Regional rainfall thresholds for landslide occurrence using a centenary database

Teresa Vaz, José L. Zêzere, Susana Pereira, Sérgio Cruz Oliveira, Ricardo A. C. Garcia and Ivânia Quaresma

Centre for Geographical Studies, Institute of Geography and Spatial Planning, Universidade de Lisboa, Portugal. tvaz@campus.ul.pt

Research Group on Environmental Hazar and Risk Assessment and Management
Centre for Geographical Studies
Institute of Geography and Spatial Planning, Universidade de Lisboa
www.riskam.ul.pt

1. Introduction

This work proposes a comprehensive methodology for rainfall triggering thresholds, using a centenary landslide database associated with a single centenary daily rainfall dataset. The specific objectives of this study are:

i) to identify the critical rainfall combination quantity-duration to landslides occurrence;
ii) to define linear and potential regression thresholds and lower limit and upper limit rainfall thresholds;
iii) to assess the thresholds performance using receiver operating characteristic (ROC) metrics;
iv) to estimate the probability of several rainfall thresholds and the probability of landslide event above each rainfall threshold;
v) to identify the area where the thresholds can be applied.

2. Methodology

The methodology applied in this study follows the steps below:

- **Identification of landslide events** from a landslide database carried out exploring newspapers published between 1865 and 2010. A landslide event was defined as a landslide/ set of landslides occurred at a unique date.
- **Selection of a rainfall gauge** considering its climatic representativeness, data resolution and available time series. Buffers (5, 10, 15, 20, 30, 40, 50, 60 km) from the selected rain gauge were applied to evaluate the area where the gauge is representative.
- **Reconstruction of accumulated rainfall** for different durations (1 to 90 consecutive days) and computation of the extreme values statistics following Gumbel (2012).
- **Identification of critical rainfall combinations for each landslide event** by selecting the pair (quantity-duration) with the highest return period (Zêzere et al., 2005, 2015). Only critical rainfall combinations with return period above 3 years were considered rainfall-triggered landslides.
- **Identification of critical distance for rain gauge** by assuming that the ratio of *non-rainfall triggered events/ total landslide events* should decrease as the distance increase from the gauge. If this relation is not identified the station is not representative.
- **Identification of "no landslide rainfall data" (*grey dots*)** by selecting the maximum values for each duration (1 to 90 consecutive days) for the years without rainfall-triggered landslides records.
- **Identification of rainfall triggering thresholds**. Critical rainfall thresholds were established by linear and potential regression. The lower limit and the upper limit thresholds were also computed.
- **Validation of thresholds based on ROC metrics** (Fawcett, 2006; Staley et al., 2013).
- **Identification of probability of rainfall thresholds and probability of landslide event above each rainfall threshold** based on the Return Period and the Positive Predictive Rate (PP_R).
- **Evaluation of the regional performance of rainfall thresholds** by calculating, to the lower rainfall threshold and to each distance buffer (15 to 60 km), the ratio *False Negative / (True positive + False Negative)*, not excluding the events with less than 3 years of return period. When this relation increase, the rain gauge is assumed no longer representative.

3. Results

The Lisboa-Geofísico was the rain gauge selected (daily rainfall records since 1864), located inside the Lisbon urban area.

10 km from the rain gauge was the critical distance, for which the lowest ratio non-rainfall triggered events/ total landslide events was obtained

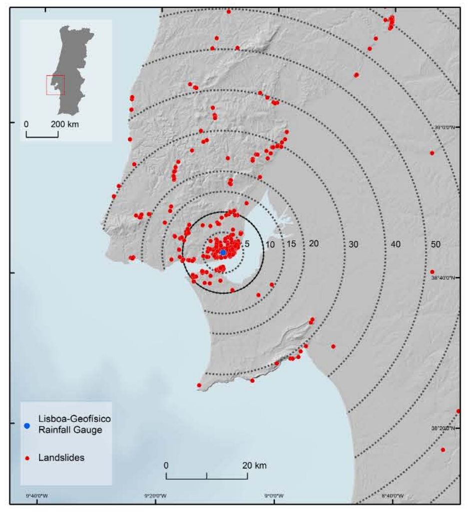

Fig. 1 Landslide cases in the Lisbon region (Note: only landslides with high spatial accuracy are represented).

3.1 Critical rainfall combination quantity-duration

Critical rainfall quantity-duration was obtained for each landslide event by selecting the pair (quantity-duration) with the highest return period and above 3 years.

82 landslides events were identified corresponding to 258 landslides cases in the Lisbon area in the period from 1865 to 2010 (Fig. 2).

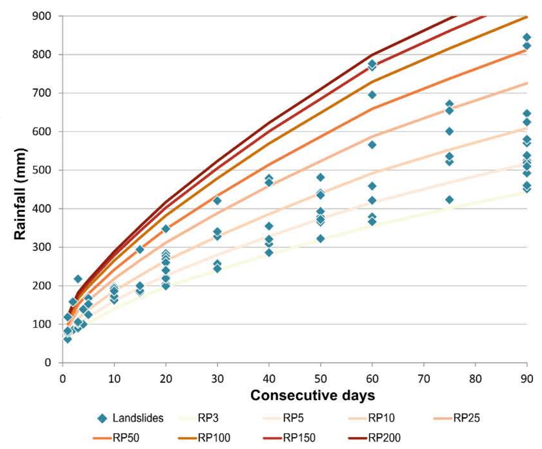

Fig. 2 Critical rainfall combinations quantity–duration that resulted in landslide events in the Lisbon region, and the return period (RP) for 3, 5, 10, 25, 50, 100, 150 and 200 years.

3.2 Rainfall Thresholds

The linear regression follows the equation: y = 5.5D + 124.6, where y is the critical rainfall in mm and D is the duration in days (Fig. 3).

Maximum values, for each duration (1 to 90 days) and for each climatological year (1865 to 2010) without rainfall-triggering landslide reported, were also computed and considered as 'negatives' to validate the rainfall thresholds.

The lower threshold is the limit below which there are not events, but only true negative. It was adjusted by the linear function: y=4.4D+56.5.

The maximum threshold is the limit above which there are only true positives and follows the function: y = 7.3D + 235.8.

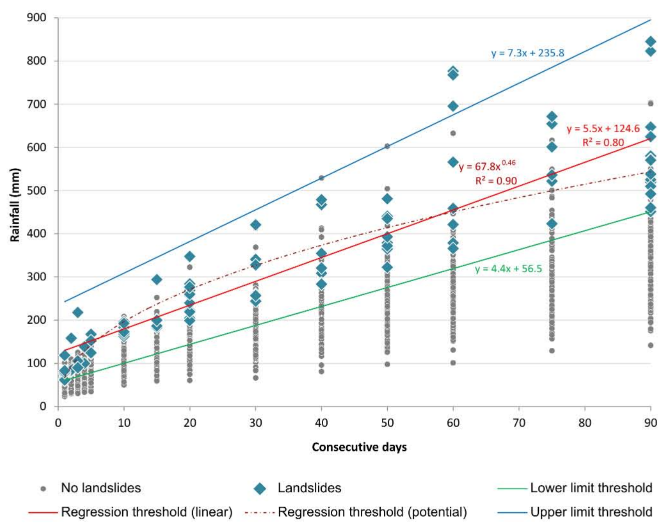

Fig. 3 Rainfall quantity–duration thresholds for the landslide events in the Lisbon region from 1865 to 2010.

3.3 Validation of thresholds based on ROC metrics

Threshold performance was assessed by ROC metrics (Table 1).

Regression threshold (linear): good results in FP_R (0.02) and in TS (0.27).

Lower limit threshold: conservative model; good results in landslide events prediction (TP_R=1); high number of "no events" above the threshold (FP_R=0.37).

Upper limit threshold: sensitive model; only surpassed by true positives occurrences, so FP_R and FA_R have the best result (0); TP_R is very low (0.03), considering the high number of FN.

Table 1 ROC metrics associated to the established rainfall thresholds in the Lisbon region.

	Regression threshold (linear)	Regression threshold (potential)	Lower limit threshold	Upper limit threshold
True Positive (TP)	35	37	96	3
False Negative (FN)	61	59	0	93
False Positive (FP)	33	46	527	0
True Negative (TN)	1395	1382	901	1428
True positive rate (TP_R=TP/(TP+FN)	0.36	0.39	1	0.03
False Positive rate (FP_R=FP/(FP+TN)	0.02	0.03	0.37	0
False alarm rate (FA_R=FP/(TP+FP)	0.49	0.55	0.85	0
Threat score (TS= TP/(TP+FN+FP)	0.27	0.26	0.15	0.03
Positive predictive rate (PP_R=TP/(TP+FP)	0.51	0.45	0.15	1.00

3.4 Probability of rainfall thresholds and probability of landslide event above each rainfall threshold

Positive predictive rate (PP_R) gives the probability of a rainfall combination quantity-duration result in a landslide event when the threshold is exceeded.

Above the lower limit threshold the probability is 0.15 (Table 1) and above the upper limit threshold the probability is 1 (landslide occurrence is certain).

In the plot area between the lower limit and upper limit thresholds the 20%, 30%, 40%, 50% and 60% thresholds probability of landslide occurrence were established based on the PP_R (Fig. 4). Further probabilities were not possible to compute due to lack of data.

Table 2 ROC metrics associated to probability of landslide events above the rainfall thresholds.

	PP_R 20%	PP_R 30%	PP_R 40%	PP_R 50%	PP_R 60%
True Positive (TP)	91	76	66	50	15
False Negative (FN)	5	20	30	46	81
False Positive (FP)	364	177	99	50	10
True Negative (TN)	1064	1251	1329	1378	1418
True positive rate (TP_R=TP/(TP+FN)	0.95	0.79	0.69	0.52	0.16
False Positive rate (FP_R=FP/(FP+TN)	0.25	0.12	0.07	0.04	0.01
False alarm rate (FA_R=FP/(TP+FP)	0.80	0.70	0.60	0.50	0.40
Threat score (TS= TP/(TP+FN+FP)	0.20	0.28	0.34	0.34	0.14
Positive predictive rate (PP_R=TP/(TP+FP)	0.20	0.30	0.40	0.50	0.60

The return periods of rainfall thresholds were computed for each duration and the results are shown in Fig. 5.

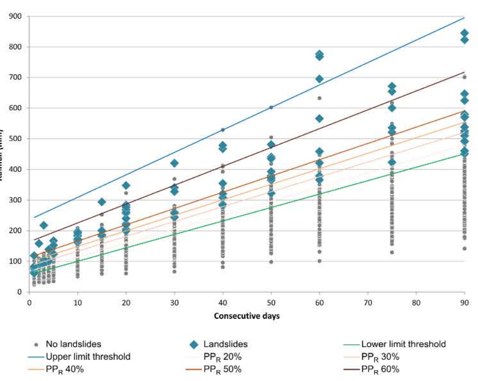

Fig. 4 Probability (20%, 30%, 40%, 50%, 60%) of landslide events above the rainfall thresholds.

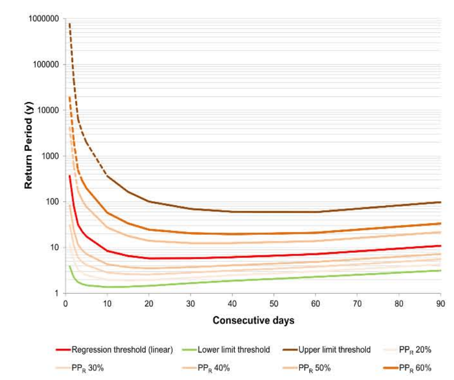

Fig. 5 Return period computed for the rainfall thresholds.

3.5 Regional performance of thresholds

We found that the ratio *False Negative / (True positive + False Negative)*, for the lower threshold, is stable up to the 50 km from the rain gauge.

Table 3 Ratio FN/(TP+FN) for the lower threshold for different buffer distances to the rain gauge.

Distance (km)	True Positive (TP)	False Negative (FN)	Ratio FN/(TP+FN)
[10-15[	20	4	0.2
[15-20[	7	2	0.2
[20-30[	26	11	0.3
[30-40[	23	7	0.2
[40-50[	8	2	0.2
[50-60[	5	5	0.5

4. Conclusions

- The critical rainfall combinations responsible for triggering the landslide events were identified by selecting the pairs (quantity-duration) with the highest return period.
- Four rainfall thresholds were firstly identified: linear regression threshold (y = 5.5D + 124.6); potential regression threshold (y = 67.8D$^{0.46}$); lower limit (y = 4.4D + 56.5) and the upper limit (y = 7.3D + 235.8).
- The thresholds performance was tested with ROC metrics with good results in Treat Score for rainfall thresholds obtained with linear and potential regression (0.27 and 0.26, respectively).
- The zone between the lower and upper thresholds (where landslide may occur) was analysed defining the probability of 20%, 30%, 40%, 50% and 60% for landslide events to occur above each rainfall threshold. These probability thresholds can be more informative to apply to early warning systems.
- The return period of rainfall thresholds is very high for the durations below 10 days. The probability of rainfall thresholds is typically highest for durations between 15 to 45 days.
- The thresholds identified for the Lisboa-Geofísico rainfall gauge can be applied with reasonable results for the area within 50 km distance.

References

Fawcett T (2006) An introduction to ROC analysis, Pattern Recognit. Lett., 27(8): 861–874.

Staley D M, Kean J W, Cannon S H, Schmidt K M, Laber JL (2013) Objective definition of rainfall intensity–duration thresholds for the initiation of post-fire debris flows in southern California. Landslides, 10(5): 547–562.

Gumbel EJ (2012) Statistics of Extremes, Dover Publications.

Zêzere JL, Trigo RM, Trigo IF (2005) Shallow and deep landslides induced by rainfall in the Lisbon region (Portugal): assessment of relationships with the North Atlantic Oscillation. Nat. Hazards Earth Syst. Sci., 5(3): 331–344.

Zêzere JL, Vaz T, Pereira S, Oliveira SC, Marques R, Garcia RAC (2015) Rainfall thresholds for landslide activity in Portugal: a state of the art. Environ. Earth Sci., 73(6): 2917–2936.

Acknowledgments

This work was financed by national funds through FCT – Portuguese Foundation for Science and Technology, I.P., under the framework of the project FORLAND – Hydro-geomorphologic risk in Portugal: driving forces and application for land use planning (PTDC/ATPGEO/1660/2014). Teresa Vaz is a PhD fellow funded by FCT [grant number SFRH/BD/74716/2010]. Sérgio Cruz Oliveira is a post-doc fellow of the FCT [grant number SFRH/BPD/85827/2012].