Landslides (2019) 16:1273-1283 DOI 10.1007/s10346-019-01169-9 Received: 10 August 2018 Accepted: 18 March 2019 Published online: 5 April 2019 © The Author(s) 2019

Nejc Bezak · Mateja Jemec Auflič · Matjaž Mikoš

Application of hydrological modelling for temporal prediction of rainfall-induced shallow landslides

Abstract In some rainfall-triggered landslides, intensity-duration thresholds can have limited prediction ability; therefore, investigation of alternative approaches that can be used for temporal prediction of rainfall-induced landslides is needed. This paper presents a methodology for predicting rainfall-induced shallow landslides based on a lumped conceptual hydrological model. The production storage level during the rainfall event and the rainfall sum during the event are used for landslide prediction. Based on these two hydro-meteorological variables a threshold is defined that could be used for rainfall-induced landslides prediction as part of an early warning system. The presented methodology is tested using the meso-scale Selška Sora River catchment case study in western Slovenia where 20 active landslides from the Slovenian National Landslide Database are used to calibrate and evaluate the methodology performance. The results are compared to three different (i.e. local, regional, and global) intensityduration thresholds. The results of the presented approach are superior in terms of several goodness-of-fit criteria compared to tested local and global ID thresholds. Because only daily rainfall, evapotranspiration, and discharge data are needed to calibrate the selected hydrological model and only daily rainfall and evapotranspiration to run the model, the presented approach could also be useful for data-scarce areas where detailed physically based landslide prediction models that require many data cannot be constructed. Moreover, we have also derived the probabilistic version of the proposed threshold for triggering of shallow landslides using copula functions.

Keywords Hydrological model · Lumped model · Shallow landslides \cdot Intensity-duration threshold \cdot Rainfall \cdot Catchment scale

Introduction

Rainfall-induced landslides and debris flows are one of the most frequent geo-hazards that can cause large economic damage and can even be responsible for human casualties (e.g. Haque et al. 2016; Mikoš et al. 2004; Petley 2012). Therefore, reliable early warning systems (EWS) are needed in order to issue warnings on time and consequently reduce the number of casualties and economic damage. Often, empirical rainfall thresholds (also known as intensity-duration (ID) curves or thresholds) are used as part of early warning systems for shallow landslides and debris flows (e.g. Huang et al. 2015; Liu et al. 2016; Mathew et al. 2014; Segoni et al. 2014). However, as recently pointed out by Bogaard and Greco (2018), ID thresholds have some limitations that should be taken into consideration and the use of these thresholds as part of the EWS should be taken with more caution. Some definitions of rainfall thresholds consider also antecedent rainfall (e.g. Aleotti 2004). However, this approach only considers rainfall in a selected time-period (e.g. 5 or 10 days) before the event without consideration of the complex hydrological processes (at least, e.g.

evapotranspiration) in the catchment area. Moreover, the majority of ID thresholds developed so far do not apply antecedent conditions as well as hillslope (hydrological) processes despite the fact that these are important for landslide initiation (Bogaard and Greco 2018). Moreover, Bogaard and Greco (2018) also argued that ID thresholds for short and long rainfall durations have limited physical meaning. The alternative for the prediction of landslide initiation is the use of spatially distributed physically based models that require large data input (e.g. Aristizabal et al. 2016; Bogaard and Greco 2018; Peres et al. 2018). Data collection and model calibration is often very complex, which consequently means that physically based models are not often incorporated in the operational EWS, especially at larger spatial scales. Thus, Bogaard and Greco (2018) concluded that more focus should be given on the development of conceptual models for regional landslide hazard assessment that would also take into consideration hydrological processes that occur during the rainfall events that trigger landslides. Some examples where more focus is on hydrological processes at larger spatial scales and the conceptual framework of the triggering mechanisms can be found in the literature. For example,, the Norwegian forecasting and warning service for rainfall- and snowmelt-induced landslides uses Hydrologiska Byråns Vattenbalansavdelning (HBV) (distributed 1 km² grid version of the conceptual HBV model is used) and S-flows models to calculate water and heat dynamics (Krøgli et al. 2018). Based on the models' results relative water supply and degree of soil water saturation are used in combination with pre-defined thresholds for landslide forecasting (Krøgli et al. 2018). However, definition of the thresholds lacks in objectivity (Krøgli et al. 2018). Moreover, Segoni et al. (2018) recently used average soil moisture to enhance the performance of the regional-scale landslide EWS where soil moisture estimates were calculated using the topographic kinematic approximation and integration (TOPKAPI) model (Ciarapica and Todini 2002). The incorporation of soil moisture in the EWS reduced the number of false and missed alarms (Segoni et al. 2018). However, a soil map and a land-use map are among other input data needed to derive the TOPKAPI parameters (Ciarapica and Todini 2002).

According to the reviewed literature and due to the need to replace empirical ID thresholds with a more hydrologically based approach (as suggested by Bogaard and Greco 2018), we propose an approach to rainfall-induced shallow landslides prediction using information obtained from a lumped conceptual hydrological model (Perrin et al. 2003; Pushpalatha et al. 2011; Valery et al. 2014a, b). Additionally, the daily time step was used to reduce the amount of required input data. Thus, the main aim of this study was to evaluate the performance of the proposed approach that applies daily rainfall data and modelled daily production store (i.e. reservoir) level for predicting rainfall-induced shallow landslides. The methodology is tested using the Slovenian National Landslide Database and the active landslide data provided for the MASPREM project (e.g. Jemec Auflič et al. 2016; Komac and Hribernik 2015; Rosi et al. 2016) on a meso-scale catchment located in the western mountainous part of Slovenia. The proposed methodology performance is compared to the ID thresholds that were developed for specific river basins in Slovenia (Rosi et al. 2016) and the global ID threshold suggested by Guzzetti et al. (2008).

Data and methods

Hydrological modelling

In this study, we tested two versions of a lumped conceptual hydrological model that has been developed by the Institut national de recherche en sciences et technologies pour l'environnement et l'agriculture (IRSTEA) hydrology group (e.g. Perrin et al. 2003; Pushpalatha et al. 2011; Valery et al. 2014a, b). We compared another simple model version named Génie Rural à 4 paramètres Journalier (GR4J) and another complex version named Cema Neige Génie Rural à 6 paramètres Journalier (CemaNeigeGR6J). Both hydrological model versions are implemented in R software airGR package (Coron et al. 2017a; Coron et al. 2017b). The GR4J model was proposed by Perrin et al. (2003) with an aim to obtain model robustness with a small number of model parameters (i.e. the model uses 4 parameters). Figure 1 shows schematic representation of the GR4J model structure where not all model steps are shown but only those relevant for this specific case study (i.e. routing steps are not shown). The GR4J model was developed based on the three-parameter Génie Rural à 3 paramètres Journalier (GR3J) version (Edijatno et al. 1999) and improved low-flow simulations (Perrin et al. 2003). The only required input data needed to model the discharge (Q) are precipitation (P) and potential evapotranspiration (Perrin et al. 2003). In this study, we used reference evapotranspiration data (PE) since reference evapotranspiration data calculated using the Penman-Monteith method and potential evapotranspiration calculated using the Oudin et al. (2005) method yielded similar behaviour. Based on these two input variables (i.e. P and PE), the model calculates net rainfall (Pn on Fig. 1). If net rainfall is not zero, a part of net rainfall is used to fill the production store (Ps on Fig. 1) (i.e. a conceptual underground reservoir) (Perrin et al. 2003). The production store is emptied either by the actual evaporation rate (Es on Fig. 1) in case that a net evapotranspiration capacity (En on Fig. 1) is not zero or by percolation (Perc on Fig. 1) from the reservoir (Perrin et al. 2003). Percolation and the difference between net rainfall and rainfall that is used to fill the production store (Pn-Ps on Fig. 1) is then used for further discharge calculations using several routing steps that are not the primary focus of this study where the routing reservoir is also used (Perrin et al. 2003). Additional description of routing can be found in Perrin et al. (2003). It should also be noted that Perc is always smaller than S and that S cannot be larger than X_1 (Perrin et al. 2003). Moreover, all the input data (PE, P, and also Q) influence the production storage level S on a given day, which means that the model structure shown in Fig. 1 takes into account several hydrological processes that occur in the catchment, while production storage can be regarded as catchment storage. As mentioned, the GR4J model uses 4 parameters, of which from this study's perspective the most important is the maximum production store capacity $(X_1 \text{ on Fig. 1})$, the other three parameters (X_2, X_3, X_4) are used for routing calculations (Perrin et al. 2003). A detailed description of the GR4J model is available in Perrin et al. (2003). Additionally, we also tested the CemaNeigeGR6J model that also includes the snow accounting routine (Valery et al. 2014a, b). This snow routine adds two more parameters to the GR6J model version (i.e. in total eight parameters are used by the CemaNeigeGR6J model since the GR6J model adds two more parameters to the GR4J model version and thus GR6J uses 6 parameters). Valéry et al. (2014a) provide detailed description of the snow routine used in the CemaNeigeGR6J model and Pushpalatha et al. (2011) provide a description of the Génie Rural à 6 paramètres Journalier (CR6J) model. The CemaNeigeGR6J model version additionally requires mean air temperature (T) data and the hypsometric curve of the catchment in order to be able to calculate Q (Coron et al. 2017a; Coron et al. 2017b; Valéry et al. 2014a). It should be noted that also the CemaNeigeGR6J model version uses the production store reservoir as the GR4J model version, which is shown in Fig. 1. Table 1 shows an overview of input data needed to run GR4J and CemaNeigeGR6J models.

Calibration of both model versions was performed using the methodology proposed by Michel (1991), which is implemented is R software airGR package (Coron et al. 2017a; Coron et al. 2017b). We selected the Nash-Sutcliffe (NS) coefficient (Nash and Sutcliffe 1970) as the efficiency criterion. Perrin et al. (2003) provide initial parameter values and their ranges needed for the model calibration. The NS criterion was also used for comparing the models' ability to predict Q.

Data

The methodology was tested on the meso-scale Selška Sora River catchment located in the mountainous western part of Slovenia (Fig. 2). The Selška Sora River is part of the Sava River catchment that drains into the Danube River. In general, the area belongs to the pre-alpine region, and it extends from 334 to 1676 m a.s.l. The study area is characterised by a diverse morphology and heterogenous geological settings, which make the area prone to landslides. Geologically, the study area is constituted mainly of Mesozoic carbonate rocks (limestone and dolomite), Mesozoic clastites

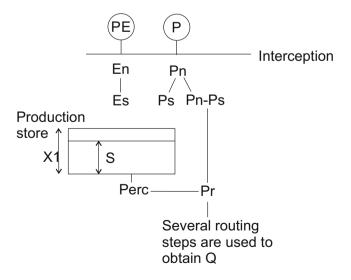


Fig. 1 Schematic representation of the conceptual hydrological model version GR4J up to the routing part of the model (adopted from Perrin et al. 2003)

Table 1 Overview of input data used in two different model versions (i.e. X indicates that these data are needed to run the model)

Variable	GR4J model	CemaNeigeGR6J model
P	Х	Χ
PE	Х	Х
T		Х
Hypsometric curve		Х

(shale, siltstone, marlstone, greywacke, sandstone, conglomerate, breccias, tuff), and Permian and Carboniferous clastites (claystone, sandstone, shale). The valleys are mainly filled by alluvial deposits, and the feet of the slopes are regularly covered by scree, both of Quaternary age (Grad and Ferjančič 1974). The soil depth varies from 1 to 3 m and, in some locations, more than 10 m. Landslides in the study area usually occur at the end of summer and in autumn, from September to November, due to particular weather conditions with heavy rain (Jemec and Komac 2013). The entire Selška Sora River catchment up to the confluence with the Poljanska Sora (P. Sora) River catchment was selected (Fig. 2). The basic characteristics of the selected area are shown in Table 2. This pre-alpine area with mountainous climate was selected since it is one of the areas with the highest landslide density in Slovenia (Zorn and Komac 2008).

Table 3 shows the basic characteristics of the data used in this study. The selected study period was from 2005 to 2014 and the daily time step was used. The data period from 2005 until 2006 was used as the warm-up period of the hydrological model. Daily Q, P, T and PE data from different stations located in the Selška Sora River catchment or close to it were used in this study, since there is no station that would have all the required data available. Therefore, we used the discharge data from the Železniki gauging station. The location of this station is also indicated in Fig. 2 and has a catchment area of 104 km². Precipitation data were used from the Davča rainfall station (station elevation is 987 m a.s.l.). Since there is no station available with air temperature and evapotranspiration data in the Selška Sora River catchment, we used the data from the closest stations with data available (e.g. Maček et al. 2018 provided a description of stations with reference evapotranspiration data). Air temperature data were obtained from the Bohinjska Češnjica meteorological station located less than 8 km from the Selška Sora River catchment boundary at 596 m a.s.l. and is not shown in Fig. 2. Moreover, evapotranspiration data were gathered from the Ljubljana meteorological station that is located less than 20 km from the Selška Sora River catchment boundary at 299 m a.s.l., and is also not shown in Fig. 2.

In order to test the performance of the ID thresholds, we also used high-frequency 5-min rainfall data measured at the location of the Davča rainfall station (shown in Fig. 2). It should be noted that pluviographic stations that are used to measure 5min rainfall data are not capable of measuring snow precipitation. Thus, snow events (i.e. three landslide events were included in the Slovenian national database) were excluded from the analysis in this study.

The landslides considered in this study were compiled from different sources (Jemec Auflič et al. 2016; Komac and Hribernik 2015; Rosi et al. 2016). It should be noted that the database

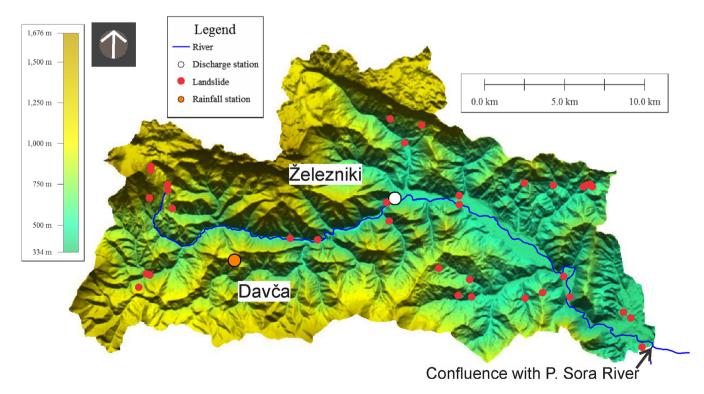


Fig. 2 The Selška Sora River catchment with the elevation map [m a.s.l.] and with indicated locations of landslides included in the Slovenian National Landslide Database, the rainfall gauging station and the discharge gauging station

Table 2 Basic characteristics of the Selška Sora River catchment

Catchment area	224.3 km ²
Minimum elevation [m a.s.l.]	334
Mean elevation [m a.s.l.]	826
Maximum elevation [m a.s.l.]	1676
Dominant land-use type	Forest (77%)
1 h, 6 h, 12 h and 24 h rainfall with 100-year return period according to the Davča rainfall station (intensity-duration-frequency (IDF) curve)	97 mm, 181 mm, 248 mm, and 274 mm

entries are not verified for this study. This could consequently mean that either the landslide location is not correctly defined or the triggering date is not valid. However, the assumption made in this study was that the landslide database entries were correct. Only shallow landslide events from the joint database were considered in this study, but not also large deep-seated landslides, debris flows, or rockfalls, which means that events can be classified as shallow landslides (Jemec Auflič et al. 2016). In the presented study, we focus on landslides that occurred in the Selška Sora River catchment up to the confluence with the P. Sora River catchment (Fig. 2). After brief pre-processing (e.g. merging duplicate events, removing events where 5-min rainfall data were not available) of the landslides' data we defined altogether 20 events that were used for comparing ID threshold performance with the methodology proposed in this study.

Prediction of rainfall-induced shallow landslides

The next steps were carried out to test the performance of the proposed approach for rainfall-induced landslide prediction:

- The hydrological model was calibrated using daily data from 2007 to 2014 (2005-2006 data were used for model warm-up).
- Rainfall events were determined based on the daily data. Two rainfall events were separated in case of no rain period (NRP) equal to 1 day.
- Based on the start and end of the rainfall event the rainfall sum during the event (Psum) and the sum of the production store level (i.e. reservoir level S as shown in Fig. 1) during the event (Rsum) were calculated.
- Psum and Rsum were plotted on a scatter diagram and a Psum-Rsum threshold was defined with the aim to maximise the critical success index (CSI) since this criterion penalises

both false positive and false negative prediction results (Formetta et al. 2016). We decided to use the linear function between two selected variables due to the simplicity of the selected threshold. The following equation was used: Psum = a*Rsum + b, where a and b (i.e. slope and intercept of the linear equation) were determined in the process of maximizing the CSI.

CSI, success index (SI), distance to perfect classification (D2PC), accuracy (ACC), positive prediction power (PPP), and negative prediction power (NPP) (detailed description is provided by Formetta et al. 2016 and Martelloni et al. 2012) were used for comparison of the proposed approach with the ID thresholds. Moreover, Table 4 shows equations for the selected performance indices, their range, and the optimal index value. Furthermore, tp, tn, fp, and fn indicate true positive (i.e. the landslide occurred and was predicted), true negative (i.e. the landslide did not occur and was not predicted), false positive (i.e. landslide did not occur and was predicted), and false negative (i.e. the landslide occurred and was not predicted) cases, respectively.

For the ID threshold evaluation, we used the high-frequency rainfall data from the Davča gauging station. Rainfall events were separated in case of NRP equal to 22 h. This value was selected since it was also used by Rosi et al. (2016) in order to not impact the ID threshold performance. However, we argue that the use of the 24 h NRP would not significantly affect the results. Rainfall Intensity Summarization Tool (RIST) was used for 5-min data preprocessing (USDA, 2014). The next ID thresholds that were proposed by Rosi et al. (2016) for the Sava River catchment and the entire Slovenia were used: $I = 53.2 \times D^{-0.84}$ (Sava River) and I = $37.7 \times D^{-0.68}$ (Slovenia), respectively. Additionally, we also tested the performance of the global ID threshold that was suggested by Guzzetti et al. (2008) and was also used by Bezak et al. (2016): I = $2.2*D^{-0.44}$ (Global Threshold).

Probabilistic Psum-Rsum threshold definition

Using the approach described in "Prediction of rainfallinduced shallow landslides" one can determine one specific Psum-Rsum threshold that can be used for predicting rainfallinduced landslides. However, if the probabilistic approach is used as part of the EWS, as this is often the case in the rainfall-induced EWS (e.g. Wei et al. 2018), and in the case of flood EWS (e.g. Petan et al. 2015), we propose a copula function-based (e.g. Salvadori et al. 2007) approach:

Table 3 Descriptive statistics of daily data used in this study

	Q [m³/s]	P [mm]	T [°C]	PE [mm]
Station	Železniki	Davč a	Bohinjska Č ešnjica	Ljubljana
Minimum	0.4	0	-13.4	0
Mean	4.1	5.3	8.8	2.2
Median	2.4	0	9.5	1.7
Standard deviation	5.5	12.5	7.9	1.8
Maximum	76.2	228	27.3	7.7

Table 4 Selected criteria indices that were used in this study

Index	Equation	Range	Optimal value
Critical success index (CSI)	$CSI = \frac{tp}{tp + fp + fn}$	[0,1]	1
Success index (SI)	$SI = \frac{1}{2} \left(\frac{tp}{tp + fn} + \frac{tn}{fp + tn} \right)$	[0,1]	1
Distance to perfect classification (D2PC)	$D2PC = \sqrt{\left(1 - \frac{tp}{tp + fn}\right)^2 + \left(\frac{fp}{fp + tn}\right)^2}$	[0,1]	1
Accuracy (ACC)	$ACC = \frac{tp + tn}{tp + fp + fn + tn}$	[0,1]	1
Positive prediction power (PPP)	$PPP = \frac{tp}{tp + fp}$	[0,1]	1
Negative prediction power (NPP)	$NPP = \frac{tn}{fn + tn}$	[0,1]	1

Fit a 2-dimensional copula function to the Psum-Rsum sample (S_{ALL}) (i.e. all determined events rather than merely the ones that are classified as landslides) using a suitable marginal distribution function. Test the adequacy of selected copulas and select the most suitable one for this application according to the selected goodness-of-fit test and the selection criterion.

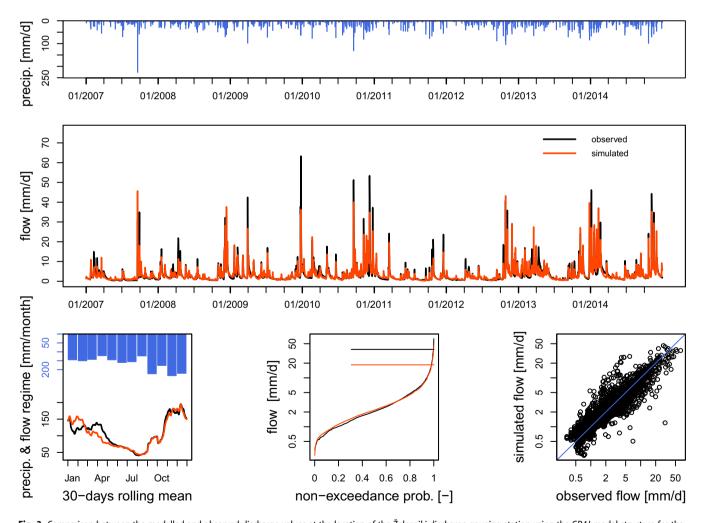


Fig. 3 Comparison between the modelled and observed discharge values at the location of the Železniki discharge gauging station using the GR4J model structure for the period from 2007 until 2014



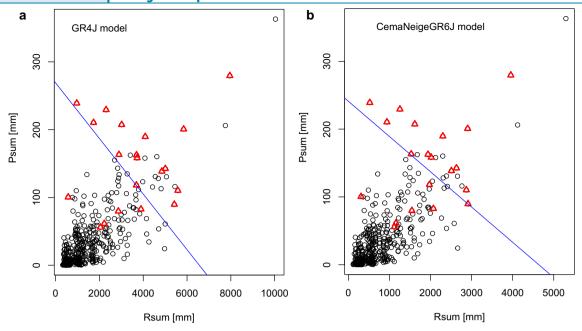


Fig. 4 Psum and Rsum values for the GR4J (left) and CemaNeigeGR6J (right) models. Red triangles indicate landslide events according to the Slovenian National Landslide Database. Black circles indicate all 463 rainfall events. Blue lines indicate the proposed Psum-Rsum threshold

Repeat the procedure for the Psum-Rsum events that are classified as landslides according to the landslides database (S_{LAND}) . Thus, two copula models are constructed using all events rather than only those that are classified as landslides.

- Using the fitted copula model generate a new S_{ALL} sample. Using the inverse marginal distribution function transform the generated values from the copula space [0, 1] to the real space. Repeat the procedure for the S_{LAND} .
- For each generated S_{LAND} , connect this value (i.e. event) to the closest generated S_{ALL} value using, for example, the nearest neighbour approach (e.g. Elseberg et al. 2012).
- Determine the Psum-Rsum threshold intercept (i.e. intercept of the linear equation since the linear equation is used for threshold definition in this study) parameter with the aim to maximise the CSI criterion. In this case study, the slope of the linear equation was determined using the procedure described in "Prediction of rainfall-induced shallow landslides" where the linear equation was used. For the definition of the probabilistic threshold we did not change the slope parameter that was determined in "Prediction of rainfall-induced shallow landslides".
- Repeat the procedure 10,000 times in order to obtain a large sample of Psum-Rsum threshold intercept values. Using the sorted sample one can estimate, for example, 10%, 50%, and 90% confidence thresholds. Since the slope of the linear equation is constant during the iterations, this will yield parallel thresholds.

In this study, we tested three copula functions from the Archimedean family, namely the Frank, Clayton, and Gumbel-Hougaard copula functions. Salvadori et al. (2007), for example, provide detailed description of these copulas and the equations. All the calculations using copula functions carried out in this study were done using the R copula package (e.g. Kojadinovic and Yan 2010). The parameters of copulas were estimated using the maximum pseudo-likelihood approach that is also implemented in the aforementioned package (Kojadinovic and Yan 2010). All three tested bivariate copula functions have one parameter (e.g. Salvadori et al. 2007). As mentioned two different copulas were fitted to the data.

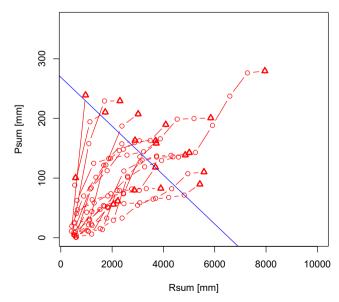


Fig. 5 Presentation of the cumulative daily *Psum* and *Rsum* values for 20 events that were characterised as landslides according to the landslide database for the GR4J model

Table 5 Several performance criteria that were derived using the defined *Psum*-Rsum threshold for the GR4J and CemaNeigeGR6J models. Performance criteria ranges and optimal values are given in Table 4

Goodness-of-fit criterion	GR4J model	CemaNeigeGR6J model
SI	0.862	0.839
CSI	0.469	0.483
D2PC	0.251	0.301
PPP	0.556	0.609
NPP	0.989	0.986
ACC	0.963	0.968

The first copula was fitted to all events (i.e. 463 events for the period from 2006 until 2014) and the second copula function was fitted to landslide events only (i.e. 20 events). The Cramér-von Mises test (Sn) was applied to test the adequacy of different copula models (Genest et al. 2009). The selected test compares empirical copula with the parametric estimate of the copula defined under the null hypothesis (Kojadinovic and Yan 2010). If more than one of the tested copula functions is not rejected by the selected goodness-of-fit test with the selected significance level, the k-fold cross-validation method can be used as a selection criterion (Grønneberg and Hort 2014). Moreover, the non-parametric distribution function defined by Hutson (2002) and Serinaldi (2009) was used in this study.

Results and discussion

Psum and Rsum threshold definition

In the first part of the study, we calibrated GR4J and CemaNeigeGR6J models using the methodology and data described in "Data and methods". Figure 3 shows the comparison between simulated and modelled discharge values for the selected period using the GR4J model structure. It also shows seasonal comparison and non-exceedance probability comparison of the measured and modelled data. The calculated Nash-Sutcliffe coefficient for the period from 2007 until 2014 using the GR4J model was 0.74 (i.e. NS ranges from $-\infty$ to 1 where NS value 1 indicates a perfect fit between the modelled and observed values). Interestingly, the more complex CemaNeigeGR6J model structure yielded a relatively similar performance as the GR4J model. The NS coefficient for the eight-parameter model was 0.77, which indicates only a slight improvement of this model version over the fourparameter model version. However, there was a relatively large difference in the calibrated model parameter that defines the maximum production storage level (i.e. X_1 on Fig. 1) despite the fact that both model versions use the same production store structure (Pushpalatha et al. 2011). For the GR4J and CemaNeigeGR6J models the estimated parameter value was 791 mm and 377 mm, respectively. However, these values are in the range of the 80% confidence intervals and the X_1 for the CemaNeigeGR6J model is similar to the median value provided by Perrin et al. (2003).

We firstly tried to use the production store level S and rainfall amount P at the triggering date of the landslide events for the

threshold definition. However, the results showed that the rainfall amounts for the triggering dates were relatively small, indicating that there could be some issues with the database triggering dates (e.g. landslides in remote areas could be spotted a few days after the actual triggering day). Moreover, if we considered the rainfall amount and production store level one day before the triggering date, the results for some cases improved but not for all. This confirmed potential database issues. Since most events were triggered during the autumn period where for the selected case study (i.e. located in the temperate continental climate), we often have long duration rainfall events and because also maximum 1-h rainfall intensities during the events were relatively small we assumed that long duration rainfall events are the main triggering mechanism of the landslides considered. Therefore, we decided to use the rainfall sum during the entire rainfall event. Consequently, we also used the production store sum during the event (i.e. *Rsum*) since such variable definition improved the threshold performance. It should be noted that Rsum is not only related to Psum (i.e. Pearson correlation coefficient 0.75) but also to discharge during the event (i.e. Pearson correlation coefficient 0.59), and evapotranspiration during the event (i.e. Pearson correlation coefficient 0.46). The positive correlation between pairs of these variables can mainly be attributed to the fact that longer duration events also yield higher absolute values. This indicates that the selected variable (i.e. Rsum) takes into account the hydrological processes occurring in the catchment rather than rainfall characteristics only (e.g. duration and amount). Moreover, this variable can be regarded as an overall indicator of the wetness increase in the catchment during the entire rainfall event.

We also checked the maximum 1-h rainfall intensity during the selected 20 landslide events. We found that for these events the intensities ranged from 2.2 mm/h to 84 mm/h. However, the median value was only 8.9 mm/h, which is significantly lower than the 2-year return period rainfall intensity characterised by the 1-h rainfall duration derived from the IDF curve (i.e. for the Davča

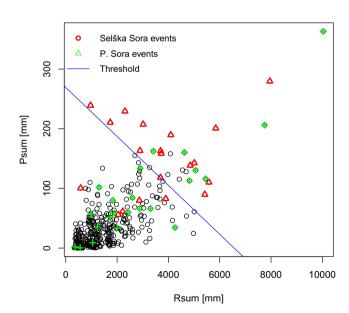


Fig. 6 Presentation of the landslide events that occurred in the adjacent P. Sora River catchment (indicated with green crosses) using the GR4J model

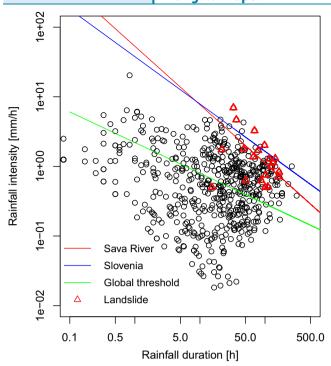


Fig. 7 Comparison between two local thresholds and one global ID threshold for the Selška Sora River catchment case study

station this value is 27 mm/h). Moreover, only for one event out of 20 the maximum 1-h rainfall intensity exceeds the 2-year return period value. Thus, we decided not to incorporate 1-h rainfall intensities in the proposed methodology.

In the next step, we defined rainfall events with the application of the selected NRP (i.e. 1 day). In the 8-year period (i.e. 2007-2014), 463 events were determined, which means approximately 58 rainfall events per year. Based on the start and end of each rainfall event we calculated the rainfall sum during the event (Psum) and the production store level sum during the event (Rsum). Figure 4 shows the calculated Psum and Rsum values in mm for both tested hydrological model structures. Based on the maximised CSI criterion, we also defined two Psum-Rsum thresholds, which are shown in Fig. 4. For the GR4J and CemaNeigeGR6J models the threshold equations are Psum = -0.041*Rsum + 269 mm and Psum =-0.052*Rsum + 241 mm, respectively. We also tested the threshold definition by maximizing the SI and D2PC criteria. However, the defined thresholds had much lower intercept values, which consequently leads to a large number of false alarms. Figure 5 shows also the cumulative daily Psum and Rsum values in mm that are characteristic of landslide event. Thus, it seems that the selected threshold is perpendicular to the trajectories determined by the cumulative Rsum and Psum values. Since an increase in Psum most likely also increases Rsum, and Rsum cannot increase without the precipitation input this could indicate that probability of events that have high Rsum and low Psum (e.g. Rsum > 6000 mm and *Psum* < 50 mm for the GR4J model version) is not high. Moreover, the threshold position (i.e. negative slope value) on the Psum-Rsum plot indicates that a rainfall-induced landslide can be

triggered either by a large rainfall input in case that Rsum is low or by a smaller rainfall input in case that Rsum is larger.

Using the 20 landslide events that were used for the threshold definition, which are also shown in Fig. 4, we calculated several performance criteria (Table 5). Similarly, as for the performance of the hydrological model in terms of its ability to predict Q values, the performance of the proposed methodology for the prediction of rainfall-induced landslides is quite similar to that of the GR4J and CemaNeigeGR6J models (Table 5). We can argue that for the Selška Sora River catchment the consideration of the snow routine and additional hydrologic model parameters do not significantly improve the performance results (Table 5).

We additionally investigated the events that were defined as false positives (11 and 9 events for the GR4J and CemaNeigeGR6J models, respectively) using the proposed approach (i.e. points located above the defined threshold that did not result in a landslide event). Seven of eleven points for the GR4J model could be associated with a landslide database entry that is related with the adjacent P. Sora River catchment (Fig. 6). This means that the landslide occurred in the adjacent catchment (Fig. 6). However, at the same time, additional entries are associated with this area (P. Sora River catchment) (Fig. 6), which would consequently lead to an increase in the false negatives (five and six events are classified as false negatives in Fig. 4 for the GR4J and CemaNeigeGR6J models, respectively). Thus, we argue that the statistics shown in Table 5 would not significantly change. If one would also want to consider the P. Sora catchment, a more complex or a separate hydrological model for this catchment should be constructed (i.e. consideration of additional rainfall and discharge data). This shows that the new proposed methodology should be developed and applied to a regional scale of at least several 100 km² rather than to a local scale of a few km² and with only few registered active landslides. At the local scale, even a single false positive event out of only a few registered landslides would lower the goodness-of-fit of the proposed threshold. More case studies of the proposed methodology are needed to test it and to clarify this issue—the methodology applying the lumped hydrologic models does not need distributed hydrological data but nevertheless needs registered landslide data to test it.

Comparison with ID thresholds

In the next step of the study, we also compared the performance of the proposed methodology that uses Psum and Rsum values in

Table 6 Several performances that were derived using two local (Sava River and Slovenia; Rosi et al. 2016) thresholds and one global (Guzzetti et al. 2008) threshold. Performance criteria ranges and optimal values are given in Table 4

Goodness-of-fit criterion	Sava River threshold	Slovenia threshold	Global threshold
SI	0.731	0.596	0.741
CSI	0.244	0.167	0.068
D2PC	0.501	0.800	0.469
PPP	0.323	0.500	0.068
NPP	0.982	0.972	0.997
ACC	0.946	0.965	0.548

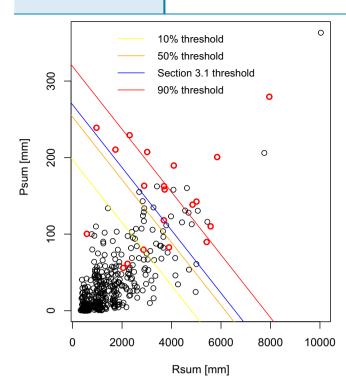


Fig. 8 Probabilistic 10%, 50% and 90% Psum-Rsum thresholds that were determined using the approach described in "Probabilistic Psum-Rsum threshold definition". The threshold that was determined in "Psum and Rsum threshold definition" is also shown. GR4J model results are shown

order to predict rainfall-induced landslides with ID thresholds. Figure 7 shows results for two local ID thresholds (Slovenia and the Sava River defined by Rosi et al. 2016) and one global (Guzzetti et al. 2008) ID threshold. For these thresholds, the same goodness-of-fit criteria were calculated as for the approach based on the conceptual hydrological model (Table 6). One can notice that the proposed approach (using Psum-Rsum threshold) generally yields better results compared to the tested ID thresholds. A similar discussion about false positive and false negative events as that provided in "Psum and Rsum threshold definition" is also valid for this part of the analysis.

Psum and Rsum probabilistic threshold definition

In the final step of the study, we also computed the probabilistic thresholds using the approach described in "Probabilistic Psum-Rsum threshold definition". We tested the performance of the one-parameter bivariate Frank, Clayton, and Gumbel-Hougaard copula functions (e.g. Salvadori et al. 2007). The Cramér-von Mises test (Sn) test (Genest et al. 2009) results indicated that only the Gumbel-Hougaard copula function could not be rejected with the selected significance level of 0.05 for S_{ALL} and S_{LAND} sub-samples (i.e. test results with the corresponding p-values in brackets were 0.04 (0.05) and 0.04 (0.11) for the S_{ALL} and S_{LAND} sub-samples, respectively). Moreover, Frank (i.e. test results with the corresponding p-values in brackets were 0.05 (0.005) and 0.05 (0.02) for the S_{ALL} and S_{LAND} sub-samples, respectively) and Clayton (i.e. test results

with the corresponding p values in brackets were 0.6 (1 \times 10⁻⁵) and 0.04 (0.04) for the S_{ALL} and S_{LAND} sub-samples, respectively) copulas were rejected at the selected significance level of 0.05. This indicates that the Frank and Clayton copulas are not suitable for the investigated sample. Moreover, test results show that the Gumbel-Hougaard copula function gives adequate fit to the data. This means this copula was used for further calculations and we did not apply the k-fold crossvalidation method in order to select the best copula among the Frank, Clayton and Gumbel-Houdaard copula functions. In case that the Frank and Clayton copulas would not be rejected by the selected statistical test, we would use the kfold cross-validation method to select the most suitable copula. Figure 8 shows probabilistic thresholds for 10%, 50% and 90% levels that were determined using the methodology described in "Probabilistic Psum-Rsum threshold definition". One can notice that the 50% threshold (i.e. median) is located slightly below the threshold that was defined in "Psum and Rsum threshold definition". The 10% threshold would slightly increase (compared to the threshold defined in "Psum and Rsum threshold definition") true positive events and, on the other hand, significantly increase false positive (i.e. false alarm) events. The 90% threshold would decrease both true and false positive events compared with "Psum and Rsum threshold definition" threshold. This can also be confirmed by the results shown in Table 7 where we also calculated the performance of the 10%, 50% and 90% thresholds. One can notice that D2PC, PPP and ACC criteria results are increasing with increasing probability. For example, this means that the higher the threshold, the higher the PPP of the threshold. On the other hand, SI and NPP results are decreasing with increasing probability, which for example means that low thresholds (i.e. smaller intercept values) have better NPP while at the same time PPP is relatively low. Moreover, the CSI that was also used to determine the deterministic threshold has the optimal value somewhere between the 10% and 50% threshold (Table 7). Thus, as pointed out also by Formetta et al. (2016), this criterion penalises both false negative and false positive predictions, which could, from the perspective of this study, yield a more robust threshold because in any EWS too many false alarms are not desired.

Table 7 Several performance criteria that were derived for the probabilistic thresholds that were defined using the copula approach. The performance criteria range and optimal value are given in Table 4

Goodness-of-fit criterion	10% threshold	50% threshold	90% threshold
SI	0.877	0.855	0.692
CSI	0.274	0.406	0.296
D2PC	0.177	0.253	0.600
PPP	0.288	0.469	0.533
NPP	0.993	0.988	0.973
ACC	0.903	0.952	0.959

Conclusions

This paper presents a methodology for predicting rainfallinduced shallow landslides, which is based on the lumped conceptual hydrological model results. The main aim was to propose an approach for the shallow landslide prediction that would not use the ID thresholds, which were recently criticised by Bogaard and Greco (2018). Therefore, we decided to use the Rsum variable that depends on the input and output data (i.e. P, PE and Q) that is used in the hydrological model and takes into account the wetness increase during the entire event. Moreover, the aim of this study was also to propose the methodology that could easily be applied to data-scarce areas where detailed data, which are needed to calibrate and apply physically based models for landslide prediction, are not available. The presented methodology is tested using the pre-alpine meso-scale Selška Sora River catchment case study in western Slovenia. Based on the presented results, the following conclusions can be made:

- We argue that the production store level, and consequently Rsum, is a relatively good proxy of the hydrological conditions in the catchment that have an important impact on landslide triggering and can be regarded as a useful variable for predicting rainfall-induced shallow landslides. The volume of water in the production store reservoir (used to calculate Rsum) is connected with rainfall, evapotranspiration, and runoff. Due to the hydrological model structure, this kind of information (i.e. proposed Psum-Rsum threshold) could also be incorporated to early warning systems (EWS). Moreover, probabilistic thresholds could be useful for this purpose. Using rainfall and evapotranspiration forecasts and a calibrated hydrological model one can calculate production store and discharge values for the forecasted period.
- For the selected Selška Sora case study where most of the analysed landslides were triggered by the long-duration rainfall events, the Psum-Rsum combination yielded meaningful results. However, in case that landslides are triggered by rainfall events with different characteristics (e.g. intense and short rainfall events), probably some other definition of the variables should be used and an hourly model type should be used instead of the daily model since an hourly model with a similar structure also exists (e.g. Coron et al. 2017a; Coron et al. 2017b).
- The GR4J model structure yielded similar performance in terms of discharge prediction (NS coefficient) and also shallow landslide prediction (Table 5) as the CemaNeigeGR6J model structure that also includes a snow routine and uses four additional parameters. Thus, for areas (i.e. catchments) similar to the Selška Sora River catchment and for lowland areas, the GR4J should be preferred due to a more simple model structure (i.e. less parameters) and a smaller amount of input data required (Table 1).
- The proposed Psum-Rsum threshold yielded better results in terms of rainfall-induced shallow landslides prediction compared to the ID threshold approach. Local and global thresholds were tested and it seems that the selected global ID threshold is not suitable for the Selška Sora River catchment since it has very low PPP. Better results were obtained using two local ID thresholds (Sava River and Slovenia).

The proposed methodology should be additionally tested using case studies where the landslide database is verified and there are no issues related to the landslide location and triggering date (e.g. a landslide could be detected a few days after the actual event in remote areas).

Acknowledgments

This project was also approved by the International Programme on Landslides (IPL) as IPL-226 Project (2017-2020); http:// iplhq.org/category/iplhq/ipl-ongoing-project/). We wish to thank the Slovenian Environment Agency (ARSO) for making data publicly available. Special thanks also to the Ministry of Defence and the Administration of the Republic of Slovenia for Civil Protection and Disaster Relief for financing the MASPREM project. The critical and useful comments of anonymous reviewers greatly improved this work, for which the authors are very grateful.

Funding information

The results of the study are part of the research project J1-8153 "Studying Landslide Movements from Source Areas to Zone of Deposition using a Deterministic Approach" that is financed by the Slovenian Research Agency (ARRS).

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http:// creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

Aleotti P (2004) A warning system for rainfall-induced shallow failures. Eng Geol 73:247-265. https://doi.org/10.1016/j.enggeo.2004.01.007

Aristizabal E, Velez JI, Martinez HE, Jaboyedoff M (2016) Shia_landslide: a distributed conceptual and physically based model to forecast the temporal and spatial occurrence of shallow landslides triggered by rainfall in tropical and mountainous basins. Landslides 13:497-517. https://doi.org/10.1007/s10346-015-0580-7

Bezak N, Šraj M, Mikoš M (2016) Copula-based IDF curves and empirical rainfall thresholds for flash floods and rainfall-induced landslides. J Hydrol 541:272-284. https://doi.org/10.1016/j.jhydrol.2016.02.058

Bogaard T, Greco R (2018) Invited perspectives: hydrological perspectives on precipitation intensity-duration thresholds for landslide initiation: proposing hydrometeorological thresholds. Nat Hazards Earth Syst Sci 18:31-39. https://doi.org/ 10.5194/nhess-18-31-2018

Ciarapica L, Todini E (2002) Topkapi: a model for the representation of the rainfall-runoff process at different scales. Hydrol Process 16:207-229. https:// doi.org/10.1002/hyp.342

Coron L, Perrin C and Michel C (2017a) airGR: suite of GR hydrological models for precipitation-runoff modelling. R package version 1.0.9.64. URL: https:// webar.irstea.fr/en/airGR/

Coron L, Thirel G, Delaigue O, Perrin C, Andreassian V (2017b) The suite of lumped GR hydrological models in an R package. Environ Model Softw 94:166-171. https:// doi.org/10.1016/i.envsoft.2017.05.002

Edijatno NND, Yang XL, Makhlouf Z, Michel C (1999) Gr3j: a daily watershed model with three free parameters. Hydrological Sciences Journal-Journal Des Sciences Hydrologiques 44:263-277. https://doi.org/10.1080/02626669909492221

Elseberg J, Magnenat S, Siegwart R, Nüchter A (2012) Comparison of nearest-neighborsearch strategies and implementations for efficient shape registration. Journal of Software Engineering for Robotics 3(1):2-12

- Formetta G, Capparelli G, Versace P (2016) Evaluating performance of simplified physically based models for shallow landslide susceptibility. Hydrol Earth Syst Sci 20:4585-4603. https://doi.org/10.5194/hess-20-4585-2016
- Genest C, Remillard B, Beadoin D (2009) Goodness-of-fit tests for copulas: a review and a power study. Insurance: Mathematics and Economics 44:199-213. https://doi.org/ 10.1016/i.insmatheco.2007.10.005
- Grad K, Ferjančič L (1974) Osnovna geološka karta SFRJ, list Kranj, 1: 100, 000 = Basic geological map of Yugoslavia, Map Kranj, scale 1: 100, 000. Zvezni geološki zavod, Belgrade
- Grønneberg S, Hort NL (2014) The copula information criteria. Scand J Stat 41(2):436-459. https://doi.org/10.1111/sjos.12042
- Guzzetti F, Peruccacci S, Rossi M, Stark CP (2008) The rainfall intensity-duration control of shallow landslides and debris flows: an update. Landslides 5:3-17. https://doi.org/ 10.1007/s10346-007-0112-1
- Haque U, Blum P, Da Silva P, Andersen P, Pilz JC, Sergey R, Male J-P, Jemec Auflič M, Andres N, Poyiadji E et al (2016) Fatal landslides in Europe. Landslides 13(6):1545-1554
- Huang J, Ju NP, Liao YJ, Liu DD (2015) Determination of rainfall thresholds for shallow landslides by a probabilistic and empirical method. Nat Hazards Earth Syst Sci 15:2715-2723. https://doi.org/10.5194/nhess-15-2715-2015
- Hutson AD (2002) A semi-parametric quantile function estimator for use in bootstrap estimation procedures. Stat Comput 12:331-338. https://doi.org/10.1023/ A:1020783911574
- Jemec Auflič M, Šinigoj J, Krivic M, Podboj M, Peternel T, Komac M (2016) Landslide prediction system for rainfall induced landslides in Slovenia (Masprem). Geologija 59(2):259-271
- Jemec Auflič M, Komac M (2013) Rainfall patterns for shallow landsliding in perialpine Slovenia. Nat Hazards 67(3):1011-1023
- Kojadinovic I, Yan J (2010) Modelling multivariate distributions with continuous margins using the copula R package. J Stat Softw 34(9)
- Komac M, Hribernik K (2015) Slovenian national landslide database as a basis for statistical assessment of landslide phenomena in Slovenia. Geomorphology 249:94-102. https://doi.org/10.1016/j.geomorph.2015.02.005
- Krogli IK, Devoli G, Colleuille H, Boje S, Sund M, Engen IK (2018) The Norwegian forecasting and warning service for rainfall- and snowmelt-induced landslides. Nat Hazards Earth Syst Sci 18:1427–1450. https://doi.org/10.5194/nhess-18-1427-2018
- Liu DL, Zhang SJ, Yang HJ, Zhao LQ, Jiang YH, Tang D, Leng XP (2016) Application and analysis of debris-flow early warning system in Wenchuan earthquake-affected area. Nat Hazards Earth Syst Sci 16:483–496. https://doi.org/10.5194/nhess-16-483-2016
- Maček U, Bezak N, Šraj M (2018) Reference evapotranspiration changes in Slovenia, Europe. Agric For Meteorol 260-261:183-192. https://doi.org/ 10.1016/j.agrformet.2018.06.014
- Martelloni G, Segoni S, Fanti R, Catani F (2012) Rainfall thresholds for the forecasting of landslide occurrence at regional scale. Landslides 9:485-495. https://doi.org/10.1007/ s10346-011-0308-2
- Mathew J, Babu DG, Kundu S, Kumar KV, Pant CC (2014) Integrating intensity-durationbased rainfall threshold and antecedent rainfall-based probability estimate towards generating early warning for rainfall-induced landslides in parts of the Garhwal Himalaya, India. Landslides 11:575-588. https://doi.org/10.1007/s10346-013-0408-2
- Michel C (1991) Hydrologie appliquée aux petits bassins ruraux, Hydrology handbook (in French), Cemagref, Antony, France
- Mikoš M, Četina M, Brilly M (2004) Hydrologic conditions responsible for triggering the Stoze landslide, Slovenia. Eng Geol 73:193-213. https:// doi.org/10.1016/i.enggeo.2004.01.011
- Nash JE, Sutcliffe JV (1970) River flow forecasting through conceptual models part I — a discussion of principles. J Hydrol 10(3):282–290. https://doi.org/ 10.1016/0022-1694(70)90255-6
- Oudin F, Hervieu F, Michel C, Perrin C, Andreassian V, Antcil F, Loumagne C (2005) Which potential evapotranspiration input for a lumped rainfall-runoff model?: Part 2-

- towards a simple and efficient potential evapotranspiration model for rainfall-runoff modelling. J Hydrol 303(1–4):290–306. https://doi.org/10.1016/j.jhydrol.2004.08.026
- Peres DJ, Cancelliere A, Greco R, Bogaard TA (2018) Influence of uncertain identification of triggering rainfall on the assessment of landslide early warning thresholds. Nat Hazards Earth Syst Sci 18:633-646. https://doi.org/10.5194/nhess-18-633-2018
- Perrin C, Michel C, Andreassian V (2003) Improvement of a parsimonious model for streamflow simulation. J Hydrol 279:275-289. https://doi.org/10.1016/ s0022-1694(03)00225-7
- Petan S, Golob A, Moderc M (2015) Hydrological forecasting system of the Slovenian environment agency. Acta hydrotechnica 28(49):119-131
- Petley D (2012) Global patterns of loss of life from landslides. Geology 40:927-930. https://doi.org/10.1130/g33217.1
- Pushpalatha R, Perrin C, Le Moine N, Mathevet T, Andreassian V (2011) A downward structural sensitivity analysis of hydrological models to improve low-flow simulation. J Hydrol 411:66-76. https://doi.org/10.1016/j.jhydrol.2011.09.034
- Rosi A, Peternel T, Jemec-Auflič M, Komac M, Segoni S, Casagli N (2016) Rainfall thresholds for rainfall-induced landslides in Slovenia. Landslides 13:1571–1577. https://doi.org/10.1007/s10346-016-0733-3
- Salvadori G, De Michele C, Kottegoda NT, Rosso R (2007) Extremes in nature: an approach using copulas. In: Springer, Dordrecht, Netherlands
- Segoni S, Rosi A, Lagomarsino D, Fanti R, Casagli N (2018) Brief communication: using averaged soil moisture estimates to improve the performances of a regional-scale landslide early warning system. Nat Hazards Earth Syst Sci 18:807-812. https:// doi.org/10.5194/nhess-18-807-2018
- Segoni S, Rosi A, Rossi G, Catani F, Casagli N (2014) Analysing the relationship between rainfalls and landslides to define a mosaic of triggering thresholds for regional-scale warning systems. Nat Hazards Earth Syst Sci 14:2637–2648. https://doi.org/10.5194/nhess-14-2637-2014
- Serinaldi F (2009) Assessing the applicability of fractional order statistics for computing confidence intervals for extreme quantiles. J Hydrol 376:528-541. https://doi.org/10.1016/j.jhydrol.2009.07.065
- USDA (2014) United States Department of Agriculture. Rainfall intensity summarization tool (RIST). Accessed from http://www.ars.usda.gov/News/docs.htm?docid=3251
- Valery A, Andreassian V, Perrin C (2014a) 'As simple as possible but not simpler': what is useful in a temperature-based snow-accounting routine? Part 1 comparison of six snow accounting routines on 380 catchments. J Hydrol 517:1166-1175. https://doi.org/10.1016/j.jhydrol.2014.04.059
- Valery A, Andreassian V, Perrin C (2014b) 'As simple as possible but not simpler': what is useful in a temperature-based snow-accounting routine? Part 2 - sensitivity analysis of the cemaneige snow accounting routine on 380 catchments. J Hydrol 517:1176-1187. https://doi.org/10.1016/j.jhydrol.2014.04.058
- Wei LW, Huang CM, Chen H, Lee CT, Chi CC, Chiu CL (2018) Adopting the I₃-R₂₄ rainfall index and landslide susceptibility for the establishment of an early warning model for rainfall-induced shallow landslides. Nat Hazards Earth Syst Sci 18:1717-1733. https://doi.org/10.5194/nhess-18-1717-2018
- Zorn M, Komac B (2008). Zemeljski plazovi v Sloveniji = landslides in Slovenia. Geografski inštitut Antona Melika ZRC SAZU, Slovenia

N. Bezak (💌) · M. Mikoš

Faculty of Civil and Geodetic Engineering, University of Liubliana. Jamova cesta 2, 1000, Ljubljana, Slovenia Email: nejc.bezak@fgg.uni-lj.si

M. Jemec Auflič

Geological Survey of Slovenia, Dimič eva ulica 14, 1000, Ljubljana, Slovenia