Natural Hazards
and Earth System
Sciences

Open Access

Analysing the relationship between rainfalls and landslides to define a mosaic of triggering thresholds for regional-scale warning systems

S. Segoni, A. Rosi, G. Rossi, F. Catani, and N. Casagli

Earth Sciences Department, University of Firenze, Via La Pira 4, 50121 Florence, Italy

Correspondence to: S. Segoni (samuele.segoni@unifi.it)

Abstract. We propose an original approach to develop rainfall thresholds to be used in civil protection warning systems for the occurrence of landslides at regional scale (i.e. tens of thousands of kilometres), and we apply it to Tuscany, Italy ($23\,000\,\mathrm{km}^2$).

Purpose-developed software is used to define statistical intensity–duration rainfall thresholds by means of an automated and standardized analysis of rainfall data. The automation and standardization of the analysis brings several advantages that in turn have a positive impact on the applicability of the thresholds to operational warning systems. Moreover, the possibility of defining a threshold in very short times compared to traditional analyses allowed us to subdivide the study area into several alert zones to be analysed independently, with the aim of setting up a specific threshold for each of them. As a consequence, a mosaic of several local rainfall thresholds is set up in place of a single regional threshold. Even if pertaining to the same region, the local thresholds vary substantially and can have very different equations. We subsequently analysed how the physical features of the test area influence the parameters and the equations of the local thresholds, and found that some threshold parameters can be put in relation with the prevailing lithology. In addition, we investigated the possible relations between effectiveness of the threshold and number of landslides used for the calibration.

A validation procedure and a quantitative comparison with some literature thresholds showed that the performance of a threshold can be increased if the areal extent of its test area is reduced, as long as a statistically significant landslide sample is present. In particular, we demonstrated that the effectiveness of a warning system can be significantly enhanced if a mosaic of site-specific thresholds is used instead of a single regional threshold.

1 Introduction

Rainfall-triggered landslides are one of the most common natural hazards, responsible for casualties and economical losses worldwide (Petley, 2012). To reduce this impact, the scientific community is working on forecasting the occurrence of landslides and setting up warning systems. When working over large areas (e.g. thousands of squared kilometres), the computational load required and the difficulty in assessing the spatial organization of geotechnical parameters prevent the application of physically based models (Baum et al., 2010; Agostini et al., 2014; Rossi et al., 2013). As a consequence, when the area of study is a large district or region, the approach used is frequently based on statistical landslide models relying on rainfall exceedance (Larsen and Simon, 1993; Farahmand and AghaKouchak, 2013; Hong et al., 2006, 2007; Bovolo and Bathurst, 2012; Segoni et al., 2014a). In particular, empirical rainfall thresholds are frequently used in operational warning systems (Brunsden, 1973; Aleotti, 2004; Hong et al., 2005; Tiranti and Rabuffetti, 2010; Cannon et al., 2011; Martelloni et al., 2012; Lagomarsino et al., 2013). Among all rainfall thresholds approaches, the one using intensity–duration (I–D) thresholds (Caine, 1980) is perhaps the most popular: it has been proved particularly valid for shallow landslides (Caine, 1980; Crosta and Frattini, 2001; Ahmad, 2003; Jakob and Weatherly, 2003; Aleotti, 2004; Guzzetti et al., 2008; Giannecchini et al., 2012), and it has also been successfully applied to

landslides in general (Zimmermann et al., 1997; Hong et al., 2005; Brunetti et al., 2010; Rosi et al., 2012).

Although widely used, this approach is currently affected by some drawbacks that hinder a fully operational application to early warning systems. One of the main problems is a certain degree of subjectivity in some state-of-the-art procedures used to obtain the $I–D$ relationship. The definition of the threshold from the $I–D$ points has long been visually drawn with manual fitting (e.g. Caine, 1980; Giannecchini et al., 2012), and only recently has this issue been solved through proposing objective and robust statistical approaches to identify a threshold with a chosen confidence level from a given cloud of $I–D$ points (Guzzetti et al., 2007, 2008; Brunetti et al., 2010; Rosi et al., 2012).

However, even the definition of the $I–D$ points themselves poses problems of subjectivity that can in turn affect the applicability of the thresholds to warning systems. In fact, especially when considering complex pluviometric rainfall paths where subsequent bursts of rain of varying intensity and duration alternate with short periods of absent or moderate rain, the whole rainfall event has to be summarized in a single $I–D$ point; this procedure is not straightforward, as the result may vary depending on the choice of the reference rain gauge and on the interpretation of the pluviometric path. In particular, the start and end point of the critical rainfall event (Aleotti, 2004) sometimes could not be clearly and univocally identified (e.g. when the hour of occurrence of the landslide is not known with sufficient precision). Most of the studies resort to subjective interpretations, but while this can influence the results (Guzzetti et al., 2008), a subjective decision in the analytical process cannot be consistently replicated by an automated warning system.

The maximum degree of objectivity, standardization and replicability is obtained when the analysis to define the threshold and the warning system is based on rainfall parameters calculated and measured in a given time span; this approach ensures that the rainfall analysis can be easily and consistently replicated by automated warning systems. Indeed, at present most operational warning systems are based on rainfall parameters as measured over a given duration (Wilson, 2000; Chleborad, 2003; Cardinali et al., 2006; Cannon et al., 2008, 2011; Lagomarsino et al., 2013).

However, $I–D$ approaches have proved very effective in defining the minimum rainfall conditions that can potentially trigger landslides (Guzzetti et al., 2008; Brunetti et al., 2010), but this aim is slightly different from the objective of an operational early warning system, where a balance between false alarms and missed alarms is usually required (Staley et al., 2013).

This work proposes an original approach to overcome the aforementioned issues: the threshold is drawn according to rigorous statistical techniques; the $I–D$ points are defined according to an automated analysis (Segoni et al., 2014b) that can be easily and consistently replicated by an automated warning system, and the proposed procedure is com-

pleted by a back analysis aimed at minimizing errors of commission (i.e. false positives). The technique of Segoni et al. (2014b) was applied to a large region (Tuscany, central Italy, 23 000 km^2), using the proposed approach of defining a mosaic of local thresholds instead of a single regional threshold in order to strengthen, according to local physical features, the empirical relationship between meteoric events and landslide triggering.

The applicability of the mosaic of thresholds to early warning systems for civil protection purposes (and thus the effectiveness of the proposed approach) has been tested by means of a validation procedure that provided satisfactory results. The validation was also extended to some literature thresholds, so as to perform a quantitative comparison for a better evaluation of the effectiveness of our approach.

Finally, we investigated the extent to which the environmental setting of a study area influences the rainfall analysis and the resulting threshold equation, with the aim of finding some physical background in the empirical intensity–duration relationship.

2 Material and methods

2.1 Test site

The proposed methodology was applied in the region of Tuscany (23 000 km^2), which is located in central Italy and is characterized by mainly hilly (66.5 %) and mountainous (25.1 %) terrain, with limited lowland areas (8.4 %) corresponding to intermontane basins and to the southern coastline (Fig. 1a).

Tuscany is characterized by a variety of lithological units with very different mechanical properties. The hilly territories are mainly constituted by granular or cohesive terrains or by soft rocks. The north and the east contain the reliefs of the Apennine folds and thrusts belt, made up of mainly flysches, while a metamorphic unit outcrops in the north-western sectors. Lastly, in smaller but still relevant portions of the territory, evaporites, carbonatic rocks, effusive rocks and intrusive rocks are present.

Tuscany has a typical Mediterranean climatic regime with mild and moist winters, hot and dry summers, and two precipitation peaks (the main one in autumn and the secondary one in spring or winter), while summer is always the driest period of the year. The wet season is characterized by rainfalls that cover large sectors of the region, while, during the warm season, convective thunderstorms give way to precipitation events with shorter duration localized in smaller spots. In both cases, the areal distribution of rainfalls is markedly influenced by the relief: in the north-western part of the region, in particular, mean annual precipitation (MAP) is about 2000 mm year^{-1} (with annual peaks of 3000 mm year^{-1}), while southern Tuscany is characterized by very lower rainfall amounts (about 600 mm year^{-1}
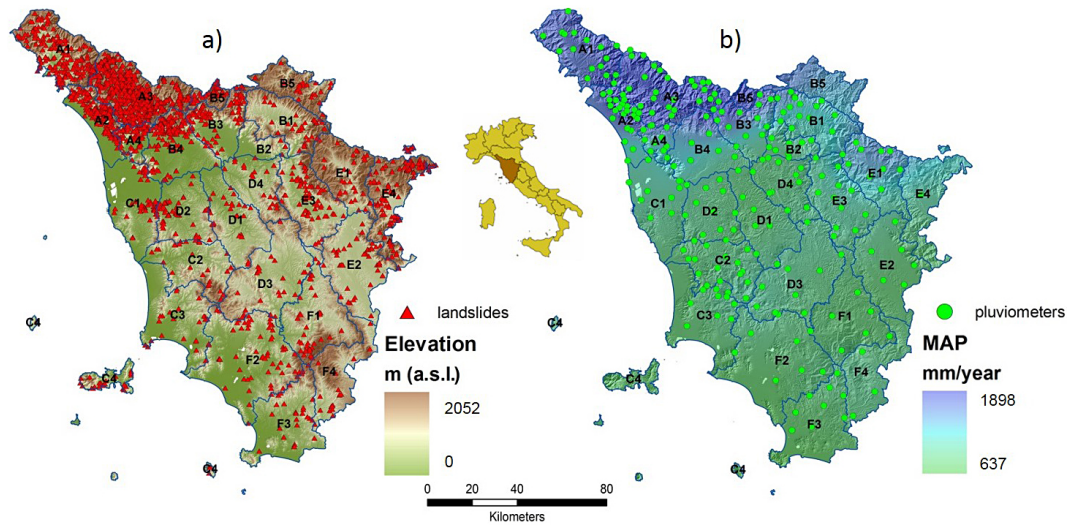
Figure 1. The Tuscany region, subdivided into 25 alert zones (AZ), with landslide inventory laid over the digital elevation model **(a)** and the rain gauge distribution laid over the mean annual precipitation map, as averaged from the measures pertaining to the data set used in this work (years 2000–2009) **(b)**.

MAP), as shown in Fig. 1b and discussed by Rapetti and Vittorini (1994) and Fatichi and Caporali (2009).

To account for the high variability of meteorological and physiographic settings encountered in the study area, and to get more accurate rainfall thresholds, the test site was partitioned into 25 alert zones (AZ) (Fig. 1).

The partition accounts for the main physical features of the territory: the AZ boundaries follow the main regional divides, thus defining catchments with homogeneous meteorological conditions (Rapetti and Vittorini, 1994; Crisci et al., 2002; Fatichi and Caporali, 2009). Furthermore, the partition discriminates AZs with different geological characteristics ensuring each AZ has homogeneous geomorphological features. Each AZ was independently analysed to devise a site-specific rainfall threshold.

2.2 Input data

To define the rainfall thresholds, data from over 2000 landslides (Fig. 1a) that occurred between the year 2000 and the beginning of 2009 were collected. These data were split into two data sets: a calibration data set (2000–2007) and a validation data set (2008–2009). Data were collected mainly from the archives of Tuscany Civil Protection Agency, but also from local authority archives, from national and local newspapers, and from existing data sets of recent research projects (Catani et al., 2013; Mercogliano et al., 2013; Rosi et al., 2013).

Every landslide was filed in a geo-database with a unique identification code, its spatial location, the main characteristics (when known), the occurrence date, and any other available information. Almost all of the landslides did not have the exact occurrence time, and instead only the day of occur-

rence was reported. More accurate timing (e.g. "during the evening" or "in the night") was available for about the 10 % of the landslides. Considering that the exact timing of occurrence is unknown in many large-scale rainfall threshold studies (Guzzetti et al., 2007), the landslide database could be considered sufficiently accurate.

Rainfall data were collected from 332 rain gauges distributed throughout the region (Fig. 1b), with a spatial density that allows for the rainstorms of interest in the region to be properly accounted for: in particular, the network is denser in those areas where convective thunderstorms of limited size are frequent (Fig. 1b). The hourly rainfall time series of the rain gauge were organized in a database and joined with other information including coordinates and AZ. The dates (day and, if available, hour) of occurrence of the landslides were used to query the rainfall database and to extract, for each landslide, the rainfall data of all the rain gauges of the corresponding AZ.

2.3 Methodology

To define the regional mosaic of thresholds, a recently proposed methodology (Segoni et al., 2014b) was applied separately to every AZ. This methodology is largely automated and thus is useful to manage a large amount of data and to carry out the large number of analyses needed to define 25 different thresholds in a single region. The methodology is explained and discussed in detail in Segoni et al. (2014b); the basic characteristics are summarized hereafter.

A software program named MaCumBA (Massive Cumulate Brisk Analyzer) (Segoni et al., 2014b) analyses the recordings of each rain gauge located in the same AZ. The software automatically carries out the following tasks, which

are otherwise traditionally performed manually and in a subjective way over a limited number of rainfall paths:

- identification, in the rainfall data, of the critical rainfall.

- Definition of the critical parameters used to describe the rainfall event (namely critical intensity I and critical duration D).

- Within all the rain gauges of the same AZ, selection of one most appropriate for the characterization of each landslide event, adopting a criterion that combines proximity to the landslide location and severity of the recorded rainfall.

- The selected I–D values are plotted in a graph, where each point represents the rainfall conditions that resulted in the triggering of a landslide in the past.

- Two thresholds are automatically defined using two different frequentist statistical approaches: the confidence interval technique and the prediction interval technique (Hahn and Meeker, 1991).

Thresholds are defined using the power law first proposed by Caine (1980), and thus they are expressed by the general equation

$$I = \alpha D^{\beta}, \tag{1}$$

where I is the rainfall intensity, D is the rainfall duration, and α (>0) and β (<0) are empirical parameters defined by means of the aforementioned statistical analysis.

The automated procedure is sensitive to some user-defined parameters. Some of them can be properly defined using GIS analyses (e.g. the maximum distance allowed between a landslide and the rain gauges to be used for the characterization of the triggering rainfall), some correspond to political decisions (e.g. the confidence level of the threshold, which in this work, according to the Tuscany Civil Protection Agency, was set to 95 %), and some cannot be decided in advance. This is the case of the parameter called "no rain gap", which accounts for the number of hours without rain needed to consider two rainfall events as separate. The no rain gap is of paramount importance for two reasons: first, it allows for a standardized analysis of the rainfall series, and second it allows for warning systems to analyse rainfall recordings/forecasts in a consistent and completely automated way. However, setting different no rain gaps produces different clouds of I–D points and different rainfall thresholds; thus an objective criterion is needed to identify the configuration that produces the most reliable results. Since the use of MaCumBA allows for thresholds to be calculated in short times, we performed several runs for each AZ using different no rain gap values; then each obtained threshold underwent a back analysis aimed at estimating its performance over the entire testing period, so as to be able to identify and

select the threshold characterized by the lowest number of false alarms, with confidence levels being equal. The details of this part of the methodology can be found in Segoni et al. (2014b), and a graphic example is shown in Fig. 2. A similar approach for choosing the threshold that minimizes errors among different possibilities can be found in Staley et al. (2013).

3 Results

Using the procedure summarized in the previous section, a rainfall threshold was defined for each AZ of the region; the equations are presented in Table 1.

In some AZs the database presented a limited number of landslides; therefore it was not possible to perform a significant statistical analysis. In such cases, we chose to group together some adjacent AZs on the basis of their characteristics (geological setting, topography, rainfall regime).

This procedure was necessary for the central coast and the archipelago (AZs C1,2,3,4), the almost flat AZ B2 (which was grouped with the landslide-rich AZ B3), the inland D1 and D3 AZs, and the southern F1,F4 and F2,F3 AZs.

Alongside the threshold equation, the automated analysis allowed for an important parameter (namely no rain gap) to be defined for each AZ that corresponds to the consecutive number of hours without rain that are needed to consider two rainfall events as separate. The no rain gap parameter is of paramount importance for the implementation of the threshold for civil protection purposes, as it provides automated early warning systems with a consistent criterion to analyse rainfall data.

3.1 Validation

To evaluate the proposed approach and the effectiveness of the thresholds' mosaic for civil protection purposes, a validation procedure was carried out using an independent data set (landslides and rainfall data from January 2008 to January 2009).

The validation procedure was performed by simulating an operational use in a civil protection warning system: if the threshold were in use, when would an alarm have been issued, and when not? And, comparing these dates with the landslide data set, how many correct predictions, false alarms and missed alarms would have been reported?

According to this approach, for each AZ, every rainfall event was classified as true positive (TP or correct prediction: some landslides were triggered in relation to a threshold being exceeded), true negative (TN: the threshold was not exceeded and no landslide was triggered), false positive (FP, or false alarm, i.e. threshold exceeded without landslides triggering), or false negative (FN, or missed alarm: the threshold was not exceeded, but some landslides were triggered).
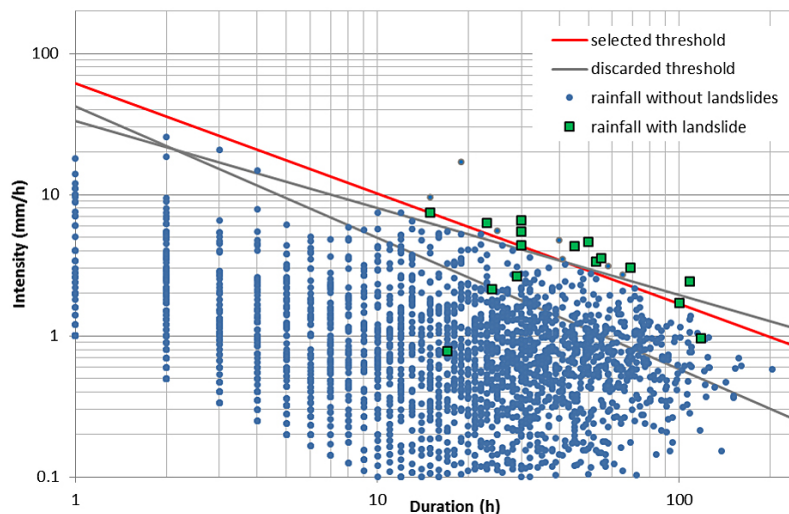
Figure 2. The rainfall threshold defined for alert zone A1, compared with two alternative thresholds that were discarded because of a larger number of false alarms (blue dots above the threshold).

Table 1. Equations and main parameters of the thresholds defined for each alert zone.

Alert zone	Threshold	No rain gap (hours)	Number of landslides	Number of rainfall events that triggered landslides
A1	$I = 61.4D^{-0.78}$	18	246	27
A2	$I = 34.0D^{-0.86}$	18	196	32
A3	$I = 52.4D^{-0.73}$	24	719	79
A4	$I = 101.5D^{-0.99}$	18	90	13
B1	$I = 33.8D^{-0.81}$	20	27	12
B2, B3	$I = 22.5D^{-0.65}$	24	61	34
B4	$I = 49.9D^{-0.73}$	24	208	34
B5	$I = 405.9D^{-1.29}$	24	44	17
C1, C2, C3, C4	$I = 49.2D^{-0.77}$	24	69	28
D1, D3	$I = 40.5D^{-0.90}$	24	39	22
D2	$I = 31.6D^{-0.76}$	12	60	23
D4	$I = 33.5D^{-0.74}$	15	12	11
E1	$I = 20.0D^{-0.66}$	12	26	8
E2	$I = 29.6D^{-0.75}$	12	40	8
E3	$I = 20.9D^{-0.78}$	10	51	13
E4	$I = 15.0D^{-0.69}$	32	166	11
F1, F4	$I = 37.2D^{-0.88}$	24	39	25
F2, F3	$I = 50.7D^{-0.78}$	36	44	20

The validation results are shown in Table 2 and are aggregated at the regional level in a contingency table (Table 3).

4 Discussion

4.1 From a single regional threshold to a regional mosaic of thresholds

The mosaic of thresholds defined for the Tuscany region was compared with two literature thresholds involving, either as a whole or in part, the same area (Fig. 3): the threshold proposed by Brunetti et al. (2010) for the whole of Italy and the threshold proposed by Rosi et al. (2012) for the whole of Tuscany. Other literature thresholds were not considered since, at local or regional scale, thresholds perform reasonably well only in the area where they were developed and cannot be easily transferred to other areas (Crosta, 1989).

The first outcome of this comparison is that the national threshold proposed by Brunetti et al. (2010) is significantly lower than any other threshold; thus it is likely to commit
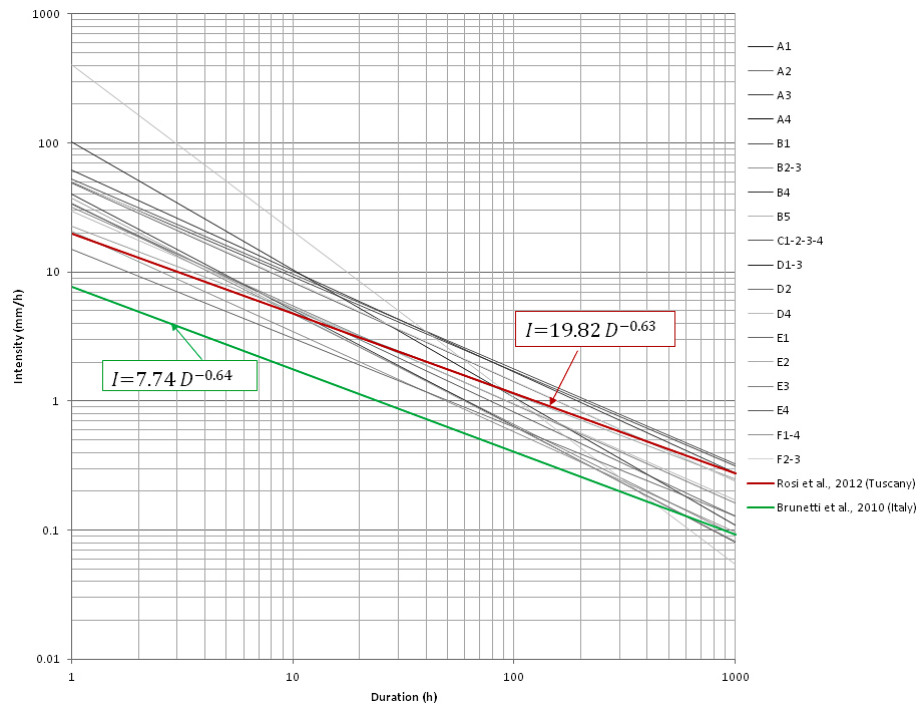
Figure 3. Comparison between the rainfall thresholds defined in this study for each Tuscany alert zone and two literature thresholds defined by Rosi et al. (2012) for the whole of Tuscany and by Brunetti et al. (2010) for Italy.

a relevant number of false-positive errors if applied to the Tuscany warning system. However, it should be stressed that the threshold proposed by Brunetti et al. (2010) is not expected to provide a balance between false positives and false negatives, because it was conceived with a different aim, i.e defining the minimum rainfall condition that can potentially lead to landsliding. In this light, the threshold performs very well, as it low-bounds the other thresholds used in this comparison.

The work of Rosi et al. (2012) and the one presented here involve the same study area (Tuscany) and have the same goal (a threshold balanced as much as possible in order to be used in a civil protection warning system); thus, a comparison between them is fully appropriate and allows for the two methodologies to be compared. In particular, we are interested in discovering whether the splitting of the region in a mosaic of local thresholds, defined using the automated routines of the MaCumBA software, could lead to a relevant improvement in the predicting capabilities of the regional warning system.

To this end, the four elements of the contingency table (Table 3) were combined to calculate some indexes that are traditionally used to quantitatively assess the performances of a model (Martelloni et al., 2012). The same statistics were calculated for the validation of the methodology proposed in this work and for the validation of a hypothetical application, in the whole Tuscany region, of the regional threshold proposed

by Rosi et al. (2012) and the national threshold (Brunetti et al., 2010) (Table 4).

The comparison between the validation statistics of Table 4 clearly shows that the effectiveness of thresholds can be increased when focusing the analysis on a smaller area. Consequently, a site-specific threshold is more precise than a general threshold applied to a single subdivision (shift from the national to the regional threshold) and a set of local thresholds is more effective than a single threshold. This proves the validity of our approach of devising a mosaic of thresholds instead of a single regional rainfall threshold. This approach is not new (see, for example, Martelloni et al., 2012) but is rarely used, as there is a preference in many studies to gather a large number of landslides for larger areas. This is partially conditioned by the necessity of increasing the landslide population to be used for statistical analyses: the larger the landslide population, the more robust the statistical analysis and the more reliable the threshold. The pros and cons of the splitting up of the study area into smaller subdivisions to be analysed independently should be carefully evaluated and counterbalanced.

On the one hand, the splitting-up of the territory brings the advantage of considering a uniform and homogeneous set of landslides, lithology and meteorological condition, thus strengthening the empirical correlation between I–D values and landslide triggering. Moreover, if a threshold pertains to a limited area, its operational use in civil protection procedures is advantaged, since a warning issued for a restricted

Table 2. Results of the validation procedure for each AZ. FP: false positives; FN: false negatives; TP: true positives; TN: true negatives; $Se = TP/(TP + FN)$; $Sp = TN/(TN + FP)$; $Lr = Se/(1-Sp)$; $Ef = (TP + TN)/(FP + FN + TP + TN)$; $Or = (TP + TN)/(FP + FN)$; $Pp = (TP)/(FP + TP)$; $Np = (TN)/(FN + TN)$.

Alert zone	FP	FN	TP	TN	Se	Sp	Lr	Ef	Or	Pp	Np
A1	1	0	5	165	1.00	0.99	166.00	0.99	170.00	0.83	1.00
A2	2	0	21	115	1.00	0.98	58.50	0.99	68.00	0.91	1.00
A3	1	3	18	184	0.86	0.99	158.57	0.98	50.50	0.95	0.98
A4	1	1	4	91	0.80	0.99	73.60	0.98	47.50	0.80	0.99
B1	1	0	5	114	1.00	0.99	115.00	0.99	119.00	0.83	1.00
B2,3	0	2	7	95	0.78	1.00	777.78	0.98	51.00	1.00	0.98
B4	0	1	7	60	0.88	1.00	875.00	0.99	67.00	1.00	0.98
B5	0	1	5	171	0.83	1.00	833.33	0.99	176.00	1.00	0.99
C1,2,3,4	0	1	8	134	0.89	1.00	888.89	0.99	142.00	1.00	0.99
D1,3	6	0	5	108	1.00	0.95	19.00	0.95	18.83	0.45	1.00
D2	0	1	3	109	0.75	1.00	750.00	0.99	112.00	1.00	0.99
D4	0	3	6	97	0.67	1.00	666.67	0.97	34.33	1.00	0.97
E1	7	0	2	106	1.00	0.94	16.14	0.94	15.43	0.22	1.00
E2	0	0	2	134	1.00	1.00	1000.00	1.00	#DIV/0!	1.00	1.00
E3	2	0	4	238	1.00	0.99	120.00	0.99	121.00	0.67	1.00
E4	3	0	4	76	1.00	0.96	26.33	0.96	26.67	0.57	1.00
F1,4	5	1	7	120	0.88	0.96	21.88	0.95	21.17	0.58	0.99
F2,3	1	0	5	127	1.00	0.99	128.00	0.99	132.00	0.83	1.00

Table 3. Contingency table summarizing the validation procedure at regional level; clockwise from the upper left corner, the numbers represent true positives, false alarms, true negatives and missed alarms.

		Observed truth	
		Landslide	No landslide
Prediction	Landslide	118	30
	No landslide	14	2244

area can be managed more easily than an alarm issued for a whole region involving dozens of cities and millions of inhabitants.

On the other hand, the splitting-up cannot be pushed too much further as a statistically significant number of landslides is needed to obtain reliable thresholds. It is not easy to establish the minimum number of landslides needed: in this work, the smallest data sets are used in the AZs D2 (12 landslides) and E1 and E2 (8 rainfall events) (Table 1). Indeed, case studies have been reported in the international literature in which significant thresholds were defined even with smaller data sets (Chen and Wang, 2014). An unexpected outcome of this research is that, although it is evident that a large landslide sample would contribute to the finding of a stronger statistical and empirical relationship between rainfall and landslides, in our case of study the dimension of the data set and the quality of the results are not closely related. It can be verified that none of the validation statistics (as reported in Table 2) can be put in close relation with the number of landslides or with the number of events (Table 1): the relationship between number of rainfall events and positive predictive power produced the highest coefficient of determination of linear regression ($R^2 = 0.1$), which, however, is rather low and supports the conclusion that acceptable results can also be obtained with limited data sets.

4.2 Relationships between thresholds and physical variables

Recently, various studies have compared rainfall thresholds with either results of physical modelling (Alvioli et al., 2014) or with geospatial analyses of the environmental variables (Rosi et al., 2012; Lagomarsino et al., 2013) with the aim of finding a physical background in the empirical intensity–duration relationship.

The definition of 18 thresholds, obtained with the same methodology in a restricted area, is a unique opportunity to make significant comparisons and to investigate what lies beyond the empirical relationship between cause (rainfall) and effect (landslide) and how different physical settings can influence threshold equations. Since the spatial distribution of the landslides in the study area is not homogeneous and depends on the physical setting (geology, geomorphology, rainfall regime, human influence), and since the main parameters of the thresholds exhibit a relevant degree of variability (Table 1), it is worth investigating whether the physical features of the various AZs can be put in relation with the parameters of their thresholds.

For this analysis, each threshold was characterized by means of the following parameters: α, β, the area under the

Table 4. Validation statistics of the mosaic of thresholds defined in this work, compared with literature thresholds proposed by Rosi et al. (2012) and Brunetti et al. (2010); as explained in the text, TP stands for true positives, TN for true negatives, FP for false positives, and FN for false negatives.

		This work	Regional threshold	National threshold
Sensitivity	TP/(TP+FN)	0.894	0.896	0.958
Specificity	TN(FP+TN)	0.987	0.732	0.692
Positive predictive power	TP/(TP+FP)	0.797	0.448	0.430
Negative predictive power	TN/(TN+FN)	0.994	0.967	0.986
Efficiency	(TP+TN)/(TP+TN+FP+FN)	0.982	0.764	0.744
Likelihood ratio	Sensitivity/(1–specificity)	67.761	3.347	3.111

threshold (AUT), and the no rain gap (NRG). The first two parameters are directly derived by the threshold equation and describe the power law relationship between intensity and duration. In a log–log plot, α represents the intercept in the y axis and defines how "high" a threshold is at low durations, while β represents the steepness of the threshold; therefore with low β values, even a threshold with high α values can become "low" for high durations. AUT defines the area under the threshold and thus quantifies how "high" or "low" a threshold is with respect to both α and β. NRG represents the minimum time gap with the absence of rainfall required to consider two rainfall events as separate, and it is a very important parameter for both the threshold analysis and its application to operational warning systems.

First, we checked the degree of correlation between the threshold parameters and found that α and β are quite correlated with each other ($R^2 = 0.71$), while NRG does not result in correlation with any of the other parameters (α, β and AUT). This last outcome can be interpreted as an indicator of the robustness of the proposed methodology: the optimal no rain gap value cannot be subjectively established in advance, it is very site-specific, and a trial and error procedure is needed to define an efficient value for the use in civil protection warning systems. The correlation between α and β means that, in general, the higher the intercept of a threshold, the higher its steepness. We can therefore infer that if an AZ has "high" α and β parameters, it is not likely to be subjected to landslides triggered by short and intense rainstorm, while prolonged rainfall events may exceed the threshold even with relatively low values of average intensity. Conversely, if α and β are relatively low, short rainstorms can trigger landslides even at relatively low intensities, while prolonged rainfall events need to reach relevant rainfall amounts before triggering landslides. This can be put in relation with the geomechanical and hydrological properties of the terrains and rock characterizing each AZ and thus suggests that the approach of sectioning the study area into independent AZs helps in finding a stronger correlation between rainfall and landslides: even if many different lithologies outcrop in Tuscany, only a limited number are present in each AZ, and thus

the response in the territory to the rainfall triggers is more homogeneous.

As a second step, we investigated the correlation between the above-mentioned threshold parameters (α, β, AUT and NRG) and the main characteristics of the physical setting of each AZ. We analysed the mean annual precipitation to account for the main triggering factor of landslides and to verify the observations of Govi and Sorzana (1980), according to which the amount of rainfall needed to trigger landslides rises with the mean annual precipitation. Slope gradient and lithology were considered in order to account for the landslide susceptibility of each area: according to recent studies on the landslide susceptibility of Tuscany (Catani et al., 2013), slope gradient and lithology are the most important predisposing factors.

Table 5 shows the degree of correlation (expressed in terms of R^2) between rainfall parameters and some basic statistics of the numerical variables that were used to characterize the physical setting (mean annual precipitation and slope gradient). Table 5 clearly shows that no R^2 reaches values higher than 0.2; therefore no significant correlation was found.

This outcome is not completely unexpected: a simple empirical correlation between cause and effect can be strengthened by the AZ subdivision, as it reduces the variability of the physical setting, but not to the point of making it possible to relate the characteristics of the rainfall threshold to just a couple of predisposing or triggering factors. This is confirmed by recent landslide susceptibility studies in the same study area: Catani et al. (2013) demonstrated that optimal susceptibility assessments can take up to 21 different parameters into account.

The influence of lithology on the threshold parameters was investigated through comparison of the prevailing lithology of each AZ with the parameters of the corresponding threshold. The 12 AZs characterized by layered rocks (e.g. flysch) exhibit a marked variability of the values of the main threshold parameters: α values range from 15.0 to 405.9, while β values range from −0.651 to −1.29. In the six AZs where terrains or soft rocks are the prevailing lithology, the same parameters have a smaller variability: α ranges from 29.6 to 50.7, while β ranges from −0.900 to −0.856. The prevailing

Table 5. R^2 values expressing the correlation between rainfall threshold parameters (α, β, combinations of α and β, area under threshold (AUT), and no rain gap (NRG)) and main numerical variables that characterize the physical setting.

		α	β	α/β	$\alpha \cdot \beta$	AUT	NRG
Mean annual precipitation	Max (a)	0.097	0.021	0.109	0.091	0.098	0.031
	Mean (b)	0.082	0.019	0.095	0.075	0.083	0.036
	SD (c)	0.067	0.008	0.062	0.071	0	0
	Min	0.140	0.074	0.162	0.127	0.141	0.084
Slope gradient	Mean (d)	0.154	0.056	0.158	0.151	0.15	0.010
	Max (e)	0.052	0.004	0.073	0.042	0.053	0.039
	SD (f)	0.020	0.007	0.022	0.019	0	0.024
Rainfall and morphology combinations	a·d	0.167	0.056	0.099	0.071	0.081	0.031
	a·e	0.095	0.021	0.112	0.087	0.096	0.030
	b·d	0.139	0.043	0.154	0.131	0.140	0.024
	b·e	0.080	0.017	0.181	0.160	0.168	0.025

Table 6. Variation of no rain gap values in Tuscany alert zones in relation to the prevailing lithology.

Prevailing lithology	No rain gap (hours)		
	Mean	Minimum	Maximum
Intrusive rocks	10	10	10
Gneiss	12	12	12
Effusive rocks	12	12	12
Terrains of mixed typology (cohesive and granular)	20.3	18	24
Flyschs	21.9	12	24
Marls	24	24	24
Granular terrains	30.4	24	36

more frequently found. This outcome, similar to other recent studies (Alvioli et al., 2014), proves that the empirical relationship between rainfall and landslides implicitly takes the physical background of the problem into account.

The use of a peculiar no rain gap value for each of the AZs allows the regional threshold mosaic to better account for both precipitation typologies encountered in Tuscany (localized convective thunderstorms and wider rainfall events): the two kinds of rainstorms are characterized by different durations, but this is indirectly taken into account by the model since, during the calibration procedure, the number of landslides triggered by one kind of thunderstorm or another in each AZ influences the determination of the optimal no rain gap value.

lithology of each AZ seems to influence the no rain gap of its rainfall threshold as well: on average, the more permeable the lithology, the higher the no rain gap value (Table 6).

A possible interpretation of this outcome is that the most permeable lithologies (granular terrains characterized by conglomerates and sands) are mainly affected by deep-seated landslides, which in turn are usually triggered by longer rainfalls, even without particularly extreme intensities. With a high no rain gap (up to 36 h), the automated algorithm used to identify triggering rainfalls is helped to focus on events with medium intensities averaged over long durations. Conversely, relatively impermeable bedrocks (e.g. tuffs, gneisses and intrusive rocks) exhibit a marked contrast of hydraulic properties with the overlying terrain: this condition predisposes to shallow landslides and debris flows, which are typically triggered by short and intense rainfalls. A short no rain gap (10 or 12 h) therefore helps the algorithm to prevalently recognize rainfall events characterized by short peaks with extreme intensity values. In those AZs where intermediate situations are present (e.g. terrains of mixed typology and flysches), intermediate values of no rain gap (e.g. 24 h) are

5 Conclusions

In this work, we propose an original approach to set up a mosaic of 18 local rainfall thresholds, in place of a single regional threshold, to be used in civil protection warning systems for the occurrence of landslides at regional scale (i.e. tens of thousands of kilometres).

The proposed approach is based on the use of a software program named MaCumBA (explained and discussed in detail in Segoni et al., 2014b), which allows for statistical intensity–duration rainfall thresholds to be identified by means of an automated and standardized analysis of rainfall data. The automation and standardization of the analysis brings several advantages that in turn have a positive impact on the applicability of the thresholds to operational warning systems.

The possibility of defining a threshold in very short times compared to traditional analyses allowed us to subdivide the study area into several AZs to be analysed independently, with the aim of setting up a specific threshold for each of them. Even if pertaining to the same region, the AZs can have

very different threshold equations; in particular, the highest variability was observed where layered rocks (i.e. flysch) are the main lithology. The subdivision into small AZs fosters the definition of robust rainfall thresholds as it circumscribes the statistical analysis to a limited and homogeneous area, thus allowing for a strong empirical relationship between cause (rainfall) and effect (landslides). Nonetheless, from a physical point of view, this linkage still remains very complex as it depends on many interplaying factors, and every attempt to relate the threshold parameters to the main numerical variables characterizing the physical setting has failed. However, results suggest that the prevailing lithology of each AZ influence the no rain gap (lapse of time without rainfall needed to consider two rainfall events as separate) of the corresponding threshold: in general, the more permeable the terrains/rocks of the AZ, the higher the no rain gap of the threshold. This outcome provides a physical background to empirical rainfall thresholds and brings us to two conclusions: on the one hand, it demonstrates the necessity of devising warning systems based on a mosaic of thresholds rather than on a single regional threshold, as the optimal criterion to be used by the warning system to analyse rainfall data and identify critical rainstorms may differ from one area to another depending of the encountered physical features; on the other hand, it stresses the necessity of using the same criterion for the rainfall analysis during both the research stage of rainfall definition and during the operational phase when the warning system performs automated computations in near-real time.

However, we come to the conclusion that the subdivision into AZs cannot be pushed too much further as it is limited by the necessity of having a statistically significant landslide sample in each AZ. Our methodology provided satisfactory results with data sets of minimum 12 and maximum 719 landslides, and we observed that the dimension of the data set was not directly related to the quality of the results. We therefore believe to have found a robust methodology and an effective compromise between AZ dimension and robustness of the landslide sample, counterbalancing the pros and cons of having small or large AZs.

Another important outcome of this work is the necessity for thresholds aimed at being employed in civil protection warning systems to be analytically validated. The proposed mosaic of thresholds was validated with an independent data set: all the rainfall events recorded from 2008 to 2009 were analysed and compared with the corresponding landslide data set. In this way, we were able to count correct predictions and errors of commission (false alarms) and omission (missed alarms); subsequently we calculated some quantitative indexes commonly used to express the effectiveness of models. This procedure allowed for the conclusion that our methodology had obtained an acceptable balance between missed alarms and false alarms that encouraged the implementation of the mosaic of thresholds in a regional civil protection warning system. Furthermore, the validation procedure was repeated for some literature thresholds, and the

quantitative comparison of the results demonstrated that the performance of a warning system can be enhanced if a specific threshold is defined for a given region rather than applying a general threshold. Moreover, this enhancement can be increased if a mosaic of site-specific thresholds is used instead of a single regional threshold.

References

Agostini, A., Tofani, V., Nolesini, T., Gigli, G., Tanteri, L., Rosi, A., Cardellini, S., and Casagli, N.: A new appraisal of the Ancona landslide based on geotechnical investigations and stability modelling, Q. J. Eng. Geol. Hydroge., 47, 29–43, doi:10.1144/qjegh2013-028, 2013.

Ahmad, R.: Developing early warning systems in Jamaica: rainfall thresholds for hydrological hazards, National Disaster Management Conf., Ocho Rios, St.Ann, Jamaica, 9–10 September 2003, available at: http://www.mona.uwi.edu/uds/rainhazards_files/frame.htm, 2003.

Aleotti, P.: A warning system for rainfall-induced shallow failures, Eng. Geol., 73, 247–265, 2004.

Alvioli, M., Guzzetti, F., and Rossi, M.: Scaling properties of rainfall induced landslides predicted by a physically based model, Geomorphology, 213, 38–87, doi:10.1016/j.geomorph.2013.12.039, 2014.

Baum, R. L., Godt, J. W., and Savage, W. Z.: Estimating the timing and location of shallow rainfall-induced landslides using a model for transient unsaturated infiltration, J. Geophys. Res., 115, F03013, doi:10.1029/2009JF001321, 2010.

Bovolo, C. I. and Bathurst, J. C.: Modelling catchment-scale shallow landslide occurrence and sediment yield as a function of rainfall return period, Hydrol. Process., 26, 579–596, 2012.

Brunetti, M. T., Peruccacci, S., Rossi, M., Luciani, S., Valigi, D., and Guzzetti, F.: Rainfall thresholds for the possible occurrence of landslides in Italy, Nat. Hazards Earth Syst. Sci., 10, 447–458, doi:10.5194/nhess-10-447-2010, 2010.

Brunsden, D.: The application of system theory to the study of mass movement, Geol. App. Idrogeoeol. 8, 185–207, 1973.

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

Cannon, S. H., Gartner, J. E., Wilson, R., Bowers, J., and Laber, J.: Storm rainfall conditions for floods and debris flows from recently burned areas in southwestern Colorado and southern California, Geomorphology, 96, 250–269, 2008.

Cannon, S., Boldt, E., Laber, J., Kean, J., and Staley, D.: Rainfall intensity-duration thresholds for postfire debris-flow emergency-response planning, Nat. Hazards, 59, 209–236, 2011.

Cardinali, M., Galli, M., Guzzetti, F., Ardizzone, F., Reichenbach, P., and Bartoccini, P.: Rainfall induced landslides in December 2004 in south-western Umbria, central Italy: types, extent, damage and risk assessment, Nat. Hazards Earth Syst. Sci., 6, 237–260, doi:10.5194/nhess-6-237-2006, 2006.

Catani, F., Lagomarsino, D., Segoni, S., and Tofani, V.: Landslide susceptibility estimation by random forests technique: sensitivity and scaling issues, Nat. Hazards Earth Syst. Sci., 13, 2815–2831, doi:10.5194/nhess-13-2815-2013, 2013.

Chen, H. X. and Wang, J. D.: Regression analyses for the minimum intensity-duration conditions of continuous rainfall for mudflows triggering in Yan'an, northern Shaanxi (China), Bull. Eng. Geol. Environ., doi:10.1007/s10064-013-0567-3, online first, 2014.

Chleborad, A. F.: Preliminary evaluation of a precipitation threshold for anticipating the occurrence of landslides in the Seattle, Washington Area, US Geological Survey Open-File Report 03, 463 pp., 2003.

Crisci, A., Gozzini, B., Meneguzzo, F., Pagliara, S., and Maracchi, G.: Extreme rainfall in a changing climate: regional analysis and hydrological implications in Tuscany, Hydrol. Process., 16, 1261–1274, 2002.

Crosta, G. B.: A study of slope movements caused by heavy rainfall in Valtellina (Italy, July 1987), in: Proc. 6th Int. Conf. and Field Workshop on Landslides ALPS 90 (Cancelli A, ed), vol. 79b., Milano: Ricerca Scientifica ed Educazione Permanente, 247–258, 1989.

Crosta, G. B. and Frattini, P.: Rainfall thresholds for triggering soil slips and debris flow, in: Proc. 2nd EGS Plinius Conf. on Mediterranean Storms, edited by: Mugnai, A., Guzzetti, F., Roth, G., Siena, 463–487, 2001.

Farahmand, A. and AghaKouchak, A.: A satellite-based global landslide model, Nat. Hazards Earth Syst. Sci., 13, 1259–1267, doi:10.5194/nhess-13-1259-2013, 2013.

Fatichi, S. and Caporali, E.: A comprehensive analysis of changes in precipitation regime in Tuscany, Int. J. Climatol., 29, 1883–1893, 2009.

Giannecchini, R., Galanti, Y., and D'Amato Avanzi, G.: Critical rainfall thresholds for triggering shallow landslides in the Serchio River Valley (Tuscany, Italy), Nat. Hazards Earth Syst. Sci., 12, 829–842, doi:10.5194/nhess-12-829-2012, 2012.

Govi, M. and Sorzana, P. F.: Landslide susceptibility as function of critical rainfall amount in Piedmont basin (Northwestern Italy), Studia Geomorphologica, Carpatho-Balcanica, 14, 43–60, 1980.

Guzzetti, F., Peruccaci, S., Rossi, M., and Stark, C. P.: Rainfall thresholds for the initiation of landslides in central and southern Europe, Meteorol. Atmos. Phys., 98, 239–267, 2007.

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

Hahn, G. J. and Meeker, W. Q.: Statistical Intervals: a guide for practitioners, Wiley-Interscience, John Wiley and Sons Inc., N.Y, ISBN:0-471-88769-2, 416 pp., 1991.

Hong, Y., Hiura, H., Shino, K., Sassa, K., Suemine, A., Fukuoka, H., and Wang, G.: The influence of intense rainfall on the activity of large-scale crystalline schist landslides in Shikoku Island, Japan, Landslides, 2, 97–105, 2005.

Hong, Y., Adler, R., and Huffman, G.: Evaluation of the potential of NASA multisatellite precipitation analysis in global landslide hazard assessment, Geophys. Res. Lett., 33, L22402, doi:10.1029/2006GL028010, 2006.

Hong, Y., Adler, R., and Huffman, G.: Satellite remote sensing for global landslide monitoring, Eos, 88, 357–358, 2007.

Jakob, M. and Weatherly, H.: A hydroclimatic threshold for landslide initiation on the North Shore Mountains of Vancouver, British Columbia, Geomorphology, 54, 137–156, 2003.

Lagomarsino, D., Segoni, S., Fanti, R., and Catani, F.: Updating and tuning a regional-scale landslide early warning system, Landslides, 10, 91–97, 2013.

Larsen, M. C. and Simon, A.: A rainfall intensity-duration threshold for landslides in a humid-tropical environment, Puerto Rico, Geogr. Ann., 75, 13–23, 1993.

Martelloni, G., Segoni, S., Fanti, R., and Catani, F.: Rainfall thresholds for the forecasting of landslide occurrence at regional scale, Landslides, 9, 485–495, 2012.

Mercogliano, P., Segoni, S., Rossi, G., Sikorsky, B., Tofani, V., Schiano, P., Catani, F., and Casagli, N.: Brief communication "A prototype forecasting chain for rainfall induced shallow landslides", Nat. Hazards Earth Syst. Sci., 13, 771–777, doi:10.5194/nhess-13-771-2013, 2013.

Petley, D.: Global patterns of loss of life from landslides, Geology, 40, 927–930, 2012.

Rapetti, F. and Vittorini, S.: Rainfall in Tuscany: observation about extreme events, Riv. Geogr. Ital., 101, 47–76, 1994 (in Italian).

Rosi, A., Segoni, S., Catani, F., and Casagli, N.: Statistical and environmental analyses for the definition of a regional rainfall thresholds system for landslide triggering in Tuscany (Italy), J. Geogr. Sci., 22, 617–629, 2012.

Rosi, A., Vannocci, P., Tofani, V., Gigli, G., and Casagli, N.: Landslide Characterization Using Satellite Interferometry (PSI), Geotechnical Investigations and Numerical Modelling: The Case Study of Ricasoli Village (Italy), International Journal of Geosciences, 4, 904–918, 2013.

Rossi, G., Catani, F., Leoni, L., Segoni, S., and Tofani, V.: HIRESSS: a physically based slope stability simulator for HPC applications, Nat. Hazards Earth Syst. Sci., 13, 151–166, doi:10.5194/nhess-13-151-2013, 2013.

Segoni, S., Lagomarsino, D., Fanti, R., Moretti, S., and Casagli, N.: Integration of rainfall thresholds and susceptibility maps in the Emilia Romagna (Italy) regional-scale landslide warning system, Landslides, doi:10.1007/s10346-014-0502-0, 2014a.

Segoni, S., Rossi, G., Rosi, A., and Catani, F.: Landslides triggered by rainfall: a semi-automated procedure to define consistent intensity-duration thresholds, Comput. Geosci., 63, 123–131, 2014b.

Staley, D. M., Kean, J. W., Cannon, S. H., Schmidt, K. M., and Laber, J. L.: Objective definition of rainfall intensity-duration thresholds for the initiation of post-fire debris flows in southern California, Landslides, 10, 547–562, 2013.

Tiranti, D. and Rabuffetti, D.: Estimation of rainfall thresholds triggering shallow landslides for an operational warning system implementation, Landslides, 7, 471–481, 2010.

Wilson, R. C.: Climatic variations in rainfall thresholds for debris-flows activity, in: Proc. 1st Plinius Conf. on Mediterranean Storms, edited by: Claps, P. and Siccardi, F., Maratea, 415–424, 2000.

Zimmermann, M., Mani, P., Gamma, P., Gsteiger, P., Heiniger, O., and Hunziker, G.: Murganggefahr und Klimaanderung – ein GIS-basierter Ansatz, in: Schlussbericht des Nationalen Forschungsprogrammes, NFP 31, Zurich: vdf Hochschulverlag AG, 161 pp., 1997.