An evaluation model for landslide and debris flow prediction using multiple hydrometeorological variables

A landslide and debris flow prediction model

Jinjin Hou [a], Ming Dou [a, b*], Yongyong Zhang [c], Jihua Wang [d], Guiqiu Li [b]

[a] School of Water Conservancy Science and Engineering, Zhengzhou University, Zhengzhou, 450001, China

[b] School of Ecology and Environment, Zhengzhou University, Zhengzhou, 450001, China

[c] Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, No. 11A Datun Road, Beijing, 100101, China

[d] Geological Environmental Monitoring Institute of Henan Province, Zhengzhou 450000, China

*Corresponding Author: Ming Dou

E-mail address: dou_ming@163.com

School of Water Conservancy Science and Engineering

Zhengzhou University

No. 100 Kexue Road,

Zhengzhou 450001, Henan, China

An evaluation model for landslide and debris flow prediction using multiple hydrometeorological variables

Abstract: Landslide and debris flows are typically triggered by rainfall-related weather conditions, including short-duration storms and long-lasting rainfall. The critical precipitation of landslide and debris flow occurrence is different under various hydrometeorological conditions. In this study, the daily hydrological states were evaluated by the SWAT model, and the trigger sensitivities of different daily hydrological variables were assessed with 50 days recorded landslide and debris flows between 2010 and 2013. Based on modeled wetness states, the event days were divided into LLR-trigger event days (long-lasting rainfall) and SDS-trigger event days (short-duration storm) with six determinate criteria. The landslide and debris flow prediction model was built using nine hydrometeorological variables and the predictive performance was tested with simulated data from 2010 to 2012. The results suggest that: Historical hydrological variables and their development provide important information for triggering debris flows, though rainfall is the most important factor for triggering debris flows. The landslides and debris flows in the selected subbasins region are triggered on 33 days by LLR and on 17 days by SDS. Specifically, LLR type landslide and debris flow account for a large proportion in July, while SDS type landslide and debris flow occur more frequently in September. The prediction model with the AUC value of 0.85, can capture most of the landslide debris flow. The temporal distribution of the two triggering-event predicted by the model is consistent with the annual distribution of precipitation. Besides, there are spatial variations of the specific trigger types in the different subbasins, which may attribute to the different land cover. Despite some uncertainty, this study thereby provides an idea of improving the landslide and debris flow prediction model.

Keywords: SWAT model; multiple hydrometeorological variables; trigger sensitivities; landslide and debris flow; prediction model

1. Introduce

Landslide and debris flow is a widespread and destructive natural disaster in the world (Huang et al., 2017; Nicolussi, Spotl, Thurner, & Reimer, 2015), which may cause a large number of casualties and economic losses. Globally, landslides cause about 1,000 casualties and about $4 billion in property damage every year (Pradhan & Youssef, 2010). In the Aranayake area, Sri Lanka, a rapid and long-traveling landslide triggered by a monsoon rainfall on 2016 has destroyed 75 houses and killed 127 persons (Tan, Sassa, Dang, Konagai, & Sato, 2020). In the mountainous areas of the Western China, landslide and debris flow leads to approximately 762 deaths and disappearances and $600 million in property losses per year recently, according to the Chinese Institution of Geological Environmental Monitoring (Zhang, Wang, Bao, & Zhao, 2019). Other research reported that landslide debris flow disasters have caused more than 1100 fatalities and $5-10 billion in China since 2000 (Hong et al., 2017). It is therefore extremely important to analysis the landslide and debris flow triggering conditions and build the disaster prediction model to minimize damage and avoid loss of life.

As the frequent hazards in mountain regions, landslide and debris flows are caused by various triggers, including earthquakes, rainfall and rapid floods, and are influenced by multiple factors, such as topography, soil and rock types, fractures and bedding planes, and moisture content (Crozier & Michael, 1986). Besides earthquakes, volcanism, precipitation is the most frequent trigger of debris flows (Mostbauer, Kaitna, Prenner, & Hrachowitz, 2018; Prenner, Hrachowitz, & Kaitna, 2019), which is widely studied by the empirical models and physical (process-based) models (Floris & Bozzano, 2008; Schiliro, Cevasco, Esposito, & Mugnozza, 2018). Using empirical models to analyze the landslide and debris flow has always been a hot topic for scholars (Aleotti, 2004; Ferro, Carollo, & Serio, 2020; Lainas, Sabatakakis, & Koukis, 2016) since Caine (1980) first proposed the global rainfall intensity duration (ID) threshold for shallow landslides. For example, Kanungo and Sharma (2014) found the rainfall threshold relationship fitted to the lower boundary of the landslide triggering rainfall events by analyzing 81 out of 128 landslides taken place in India from 2009 to 2012. Although the empirical model is relatively simple and suitable for large areas, this method requires higher quality precipitation data and does not take into account the different trigger sensitivities of different regions, which leads to many false positives and reduces the accuracy of the prediction model. To reduce the negative effects of the lack of geological parameters, a rainfall threshold is usually only used for a particular geographical region. Martinotti et al. (2017) designed a new ensemble-non-exceedance probability algorithm for the quantitative evaluation of the possible occurrence of rainfall-induced landslides in karst areas, providing better diagnostics than the single metrics often used for landslide forecasting. Vessia, Curzio, Chiaudani, and Rusi (2020) had taken the local geo-morphological characters into account by means of the co-kriging technique to constrain the cumulated and duration mean values of a regional Empirical Rainfall Thresholds and their confidence intervals. Furthermore, as for the poor quality data on which empirical methods, Frattini, Crosta, and Sosio (2009) used logistic regression to define the ID thresholds associated with different precipitation regression periods that trigger landslide and debris flow. And Jaiswal and van Westen (2009) used a control data set that was not used in the empirical model to verify the threshold value of the visual drawing to estimate the conditional probability of the landslide and the overall time probability of the occurrence of the landslide. Although many measures were used to improve the empirical model, the direct contributions of that to predicting disasters accurately are limited due to its limited data and parameters.

Physical (process-based) models are also used to study the triggering mechanisms of landslide debris flows by considering the relevant geo-morphological features (Bui et al., 2019; Dou et al., 2020; Schiliro et al., 2018). For example, Dai and Lee (2002) described the physical characteristics of landslides and the statistical relations of landslide frequency with the physical parameters contributing to the initiation of landslides on Lantau Island in Hong Kong using the Geographical Information Systems (GIS) database. Hu, Li, Chen, and Zhang (2007) used nine evaluation factors for the landslide prediction and demonstrated that the Support vector machine model is efficient and accurate for landslide hazard evaluation and spatial prediction. However, most of the studies focused on a single slope or landslide event in a relatively small catchment, which is also a prerequisite to validate model results against observed landslides from inventories (Tian, Xiao, Liu, & Wu, 2008). In order to apply physically-based landslide prediction models in a large region, Wang et al. (2020) provided an effective method by coupling a hydrological model and a slope stability model to predict landslides over large regions in which fine-scale topographical information is incorporated. Even so, the physical (process-based) models that use physical thresholds typically predict pore pressure based on spatial variable characteristics are difficult to apply to those areas where their key parameters (such as soil thickness, groundwater conditions or shear strength) are difficult to obtain (Camilo, Lombardo, Mai, Dou, & Huser, 2017).

By considering the hydrological history of the basin, Crozier and Michael (1999) used the Antecedent Water Level (AWS) model to verify the hydrological sensitivity of landslides in the basin, indicating that different hydrological basin states may affect the critical water input required to trigger the landslide. Although the occurrence of landslide and debris flow is a local phenomenon in the basin, the hydrometeorological process ensures sufficient water input in the basin, so that the landslide and debris flow can be identified on a larger scale. Prenner, Kaitna, Mostbauer, and Hrachowitz (2018) quantitatively determined different trigger types for historical debris flows and used four Naive Bayes classifier models, ranging from a simple rainfall-only model to a multi-parameter hydrometeorological model differentiating between trigger types, to predict the days susceptible for debris flow occurrence in the region, which improved understanding of the hydrometeorological impact on debris flow initiation in a mountain watershed. Therefore, using hydrometeorological variables to predict landslide debris flow can not only make up for the defect of single precipitation data in the empirical model, but also avoid the need for many geological parameters in the physical model.

Following these research ideas and theoretical basis above (Ciavolella, Bogaard, Gargano, & Greco, 2016; Prenner et al., 2019; Prenner et al., 2018), this paper aim to build an evaluation model for landslide and debris flow prediction using multiple hydrometeorological variables. The upper Han River basin, China is selected as this case study. The specific objectives are to: (1) comprehensively assess the trigger sensitivities of different daily hydrological variables simulated by the SWAT model; (2) divide the hydrometeorological conditions triggering landslide and debris flows into LLR (long-lasting rainfall) and SDS (short-duration storms) and build a landslide and debris flow prediction model with multiple hydrometeorological variables based on the naive Bayesian model; (3) test the applicability of the landslide and debris flow predictive model using simulated data from 2010 to 2012. The results are expected to gain insights into the effect of hydrological state on triggering landslide and debris flows, and give implications for improving the landslide and debris flow prediction model.

2. Materials and Methods

2.1 Materials

2.1.1. Study area

As the largest tributary of the Yangtze River, Han River originates from the Luzhang Mountain in Ningqiang County, Shaanxi Province of China, runs across the Qinling Mountains and Daba Mountains from west to east, flows through Shaanxi and Hubei provinces, and injects the Yangtze River in Hankou City, Hubei Province. The upper Han River basin (30°8' - 34°11' N, 106°12' - 114°14'E) is located between the Western Plains and the Tibet Plateau, with a total catchment area of about 95,200 km^2 and an elevation of 82 - 3545 m, about 925 km long. The study area is a subtropical area with a monsoon climate, with an average annual temperature of 15-17℃. The annual average rainfall is about 700 - 1100 mm, which is very uneven throughout the year, of which 70-80% is concentrated in the rainy season from May to October. About 48% of the upper Han River is covered by woodland, 36% by grassland and 9% by cultivated land. In the region, there were 93 landslide and debris flows caused significant damage between 2010 and

2013. Among them, landslide and debris flows that occurred in the same subbasin at the same time were regarded as the same disaster. In the end, 50 landslide and debris flows were used for research.

2.1.2. Available data

The data required for research and analysis are mainly divided into attribute data and map data: attribute data are mainly hydrometeorological data including precipitation, temperature, and runoff data; map data are mainly digital elevation model (DEM) data, landuse data, and soil type data. As input meteorological data, precipitation and temperature data are daily average data observed by four hydrological stations (i.e., Hanzhong station, Ankang station, Shangluo station and Xixia station) in the upstream of the Han River from 1961 to 2013 (see Figure 1). Runoff data from 2000 to 2013 is monthly average data monitored by Danjiangkou station at the outlet of the basin located 800 m downstream of the confluence of Han River and Dan River, which was used to calibrate and verify the performance of the hydrological model. The DEM data was obtained from the GDEMDEM 30 m resolution digital elevation data downloaded from the geographic spatial data cloud (http://www.gscloud.cn), and the geographic coordinates is WGS_1984 (Figure 1). The landuse data and soil type data are downloaded from the data set of Resources and Environmental Data Cloud Platform (http://www.resdc.cn), with a spatial resolution of 1 km (Figure 2). The landslide and debris flow data analyzed in the study was obtained from the Shaanxi Disaster Relief Yearbook, including the time, location, and casualties of the incident.

2.2. Methods

This paper contains four parts (Figure 3): First, the hydrological model SWAT was used to simulate the hydrological processes in the upstream of the Han River Basin and estimate the values of hydrological state variables. Second, the landslide and debris flow triggering probability and characteristics conditional on the multiple hydrometeorological variables were analyzed by using the Bayesian method in the area with the landslide debris flow frequently occurring. Third, the observed landslide and debris flows were divided into two trigger classes LLR and SDS based on the triggering probability and characteristics, and the differences, as well as the distribution characteristics between two classes were compared. Finally, a landslide and debris flows prediction model was built with the Naive Bayesian probability to predict the debris flows occurrence on a specific day as a function of a range of hydrometeorological variables.

2.2.1. SWAT model

The process-based SWAT model was used to simulate the hydrological catchment state to obtain estimates of the daily study region state and flux variables such as soil water content, actual evapotranspiration, potential evapotranspiration, underground runoff, surface runoff and net water production, which were used for subsequent analysis and prediction model establishment. As a semi-distributed hydrological model, the SWAT model divides the study area into several subbasins according to topography and water system, and then subdivide the subbasins into different hydrologic response units (HRUs) based on soil properties and landuse types. Thus, it can take into account the comprehensive effects of weather, soil properties, topography and land cover so as to realize more accurate simulation of the basins.

In this study, the SWAT model was used to simulate the hydrological processes for the period of 2006-2013. The 2006-2010 period is used to calibrate the model, while the 2011-2013 as the validation period. The model was calibrated and validate with manual discharge data set of the basin outlet. Based on the parameter sensitivity analysis of potential influencing factors, the sensitive parameters were calibrated using the automatic calibration software SWAT-CUP under SUFI-2 optimization algorithm. The performance of model simulation was quantitatively evaluated by using three indicators: Nash-Sutcliffe Efficiency Coefficient (NSE), the Percent Bias (PBIAS) and Coefficient of Determination (R2).

$$NSE = 1 - \frac{\sum_{i=1}^{n}\left(Q_{o,i} - Q_{p,i}\right)^2}{\sum_{i=1}^{n}\left(Q_{o,i} - \overline{Q_o}\right)^2} \tag{1}$$

$$PBIAS = \frac{\sum_{i=1}^{n}\left(Q_{o,i} - Q_{p,i}\right)}{\sum_{i=1}^{n} Q_{o,i}} \times 100\% \tag{2}$$

$$\mathrm{R}^2 = \frac{\left[\sum_{i=1}^{n}\left(Q_{o,i} - \overline{Q_o}\right)\left(Q_{p,i} - \overline{Q_p}\right)\right]^2}{\sum_{i=1}^{n}\left(Q_{o,i} - \overline{Q_o}\right)^2 \sum_{i=1}^{n}\left(Q_{p,i} - \overline{Q_p}\right)^2} \tag{3}$$

Where, $Q_{p,i}$ is the i^{th} estimated flow value, $Q_{o,i}$ is the i^{th} observed flow value, $\overline{Q_p}$ is the mean of estimated flow value, $\overline{Q_o}$ is the mean of observed flow value.

Except for precipitation and temperature, the hydrological variables required for analysis in this paper were estimated data of SWAT model, which included precipitation, soil water content, potential evapotranspiration, actual evapotranspiration, surface runoff, groundwater discharge, and net water yield. Among them, the daily hydrological variables were extracted from the SWAT hydrological model. Considering the hysteresis of time and the treatment of effective precipitation, the calculation of the cumulative hydrological variables was defined as follows (LIU et al., 2016):

$$X_k = X_0 + \prod_{k=1}^{n} X_k \cdot \alpha^k \tag{4}$$

Where, X_k is the cumulate hydrological variable X of the k^{th} days before the event; X_0 is the value of the hydrological variable X on the day of the event; α is the effective coefficient, here $\alpha = 0.84$ (Cong, Pan, Li, Ren, & Li, 2006; Glade, Crozier, & Smith, 2000).

2.2.2 Trigger probability of hydrometeorological variables

Bayes' theorem computes the probability that a landslide debris flow (E) occurs given a hydrometeorological variable (X_i), including the time series of hydrological model states and flux variables, and their temporal derivatives and accumulations:

$$P\left(E \middle| X_i\right) = \frac{P(E) \times P\left(X_i \middle| E\right)}{P(X_i)} \tag{5}$$

Bayesian probability is usually computed in terms of relative frequencies. Thus, equation (5) can be simplified to equation (2), which expresses the ratio between the number of occurrences of magnitude X_i in connection to debris flow events $N(X_i|E)$ and the total number of occurrences $N(X_i)$:

$$P\left(E \middle| X_i\right) = \frac{N\left(X_i \middle| E\right)}{N(X_i)} \tag{6}$$

In this study the Bayesian probability is used to identify the effects of the crucial hydrological indicators on triggering landslide and debris flows, and provides a basis for determining the trigger type classification criteria. Here, six hydrological variables including soil water content (SW), groundwater discharge (GW_Q), net water yield (WYLD), potential evapotranspiration (PE), actual evapotranspiration (E), and surface runoff (SURQ) are involved due to the potential impact of rainfall on the hydrological process, and the specific value of them are derived from the SWAT model simulation results. According to the SWAT model spatial analysis, the study area was divided into 109 subbasins, and each subbasin used the rainfall and temperature data from the nearest weather station to the center of the subbasin. In order to avoid data splitting while retaining the meteorological spatial characteristics, we selected 29 subbasins where landslides and debris flows occurred as the research subjects, assigned each historical landslide debris flow to the subbasins, and combined the hydrometeorological data of the 29 subbasins into one hydrometeorological data set for analysis (Berti et al., 2012).

2.2.3 Identification of landslide and debris flow trigger type

To identify the different hydrometeorological conditions when the landslide debris flow occurs, we divided the trigger types of landslide and debris flow in the upper Han River into LLR and SDS with a combined analysis of multiple hydrometeorological variables. Based on the above holistic trigger probability analysis of the temporal development of watershed state before event occurrence, we formulated three individual criteria for the respective classification of LLR and SDS trigger types with soil moisture, potential evapotranspiration and air temperature span (Prenner et al., 2019). An overview of the criteria C1–C6 including the respective absolute values associated with the above percentile ranges are given in Table 1. Note that, to some degree, the epistemic uncertainties from point precipitation measurements and exploiting the low-pass filter properties of watersheds (Euser, Hrachowitz, Winsemius, & Savenije, 2015), precipitation is here not directly used as a criterion. Importantly, the actual threshold values for these criteria were not arbitrarily defined a priori but sampled from a uniform distribution within a range of respective percentiles that were selected from an explorative-iterative modeling process, guided by the outcomes of the probability analysis. For comparison and calculation, the specific classification thresholds were expressed in percentiles. The trigger type of event was classified as LLR or SDS, depending on which trigger met more corresponding criteria (Prenner et al., 2018). Moreover, except for classification of the triggering event days, these criteria also used to identify the non-triggering event days, which is helpful to understand watershed characteristics and predict the occurrence of landslide and debris flows of the prediction model.

2.2.4 The prediction model of landslide and debris flow

Since the occurrence of debris flows in hydrometeorology is time-sensitive (variable in time) (Cardinali, 2000), the critical water input required to trigger the disasters was affected by historical hydrometeorological state. In order to predict the landslide and debris flow on a certain day, we built a predictive model with the Naive Bayes classifier by dividing the trigger classes into LLR type of triggering event, LLR type of non-triggering event, SDS type of triggering event and SDS type of non-triggering event, and collecting corresponding hydrometeorological data sets. The Naive Bayes classifier was used to calculate the relative probability that a certain day during the study period belongs to a certain trigger class with different predictor variables. The type with the highest probability is the type of trigger event on the predicted day.

The Naive Bayes classifier (Heiser, Scheidl, Eisl, Spangl, & Hübl, 2015; Perez, Larranaga, & Inza, 2009; Tsangaratos & Ilia, 2016) is given by equation (7):

$$P\left(c_j \middle| x_1, x_2, \ldots, x_n\right) = \frac{P(c_j) \times \prod_{i=1}^{n} P\left(x_i \middle| c_j\right)}{\sum_{j=1}^{k} P(c_j) \times \prod_{i=1}^{n} P\left(x_i \middle| c_j\right)} \tag{7}$$

where $P(c_j|x_{1, \ldots} x_n)$ is the probability that a signature of the subbasins state, described by n hydrometeorological predictor variables $x_{1, \ldots} x_n$, corresponds to the trigger class c_j. The prior probability $P(c_j)$ is the probability for a given trigger class c_j to occur, regardless of the catchment state. $P(x_i|c_j)$ describes the likelihood that the magnitude of predictor variable x_i was observed in connection with a debris flow event from that trigger class c_j. The denominator acts as a normalizing constant, which ensures that the determined probabilities for the k trigger classes integrate to unity.

To describe the characteristics of the trigger classes (which are a combination of catchment state and rainfall input) as holistically as possible, we used the following set of predictor variables for the predict models: (1) effective precipitation of the day (mm), (2) soil moisture of the day (mm), (3) moisture gradient to the previous day, (4) soil moisture gradient between the first and the second previous day, (5) soil moisture gradient between the second and the third previous day, (6) potential evaporation of the day (mm), (7) potential evaporation of the previous day (mm), (8) potential evaporation gradient between the second and the third previous day, (9) air temperature span of the day (℃). To ensure the accuracy of using the trigger classes, we examined the correlations between any selected variables (Chawla, 2005). The result showed that the highest correlation between individual predictor variables was computed between air temperature span and soil moisture with a R^2 of 0.42, followed by an R^2 of 0.37 for effective precipitation and soil moisture, while all others did not exhibit statistically significant correlations.

The prediction model was trained and evaluated with the above hydrometeorological data sets, which contained 50 event days and 250 non-event days selected randomly to represent the distribution of watershed states over our study period better. Here, we used 40 days with landslide and debris flow occurrence and 200 non-event days as the training data, while the remaining 10 event days and 50 non-event days was the validation data. Each hydrometeorological variable conditional probability $P(x_i|c_j)$ was calculated by the maximum likelihood fitting to the training data, except the precipitation applied a general extreme value distribution for better performance. For training, we repeated this procedure 1,000 times to consider the uncertainties by varying the model training data sets. We evaluated the model's predictive performance by comparing the predicted landslide and debris flow trigger classes with actually observed landslide and debris flow.

To quantify and summarize the model skill, receiver operating characteristic (*ROC*) statistics of true positive rates (TPR) and false positive rates (FPR) were used (Fawcett, 2006). The performance of the prediction model was evaluated based on the ROC curve formed by TPR and FPR computed by comparing the model predictions with the verification data set. As a measure for model performance we use the area under curve (AUC), and a perfect model yields unity for TPR with a value of 0 for FPR. Finally, we selected the best model to evaluate the whole days of the study period to verify the model performance.

3. Results

3.1 SWAT model and hydrometeorological states of the event day

Because of the limitation of the observed runoff data, The SWAT model was calibrated and validated for monthly streamflow only at the basin outlet (Danjiangkou station). After 3000 times optimization, the twelve sensitive parameters with detailed descriptions and fitted values are listed in Table 2. The values of Nash-Sutcliffe Efficiency Coefficient, the Percent Bias and Coefficient of Determination in calibration are 0.81, 7.3 and 0.84, respectively, while the values of Nash-Sutcliffe Efficiency Coefficient, the Percent Bias and Coefficient of Determination in calibration are 0.74, 3.3 and 0.87, respectively. The processes of modeled and observed monthly discharge for eight years (a total of 96 months) are shown in figure 4. The simulated discharges show generally good agreement with the observations implying that the calibrated SWAT model is good enough to describe the hydrological condition of the study area and the optimized model parameters can be used to simulate the daily flow for the following research.

In figure 5a-b, we provide a detailed view of the subbasin state around the two selected rainstorm events in the study area. The X-coordinate shows the relative sequence of events with respect to the event date (unit: day); minus sign means days before the event date while 0 refers to the event date; y-coordinate refers to simulation values of corresponding hydrological variables. For the event on 16th September 2011 described as LLR trigger event (Figure 5a), it rained for four days before the event, including two days with high precipitation intensity (above 50 mm d^{-1}) and one day with moderate precipitation intensity (about 20 mm d^{-1}), which caused a large amount of precipitation input before the disaster occurs. On the contrary, evaporation in the region had been at a low level with no more than 2.5 mm. As heavy rain continues pounding the subbasin as well as low evaporation, the soil moisture level was still rising and up to 100 mm d^{-1}, and the mean value of soil water content was above 95.4 mm. Therefore, the landslide and debris flows would happen even though almost no rainfall was observed on 16th September 2011.

Another example is the event 9th July 2012 described as SDS trigger event (Figure 5b), there was little rain in the days before the event and the total rainfall in the four days before the event was 48 mm with the low intensity below 30 mm d^{-1}. In addition, the evaporation was generally at a high level although the trend was down and the total evaporation in the four days before the event was 17.44 mm with over 4.36 mm on average. Because of the little rain but the high evaporation, the soil moisture of the subbasin kept a relatively low level, which the mean value of soil water content was about 68.3 mm in the four days before the event. However, the rainfall with high intensity more than 80 mm d^{-1} also could cause the landslide and debris flows although not so much precipitation accumulation before 9th July 2012. Since no snowmelt was found in the subbasin debris flow occurs during the study period, the type of intense snowmelt trigger (SM) was not considered in this paper (Prenner et al., 2018).

3.2 The trigger probability of hydrometeorological variables

The triggering probabilities based on the magnitude of different hydrometeorological variables of the day with a landslide and debris flow occurrence in the event area (29 subbasins) and their cumulative hydrometeorological variables on the 12 days before the day of the event are analyzed. The trigger probability shows a volatility increasing trend as the value of the variable increases (Figure 6). For example, the actual evapotranspiration reaches a high probability of more than 1% at the 20th percentile, after which the probability decreases, and a high probability again at the 70th percentile. The similar trends are observed in soil water content, potential evapotranspiration and groundwater discharge, while a single increasing trend of trigger probabilities is observed with increasing net water yield and surface runoff. In addition, with the accumulation of time, the probability distribution of triggering is gradually stable, such as soil water content, potential evapotranspiration, actual evapotranspiration, etc.

Trigger probabilities for landslide and debris flows conditional to gradients (divided into positive and negative absolute gradients) of potential evapotranspiration (PE), soil moisture (SW), and groundwater discharge (GW_Q) fraction up to 12 days prior to the event (always with respect to the event day) are shown in figure 7. The trigger probability either in positive gradient or negative gradient increase as the percentile of each variable increase to a higher range. For soil water content, we find increased trigger probabilities from 0.5% to 1% when positive gradients 2 days before the event are in the high percentile range of 60-80%, as well as the negative gradient with the same tendency but a lower probability. At the same time, the trigger probability of potential evapotranspiration is increased significantly as the positive and negative gradients of the four days prior to the event are in the high percentile range. In summary, it suggests that there are different watershed states developed about four days before the event. The evapotranspiration was significantly reduced to increase the soil moisture pointing to LLR, while the evapotranspiration was increased to reduce the soil moisture pointing to SDS. On the whole, these sometimes opposing trends of increased landslide and debris flows probabilities conditional to different hydrometeorological variables show that different weather conditions are connected with the landslide and debris flows occurrence in our study region.

3.3 Temporal characteristics of triggers and their watershed states at the event day

The procedure described in section 3.3 was used to separate the different trigger type, which allows us to identify only the most likely trigger but the distinct threshold values for the different criteria. The results suggest that landslides and debris flows in the selected subbasins region were triggered on 33 days by LLR and on 17 days by SDS, and temporal occurrence characteristic is obvious (Figure 8). In a year, landslide and debris flows mainly occur in July, August and September, with the most occurring in July, accounting for 68% (34 days), followed by September, accounting for 22% (11 days). In addition, triggers show different dominant type in time, such as the LLR in July was up to 79%, while in August and September, SDS gradually became the dominant trigger factor, accounting for 64% in September. This is basically consistent with the rule found in Prenner's previous research (Prenner et al., 2019).

The differences (p-value of Wilcoxon rank sum test < 5%) in median value of each variable per trigger class are statistically significant for the study regions (Figure 9). Here, the data of all variables were normalized for the convenience of comparison. For example, the median accumulated precipitation of LLR-trigger events was higher than those of SDS-trigger events, especially the median accumulative precipitation of the event day was 127% higher than those of SDS-type events. Meanwhile the median accumulated soil water content a day before the event was 57% higher than that of SDS-type events. In addition, the similar trends were displayed on the actual evapotranspiration, mean temperature and potential evapotranspiration. On the contrary, the average potential evaporation on the day of LLR-type events was 94% lower than that of SDS-type events. It was noted that the differences between the two trigger types gradually decrease, as the time before the event prolongs. These findings indicate that the classification of different trigger types in the study area can better analyze the hydrologic state when landslide debris flows occur.

3.4 Validation and application of the landslide and debris flow predictive model

Through the most likely trigger can be identified based on the procedure described in section 3.3, the threshold values for the different criteria are not certain to predict whether landslide and debris flow will occur. Thus, the prediction models described in section 3.4 is used to help predict landslide and debris flow. After optimization with 1000 sets of calibration and verification samples obtained by random sampling, the best performance model determined by comparing the true positive rate and false positive rate was selected to perform experiments to test predictions. The results show that the AUC value is up to 0.85 and the true positive rate TPR and false positive rate FPR of the model with the best performance are 0.9 and 0.177 respectively, indicating that this model possesses good simulation capability.

3.4.1 The landslide and debris flow prediction and their temporal characteristics

The selected prediction model was applied to typical subbasins of landslide and debris flow respectively to predict days with landslide and debris flow occurrence for 2010-2012. The results are shown in the figure 10，where the horizontal axis is the day in a year, and the vertical axis shows the typical subbasin ID numbers of landslide and debris flow. Comparing with the observed events, the true positive rate and false positive rate of landslide and debris flow predicted by the model reach 0.64 and 0.023 respectively. For 2010, 21 landslides and debris flows are predicted successfully, while nine recorded landslides and debris flows are not recognized by the prediction model. Among the nine observed but not predicted landslides and debris flows, some events do not have the hydrological conditions that trigger landslides and debris flows, like the event on 4th August 2010 which occurred near the exit of the Xikang Railway Tunnel and inside the number 103 provincial highway in Liangheguan Village, Xiaohe Town, Xunyang County, Ankang City, within subbasin number 39 in our study area. The total rainfall in the six days before that event was 20.91 mm with the high intensity evaporation up to 6.24 mm on average and decreasing soil moisture content. What's more, no rainfall was observed that day implying that the possibility of landslide debris flow is very low. Thus, the non-rainfall factors (such as human activity) may be the main triggers for such events. Besides, the most of the prediction errors lie in the earlier or later occurrence time predicted by the model, which indicates that the model is not sensitive enough to identify the hydrological state of the landslide debris flows due to the limited historical event data.

The days of landslide and debris flow mainly gathered between May (about the 120th day) and mid-September (about the 260th day), among which the number of days of landslide and debris flow occurred from July to September is the largest. Moreover, the LLR-trigger days are mainly concentrated in August (about 220th day), while the SDS-trigger days are mainly concentrated in July and September (about 181th and 243th day). This is consistent with the annual distribution of observed landslides and debris flows and the precipitation in the study area from July to September accounts for about 57% of the annual precipitation (Zhang et al., 2019).

3.4.2 Spatial distribution of predicted landslide and debris flow

Besides temporal distribution, the spatial variations among the subbasins for the occurrence of specific trigger types were analyzed by choosing four subbasins (the numbers are sub_41, sub_42, sub_50 and sub_87) with high frequency of landslide debris flows which are located in different geographical locations. For number sub_41 covering a total area of nearly 128.3 km^2 and 495 -1894 meters in height, the main land use type is the agricultural land, account for about 66.2% of the subbasin area, as well as other land use types are urban land, account for about13.7%. For number sub_42 covering a total area of nearly 1152.97 km^2 and 155-1490 meters in height, the three main land use types are forest, pasture and agricultural land, account for about 58.0%, 25.7% and 15.4%, respectively. For number sub_50 covering a total area of nearly 646.88 km^2 and 494-2328 meters in height, it also has three main land use types including pasture, agricultural and forest land, account for about 39.4%, 31.0% and 28.0% of the subbasin area, respectively. And sub_87 with a total area of about 63.18 km^2 and 318 - 1300 meters in height have two main land use types including pasture and agricultural land, account for about 61.7% and 36.4% of the subbasin area, respectively. Above all, the four subbasins have different characteristics in land use and soil type.

The results are shown in the figure 11, the LLR-trigger days are dominant in sub_41, sub_50, while the SDS-trigger days make up the majority in sub_42 and sub_87. As the different proportions of the main land use, we attribute the more LLR-trigger days to the larger forest area than agricultural area, while the more SDS-trigger days may due to the larger agricultural area than forest area, which needs further proof through more detailed evidence. On the other side, we found that the number of landslide and debris flow day are different with the same rainfall scenarios, which implies that the geological conditions in different areas have a certain influence on triggering landslides and debris flows. That is, although the landslide and debris flow prediction model of multi-meteorological hydrology does not directly use geological parameters for modeling, the geographical distribution characteristic of landslide and debris flows occurrence still could be identified by using the semi-distributed hydrological model (SWAT) and hydrometeorological variables, making up for the lack of geological parameters to some extent.

4. Discussion

China has been suffering a lot of loss of life and property due to landslides and debris flows, and the number of disasters will increase further (Gariano & Guzzetti, 2016) because of the excessive exploitation of natural resources and vegetation damage (Nadim, Kjekstad, Peduzzi, Herold, & Jaedicke, 2006), land urbanization, especially mountainous urbanization (LiGerui, LeiYalin, YaoHuajun, WuSanmang, & GeJianping, 2017), as well as the increase in extreme precipitation (Fu et al., 2013). The interacting factors between the precipitation and topography, soil, lithology, vegetation and population density are more closely related to the spatial distribution of fatal landslides than each individual factor (Lin & Wang, 2018). This paper attempt predict the occurrence of the landslides and debris flows based on the different hydrological subcatchment states evaluated by the SWAT model. The hydrologic response unit divided by spatial analysis used in SWAT model for hydrological calculation can make up for the lack of geological parameters to some extent.

The differences of most hydrometeorological variables between the trigger types are statistically significant, implying that the hydrometeorological formation processes (Ford, Rapp, Quiring, & Blake, 2015; Rulfova & Kysely, 2013) leading to landslides and debris flows can be clearly identifiable at subbasin scales. In addition, the annual time distribution of event days in different trigger types is consistent with local meteorological and precipitation trends (Qiu, Cui, Hu, et al., 2019; Qiu, Cui, Yang, et al., 2019). With 1000 times training, the AUC value of the prediction model is up to 0.85, and the true positive rate TPR and false positive rate FPR of the model with the best performance are 0.9 and 0.177 respectively, indicating that this model possesses good prediction capability. Therefore, combined with the hydrological model, it can better reflect the hydrological basin state with landslides and debris flows and give early warning of possible disaster without too many geological parameters. Our application test also show that the land cover has a certain effect on the triggering of landslide debris flow (Lin & Wang, 2018) and the specific influence mechanism can be further explored in subsequent studies.

For our study, we mainly study the hydrometeorological variables with similar geographical conditions, but the other factors like the slope failure (Fan et al., 2019), intensive channel erosion and human activity were not considered and may partially explain that some observed landslides and debris flows are not predicted by our model. Furthermore, the limited data used in this study will also increase the uncertainty of prediction model. On the one hand, the 50 landslide and debris flow information used in this study were obtained the disaster relief yearbook of Shaanxi province, where only with large scale or casualty disasters are recorded. But according to statistics, there were 1,891 landslide and debris flow events without specifying the date and place in the study area during the study period. So the incomplete landslide and debris flows may affect the prediction performance of the model. On the other hand, we only have hydrological stations all over the study region with a total area of 95200 km^2, so the spatial distribution of precipitation is not accurate enough. Thus, adding more information about small and medium landslide and debris flows into the model calibration data set as well as improving the resolution of precipitation data can improve the accuracy of model prediction.

5. Conclusion

In this study, the hydrometeorological conditions of the upper and middle reaches of the Han River were investigated by the process-based, semidistributed hydrological model SWAT. Then, based on this information of 29 typical subbasins, we divide the landslides and debris flows into LLR-trigger and SDS-trigger events. In order to predict the occurrence of landslides and debris flows, we established a predictive model with the nine hydrometeorological variables of different type trigger or non-trigger event day using a Naive Bayes classifier. Finally, we used the model to evaluate the occurrence of landslide and debris flow in the typical subbasin of landslide and debris flow from 2010 to 2012 and carried out verification analysis. The main findings can be summarized as follows:

(1) Validation of the SWAT model suggests that it performs well in simulating streamflow, and can thus be used to estimate the hydrological condition for the following research.

(2) We found that there is no single watershed state when landslides and debris flows occur, and historical hydrological variables and their development provide important information for triggering debris flows, through rainfall is the most important factor for triggering debris flows. Furthermore, there are significant differences in various hydrological variables between two trigger types, which may explain the uncertainty of the traditional I-D threshold prediction.

(3) The landslides and debris flows in the selected subbasins regions are triggered on 33 days by LLR and on 17 days by SDS. Specifically, landslides and debris flows are mainly concentrated between July and September, with the most occurrences in July, followed by September. The seasonal distribution of different trigger events is different. LLR type landslide

and debris flow account for a large proportion in July, while SDS type landslide and debris flow occur more frequently in September.

(4) The AUC value of the prediction model is up to 0.85, and it can capture most of the landslide debris flow with the true positive rate and false positive rate of 0.64 and 0.023 respectively. The temporal distribution of the two triggering-event predicted by model is consistent with the annual distribution of precipitation. Besides, there are spatial variations of the specific trigger types in the different subbasins, which may attribute to the different land over.

Acknowledgements

Funding: This work was supported by the National Natural Science Foundation of China [grant numbers 51679218, 51879239, 51879252, 51709238].

Data Availability Statement

The acquisition of datasets has been introduced in “Available data” and is available from the corresponding author upon reasonable request.

Reference

Aleotti, P. (2004). A warning system for rainfall-induced shallow failures. *Engineering Geology, 73*(3), pp. 247-265.

Berti, M., Martina, M. L. V., Franceschini, S., Pignone, S., Simoni, A., & Pizziolo, M. (2012). Probabilistic rainfall thresholds for landslide occurrence using a Bayesian approach. *Journal of Geophysical Research-Earth Surface, 117*, p. 20.

Bui, D. T., Shirzadi, A., Shahabi, H., Geertsema, M., Omidvar, E., Clague, J. J., . . . Lee, S. (2019). New Ensemble Models for Shallow Landslide Susceptibility Modeling in a Semi-Arid

Watershed. *Forests*, *10*(9).

Caine, N. (1980). The Rainfall Intensity: Duration Control of Shallow Landslides and Debris Flows. *Geografiska Annaler. Series A, Physical Geography, 62*(1/2), pp. 23-27.

Camilo, D. C., Lombardo, L., Mai, P. M., Dou, J., & Huser, R. (2017). Handling high predictor dimensionality in slope-unit-based landslide susceptibility models through LASSO-penalized Generalized Linear Model. *Environmental Modelling & Software, 97*, pp. 145-156.

Cardinali, A. (2000). *Complex ray tracing method in high harmonic fast wave propagation and absorption*. Bologna: Editrice Compositori.

Chawla, N. V. (2005). Data Mining for Imbalanced Datasets: An Overview. In O. Maimon & L. Rokach (Eds.), *Data Mining and Knowledge Discovery Handbook* (pp. 853-867). Boston, MA: Springer US.

Ciavolella, M., Bogaard, T., Gargano, R., & Greco, R. (2016). Is there Predictive Power in Hydrological Catchment Information for Regional Landslide Hazard Assessment? *Procedia Earth and Planetary Science, 16*, pp. 195-203.

Cong, W., Pan, M., Li, T., Ren, Q., & Li, R. (2006). Quantitative analysis of critical rainfall-triggered debris flow (in chinese). *Chinese Journal of Rock Mechanics and Engineering*(S1), pp. 2808-2812.

Crozier, & Michael, J. (1986). *Landslides : causes, consequences & environment*. London: Croom Helm Pub.

Crozier, & Michael, J. (1999). Prediction of rainfall-triggered landslides: A test of the antecedent water status model. *Earth Surface Processes and Landforms, 24*(9), pp. 825-833.

Dai, F. C., & Lee, C. F. (2002). Landslide characteristics and slope instability modeling using GIS, Lantau Island, Hong Kong. *Geomorphology, 42*(3), pp. 213-228.

Dou, J., Yunus, A. P., Bui, D. T., Merghadi, A., Sahana, M., Zhu, Z. F., . . . Pham, B. T. (2020). Improved landslide assessment using support vector machine with bagging, boosting, and stacking ensemble machine learning framework in a mountainous watershed, Japan. *Landslides, 17*(3), pp. 641-658.

Euser, T., Hrachowitz, M., Winsemius, H., & Savenije, H. (2015). The effect of forcing and landscape distribution on performance and consistency of model structures: DISTRIBUTION OF FORCING AND MODEL STRUCTURES. *Hydrological Processes*.

Fan, W., Lv, J., Cao, Y., Shen, M., Deng, L., & Wei, Y. (2019). Characteristics and block kinematics of a fault-related landslide in the Qinba Mountains, western China. *Engineering Geology, 249,* pp. 162-171.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters, 27*(8), pp. 861-874.

Ferro, V., Carollo, F. G., & Serio, M. (2020). Establishing a threshold for rainfall induced landslides by a kinetic Energy - Duration relationship. *Hydrological Processes*.

Floris, M., & Bozzano, F. (2008). Evaluation of landslide reactivation: A modified rainfall threshold model based on historical records of rainfall and landslides. *Geomorphology, 94*(1-2), pp. 40-57.

Ford, T. W., Rapp, A. D., Quiring, S. M., & Blake, J. (2015). Soil moisture-precipitation coupling: observations from the Oklahoma Mesonet and underlying physical mechanisms. *Hydrology and Earth System Sciences, 19*(8), pp. 3617-3631.

Frattini, P., Crosta, G., & Sosio, R. (2009). Approaches for defining thresholds and return periods for rainfall-triggered shallow landslides. *Hydrological Processes, 23*(10), pp. 1444-1460.

Fu, G., Yu, J., Yu, X., Ouyang, R., Zhang, Y., Wang, P., . . . Min, L. (2013). Temporal variation of extreme rainfall events in China, 1961-2009. *Journal of Hydrology, 487*, pp. 48-59.

Gariano, S. L., & Guzzetti, F. (2016). Landslides in a changing climate. *Earth-Science Reviews, 162*, pp. 227-252.

Glade, T., Crozier, M., & Smith, P. (2000). Applying probability determination to refine landslide-triggering rainfall thresholds using an empirical "Antecedent Daily Rainfall Model". *Pure and Applied Geophysics, 157*(6-8), pp. 1059-1079.

Heiser, M., Scheidl, C., Eisl, J., Spangl, B., & Hübl, J. (2015). Process type identification in torrential catchments in the eastern Alps. *Geomorphology, 232,* pp. 239-247.

Hong, H. Y., Liu, J. Z., Zhu, A. X., Shahabi, H., Pham, B. T., Chen, W., . . . Bui, D. T. (2017). A novel hybrid integration model using support vector machines and random subspace for weather-triggered landslide susceptibility assessment in the Wuning area (China). *Environmental Earth Sciences, 76*(19).

Hu, D. Y., Li, J., Chen, Y. H., & Zhang, J. S. (2007). GIS-based Landslide Spatial Prediction Methods, a Case Study in Cameron Highland, Malaysia (in Chinese). *Journal Of Remote Sensing, 11*(6), pp. 852-859.

Huang, X. H., Li, Z. Y., Yu, D., Xu, Q., Fan, J. Y., Hao, Z., & Niu, Y. P. (2017). Evolution of a giant debris flow in the transitional mountainous region between the Tibetan Plateau and the Qinling Mountain range, Western China: Constraints from broadband seismic records. *Journal of Asian Earth Sciences, 148*, pp. 181-191.

Jaiswal, P., & van Westen, C. J. (2009). Estimating temporal probability for landslide initiation along transportation routes based on rainfall thresholds. *Geomorphology, 112*(1-2), pp. 96-105.

Kanungo, D. P., & Sharma, S. (2014). Rainfall thresholds for prediction of shallow landslides around Chamoli-Joshimath region, Garhwal Himalayas, India. *Landslides, 11*(4), pp. 629-638.

Lainas, S., Sabatakakis, N., & Koukis, G. (2016). Rainfall thresholds for possible landslide initiation in wildfire-affected areas of western Greece. *Bulletin of Engineering Geology and the Environment, 75*(3), pp. 883-896.

LiGerui, LeiYalin, YaoHuajun, WuSanmang, & GeJianping. (2017). The influence of land urbanization on landslides: An empirical estimation based on Chinese provincial panel data. *Science of the Total Environment, 595*, pp. 681-690.

Lin, Q., & Wang, Y. (2018). Spatial and temporal analysis of a fatal landslide inventory in China from 1950 to 2016. *Landslides, 15*(12), pp. 2357-2372.

Martinotti, M. E., Pisano, L., Marchesini, I., Rossi, M., Peruccacci, S., Brunetti, M. T., . . . Guzzetti, F. (2017). Landslides, floods and sinkholes in a karst environment: the 1-6 September 2014 Gargano event, southern Italy. *Natural Hazards and Earth System Sciences, 17*(3), pp. 467-480.

Mostbauer, K., Kaitna, R., Prenner, D., & Hrachowitz, M. (2018). The temporally varying roles of rainfall, snowmelt and soil moisture for debris flow initiation in a snow-dominated system. *Hydrology and Earth System Sciences, 22*(6), pp. 3493-3513.

Nadim, F., Kjekstad, O., Peduzzi, P., Herold, C., & Jaedicke, C. (2006). Global landslide and avalanche hotspots. *Landslides, 3*(2), pp. 159-173.

Nicolussi, K., Spotl, C., Thurner, A., & Reimer, P. J. (2015). Precise radiocarbon dating of the giant Weis landslide (Eastern Alps, Austria). *Geomorphology, 243*, pp. 87-91.

Perez, A., Larranaga, P., & Inza, I. (2009). Bayesian classifiers based on kernel density estimation: Flexible classifiers. *International Journal of Approximate Reasoning, 50*(2), pp. 341-362.

Pradhan, B., & Youssef, A. M. (2010). Manifestation of remote sensing data and GIS on landslide hazard analysis using spatial-based statistical models. *Arabian Journal of Geosciences, 3*(3), pp. 319-326.

Prenner, D., Hrachowitz, M., & Kaitna, R. (2019). Trigger characteristics of torrential flows from high to low alpine regions in Austria. *Science of the Total Environment, 658*, pp. 958-972.

Prenner, D., Kaitna, R., Mostbauer, K., & Hrachowitz, M. (2018). The Value of Using Multiple Hydrometeorological Variables to Predict Temporal Debris Flow Susceptibility in an Alpine Environment. *Water Resources Research, 54*(9), pp. 6822-6843.

Qiu, H., Cui, Y., Hu, S., Yang, D., Pei, Y., & Yang, W. (2019). Temporal and spatial distributions of landslides in the Qinba Mountains, Shaanxi Province, China. *Geomatics, Natural Hazards and Risk, 10*(1), pp. 599-621.

Qiu, H., Cui, Y., Yang, D. D., Pei, Y. Q., Hu, S., Ma, S. Y., . . . Liu, Z. J. (2019). Spatiotemporal Distribution of Nonseismic Landslides during the Last 22 Years in Shaanxi Province, China. *Isprs International Journal of Geo-Information*, *8*(11).

Rulfova, Z., & Kysely, J. (2013). Disaggregating convective and stratiform precipitation from station weather data. *Atmospheric Research, 134*, pp. 100-115.

Schiliro, L., Cevasco, A., Esposito, C., & Mugnozza, G. S. (2018). Shallow landslide initiation on terraced slopes: inferences from a physically based approach. *Geomatics Natural Hazards & Risk*, *9*(1), pp. 295-324.

Tan, Q., Sassa, K., Dang, K., Konagai, K., & Sato, G. (2020). Estimation of the past and future landslide hazards in the neighboring slopes of the 2016 Aranayake landslide, Sri Lanka. *Landslides*(3–4).

Tian, Y., Xiao, C., Liu, Y., & Wu, L. (2008). Effects of raster resolution on landslide susceptibility mapping: A case study of Shenzhen. *Science in China Series E-Technological Sciences, 51*, pp. 188-198.

Tsangaratos, P., & Ilia, I. (2016). Comparison of a logistic regression and Naive Bayes classifier in landslide susceptibility assessments: The influence of models complexity and training dataset size. *Catena*, *145*, pp. 164-179.

Vessia, G., Curzio, D. D., Chiaudani, A., & Rusi, S. (2020). Regional rainfall threshold maps drawn through multivariate geostatistical techniques for shallow landslide hazard zonation. *Science of the Total Environment, 705*.

Wang, S., Zhang, K., Beek, v., H., L. P., Tian, X., & Bogaard, T. A. (2020). Physically-based landslide prediction over a large region: Scaling low-resolution hydrological model results for high-resolution slope stability assessment. *Environmental Modelling & Software, 124*, p. 104607.

Zhang, K., Wang, S., Bao, H., & Zhao, X. (2019). Characteristics and influencing factors of rainfall-induced landslide and debris flow hazards in Shaanxi Province, China. *Natural Hazards and Earth System Sciences, 19*(1), pp. 93-105.

Figure legends

Figure 1. Overview of the upper Han River basin.

Figure 2. Landuse (a) and soil map (b) of the upper Han River basin.

Figure 3. Structure of the research methods.

Figure 4. The monthly rainfall and runoff (measured runoff (black) and simulated runoff (red) during calibration (2006-2010 (0-60 month)) and validation (2011-2013 (61-96 month)) periods.

Figure 5. Precipitation, actual evapotranspiration, and soil water content around the landslides and debris flow event on 9th July 2012, which may be interpreted as SDS (short-duration storm) trigger type (a), on 16th September 2011, which suggests being an SDS (long-lasting rainfall) trigger type (b).

Figure 6. Mean trigger probability of cumulative hydrometeorological variables in previous t days including soil water content (SW), groundwater discharge (GW_Q), net water yield (WYLD), potential evapotranspiration (PE), actual evapotranspiration (E), and surface runoff (SURQ).

Figure 7. Mean trigger probability of the cumulative gradient of the variable over t days. A positive gradient means that the parameter value increases over previous t days; a negative gradient means that the parameter value decreases over previous t days.

Figure 8. Temporal distribution of LLR and SDS triggers in study regions.

Figure 9. The distribution of hydrometeorological variables between triggers.

Figure 10. Daily predicted trigger classes exemplarily for 29 selected subbasins from 2010 to 2012.

Figure 11. Temporal distribution of LLR and SDS triggers in different regions.