consequently the study of the landslide triggering mechanisms over large areas is based on empirical approaches, mainly exploiting direct rainfall measures associated to landslide phenomena. Empirical rainfall thresholds have been largely used to study the landslides initiation mechanisms (**Guzzetti et al., 2007; 2008**) and to predict landslides over a territory. Different rainfall measures have been exploited to define empirical rainfall thresholds, but commonly mean intensity, cumulative and duration of rainfall are used to define empirical rainfall thresholds. Those type of thresholds are defined as low boundary of rainfall conditions associated to landslides, and they require the identification of (i) the beginning, (ii) the end, and (iii) the cumulative rainfall of events that triggered landslides. Rainfall during an events is considered the responsible of landslide triggering and indirectly reflects the infiltration process. Commonly rainfall events are identified by experts following heuristic procedures (**Rossi et al., 2012**), but often due to their subjectivity they introduce a source of uncertainty (due to the subjectivity) in rainfall threshold definition. To overcome this problem we implemented an automatic procedure to identify rainfall events starting from raw rainfall series.

4.1 Procedure for the identification of rainfall events

The automatic procedure for the identification of rainfall events was implemented in R (**R Development Core Team, 2011**). A rainfall event is defined as the cumulative rainfall occurring during a period in hours with a continuous rainfall series. To separate rainfall events we consider a minimum period in hours with cumulative rainfall below a given thresholds. Essentially two rainfall events are separated, if between them there is a period of a given length without rainfall or with rainfall below a given threshold. The procedure iteratively determine this condition using a moving window. The procedure output consists of the identification of the beginning, the end and the cumulative rainfall associated to rainfall events in a given rainfall series. The automatic procedure produces plots and tables identifying the rainfall events. The procedure was applied in the Umbria region where rainfall events associated to landslides, collected in the period 2003-2011 (**Figure 10**), were identified using the heuristic procedure above mentioned. The automatic procedure allowed the identification of rainfall events in each rain gauge and in each TRMM cell centroid inside the Umbria regional boundary. Two example of the rainfall event series identified for one rain gauge and one TRMM cell centroid are shown in **Figure 11**. In the analysis, to identified rainfall events, we use a 24-hours period and two rainfall thresholds corresponding to 0.2 mm for rain gauge measurements and 0 mm for satellite rainfall estimates.

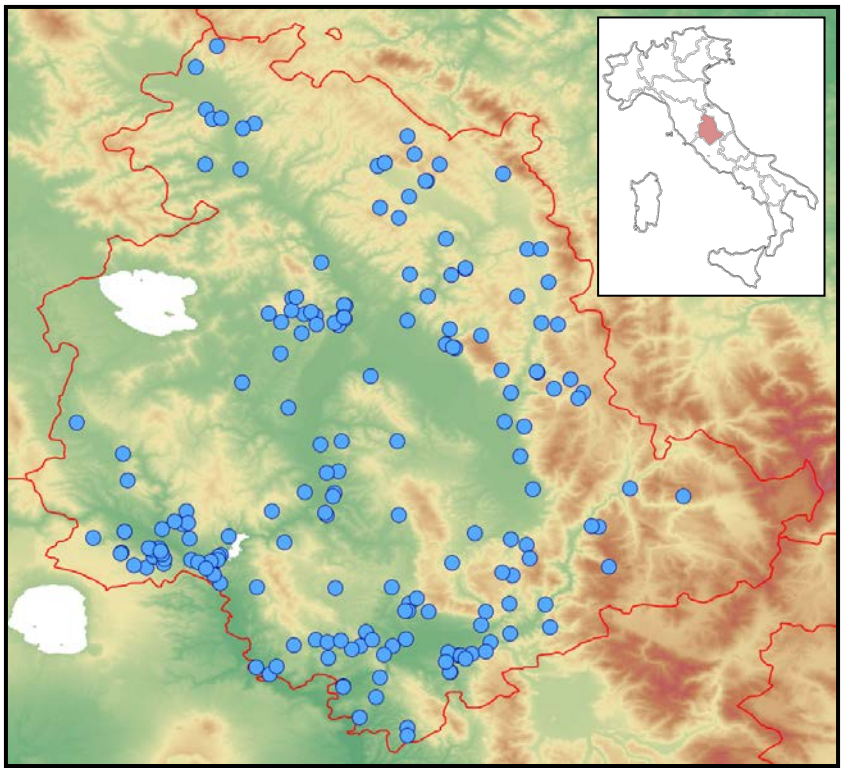

Figure 10. Location of landslide events recorded in the Umbria region (inset) in the period 2003-2011.

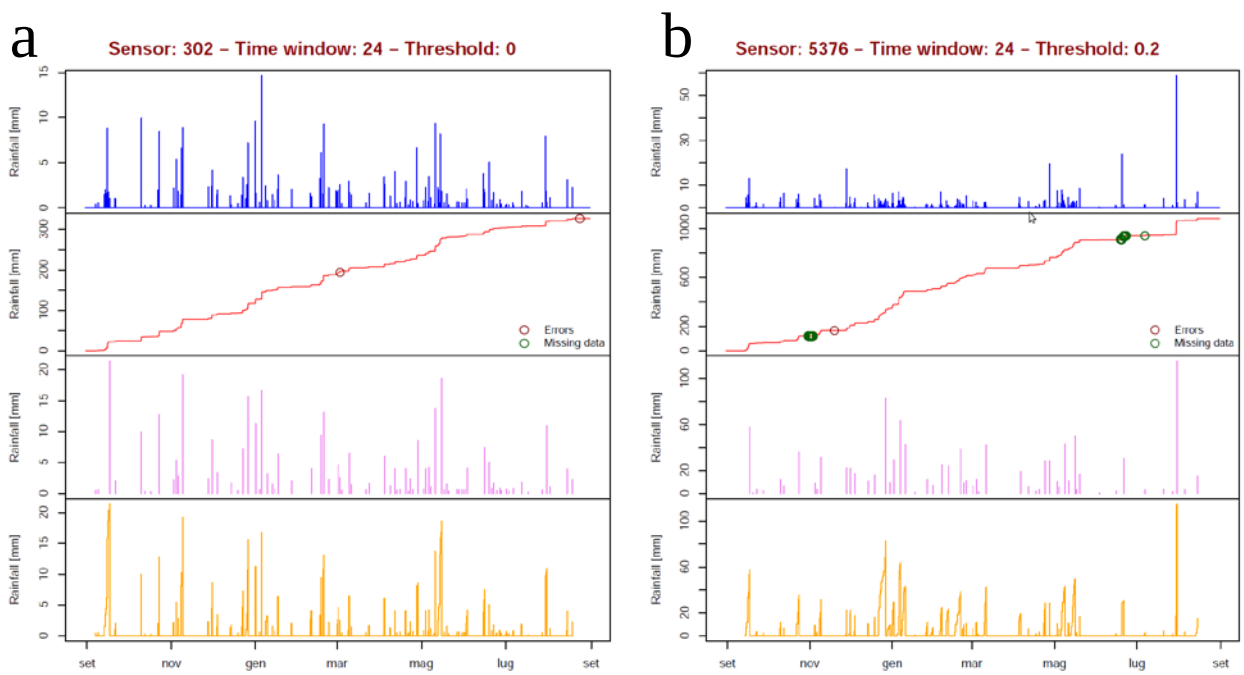

Figure 11. Rainfall events determined by an automatic procedure for rain gauge measures *R* (a) and for TRMM rainfall estimates *T* (b). Blue bars: raw rainfall data; red lines: cumulative rainfall data; pink bars: rainfall events; yellow lines: cumulative during rainfall events.

4.2 Rainfall event comparison

Rainfall events calculated by the procedure starting from rain gauge and satellite rainfall series, were compared with the distribution of rainfall events associated to landslide using a standard heuristic procedure (**Rossi et al., 2012**). Cumulative and duration associated to rainfall events, identified in the Umbria region, were plotted in a bi-logarithmic plot (**Figure 12**). Finally we calculated descriptive statistics of these events including: (i) the number of rain gauges and TRMM cell centroids used in the Umbria region for the rainfall event identification; (ii) the average number of events identified by the procedure; (iii) the mean duration; (iv) the mean cumulative; and (v) the mean intensity of the identified events (**Table 3**).

Table 3. Descriptive statistics calculated for rainfall events determined automatically analysing rain gauge and TRMM rainfall data series in the Umbria region in the period 01/09/2009 to 31/08/2010.

DATA TYPE	SENSORS CENTROIDS (#)	RAINFALL EVENTS (#)	MEAN DURATION (hr)	MEAN CUMULATIVE (mm)	MEAN INTENSITY (mm/hr)
Rain gauge data	52	56	28.07	17.70	1.24
TRMM data	17	56	17.47	6.07	0.59

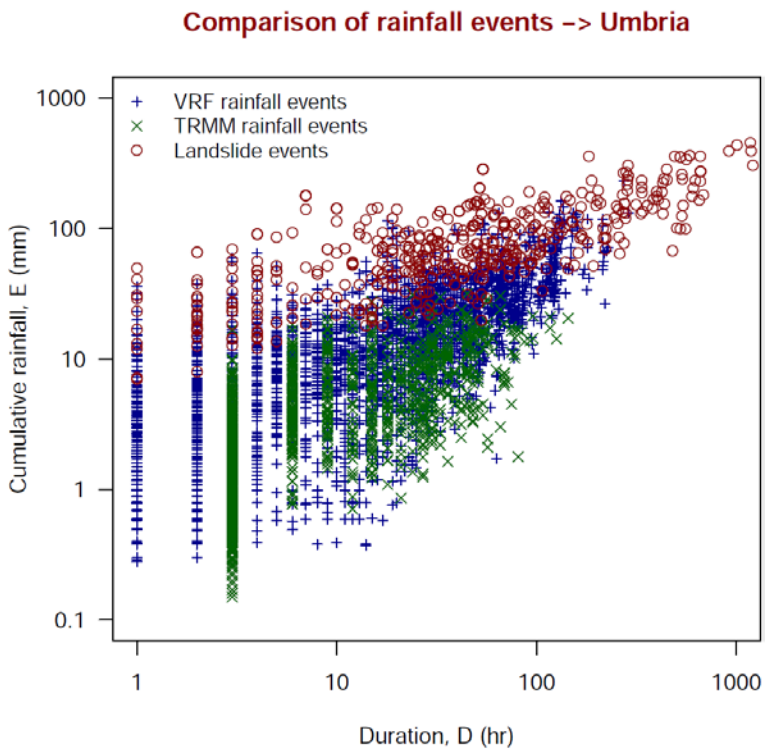

Figure 12. Cumulative rainfall and duration of: (i) rainfall events associated to landslides in Umbria region (red circles) determined heuristically analysing rain gauges rainfall in the period 2003-2010; (ii) rainfall events determined automatically from rain gauge measures (blue pluses); and (iii) rainfall events determined automatically from TRMM rainfall estimates (green crosses). In (ii) and (iii) rainfall events were determined in the Umbria region in the period 01/09/2009 to 31/08/2010.

5. Discussion and conclusion

In this work we performed an extensive validation of NASA TRMM v6 Algorithm 3B42 rainfall estimates in Italy, exploiting rainfall data from the Italian rain gauge network available through the Experience system of the Italian Civil Protection Department. To consider the variability data associated to different morphological and climatological condition, the validation activities were performed aggregating data in different Italian morphological subdivisions derived from **Guzzetti and Reichenbach (1994)** analysing a 230m DTM obtained mosaicking the entire Italian mean elevation archive. The results allowed also to evaluate the use of satellite rainfall estimates in the analysis of rainfall-induced trigger mechanisms, and in empirical landslide prediction models.

Primarily we analyse the correlation between cumulative rain gauge and satellite rainfall relative to 24 and 72-hours durations (**§3.1**) using linear and power-law fitting models. We repeat the analysis for rainfall above different thresholds (peak over threshold approaches) in different morphological and climatological conditions. Boxplot in **Figure 6c,d,e,f** describe the variation of parameters and of R^2 squared estimated for different rainfall thresholds in each Italian morphological zones. For both fitting models (linear and power law), the regression parameters changes in different morphological zones, indicating a complex correlation between satellite and rain gauge rainfall. In general results show that TRMM data underestimates rain gauge data. In particular in high elevation areas (Alps, ApeU) TRMM data tends to underestimate rain gauge data much more (given the higher correlation coefficients) than in low altitude areas (Sici, Adri, Sard). Results for some morphological subdivision (Popl, ApeC, Lang) seem to don't support this finding probably because the rough aggregation procedure. This finding has a direct impact on the use of TRMM satellite estimates in the analysis of rainfall triggering mechanisms and in the use of this data in landslide prediction models. For instance TRMM rainfall estimates cannot be used directly in landslide early warning systems based on the comparison of rainfall data and empirical rainfall thresholds derived analyzing rain gauge rainfall measures associated to landslides (**Guzzetti et al., 2007; 2008; Rossi et al., 2012**). Power law regression models gave always the best fitting performance (mean $R^2 \approx 0.95$) compared to the linear regression models (mean $R^2 \approx 0.7$) but at the expenses of high prediction intervals in particular for high rainfall values (**Figure 6a,b**). Best fitting results are obtained in correlating high rainfall values (higher R^2 values for high rainfall thresholds) as shown in **Figure 6**.

We further analyzed the statistical distribution of satellite and rain gauge 72-hours cumulative rainfall data using non parametric (HDE, KDE) and parametric (KDE) estimation methods (**§3.2**). Empirical frequency distributions (**Figure 7**) of the logarithm of rain gauge and satellite rainfall (obtained by Histogram Density Estimation methods) are different, as shown by the difference in the mode (more frequent rainfall values) and by the different distribution tails. Similar differences can be observed analyzing the probability densities of the logarithm of rain gauge and satellite rainfall calculated using the Kernel Density Estimation method (**Figure 8**). Using the maximum likelihood method (MLE) we then estimated exponential model parameters starting from raw cumulative rainfall data series, and we use QQplots to compare observed and modeled data. Results indicate that an exponential model can be appropriate to describe the 72-hours TRMM cumulative rainfall data (QQplot values along bisector, see inset **Figure 9b**), while a different distribution, characterized by an heavier tail for high rainfall values, must be used to model rain gauge rainfall (positive difference between QQplot values and bisector for high

rainfall values, see inset **Figure 9a**). Similar results can be observed analyzing the differences of the cumulative rainfall distributions obtained for rain gauge and satellite data for different morpho-climatological zones and for 24-hours rainfall duration.

Finally we compared rainfall events identified from rain gauge and satellite raw rainfall series using and automatic procedure, with rainfall events associated to landslide calculated using an heuristic procedure. For these analysis we selected rain gauges and TRMM cell centroids in the Umbria region where landslide information (**Figure 10**) and the associated rainfall were already available. The comparison (**Figure 12**) showed that a different distribution of cumulative and duration values characterizes rainfall events obtained using satellite and rain gauge rainfall. In general rainfall events identified from rain gauge data series are longer and more intense than those identified from satellite rainfall series (**Table 3**). A similar number of rainfall events have been identified by the automatic procedure from 52 rain gauges and from 17 TRMM cell centroids in the Umbria region in the period 01/09/2009 to 31/08/2010. Both rainfall event distributions are distinct from the distribution characterizing rainfall events associated to landslides. These important findings suggest once more that TRMM satellite estimates are incomparable with rain gauge data and hence with empirical rainfall thresholds for landslide prediction derived from rain gauge data. Nevertheless the similar number of rainfall events identified from rain gauge and satellite data suggests that TRMM data can be used to identify possible triggering rainfall events and hence to define new rainfall thresholds from satellite data. Those thresholds could be potentially integrated in existing early warning system to predict landslides exploiting satellite rainfall estimates.

The distributions characterizing rainfall events, derived from rain gauge data, associated and not associated to landslide are significantly different (**Figure 12**). This suggests that is possible to determine empirical rainfall thresholds starting from the rainfall event conditions associated or not associated to landslide in a given area. In this context the procedure for rainfall event determination could be a fundamental tool to identify and analyze rainfall conditions triggering landslides exploiting both rain gauge and satellite rainfall data. The procedure can also potentially allow to diminish the uncertainty related to the heuristic identification of rainfall event associated to landslides.

In conclusion a complex relation exists between rain gauge and satellite rainfall data in different morpho-climatological settings characterizing the Italian territory. Satellite rainfall estimates are not straightforward comparable with rain gauge rainfall measures, requiring a detailed local investigation of degree of correlation. Nevertheless satellite rainfall data can be exploited to analyze rainfall conditions triggering landslides and hence to define empirical rainfall threshold for the landslide initiation. Those could potentially be integrated in early warning system to predict landslide over large areas.

Acknowledgements

MR was supported by a grant of the Italian national Department for Civil Protection (DPC).

References

Caine, N., (1980). The rainfall intensity-duration control of shallow landslides and debris flows. Geogr Ann A 62: 23–27.

Fisher, R.A. (1922a). On the mathematical foundations of theoretical statistics. Philosophical Transactions of the Royal Society of London. Series A 222: 309–368.

Fisher, R.A. (1922b). The goodness of fit of regression formulae, and the distribution of regression coefficients. Journal of the Royal Statistical Society 85(4): 597–612.

Guzzetti, F., Peruccacci, S., Rossi, M., Stark, C.P. (2007). Rain-fall thresholds for the initiation of landslides in central and southern Europe. Meteorology and Atmospheric Physics 98: 239-267.

Guzzetti, F., Peruccacci, S., Rossi, M., Stark, C.P. (2008). The rainfall intensity-duration control of shallow landslides and debris flows: an update. Landslides 5(1): 3-17.

Guzzetti, F. and Reichenbach, P. (1994) Towards a definition of topographic divisions for Italy. Geomorphology 11(1): 57-74.

Kirschbaum, D.B., Robert Adler, Yang Hong, S.V. Kumar, C. Peters-Lidard, and Lerner-Lam, A. (2011). Advances in landslide nowcasting: evaluation of a global and regional modeling approach Environmental Earth Sciences10.1007/s12665-011-0990-3.

Kolmogorov, A. (1933). Grundbegriffe der Wahrscheinlichkeitsrechnung. Julius Springer: Berlin.

Innes, J.L. (1983). Debris flows. Prog Phys Geog 7: 469–501.

Iverson, R.M. (2000) Landslide triggering by rain infiltration. Water Resources Research, 36, 1897-1910.

Montgomery, D.R., and Dietrich, W.E. (1994). A physically based model for the topographic control of shallow landsliding. Water Resources Research, 30:4 1153-1171.

Neteler, M.; Mitasova, H. (2008). Open Source GIS : a GRASS GIS approach, 3rd edition. New York: Springer. ISBN 978-0-387-35767-6.

Parzen, E. (1962). On estimation of a probability density function and mode. Annals of Mathematical Statistics 33: 1065–1076. DOI. 10.1214/aoms/1177704472.

R Development Core Team (2011). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. http://www.R-project.org.

Rossi, M., Peruccacci, S., Brunetti, M.T., Marchesini I., Luciani, S., Ardizzone, F., Balducci, V., Bianchi, C., Cardinali, M., Fiorucci, F., Mondini, P. Reichenbach, A.C., Salvati, P., Santangelo, M., Bartolini, D., Gariano, S.L., Palladino, M., Vessia, G., Viero, A., Antronico, L., Borselli, L., Deganutti, A.M., Iovine, G., Luino, F., Parise, M., Polemio, M., Guzzetti, F., Tonelli, G. (Accepted for publication). SANF: a national warning system for rainfall-induced landslides in Italy. Proceeding of the 11th International Symposium on Landslides (ISL) and the 2nd North American Symposium on Landslides (NASL) 2012. Banff, Alberta, Canada, 2-8 June 2012.

Salvati, P., Bianchi, C., Rossi, M., Guzzetti, F. (2010). Societal landslide and flood risk in Italy. Natural Hazards and Earth System Sciences 10: 465-483.

Smirnov, N. (1933). Estimate of deviation between empirical distribution functions in two independent samples. Bulletin of Moscow University 2(2): 3–16.

Venables, W.N. and Ripley, B.D. (1999) Modern Applied Statistics with S-PLUS, Springer, New York.

White, E.P., Enquist, B.J., Green, J.L. (2008). On estimating the exponent of power-law frequency distributions. Ecology 89(4): 905–912. DOI. 10.1890/07–1288.1.

Wieczorek, G.F. (1996). Landslide triggering mechanisms. In: Landslides: investigation and mitigation (Turner AK, Schuster RL, eds). Washington DC: Transportation Re- search Board, National Research Council, special report, pp 76–90.

Wilk, M.B. and Gnanadesikan R. (1968). Probability plotting methods for the analysis of data. Biometrika 55(1): 1–17.

Wu, W., Sidle, R.C. (1995) A distributed slope stability model for steep forested basins. Water Resour Res 31: 2097–2110.