**Jung-Hyun Lee · Hanbeen Kim · Hyuck-Jin Park · Jun-Haeng Heo**

# Temporal prediction modeling for rainfall-induced shallow landslide hazards using extreme value distribution

**Abstract** As the frequency and intensity of heavy rainfall increase, the frequency of extreme rainfall-induced landslides also increases. Thus, the importance of accurate assessment of extreme rainfall-induced landslide hazard increases. Landslide hazard assessment requires estimations of two components: spatial probability and temporal probability. While various approaches have been successfully used to estimate spatial landslide susceptibility, fewer studies have addressed temporal probability and, consequently, a commonly accepted method does not exist. Prior approaches have estimated temporal probability using frequency analysis of past landslides or landslide triggering rainfall events. Hence, a large amount of information was required: sufficiently complete historical data on recurrent landslides and repetitive rainfall events. However, in many cases, it is difficult to obtain such complete historical data. Therefore, this study developed a new approach that can be applied to an area where incomplete data are available or where nonrepetitive landslide events have occurred. To evaluate the temporal probability of landslide occurrence, the developed approach adopted extreme value analysis using the Gumbel distribution. The exceedance probability of a rainfall threshold was evaluated, using the Gumbel model, with 72-h antecedent rainfall threshold. This probability was then considered to be the temporal probability of landslide occurrence. The temporal probability of landslides was then integrated with landslide susceptibility results from a multi-layer perceptron model. Consequently, the landslide hazards for different future time periods, from 1 to 200 years, were estimated.

**Keywords** Temporal probability · Gumbel distribution · Extreme value distribution · Rainfall-induced landslides · Multi-layer perceptron · Landslide hazard

## Introduction

As the frequency and intensity of heavy rainfall increase with the changing climate, extreme rainfall events, which are characterized by short duration and high intensity, become one of the major triggers of shallow landslides and debris flows, in Korea and other countries (Chou et al. 2012; Wu et al. 2011; Dahal et al. 2009; Pradhan and Kim 2015; Tsuchida et al. 2015; Vasu et al. 2016;

**J.-H. Lee · H.-J. Park (✉)**
Department of Geoinformation Engineering,
Sejong University,
Seoul, Republic of Korea
Email: hjpark@sejong.ac.kr

**H. Kim · J.-H. Heo**
School of Civil and Environmental Engineering,
Yonsei University,
Seoul, Republic of Korea

Kumar et al. 2017; Wu 2017). Extreme rainfall-induced landslides are powerful and can cause enormous damage and catastrophic fatalities. In Korea, shallow landslides induced by extreme rainfall events have, in recent decades, caused severe damage to property, economic losses, and casualties (National Institute for Disaster Prevention 2009; Park et al. 2013; Oh and Park 2014). Therefore, a process for recognizing landslide-prone areas and evaluating the landslide hazard caused by extreme rainfall is badly needed.

Landslide hazard is defined as the probability of a potentially damaging landslide occurrence in a specific period of time and in a given area (Varnes 1984; van Westen et al. 2006). It means that in the process of assessing landslide hazards, one has to predict "where" a landslide may occur (spatial probability) and "when" it may occur (temporal probability). Landslide susceptibility is the likelihood of a landslide occurring in an area. It depends on local terrain conditions and estimates "where" landslides are likely to occur (Guzzetti et al. 2005; Reichenbach et al. 2018). Various approaches have been developed, and extensively published, for the spatial assessment of landslide susceptibility. Compared with susceptibility analysis, only limited research has sought to estimate the temporal probability of landslide occurrence (Guzzetti et al. 2005, 2006; Corominas and Moya 2008; Zêzere et al. 2008; Jaiswal and van Westen 2009; Jaiswal et al. 2010, 2011; Das et al. 2011; Martha et al. 2013; Tien Bui et al. 2013; Afungang and Bateira 2016; Vasu et al. 2016; Dikshit et al. 2020). Two main approaches have been commonly used to assess the temporal probability of the future occurrence of landslides: the physically based method (or geomechanical approach) and statistical analysis of past landslide events (Corominas and Moya 2008; Corominas et al. 2014). The physically based method considers the present slope conditions and evaluates landslide potential using stability analysis and numerical modeling. This approach may also couple the slope stability analysis with hydrological models to estimate the effect of rainfall on slope stability. It has been applied to regional-scale analyses of rainfall-induced shallow landslide susceptibility with simplified hydrological methods in a GIS platform (Crosta and Frattini 2003; Godt et al. 2008; Park et al. 2013; Chen and Zhang 2014; Corominas et al. 2014; Lee and Park 2016; Salciarini et al. 2017; Gutiérrez-Martín 2020). In addition, a probabilistic approach has been used to properly deal with uncertainties in the geotechnical and hydrological input parameters, caused by the limited and incomplete information common in a regional study. However, the deterministic approach does not easily accommodate such uncertainties especially when large spatial datasets must be obtained and processed (Park et al. 2013; Rossi et al. 2013; Raia et al. 2014; Lee and Park 2016; Chae et al. 2017; Salciarini et al. 2017). In the evaluation of temporal probability using a physically based model, the computational demands increase significantly when applied in a dynamic time-dependent modeling framework, especially in a

broad study area where high-resolution DEM and time-variant rainfall patterns are used in the modeling. These extremely high computational requirements bring forth the reluctance to use the physically based models for temporal probability estimation over large areas (Rossi et al. 2013; Vasu et al. 2016; Canli et al. 2018a, b).

Statistical analysis of historical landslide records covering a long time span support estimation of the temporal probability of landslide occurrence (Brabb 1985; Guzzetti et al. 2002, 2005, 2006; Jaiswal et al. 2011; Das et al. 2011; Motamedi and Liang 2014). In this approach, the return periods of landslide events are determined from the historical landslide records, and then, the expected frequency of future landslides is obtained. Using historical data such as a multi-temporal landslide inventory, the annual probability of landslide events and the exceedance probability, which is the probability of occurrence of one or more landslides at any given time, can be calculated. The statistics involved in the approach are simple, and results are easy to implement (Jaiswal et al. 2010). However, it is quite difficult to acquire sufficiently complete multi-temporal landslide inventories for long time spans at a regional scale. In addition, because this approach is based on the historical recurrences of landslides, it is only valid for repetitive landslide conditions and is not appropriate for the prediction of unique events (Guzzetti et al. 1999; Cascini et al. 2005).

Therefore, an alternative indirect method, using the frequency of landslide triggers (such as extreme rainfall), has been suggested to estimate the temporal probability of landslides. In this approach, the temporal probability of a landslide triggering rainfall event is considered to also be the temporal probability of landslide occurrence. The advantage of this approach is that it does not require a complete multi-temporal landslide inventory because the return period of the landslides is assumed to be that of the rainfall, which can be evaluated from historical rainfall records. In this approach, a reliable relationship between rainfall and the occurrence of landslides, which is a rainfall threshold, needs to be established (Jaiswal and van Westen 2009; Jaiswal et al. 2010). The rainfall threshold is the minimum amount or duration of rainfall required to trigger landslides. Determination of the rainfall thresholds for landslide occurrence is a basic and prerequisite step for the evaluation of exceedance probability. Various methods have been proposed to determine rainfall thresholds; these tend to be either process based (physical) or empirically based (historical) (Corominas 2000; Crosta and Frattini 2001; Aleotti 2004; Wieczorek and Glade 2005; Guzzetti et al. 2007). Process-based models use a physical slope model coupled with a hydrological model to calculate the rainfall amount needed to trigger slope failures, and hence the anticipated location and time of landslides. On the other hand, empirical rainfall thresholds are defined by analysis of past rainfall events that have resulted in landslides. Recently, automated methods have been used to determine empirical rainfall thresholds for shallow landslide initiation (Vessia et al. 2014, 2016; Melillo et al. 2015), while geostatistical tools have been used to evalaute map-based empirical rainfall thresholds (Vessia et al. 2020).

Once the rainfall threshold is determined, the probability that a rainfall event will exceed the threshold can be obtained (Chleborad et al. 2006). Subsequently, the exceedance probability of rainfall events is evaluated using binomial or Poisson distribution models. However, this approach also requires the assumption that recurrent rainfall events trigger landslides. Thus, the mean recurrence interval, which is a required parameter for the evaluation of the exceedance probability using either the binomial or Poisson distribution model, should be estimated using the historical rainfall recurrence records for periods between landslide triggering events. In addition, the threshold for landslide initiation can be estimated using the correlation between rainfall events and recurrent landslide events. However, in areas where historical data are limited, for example, where an initial landslide event or a nonrecurrent landslide event has occurred under extreme rainfall, it is difficult to estimate either the mean recurrence interval or the rainfall threshold. Consequently, it is impossible to obtain the temporal probability of future landslide occurrence in such an area.

The main objective of this study was to assess the temporal probability of future landslide occurrence for a study area with limited historic landslide records. The Jinbu area, the study area of this research, experienced a large number of landslides, under extreme rainfall, in 2006. However, there are no landslide records from before or after the 2006 events. Therefore, a multi-temporal landslide inventory is not available in this area, and there are no records of recurrent landslide triggering rainfall events. Consequently, the previously mentioned approach to temporal probability is not appropriate for this study area. This study, therefore, developed a new approach to assessment of the temporal probability of future landslide occurrence using the limited rainfall records and landslide inventory. In particular, extreme value theory, which is commonly used for the analysis of extreme events, was used to evaluate the temporal probability of the future landslide occurrence. Then, the spatial probability of landslide was evaluated using a multi-layer perceptron; the temporal probability was then integrated with this spatial probability in order to estimate landslide hazard.

### Study area and landslide inventory

To apply the proposed approach to a practical situation, a study area in Jinbu, Gangwon-do, Korea, was selected (Fig. 1). This area experienced a great number of landslides related to an extreme rainfall event on July 14–16, 2006. The study area lies between 37° 34′ 42.93″ and 37° 39′ 16.01″ north latitude, and between 128° 29′ 9.73″ and 128° 36′ 35.95″ east longitude. The total area is 64.74 km², and the extents are 10.81 km east–west and 8.6 km south–north. The altitude of the study area is between 495 m and 1385 m above sea level, and more than 78% of it is mountainous terrain.

Geologically, this area is based on Precambrian biotite gneiss, which is exposed in the southwest and on which discontinuous Jeongseon limestone and Nokan series sandstone are located. There has been an extensive intrusion of Jurassic Imgye granite, widely distributed in the study area, with deposits of quaternary alluvium along the streams and in lowlands (Geological Society of Korea 1962) (Fig. 2).

In this study, landslide inventory data were obtained and used for the spatial landslide susceptibility analysis. Because the study area is too large for a complete field survey, landslide locations were determined using aerial photographs with 0.5-m resolution, taken before and after the landslides. Related information is available in the authors' prior research (Park et al. 2013). Based on these photographs, a point vector layer was constructed by digitizing the landslide locations. A total of 1035 landslide locations were digitized, and a landslide inventory map was compiled (Fig. 1).
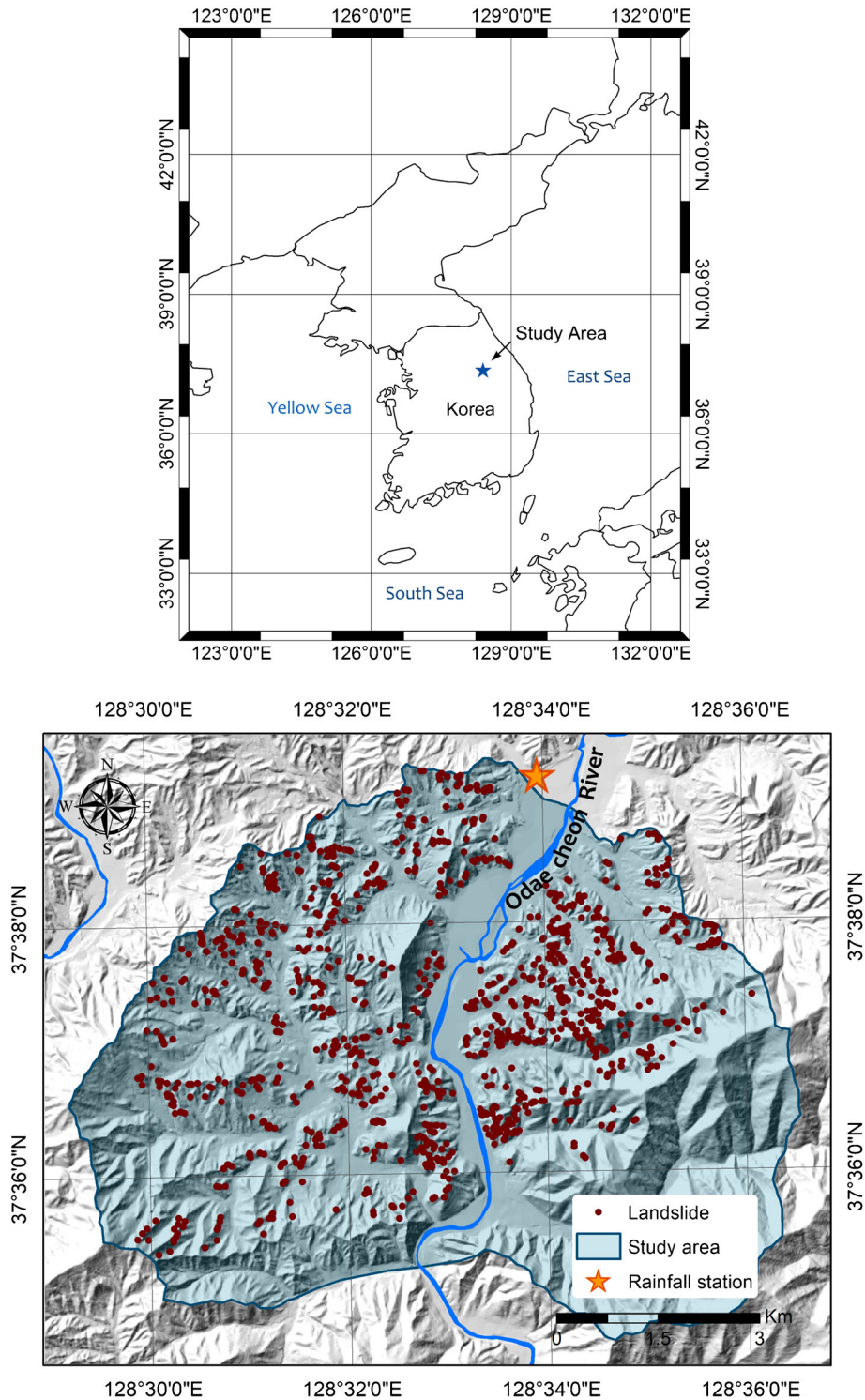
**Fig. 1** The study area

To evaluate the temporal probability of landslide occurrence using rainfall event records, hourly rainfall data, from 1986 to 2018, were obtained from the Sangjinburi rainfall station, which is located at 37° 39′ 32″ north and 128° 34′ 41″ east (Fig. 1). This is the closest and most representative station in the area where the rainfall-induced landslides occurred. The hourly rainfall data are available via the Water Resources Management Information System of South Korea (http://wamis.go.kr/wkw/rf_dubrfobs.aspx). In addition to gauge-based rainfall data, satellite-derived rainfall data can also be used to analyze rainfall triggered landslides (Guzzetti et al. 2008; Posner and Georgakakos 2015; Robbins 2016; Segoni et al. 2018). To compare the gauge-based rainfall data with the
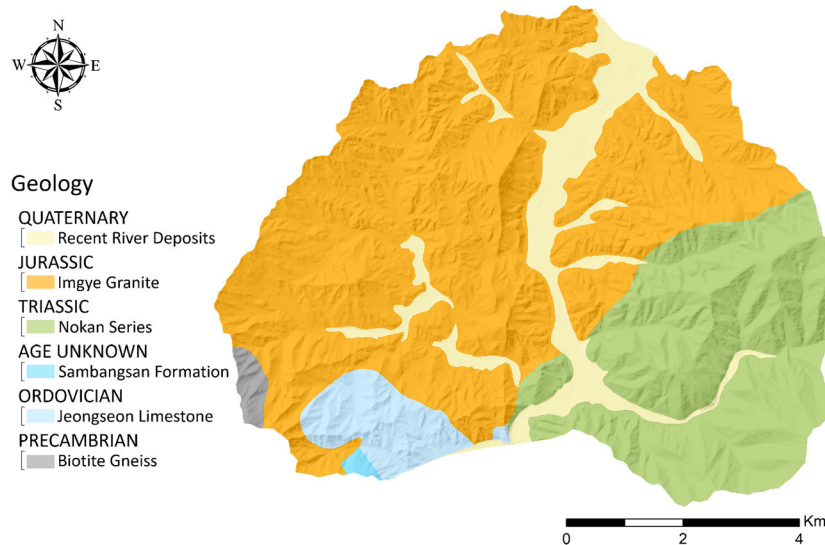
**Fig. 2** Geological map of the study area

satellite-derived rainfall data, we obtained the Tropical Rainfall Measurement Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA, Huffman et al. 2007) 3B42 version 7 rainfall product with fine temporal and spatial scales (3-hourly and 0.25°×0.25°) from 1998 to 2018. TRMM data are available from the Goddard Earth Sciences Data and Information Services Center (GES DISC) (https://disc.gsfc.nasa.gov/datasets/ TRMM_3B42 _7/summary).

### Temporal probability of landslide occurrence

#### Determination of rainfall threshold

The most commonly used empirical threshold in previous temporal probability evaluations is the rainfall intensity–duration threshold (Zêzere et al. 2004; Chleborad et al. 2006; Corominas and Moya 2008; Brunetti et al. 2010; Segoni et al. 2014, 2018; Afungang and Bateira 2016; Bogaard and Greco 2018). However, the intensity–duration threshold requires extensive rainfall records and landslide occurrence data with high-quality and temporal resolution, which were not available in this study area, or many other areas. Therefore, an antecedent threshold has been adopted for evaluation of temporal probability in several studies (Jaiswal and van Westen 2009; Jaiswal et al. 2010; Tien Bui et al.

2013; Afungang and Bateira 2016; Dikshit et al. 2020). An antecedent threshold is based on the amount of antecedent rainfall, which is relatively simple to measure over a large area. This approach was employed here because the study area has experienced only one landslide occurrence by an extreme rainfall event, in 2006, and consequently, sufficient rainfall and landslide data are not available. However, it is difficult to determine the number of days that should be used when using an antecedent rainfall threshold (Guzzetti et al. 2007). Because landslides have different types and geological conditions, the number of antecedent days for the rainfall threshold may appropriately range from 3 days for shallow landslides to 30 days for deep landslides (Kim et al. 1992; Aleotti 2004; Zêzere et al. 2004; Chleborad et al. 2006; Jaiswal and van Westen 2009; Tien Bui et al. 2013). Based on previous landslide studies in Korea (Kim et al. 1992; Yoo et al. 2012; Lee et al. 2015), as well as in other countries (Dahal and Hasegawa 2008; Jaiswal and van Westen 2009; Shou and Yang 2015), and because the landslides in the study area were mostly shallow, a 3-day (72-h) antecedent rainfall threshold was used in this study.

In the Jinbu area, landslides occurred sporadically during the heavy rainstorm of July 14–16, 2006, resulting in seven deaths (National Institute for Disaster Prevention 2006). Figure 3 shows
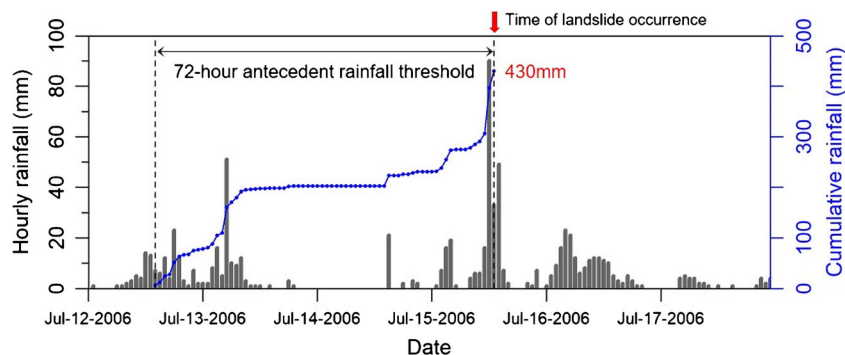


**Fig. 3** Hourly rainfall records in the study area in the period July 12–17, 2006
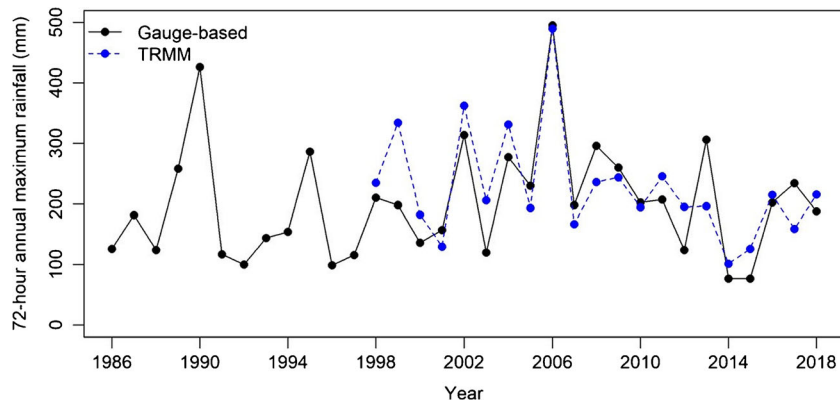
**Fig. 4** 72-h annual maximum rainfall events at Sangjinburi rainfall station (1986–2018)

the hourly rainfall time series for July 12–16, 2006 from Sangjinburi rainfall station in the study area. A rainfall event of 90 mm/h, which is the maximum rainfall intensity during the overall observation period, occurred on 15 July at the time when the most casualties were incurred. Therefore, a rainfall event of 90 mm/h is considered a key trigger of initial landslide occurrence. In addition, based on the government report on the landslide events in 2006 (National Institute for Disaster Prevention 2006), the 72-h rainfall event preceding the landslides totaled 430 mm (blue line in Fig. 3).

**Estimation of the temporal probability of rainfall-induced landslides**
Estimation of exceedance probability, using the Poisson distribution model, is one of the most commonly used methods to calculate the temporal probability of landslide occurrence (Coe et al. 2000; Crovelli 2000; Guzzetti et al. 2005; Romeo et al. 2006; Jaiswal and van Westen 2009; Jaiswal et al. 2010; Tien Bui et al. 2013; Afungang and Bateira 2016). The exceedance probability, using the Poisson distribution, was evaluated using

$$P\{N(t){\geq}1\} = 1-e^{-\lambda t} = 1-e^{-t/\mu}, \tag{1}$$

where $N(t)$ is the number of landslides, in $t$ years, whose probability is being estimated; $\lambda t$ is the expected number of landslide occurrences over $t$ years, and $\mu$ $(=1/\lambda)$ is the anticipated mean recurrence interval.

However, if landslides have only occurred once so far in an area or there are insufficient historical rainfall records, it is difficult to apply this approach. In these circumstances, it is not feasible to determine the mean recurrence interval of landslide triggering rainfall events (or threshold exceedance). Therefore, a new approach is necessary to estimate the temporal probability of future landslide occurrence in an area where no recurrent landslide triggering rainfall event has ever been recorded.

Extreme value analysis enables inferences about future probabilities of extreme event occurrences based on past observations (El Adlouni et al. 2007). Therefore, it has been widely used for engineering design and risk management related to hydrological extreme events such as heavy rainfall and flooding. The block maxima approach, one of the techniques used for extreme value

analysis, consists of modeling a sequence of maximum values obtained from blocks (periods) of equal length. If the period is a year, the annual maximum (AM) event data are used. Under the assumption that the maximum values are stationary, independent, and identically distributed, they follow a generalized extreme value (GEV) distribution (see Coles et al. (2001) for more details). The GEV distribution is classified into three types: extreme value type I (Gumbel distribution), extreme value type II (Fréchet distribution), and extreme value type III (Weibull distribution). The cumulative density function (CDF) for the GEV is

$$F_{\text{GEV}}(x) = \exp\left[-\left(1-k\frac{x-u}{\alpha}\right)^{\frac{1}{k}}\right], \tag{2}$$

where $u$, $\alpha$, and $k$ are the location, scale, and shape parameters, respectively. The three cases are (1) for $k = 0$, the extreme value type I distribution, (2) for $k < 0$, extreme value type II distribution, and (3) for $k > 0$, extreme value type III distribution (Chow et al. 1988). In all three cases, $\alpha$ is greater than zero.

The Gumbel (extreme value type I) distribution has been used for extreme value analysis of heavy rainfall events to estimate the temporal probability of rainfall-induced landslides (Finlay et al. 1997; Zêzere et al. 2008; Frattini et al. 2010; Jaiswal et al. 2011; Nefeslioglu et al. 2011; Martha et al. 2013; Lee et al. 2015; Afungang and Bateira 2016). In addition, the Gumbel distribution is recommended for estimation of the occurrence frequency and magnitude of extreme rainfall events in Korea (ME 2019). However, while previous studies have used the Gumbel distribution to evaluate rainfall return periods or probabilities of maximum daily rainfall events, the Gumbel distribution has not been used to evaluate the temporal probability of landslide occurrence. This study used the Gumbel distribution to evaluate the temporal probability that one or more rainfall events, which exceed the rainfall threshold, will occur during a specified time. Unlike the Poisson distribution, the Gumbel distribution does not require the mean recurrence interval, normally obtained from historical rainfall records. This is an advantage when the exceedance probability must be evaluated in an area with limited data on rainfall and landslides. The CDF of the Gumbel distribution is given by

$$F_{\text{GUM}}(x) = \exp\left\{-\exp\left(-\frac{x-u}{\alpha}\right)\right\} \quad -\infty < x < \infty, \tag{3}$$

**Table 1** 72-h annual maximum rainfall series and its sample PWMs at Sangjinburi rainfall station ($n = 33$)

| Observations Year | 72-h AM rainfall (mm) | Probability-weighted moments Order ($j$) | Ordered samples ($x_{j:n}$) | $\frac{(j-1)}{(n-1)}x_{j:n}$ |
|---|---|---|---|---|
| 1986 | 126 | 1 | 77 | 0.000 |
| 1987 | 182 | 2 | 77 | 2.406 |
| 1988 | 124 | 3 | 99 | 6.188 |
| 1989 | 258 | 4 | 100 | 9.375 |
| 1990 | 426 | 5 | 116 | 14.500 |
| 1991 | 117 | 6 | 117 | 18.281 |
| 1992 | 100 | 7 | 120 | 22.500 |
| 1993 | 144 | 8 | 124 | 27.125 |
| 1994 | 154 | 9 | 124 | 31.000 |
| 1995 | 286 | 10 | 126 | 35.438 |
| 1996 | 99 | 11 | 136 | 42.500 |
| 1997 | 116 | 12 | 144 | 49.500 |
| 1998 | 210 | 13 | 154 | 57.750 |
| 1999 | 198 | 14 | 157 | 63.781 |
| 2000 | 136 | 15 | 182 | 79.625 |
| 2001 | 157 | 16 | 188 | 88.125 |
| 2002 | 314 | 17 | 198 | 99.000 |
| 2003 | 120 | 18 | 198 | 105.188 |
| 2004 | 277 | 19 | 202 | 113.625 |
| 2005 | 230 | 20 | 202 | 119.938 |
| 2006 | 495 | 21 | 207 | 129.375 |
| 2007 | 198 | 22 | 210 | 137.813 |
| 2008 | 296 | 23 | 230 | 158.125 |
| 2009 | 260 | 24 | 234 | 168.188 |
| 2010 | 202 | 25 | 258 | 193.500 |
| 2011 | 207 | 26 | 260 | 203.125 |
| 2012 | 124 | 27 | 277 | 225.063 |
| 2013 | 306 | 28 | 286 | 241.313 |
| 2014 | 77 | 29 | 296 | 259.000 |
| 2015 | 77 | 30 | 306 | 277.313 |
| 2016 | 202 | 31 | 314 | 294.375 |
| 2017 | 234 | 32 | 426 | 412.688 |
| 2018 | 188 | 33 | 495 | 495.000 |
| Sample PWMs | | | $\widehat{b}_0 = n^{-1} \sum\limits_{j=1}^{n} x_{j:n}$ | $\widehat{b}_1 = n^{-1} \sum\limits_{j=1}^{n} \frac{(j-1)}{(n-1)} x_{j:n}$ |
| | | | $= 201.212$ | $= 126.689$ |

where $u$ and $\alpha$ ($>0$) are location and scale parameters, respectively. Then, the exceedance probability ($p$) of a certain rainfall event in a given year can be calculated by

$$p = 1 - F_{\text{GUM}}(x) = 1 - \exp\left\{-\exp\left(-\frac{x-u}{\alpha}\right)\right\}. \tag{4}$$

To determine the exceedance probability using the Gumbel distribution, the location and scale parameters must be estimated.
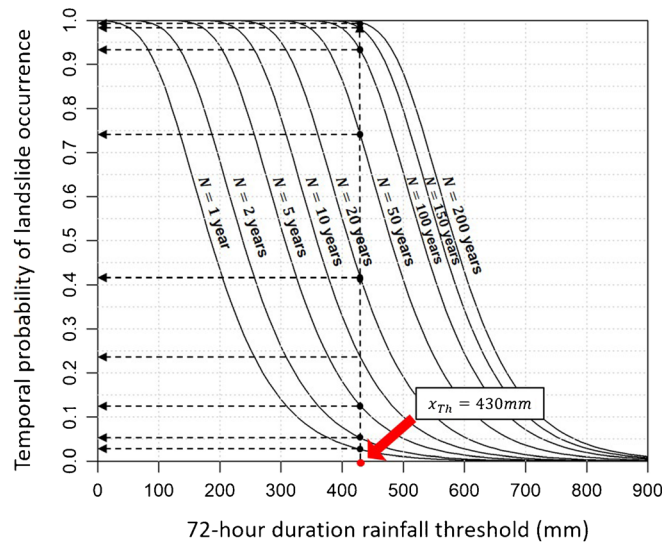
**Fig. 5** Temporal landslide probabilities for different time periods (1 year 0.027, 2 years 0.052, 5 years 0.126, 10 years 0.236, 20 years 0.416, 50 years 0.739, 100 years 0.932, 150 years 0.982, and 200 years 0.995)

Several methods estimate the statistical parameters of a probability distribution, including the method of moments (Wallis et al. 1974; Jaiswal et al. 2011; Martha et al. 2013), the method of maximum likelihood, and the method of probability-weighted moments (PWM) (or L-moments). The PWM method (Greenwood et al. 1979), in particular, provides efficient and computationally convenient estimates, even for small samples (Hosking 1990). In addition, the L-moments method, which is defined as linear combinations of PWMs, is widely used to estimate the parameters of various probability distributions, including the GEV and Gumbel distributions because parameter estimates based on L-moments are unbiased and less sensitive to outliers and observational errors than other conventional methods (Hosking and Wallis 1997; Rao and Hameed 2000; Gubareva and Gartsman 2010; Drissia et al. 2019). Therefore, the L-moments method was employed to estimate the parameters of the Gumbel distribution.

The first and second population L-moments are defined by the first and second population PWMs ($b_0$ and $b_1$) of the Gumbel

distribution as follows (Hosking and Wallis 1997):

$$\lambda_1 = b_0 = u + \gamma\alpha, \tag{5}$$

$$\lambda_2 = 2b_1 - b_0 = \alpha \cdot ln2, \tag{6}$$

where $\gamma$ is Euler's constant (=0.577215 ...). Then, the L-moment estimators, $\widehat{\alpha}$ and $\widehat{u}$, of the Gumbel distribution can be derived from Eqs. (5) and (6).

$$\widehat{\alpha} = \left(2\widehat{b}_1 - \widehat{b}_0\right)/ln2, \tag{7}$$

$$\widehat{u} = \widehat{b}_0 - \gamma\widehat{\alpha}, \tag{8}$$

where the $r$-th sample PWM, $\widehat{b}_r$, can be obtained by (Hosking et al. 1985)

$$\widehat{b}_r = n^{-1}\sum_{j=r+1}^{n}\frac{(j-1)(j-2)\cdots(j-r)}{(n-1)(n-2)\cdots(n-r)}x_{j:n} \quad \text{for } r > 0, \tag{9}$$

and $x_{1:n} \leq x_{2:n} \leq \cdots x_{n:n}$ are the ordered AM rainfall series $x_i$ ($i = 1, 2, \cdots, n$) where $n$ is the sample size. Note that $\widehat{b}_0$ is the sample mean ($\frac{1}{n}\sum_{i=1}^{n}x_i$).

The value of AM (annual maximum) rainfall event for the $i$-th year ($x_i$) is given by the largest rainfall event, within a specific duration that occurred in that year. Since the duration of the rainfall threshold was determined as 72-h in "Determination of rainfall threshold" section, Gumbel distribution modeling was performed using the 72-h AM rainfall series. To determine the 72-h AM rainfall event value, a moving window procedure was used with hourly rainfall observations. For each year, all 72-h rainfall event values were extracted by moving the 72-h window on an hourly basis. The AM rainfall event for that year

**Table 2** Temporal probabilities, in the study area, of rainfall-induced landslide occurrences over period N

| Period $N$ (year) | Temporal probability of rainfall-induced landslide occurrences ($Pr_{GUM}(x_{Th}, N)$) |
|---|---|
| 1 | 0.0265 (2.65%) |
| 2 | 0.0523 (5.23%) |
| 5 | 0.1257 (12.57%) |
| 10 | 0.2355 (23.55%) |
| 20 | 0.4156 (41.56%) |
| 50 | 0.7389 (73.89%) |
| 100 | 0.9318 (93.18%) |
| 150 | 0.9822 (98.22%) |
| 200 | 0.9954 (99.54%) |

**Table 3** The landslide conditioning factors used in this study

| Source maps (scale) | Conditioning factors | Source maps (scale) | Conditioning factors |
|---|---|---|---|
| Topographic map (1:5000) | Elevation<br>Slope<br>Aspect<br>Standard curvature<br>Profile curvature<br>Planform curvature<br>SCA<br>TWI<br>SPI | Forest map (1:25,000) | Forest type<br>Timber diameter<br>Timber age<br>Forest density |
| | | Soil map (1:25,000) | Soil series<br>Soil texture<br>Soil sub-texture<br>Soil drainage<br>Soil thickness |
| Geological map (1:50,000) | Geology | Land use map (1:5000) | Land use |

was set to the largest extracted value. Figure 4 presents the 72-h AM rainfall series obtained both from hourly observations at Sangjinburi rainfall station ($n = 33$ years) and from the 3-hourly TRMM gridded data ($n = 21$ years). The 72-h AM rainfall series from the station (black solid line) and from TRMM data (blue dotted line) show a similar pattern. In particular, the gauged and TRMM AM values in 2006, the year of the first landslide occurrence in the study area, are almost the same and have the largest value over the entire observation period. However, since the rainfall data from the rainfall station have a longer time span and a higher sampling rate, the temporal probability analysis was carried out using these gauge-based rainfall data. Using these data, sample PWMs were calculated as shown in Table 1. Based on Eqs. (7) and (8), the location and scale parameters were estimated as $\hat{\mu} = 157.772$ and $\hat{\sigma} = 75.258$, respectively. The exceedance probability of the determined rainfall threshold ($x_{Th} = 430$ mm) was calculated as $p = 0.0265$.

From this, the probability of experiencing at least one extreme event, greater than the threshold, over $N$ years, $Pr(N)$, can be calculated as follows:

$$Pr(N) = \sum_{t=1}^{N} (1-p)^{t-1} p = 1 - (1-p)^N. \tag{10}$$

The concept of the $Pr(N)$ is referred to as "risk of failure" defined by Yen (1970). It is described in most introductory textbooks and recommended as an essential metric for assessing the actual probability of an extreme event during a given period (Serinaldi 2015; Serinaldi and Kilsby 2015). In this study, we assumed that extreme rainfall events over the rainfall threshold $x_{Th}$ lead to landslide occurrences. Then, the concept of the $Pr(N)$ was employed to calculate the temporal probability of landslide occurrences. The probability of experiencing at least one extreme rainfall event greater than $x_{Th}$ in $N$ years, $Pr_{GUM}(x_{Th}, N)$, can be defined by substituting Eq. (4) into Eq. (10):

$$Pr_{GUM}(x_{Th}, N) = 1 - \{F_{GUM}(x_{Th})\}^N. \tag{11}$$

Figure 5 illustrates the temporal probability of landslide occurrences for $N = 1, 2, 5, 10, 20, 50, 100, 150, 200$ years in the study area. Because the graph of $Pr_{GUM}(x_{Th}, N)$ moves rightward as $N$ increases, for a certain rainfall threshold, the temporal probability of landslide occurrences increases with $N$. The temporal probabilities of landslide occurrences for the determined rainfall threshold ($x_{Th} = 430$ mm) are presented in Table 2. As an example, the temporal probability of

landslide occurrences during any 1 year is equal to the exceedance probability of the rainfall threshold ($p = 0.0265$).

### Landslide susceptibility analysis

Landslide susceptibility analysis (or spatial probability analysis) estimates the potential for a landslide by considering landslide inducing factors (or conditioning factors) and analyzing the spatial distribution of these factors (Chae et al. 2017). However, landslide susceptibility assessment is not straightforward because the occurrence of landslide is controlled by diverse spatial conditioning factors, including geology, topography, and vegetation. Several assessment approaches have been suggested. Among them, data-driven landslide susceptibility analysis has been widely used and evaluates the statistical relationships between the locations of landslides and the landslide inducing factors. However, because the relationship between the independent variables (landslide affecting factors) and a dependent variable (landslide occurrence) is often complex and nonlinear, it is not easy to determine the relationship through simple statistical analysis. Therefore, various machine learning approaches, such as neurofuzzy methods, support vector machines, evidential belief functions, and artificial neural networks, have been developed to evaluate the susceptibility of landslide occurrence. Multi-layer perceptron (MLP) neural networks, a type of artificial neural network (ANN), have been applied successfully and efficiently in landslide prediction (Gomez and Kavzoglu 2005; Zare et al. 2013; Pham et al. 2017). MLP is a feedforward back error propagation neural network with one or more hidden layers between the input layers and output layers. It has several advantages including that the distribution of training dataset is not dependent on pre-assumptions, no decision is required about the relative importance of input factors, and most input factors are selected based on weightings adjustment during the training process (Gardner and Dorling 1998). In addition, ANN can process input data with different measurement scales and units, such as continuous, categorical, and binary data, a significant advantage for landslide susceptibility analysis (Garrett 1994).

### Spatial database construction

For the susceptibility analysis, various digital maps were used to construct a spatial database including the following landslide conditioning factors (Table 3): elevation, slope angle, slope aspect, planform curvature, profile curvature, standard curvature, stream power index (SPI), topographic wetness index (TWI), specific catchment area (SCA), geology, forest type, forest density, timber
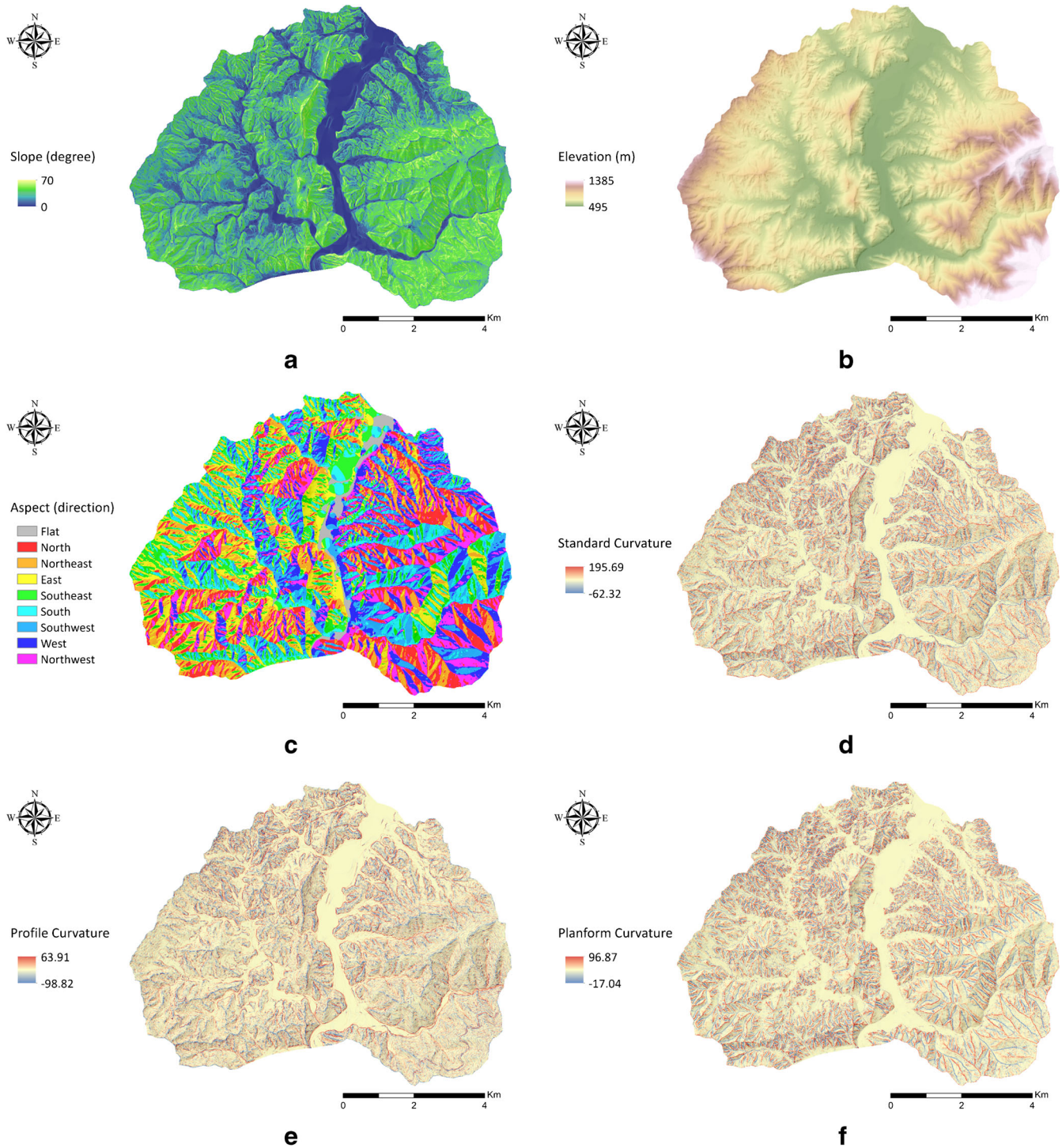
**Fig. 6** Thematic maps of landslide conditioning factors: **a** slope, **b** elevation, **c** aspect, **d** standard curvature, **e** profile curvature, **f** planform curvature, **g** specific catchment area, **h** topographic wetness index, **i** stream power index, **j** forest type, **k** timber diameter, **l** timber age, **m** forest density, **n** soil series, **o** soil texture, **p** soil sub-texture, **q** soil drainage, **r** soil thickness, and **s** land use

age, timber diameter, soil series, soil texture, soil sub-texture, soil drainage, soil thickness, and land use. In total, 20 factors were used in the analysis. The relevant digital maps are listed in Table 3, and they are available at different scales and formats in the Korean archives. After acquisition from several governmental institutes,

the digital maps were converted into grid-based raster layers with 10-m resolution to create thematic maps for analysis.

The geomorphological database of the study area was based on a 1:5000 scale digital topographic map, provided by the National Geographic Information Institute. The contour polyline vector
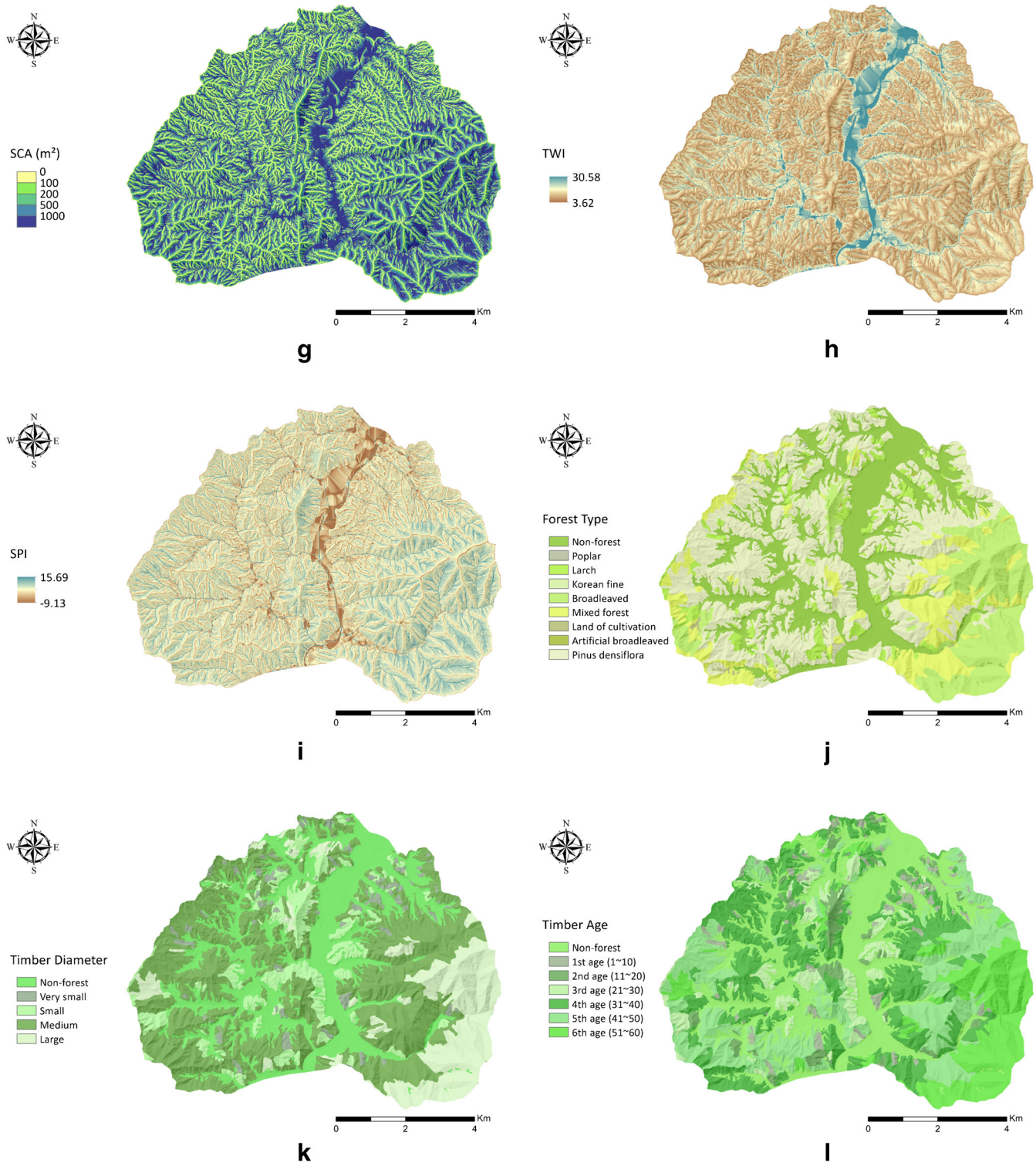
**Fig. 6** (continued)

layer, including elevation attributes, was extracted from the digital map. A TIN (triangulated irregular network) was generated from the contour layer and converted into a grid-based raster layer as a digital elevation model (DEM). The DEM was used to compute the geomorphological factors (Fig. 6a–i): elevation, slope angle, slope aspect, standard curvature, profile curvature, planform curvature,

SCA, TWI, and SPI. SCA, defined as the area of land upslope of a width of the contour, is commonly used to describe complex terrain when analyzing water flow on hill slopes (Tretkoff 2011). The TWI is a runoff-based model that describes the relationship between the water inclination that accumulates in any location and the gravitational force inclination in the
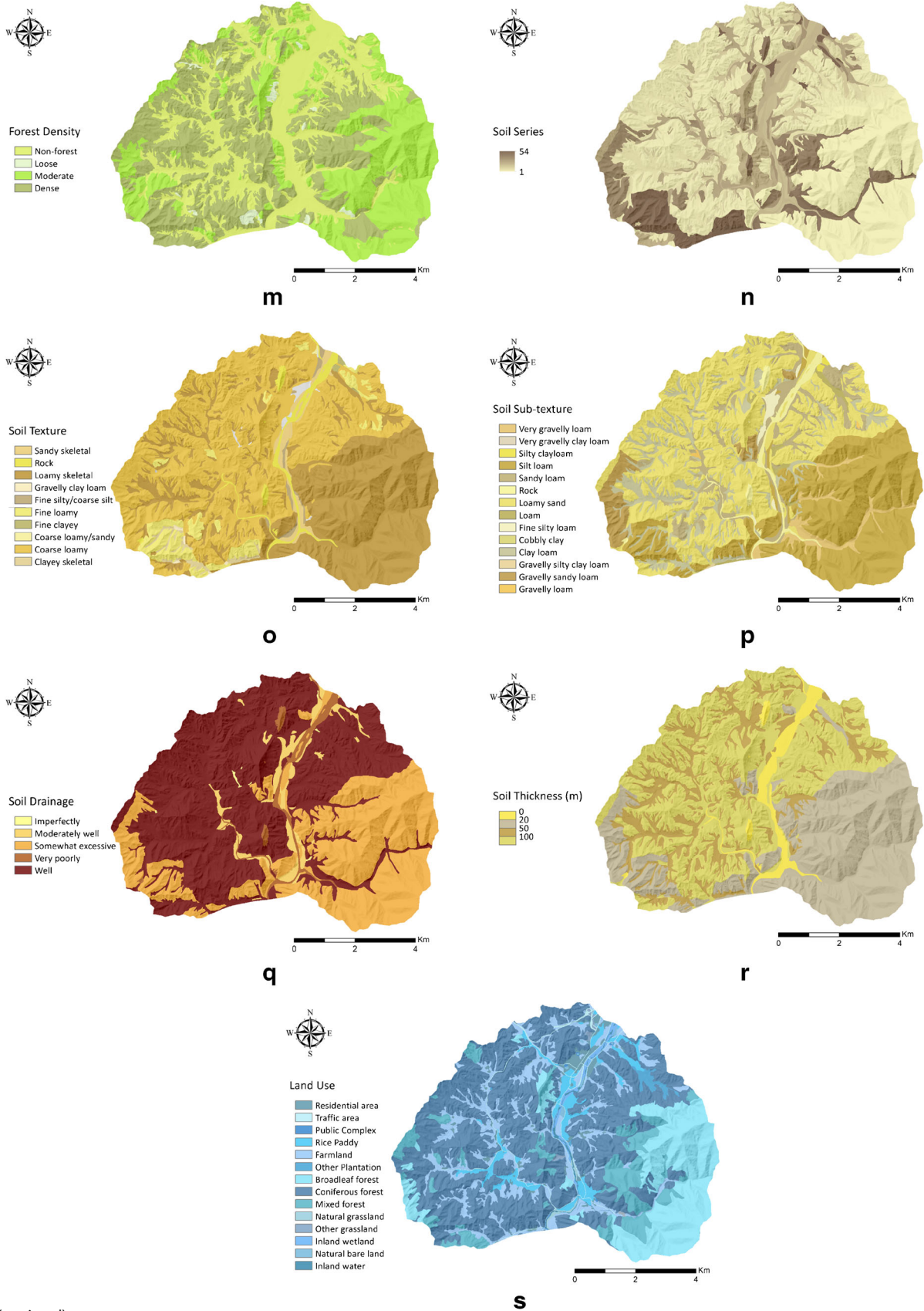
Forest Density
- Non-forest
- Loose
- Moderate
- Dense

0 2 4 Km

**m**

Soil Series
- 54
- 1

0 2 4 Km

**n**

Soil Texture
- Sandy skeletal
- Rock
- Loamy skeletal
- Gravelly clay loam
- Fine silty/coarse silt
- Fine loamy
- Fine clayey
- Coarse loamy/sandy
- Coarse loamy
- Clayey skeletal

0 2 4 Km

**o**

Soil Sub-texture
- Very gravelly loam
- Very gravelly clay loam
- Silty clayloam
- Silt loam
- Sandy loam
- Rock
- Loamy sand
- Loam
- Fine silty loam
- Cobbly clay
- Clay loam
- Gravelly silty clay loam
- Gravelly sandy loam
- Gravelly loam

0 2 4 Km

**p**

Soil Drainage
- Imperfectly
- Moderately well
- Somewhat excessive
- Very poorly
- Well

0 2 4 Km

**q**

Soil Thickness (m)
- 0
- 20
- 50
- 100

0 2 4 Km

**r**

Land Use
- Residential area
- Traffic area
- Public Complex
- Rice Paddy
- Farmland
- Other Plantation
- Broadleaf forest
- Coniferous forest
- Mixed forest
- Natural grassland
- Other grassland
- Inland wetland
- Natural bare land
- Inland water
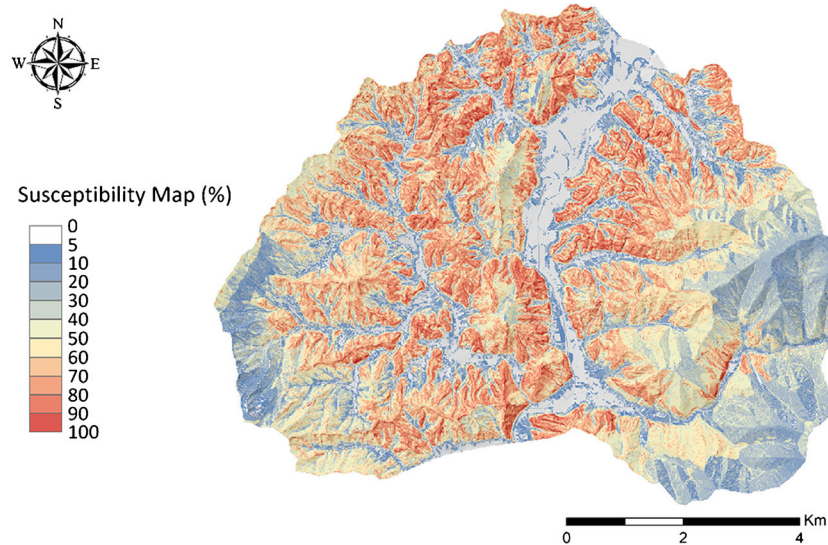
0 2 4 Km

**s**

**Fig. 6** (continued)

**Fig. 7** Landslide susceptibility map obtained using MLP

stream that accumulates down the slope (Althuwaynee et al. 2014). The SPI is the erosive power of the water flow based on an assumption that the discharge is proportional to the SCA. This factor predicts erosion along the slope of the profile's convexity and deposition in the profile's concavity. Among the other factors, curvature refers to the curvature of a line formed by intersecting the ground surface and a plane (horizontal or vertical). Profile curvature is parallel to the direction of the slope angle and indicates the direction of the slope. It affects the acceleration and deceleration of flow through the surface, and associates with erosion and deposition of slope material. Planform curvature is perpendicular to
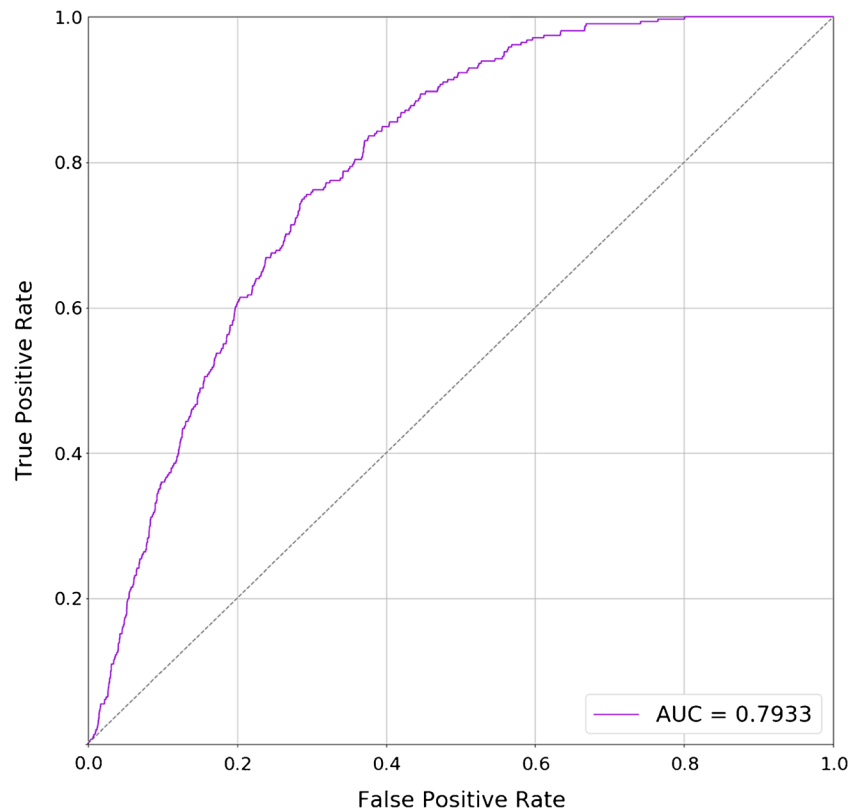


**Fig. 8** ROC (receiver operating characteristics) graph. (AUC: area under the curve)

**Table 4** Landslide hazard value proportions (%) for different time periods

| Time periods | Landslide hazard value 0 – 0.1 | 0.1 – 0.2 | 0.2 – 0.3 | 0.3 – 0.4 | 0.4 – 0.5 | 0.5 – 0.6 | 0.6 – 0.7 | 0.7 – 0.8 | 0.8 – 0.9 | 0.9 – 1.0 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 100 | – | – | – | – | – | – | – | – | – |
| 2 | 100 | – | – | – | – | – | – | – | – | – |
| 5 | 91.38 | 8.62 | – | – | – | – | – | – | – | – |
| 10 | 45.10 | 50.44 | 4.46 | – | – | – | – | – | – | – |
| 20 | 25.59 | 25.99 | 31.29 | 16.35 | 0.78 | – | – | – | – | – |
| 50 | 16.61 | 12.37 | 14.10 | 16.11 | 17.59 | 16.09 | 6.14 | 0.99 | – | – |
| 100 | 14.63 | 8.32 | 11.65 | 11.02 | 12.97 | 13.75 | 14.24 | 9.49 | 3.20 | 0.73 |
| 150 | 14.24 | 7.72 | 10.89 | 10.35 | 11.86 | 12.88 | 13.72 | 11.40 | 5.30 | 1.64 |
| 200 | 14.14 | 7.57 | 10.68 | 10.24 | 11.56 | 12.73 | 13.43 | 11.78 | 5.88 | 1.98 |

the slope and affects the convergence and divergence of flow on the surface. Standard curvature is a combination of profile and planform curvature. Using these factors, it is possible to understand the flow from the surface when rainfall occurs (Yilmaz et al. 2012; Chen et al. 2017).

The geological data (Fig. 2) were derived from a 1:50,000 scale digital geological map provided by the Korea Institute of Geoscience and Mineral Resources. Forest-related data drew on the Korea Forest Service's 1:25,000 digital forest map. Forest type, timber diameter, timber age, and forest density were isolated as thematic maps (Fig. 6j–m). Based on the 1:25,000 digital soil map provided by the National Institute of Agricultural Sciences, the soil series maps—texture, sub-texture, drainage, and thickness—were derived (Fig. 6n–r). Finally, the land use map (Fig. 6s) was based on the 1:5000 digital land cover map provided by the Ministry of Environment.

**Evaluation of landslide susceptibility using a multi-layer perceptron**
Machine learning is related to pattern recognition and computational learning techniques in artificial intelligence, a subfield of computer science. Neural networks, one of the machine learning approaches, have been used successfully for classification and prediction in various problem domains (Paliwal and Kumar 2009), and MLP is the most widely used neural network architecture for classification or regression problems (Yilmaz and Kaynar 2011). MLP uses a neural network of at least three node layers: input, hidden, and output. Input layers are the landslide inducing factors, regarded as independent variables, and the output layer has a result classified as either a landslide or nonlandslide class. The hidden layers are the classifying layers that transform inputs into outputs (Pham et al. 2017). The hidden layer computes the expected output through nonlinear activation functions and determines the difference from the actual output in the training dataset (feedforward propagation). Using the differences, the number of layers and nodes in hidden layer is adjusted. Each node is a neuron that uses its nonlinear activation function to distinguish data that are not linearly separable, or separable by a hyperplane (Cybenko 1989; Khalil Alsmadi et al. 2009). In the MLP, the connection weights between the input neuron and the hidden neurons as well as between the hidden neurons and the outputs are initialized and then trained using a backpropagation algorithm. There are two steps in the training, i.e., feedforward and backpropagation (Tien Bui et al. 2016): (i) in the feedforward, the input is propagated forward through the layers resulting in a response at the output layer. The output response values are compared with actual values, and the differences are assessed; (ii) in the backpropagation, the connection weights are updated to minimize the differences. In this study, the weights were calculated using rectified linear unit (ReLU) activation functions, the most commonly used activation function in neural networks. The goal of MLP is to derive a model that minimizes the difference between the actual and estimated values. In this study, we set the number of times the model repeats the learning (epoch) as 1000 and the rate at which to approach the minimum target error (learning rate) as 0.005.

First, it was necessary to preprocess the data for machine learning. The database of the study area, including landslide locations and landslide conditioning factors, was reconstructed to a landslide inventory map and thematic maps of conditioning factors using ArcGIS and Python. The grid pixels in the landslide inventory map were divided into two classes: landslide pixels and
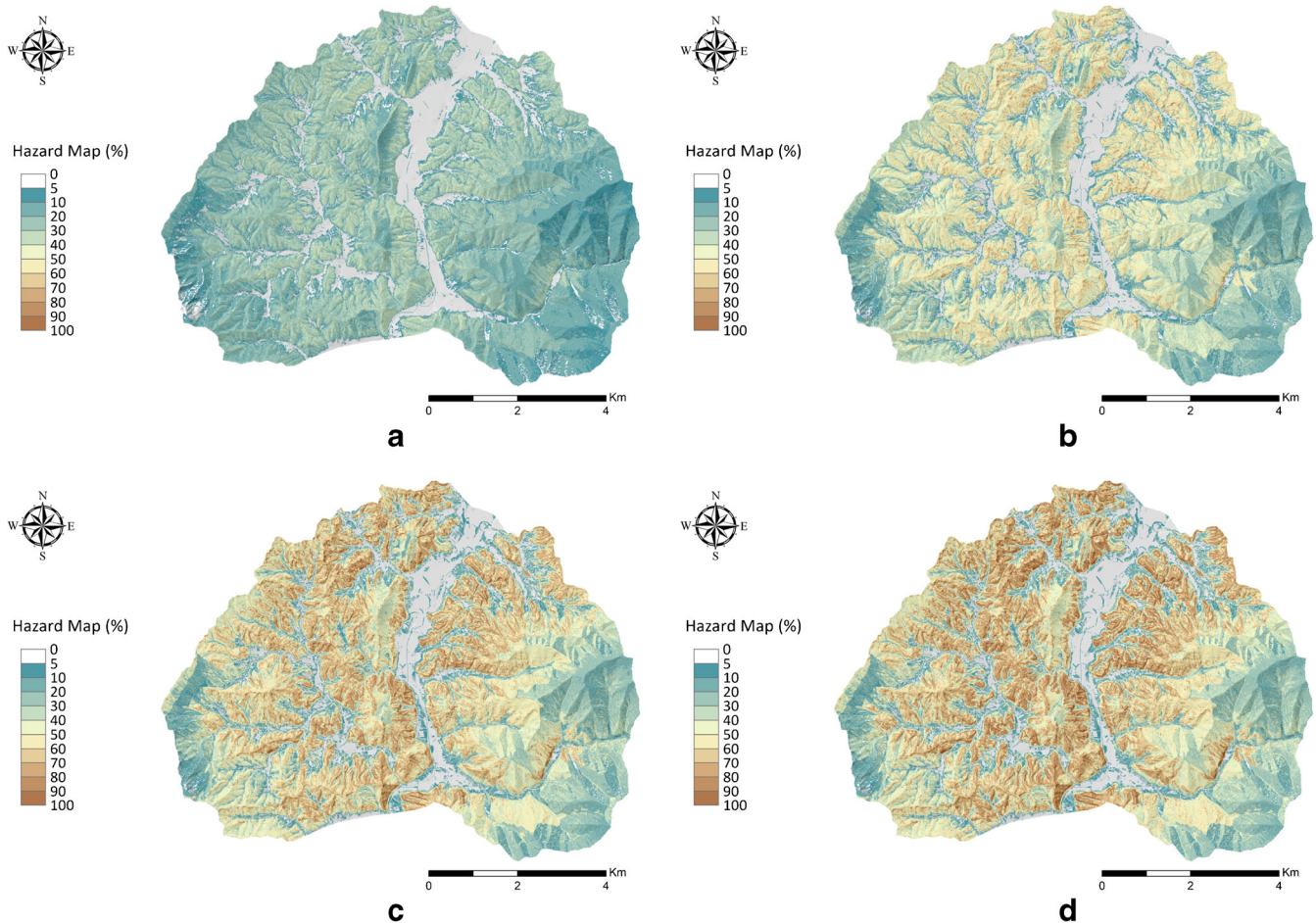
**Fig. 9** Examples of landslide hazard maps for four different periods: **a** 20 years, **b** 50 years, **c** 100 years, and **d** 200 years

nonlandslide pixels (Tien Bui et al. 2016). A grid cell where a landslide has occurred is assigned "1," and a cell without the landslide is assigned "0." The landslide dataset was combined with the attribute data of the conditioning factors and used in training the MLP model and then to calculate the probability of landslide occurrence. The landslide susceptibility analysis results are expressed in terms of the probability of occurrence, which ranged from "0" (landslide less likely to occur) to "100%" (landslide more likely to occur). In general, because the number of landslide pixels is very small compared with the total number of pixels in the study area, the amount of landslide occurrence data is smaller than the amount of nonlandslide data. This study had 647,433 pixels, but the number of landslide pixels was only 1035. Consequently, the data were significantly unbalanced. This data unbalance can cause serious errors in the analysis. Specifically, the evaluated probability values in all pixels of the study area can converge to "0," which meant that the entire study area would be classified as stable. Therefore, nonlandslide data representing just one or two times the number of landslide pixels should be selected to overcome the data unbalance problem described in previous studies (Pradhan 2013; Pham et al. 2016; Chen et al. 2018). In this case, we trained the model using the same number of the nonlandslide and landslide

pixels selected randomly. The randomly selected dataset was partitioned into two subsets as training and test datasets. Training data were used to build the model while test data were used to validate the model and evaluate its accuracy. In this process, the test data should be different from those that trained the model (Chung and Fabbri 2008; Pradhan and Lee 2010). However, no mathematical rule to determine the relative partitioning of these subsets exists (Pham et al. 2017). In this study, 70% of the data were used for training the network, and the rest were used to test the final MLP model, as previous studies have suggested (Tien Bui et al. 2016; Chen et al. 2017). A confusion matrix was used to evaluate the performance of the prediction model using the test data. Finally, the remaining data, which were not used for training or testing of the prediction model, were used to evaluate the spatial probability of landslide susceptibility throughout the study area. Figure 7 shows the resulting landslide susceptibility map. A receiver operating characteristics (ROC) graph and its area under the graph (AUC) were used to verify the accuracy of the susceptibility map by comparing actual landslide locations with the analysis results (Fig. 8). The AUC value of the landslide susceptibility map in this study was 79.33%.

## Landslide hazard assessment

The landslide hazard in the study area was then computed by multiplying the temporal probability of landslide occurrence, derived using the Gumbel distribution, with the landslide susceptibility obtained using MLP. To compare landslide hazards for different time periods, nine future time periods (1, 2, 5, 10, 20, 50, 100, 150, 200 years) were used, as described in "Estimation of the temporal probability of rainfall-induced landslides" section. The hazard assessments for the different periods are given in Table 4, which shows that the calculated hazard values for most pixels of the study area for 1, 2, and 5-year periods were less than 0.3 (30%). In each of the relatively short periods (1, 2, 5, 10, and 20 years), the hazard values of all pixels were less than 0.5 (50%). This is because the mean AM rainfall value is much smaller than the landslide rainfall trigger threshold. Statistical analysis of the time-series rainfall data for 33 years (from 1986 to 2018) showed that the mean AM rainfall amount was 201.2 mm/72 h. However, the landslide trigger rainfall in 2006 was 430 mm/72 h, and this was used as the threshold for the temporal probability analysis. Consequently, the temporal probabilities for the five periods (1, 2, 5, 10, 20 years) were very small, and so, the hazard values were also generally small for these periods. In contrast, the evaluated hazard values in the area were much higher for time periods greater than 50 years. That is, the proportion of hazard values greater than 0.5 (50%) in the study area increased from 23.22% over 50 years to 45.8% over 200 years, which means that the landslide hazard is greater when longer periods are considered, as is evident in the landslide hazard maps for 20, 50, 100, 200-year periods (Fig. 9). For presentation purpose, only four landslide hazard maps for 20, 50, 100, 200-year periods are shown in Fig. 9.

## Discussion and conclusions

In this study, we developed an approach that can be applied in an area where the landslide inventory and landslide triggering rainfall events data are incomplete or where a nonrepetitive landslide event has occurred. This approach allowed us to estimate the temporal probability of landslide occurrence in an area where a multi-temporal landslide inventory and historical rainfall records for repetitive landslides were incomplete. The study area had experienced landslide occurrence just once, in 2006, which means that the recurrent rates of landslide or triggering rainfall events could not be determined. Instead, this study evaluated temporal probability using the extreme value analysis. Extreme value analysis has been widely used for hydrological extreme events such as heavy rainfall. To evaluate the future probability of extreme event occurrence (e.g., extreme rainfall) based on extreme value theory, the Gumbel (extreme value type I) distribution was applied to the available time series rainfall data to estimate the probability of extreme events over several future time periods. This probability was then, indirectly, considered as the temporal probability of extreme rainfall-induced landslides.

The described approach was applied to the study area, which had experienced a large number of landslides following extreme rainfall in 2006. Based on previous landslide studies and rainfall records in the study area, a 3-day (72-h) antecedent rainfall threshold was adopted and estimated as 430 mm. AM values, for each 72-h period, were calculated using a moving window and hourly rainfall records from 1986 to 2018. Then, using the AM data, the location and scale parameters of the Gumbel distribution were found to be 157.77 and 75.25, respectively. Subsequently, the rainfall threshold's exceedance probability for any year was calculated as 0.0265. From this, the probabilities of at least one extreme event, greater than the threshold, over 1, 2, 5, 10, 100, 150, and 200 years were 0.0265, 0.0523, 0.1257, 0.2355, 0.4156, 0.7389, 0.9318, 0.9822, and 0.9954, respectively.

The temporal probability of landslide occurrence was then combined with the spatial probability of landslides to determine landslide hazard. The landslide susceptibility map was constructed after MLP modeling in which 20 conditioning factors and the landslide inventory were used to analyze the relationship between landslide locations and conditioning factors. The MLP was trained through backpropagation, the most frequently used neural network method. Validation showed a prediction accuracy of 79.3%. Landslide hazard maps were then obtained by integrating the temporal probability of landslides with the landslide susceptibility maps.

The proposed approach improves on previous studies. Determination of the exceedance probability of a rainfall threshold has previously depended on frequency analysis of past repetitive events, with a Poisson distribution. However, when the historical data for repetitive events are limited or there is only one landslide event, as in our study area, temporal probability cannot be estimated using the previous approach. In this case, the Gumbel distribution can be used for temporal probability evaluation provided continuous time-series rainfall data are available in the area. Thus, the proposed approach can be used even if a Poisson model cannot be adopted. Several previous studies have adopted the Gumbel model, but they did not use it to evaluate the temporal probability of a landslide. The previous studies utilized the Gumbel distribution to calculate the return period of extreme rainfall events or to evaluate the probability of daily rainfall occurrence. The developed approach, on the other hand, was able to estimate the temporal probability of landslide occurrence using the Gumbel model by adopting extreme value theory.

Our approach also has some limitations. The importance of the spatial variability of rainfall is well recognized in landslide studies. However, precise prediction of rainfall across an area is difficult when there are few available rainfall gauges. Therefore, the simplified approach to determining the rainfall amount for a certain region was utilized with single rain gauge near a specified landslide site or the selected rain gauge as representative locations for a predefined region (Chiang and Chang 2009; Canli et al. 2018a; b). In this study, we used rainfall records obtained from only one rainfall station (Sangjinburi) to evaluate the temporal probability of landslide triggering rainfall. Sangjinburi is the station closest to the landslide locations, and best represents rainfall conditions in the study area. Consequently, we could not consider spatial variation of rainfall. If there is no representative rainfall station, or the rainfall record is too short to conduct the extreme value analysis, satellite-based rainfall data, such as TRMM, can be used to estimate the temporal probability of landslide occurrences. In fact, gauged data and TRMM data produced differences in the temporal probabilities of rainfall-induced landslide occurrences of less than 3% for all the analysis period. Therefore, when the gauge-based rainfall data are not available, TRMM rainfall data can be a viable alternative.

In addition, this approach considered stationary rainfall data, which means that the conditions of rainfall and its essential statistical characteristics (e.g., mean and variance of the probability distribution) are invariant with respect to time. Consequently, the Gumbel distribution used in this analysis is also based on the stationary condition. However, as IPCC (2014) pointed out, global warming is expected to lead to a greater frequency and magnitude of heavy precipitation. This means that the stationary rainfall assumption will not be valid. Recently, in the field of hydrology, nonstationary approaches have been suggested for most hydrological processes including rainfall. Because the probability distribution parameters will change with the time due to climate change, we need a different approach that can analyze nonstationary rainfall time series data.

## References

Afungang RN, Bateira CV (2016) Temporal probability analysis of landslides triggered by intense rainfall in the Bamenda Mountain Region, Cameroon. Environ Earth Sci 75(12):1032

Aleotti P (2004) A warning system for rainfall-induced shallow failures. Eng Geol 73(3–4):247–265

Althuwaynee OF, Pradhan B, Park HJ, Lee JH (2014) A novel ensemble bivariate statistical evidential belief function with knowledge-based analytical hierarchy process and multivariate statistical logistic regression for landslide susceptibility mapping. Catena 114:21–36

Bogaard T, Greco R (2018) Invited perspectives: hydrological perspectives on precipitation intensity-duration thresholds for landslide initiation: proposing hydro-meteorological thresholds. Nat Hazards Earth Syst Sci 18(1):31–39

Brabb EE (1985) Innovative approaches to landslide hazard and risk mapping. Proceedings of the International Landslide Symposium, Toronto, Canada 1:17–22

Brunetti MT, Peruccacci S, Rossi M, Luciani S, Valigi D, Guzzetti F (2010) Rainfall thresholds for the possible occurrence of landslides in Italy. Nat Hazards Earth Syst Sci 10(3):447–458

Canli E, Lounge B, Glade T (2018a) Spatially distributed rainfall information and its potential for regional landslide early warning systems. Nat Hazards 91:S103–S127

Canli E, Mergili M, Thiebes B, Glade T (2018b) Probabilistic landslide ensemble prediction systems: lessons to be learned from hydrology. Nat Hazards Earth Syst Sci 18:2183–2202

Cascini L, Bonnard C, Corominas J, Jibson R, Montero-Olarte J (2005) Landslide hazard and risk zoning for urban planning and development. In: Hungr O, Fell R, Couture R, Eberhardt E (eds) Landslide risk management. CRC Press, pp 209–246

Chae BG, Park HJ, Catani F, Simoni A, Berti M (2017) Landslide prediction, monitoring and early warning: a concise review of state-of-the-art. Geosci J 21(6):1033–1070

Chen HX, Zhang LM (2014) A physically based distributed cell model for predicting regional rainfall-induced shallow slope failures. Eng Geol 176:79–92

Chen W, Xie X, Wang J, Pradhan B, Hong H, Tien Bui D, Duan Z, Ma J (2017) A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility. Catena 151:147–160

Chen W, Zhang S, Li R, Shahabi H (2018) Performance evaluation of the GIS-based data mining techniques of best-first decision tree, random forest, and naïve Bayes tree for landslide susceptibility modeling. Sci Total Environ 644:1006–1018

Chiang SH, Chang KT (2009) Application of radar data to modeling rainfall-induced landslides. Geomorphology 103:299–309

Chleborad AF, Baum RL, Godt JW (2006) Rainfall thresholds for forecasting landslides in the Seattle, Washington, area: exceedance and probability. US Geological Survey Open-File Report, 1064

Chou HT, Lee CF, Lo CM, Lin CP (2012) Landslide and alluvial fan caused by an extreme rainfall in Suao, Taiwan. Proceedings of the 11th International Symposium on Landslides (ISL) and the 2nd North American Symposium on Landslides, Banff, Alberta, Canada

Chow VT, Maidment DR, Mays LW (1988) Applied hydrology. McGraw Hill Series in Water Resources and Environmental Engineering

Chung CJ, Fabbri AG (2008) Predicting landslides for risk analysis—spatial models tested by a cross-validation technique. Geomorphology 94(3–4):438–452

Coe JA, Michael JA, Crovelli RA, Savage WZ (2000) Preliminary map showing landslide densities, mean recurrence intervals, and exceedance probabilities as determined from historic records, Seattle, Washington. US Geological Survey Open-File Report, 303

Coles S, Bawa J, Trenner L, Dorazio P (2001) An introduction to statistical modeling of extreme values. Springer, London

Corominas J (2000) Landslides and climate. Proceedings of the 8th International Landslide Symposium, Cardiff, UK 4:1–33

Corominas J, Moya J (2008) A review of assessing landslide frequency for hazard zoning purposes. Eng Geol 102(3–4):193–213

Corominas J, van Westen C, Frattini P, Cascini L, Malet JP, Fotopoulou S, Catani F, Van Den Eeckhaut M, Mavrouli O, Agliardi F, Pitilakis K, Winter MG, Pastor M, Ferlisi S, Tofani V, Hervás J, Smith JT (2014) Recommendations for the quantitative analysis of landslide risk. Bull Eng Geol Environ 73(2):209–263

Crosta GB, Frattini P (2001) Rainfall thresholds for triggering soil slips and debris flow. Proceedings of the 2nd EGS Plinius Conference on Mediterranean Storms, Siena 1:463–487

Crosta GB, Frattini P (2003) Distributed modelling of shallow landslides triggered by intense rainfall. Nat Hazards Earth Syst Sci 3:81–93

Crovelli RA (2000) Probability models for estimation of number and costs of landslides. US Geological Survey Open-File Report, 249

Cybenko G (1989) Approximation by superpositions of a sigmoidal function. Math Control Signals Syst 2(4):303–314

Dahal RK, Hasegawa S (2008) Representative rainfall thresholds for landslides in the Nepal Himalaya. Geomorphology 100(3–4):429–443

Dahal RK, Hasegawa S, Nonomura A, Yamanaka M, Masuda T, Nishino K (2009) Failure characteristics of rainfall-induced shallow landslides in granitic terrains of Shikoku Island of Japan. Environ Geol 56(7):1295–1310

Das I, Stein A, Kerle N, Dadhwal VK (2011) Probabilistic landslide hazard assessment using homogeneous susceptible units (HSU) along a national highway corridor in the northern Himalayas, India. Landslides 8(3):293–308

Dikshit A, Sarkar R, Pradhan B, Jena R, Drukpa D, Alamri AM (2020) Temporal probability assessment and its use in landslide susceptibility mapping for eastern Bhutan. Water 12(1):267

Drissia TK, Jothiprakash V, Anitha AB (2019) Flood frequency analysis using L moments: a comparison between at-site and regional approach. Water Resour Manag 33(3):1013–1037

El Adlouni S, Ouarda TB, Zhang X, Roy R, Bobée B (2007) Generalized maximum likelihood estimators for the nonstationary generalized extreme value model. Water Resour Res 43(3):W03410

Finlay PJ, Fell R, Maguire PK (1997) The relationship between the probability of landslide occurrence and rainfall. Can Geotech J 34(6):811–824

Frattini P, Crosta G, Carrara A (2010) Techniques for evaluating the performance of landslide susceptibility models. Eng Geol 111(1–4):62–72

Gardner MW, Dorling SR (1998) Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. Atmos Environ 32(14–15):2627–2636

Garrett JH (1994) Where and why artificial neural networks are applicable in civil engineering. J Comput Civ Eng 8(2):129–130

Geological Society of Korea (1962) Changdong–Hajinburi geological map sheet. Korea Institute of Geoscience and Mineral Resources

Godt JW, Baum RL, Savage WZ, Salciarini D, Schulz WH, Harp EL (2008) Transient deterministic shallow landslide modeling: requirements for susceptibility and hazard assessments in a GIS framework. Eng Geol 102:214–226

Gomez H, Kavzoglu T (2005) Assessment of shallow landslide susceptibility using artificial neural networks in Jabonosa River Basin, Venezuela. Eng Geol 78(1–2):11–27

Greenwood JA, Landwehr JM, Matalas NC, Wallis JR (1979) Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form. Water Resour Res 15(5):1049–1054

Gubareva TS, Gartsman BI (2010) Estimating distribution parameters of extreme hydro-meteorological characteristics by L-moments method. Water Resour 37(4):437–445

Gutiérrez-Martín A (2020) A GIS-physically-based emergency methodology for predicting rainfall-induced shallow landslide zonation. Geomorphology 359:107121

Guzzetti F, Carrara A, Cardinali M, Reichenbach P (1999) Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. Geomorphology 31(1–4):181–216

Guzzetti F, Malamud BD, Turcotte DL, Reichenbach P (2002) Power-law correlations of landslide areas in Central Italy. Earth Planet Sci Lett 195(3–4):169–183

Guzzetti F, Reichenbach P, Cardinali M, Galli M, Ardizzone F (2005) Probabilistic landslide hazard assessment at the basin scale. Geomorphology 72(1–4):272–299

Guzzetti F, Galli M, Reichenbach P, Ardizzone F, Cardinali MJNH (2006) Landslide hazard assessment in the Collazzone area, Umbria, Central Italy. Nat Hazards Earth Syst Sci 6:115–131

Guzzetti F, Peruccacci S, Rossi M, Stark CP (2007) Rainfall thresholds for the initiation of landslides in central and southern Europe. Meteorol Atmospheric Phys 98(3–4):239–267

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

Hosking JR (1990) L-moments: analysis and estimation of distributions using linear combinations of order statistics. J R Stat Soc Series B Stat Methodol 52(1):105–124

Hosking JRM, Wallis JR (1997) Regional frequency analysis: an approach based on L-moments. Cambridge University Press

Hosking JRM, Wallis JR, Wood EF (1985) Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. Technometrics 27(3):251–261

Huffman GJ, Bolvin DT, Nelkin EJ, Wolff DB, Adler RF, Gu G, Hong Y, Bowman KP, Stocker EF (2007) The TRMM multi-satellite precipitation analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. J Hydrometeorol 8:38–55

IPCC (2014) Climate Change 2014. Firth assessment report of IPCC (Intergovernmental Panel on Climate Change), Geneva, Switzerland

Jaiswal P, van Westen CJ (2009) Estimating temporal probability for landslide initiation along transportation routes based on rainfall thresholds. Geomorphology 112(1–2):96–105

Jaiswal P, van Westen CJ, Jetten V (2010) Quantitative landslide hazard assessment along a transportation corridor in southern India. Eng Geol 116(3–4):236–250

Jaiswal P, van Westen CJ, Jetten V (2011) Quantitative estimation of landslide risk from rapid debris slides on natural slopes in the Nilgiri hills, India. Nat Hazards Earth Syst Sci 11(6):1723–1743

Khalil Alsmadi M, Omar KB, Noah SA, Almarashdah I (2009) Performance comparison of multi-layer perceptron (back propagation, delta rule and perceptron) algorithms in neural networks. Proceedings of the 2009 IEEE International Advance Computing Conference, IEEE, pp 296-299

Kim, S. K., Hong, W. P., & Kim, Y. M. (1992). Prediction of rainfall-triggered landslides in Korea. Proceedings of the 6th international symposium on landslides, Balkema, Rotterdam, Netherlands, pp 989-994

Kumar A, Asthana AKL, Priyanka RS, Jayangondaperumal R, Gupta AK, Bhakuni SS (2017) Assessment of landslide hazards induced by extreme rainfall event in Jammu and Kashmir Himalaya, northwest India. Geomorphology 284:72–87

Lee JH, Park HJ (2016) Assessment of shallow landslide susceptibility using the transient infiltration flow model and GIS-based probabilistic approach. Landslides 13:885–903

Lee S, Won JS, Jeon SW, Park I, Lee MJ (2015) Spatial landslide hazard prediction using rainfall probability and a logistic regression model. Math Geosci 47(5):565–589

Martha TR, van Westen CJ, Kerle N, Jetten V, Kumar KV (2013) Landslide hazard and risk assessment using semi-automatically created landslide inventories. Geomorphology 184:139–150

ME (2019) Standard guidelines for design flood estimation. Ministry of Environment

Melillo M, Brunetti MT, Peruccacci S, Gariano S, Guzzetti F (2015) An algorithm for the objective reconstruction of rainfall events responsible for landslides. Landslides 12:311–320

Motamedi M, Liang RY (2014) Probabilistic landslide hazard assessment using copula modeling technique. Landslides 11(4):565–573

National Institute for Disaster Prevention (2006) A study on the monitoring & detection of slope failure (III). Research Report, NIDP-2006-01

National Institute for Disaster Prevention (2009) Study on the steep-slope early warning and evacuation system using rainfall data(II). Research Report, NIDP-2009-07-02

Nefeslioglu HA, Gokceoglu C, Sonmez H, Gorum T (2011) Medium-scale hazard mapping for shallow landslide initiation: the Buyukkoy catchment area (Cayeli, Rize, Turkey). Landslides 8(4):459–483

Oh J, Park HJ (2014) Analysis of landslide triggering rainfall threshold for prediction of landslide occurrence. J Korean Soc Hazard Mitig 14(2):115–129

Paliwal M, Kumar UA (2009) Neural networks and statistical techniques: a review of applications. Expert Syst Appl 36(1):2–17

Park HJ, Lee JH, Woo I (2013) Assessment of rainfall-induced shallow landslide susceptibility using a GIS-based probabilistic approach. Eng Geol 161:1–15

Pham BT, Pradhan B, Tien Bui D, Prakash I, Dholakia MB (2016) A comparative study of different machine learning methods for landslide susceptibility assessment: a case study of Uttarakhand area (India). Environ Model Softw 84:240–250

Pham BT, Tien Bui D, Prakash I, Dholakia MB (2017) Hybrid integration of multilayer perceptron neural networks and machine learning ensembles for landslide susceptibility assessment at Himalayan area (India) using GIS. Catena 149:52–63

Posner AJ, Georgakakos KP (2015) Soil moisture and precipitation thresholds for real-time landslide prediction in El Salvador. Landslides 12:1179–1196

Pradhan B (2013) A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS. Comput Geosci 51:350–365

Pradhan AMS, Kim YT (2015) Application and comparison of shallow landslide susceptibility models in weathered granite soil under extreme rainfall events. Environ Earth Sci 73(9):5761–5771

Pradhan B, Lee S (2010) Delineation of landslide hazard areas on Penang Island, Malaysia, by using frequency ratio, logistic regression, and artificial neural network models. Environ Earth Sci 60(5):1037–1054

Raia S, Alveoli M, Rossi M, Baum RM, Godt JW, Guzzetti F (2014) Improving predictive power of physically based rainfall-induced shallow landslide models: a probabilistic approach. Geosci Model Dev 7:495–514

Rao AR, Hameed KA (2000) Flood frequency analysis. CRC press, Washington

Reichenbach P, Rossi M, Malamud BD, Mihir M, Guzzetti F (2018) A review of statistically-based landslide susceptibility models. Earth-Sci Rev 180:60–91

Robbins J (2016) A probabilistic approach for assessing landslide-triggering event rainfall in Papua New Guinea, using TRMM satellite precipitation estimates. J Hydrol 541(part a):296–309

Romeo RW, Floris M, Veneri F (2006) Area-scale landslide hazard and risk assessment. Environ Geol 51(1):1–13

Rossi G, Catani F, Leoni L, Segoni S, Tofani V (2013) HIRESS:a physically based slope stability simulator for HPC applications. Nat Hazards Earth Syst Sci 13:151–166

Salciarini D, Fanelli G, Tamagnini C (2017) A probabilistic model for rainfall-induced shallow landslide prediction at the regional scale. Landslides 14:1731–1746

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

Segoni S, Piciullo L, Gariano SL (2018) A review of the recent literature on rainfall thresholds for landslide occurrence. Landslides 15(8):1483–1501

Serinaldi F (2015) Dismissing return periods! Stoch Environ Res Risk Assess 29(4):1179–1189

Serinaldi F, Kilsby CG (2015) Stationarity is undead: uncertainty dominates the distribution of extremes. Adv Water Resour 77:17–36

Shou KJ, Yang CM (2015) Predictive analysis of landslide susceptibility under climate change conditions—a study on the Chingshui River Watershed of Taiwan. Eng Geol 192:46–62

Tien Bui D, Pradhan B, Lofman O, Revhaug I, Dick ØB (2013) Regional prediction of landslide hazard using probability analysis of intense rainfall in the Hoa Binh province, Vietnam. Nat Hazards 66(2):707–730

Tien Bui D, Tuan TA, Klempe H, Pradhan B, Revhaug I (2016) Spatial prediction models for shallow landslide hazards: a comparative assessment of the efficacy of support vector machines, artificial neural networks, kernel logistic regression, and logistic model tree. Landslides 13(2):361–378

Tretkoff E (2011) Calculating specific catchment area. Eos 92(27):232–232

Tsuchida T, Athapaththu AMRG, Hanaoka T, Kawaguchi M (2015) Investigation of landslide calamity due to torrential rainfall in Shobara City, Japan. Soils Found 55(5):1305–1317

van Westen CJ, van Asch TWJ, Soeters R (2006) Landslide hazard and risk zonation—why is it still so difficult? Bull Eng Geol Environ 65(2):167–184

Varnes DJ (1984) Landslide hazard zonation: a review of principles and practice. Commission on landslides of the IAEG, natural hazards 3, the UNESCO press, Paris

Vasu NN, Lee SR, Pradhan AMS, Kim YT, Kang SH, Lee DH (2016) A new approach to temporal modelling for landslide hazard assessment using an extreme rainfall induced-landslide index. Eng Geol 215:36–49

Vessia G, Parise M, Peruccacci S, Brunetti MT, Rossi M, Vennari C, Guzzetti F (2014) An automated method for the identification of rainfall events responsible for shallow landslides. Nat Hazards Earth Syst Sci 14:2399–2408

Vessia G, Pisano L, Vennari C, Rossi M, Parise M (2016) Mimic expert judgement through automated procedure for selecting rainfall events responsible for shallow landslide: a statistical approach to validation. Comput Geosci 86:146–153

Vessia G, Di Curzio D, Chiaudani A, Rusi S (2020) Regional rainfall threshold maps drawn through multivariate geostatistical techniques for shallow landslide hazard zonation. Sci Total Environ 705:135815

Wallis JR, Matalas NC, Slack JR (1974) Just a moment! Water Resour Res 10(2):211–219

Wieczorek GF, Glade T (2005) Climatic factors influencing occurrence of debris flows. In: Jakob M, Hungr O, Jakob DM (eds) Debris-flow hazards and related phenomena. Springer, Berlin, pp 325–362

Wu CH (2017) Comparison and evolution of extreme rainfall-induced landslides in Taiwan. ISPRS Int J Geo-Inf 6(11):367

Wu CH, Chen SC, Chou HT (2011) Geomorphologic characteristics of catastrophic landslides during typhoon Morakot in the Kaoping Watershed, Taiwan. Eng Geol 123(1–2):13–21

Yen BC (1970) Risks in hydrologic design of engineering projects. J Hydraul Div 96(4):959–966

Yilmaz I, Kaynar O (2011) Multiple regression, ANN (RBF, MLP) and ANFIS models for prediction of swell potential of clayey soils. Expert Syst Appl 38(5):5958–5966

Yilmaz C, Topal T, Süzen ML (2012) GIS-based landslide susceptibility mapping using bivariate statistical analysis in Devrek (Zonguldak-Turkey). Environ Earth Sci 65(7):2161–2178

Yoo N, Yoon D, Um J, Kim D, Park B (2012) Analysis of rainfall characteristics and landslides at the west side area of Gangwon Province. J Korean Geo-Environ Soc 13(9):75–82

Zare M, Pourghasemi HR, Vafakhah M, Pradhan B (2013) Landslide susceptibility mapping at Vaz watershed (Iran) using an artificial neural network model: a comparison between multilayer perceptron (MLP) and radial basic function (RBF) algorithms. Arab J Geosci 6(8):2873–2888

Zêzere J, Reis E, Garcia RA, Oliveira S, Rodrigues ML, Vieira G, Ferreira AB (2004) Integration of spatial and temporal data for the definition of different landslide hazard scenarios in the area north of Lisbon (Portugal). Nat Hazards Earth Syst Sci 4(1):133–146

Zêzere JL, Garcia RAC, Oliveira SC, Reis E (2008) Probabilistic landslide risk analysis considering direct costs in the area north of Lisbon (Portugal). Geomorphology 94(3–4):467–495

**J.-H. Lee · H.-J. Park** (✉)
Department of Geoinformation Engineering,
Sejong University,
Seoul, Republic of Korea
Email: hjpark@sejong.ac.kr

**H. Kim · J.-H. Heo**
School of Civil and Environmental Engineering,
Yonsei University,
Seoul, Republic of Korea