CrossMark

ORIGINAL ARTICLE

A comparative study of landslide susceptibility mapping using weight of evidence, logistic regression and support vector machine and evaluated by SBAS-InSAR monitoring: Zhouqu to Wudu segment in Bailong River Basin, China

Zhengtuan Xie[1,2] · Guan Chen[1,2] · Xingmin Meng[1,2] · Yi Zhang[1,2] ·
Liang Qiao[1,2] · Long Tan[3]

Abstract The determining of landslide-prone areas in mountainous terrain is essential for land planning and hazard mitigation. In this paper, a comparative study using three statistical models including weight of evidence model (WoE), logistic regression model (LR) and support vector machine method (SVM) was undertaken in the Zhouqu to Wudu segment in the Bailong River Basin, Southern Gansu, China. Six conditionally independent environmental factors, elevation, slope, aspect, distance from fault, lithology and settlement density, were selected as the explanatory variables that may contribute to landslide occurrence based on principal component analysis (PCA) and Chi-square test. The relation between landslide distributions and these variables was analyzed using the three models and the results then used to calculate the landslide susceptibility (LS). The performance of the models was then evaluated using both the highly accurate deformation signals produced by using the Small Baseline Subset Interferometric Synthetic Aperture Radar technique and Receiver Operating Characteristic (ROC) curve. Results show more deformation points in areas with high and very high LS levels, and also more stable points in areas with low and very low LS levels for the SVM model. In

addition, the SVM has larger area under the ROC curve. It indicates that the SVM has better prediction accuracy and classified ability. For the interpretability, the WoE derives the class of factors that most contributed to landsliding in the study area, and the LR reveals that factors including elevation, settlement density and distance from fault played major roles in landslide occurrence and distribution, whereas the SVM cannot provide relative weights for the variables. The outperformed SVM could be employed to determine potential landslide zones in the study area. Outcome of this research would provide preliminary basis for general land planning such as choosing new urban areas and infrastructure construction in the future, as well as for landslide hazard mitigation in Bailong River Basin.

Keywords Landslide susceptibility · Weight of evidence · Logistic regression · Support vector machine · SBAS-InSAR

Introduction

Landslides are one of the most destructive natural hazards and very sensitive to climate change and urban expansion, particularly in mountainous terrain (Keefer and Larsen 2007). The Zhouqu to Wudu segment along Bailong River basin is located at the eastern edge of Tibetan Plateau, in Southern Gansu, China. In response to plateau uplift and river erosion, high mountains with steep slopes, active tectonics, weak rock types and intense rainfall, the Zhouqu to Wudu segment along Bailong River basin is one of the areas most severely affected by landslide and debris flow disasters in China (Derbyshire et al. 2000). As the increased extreme climate events, rapid expansion of urbanization and enhanced tectonic activity in recent years,

✉ Guan Chen
 gchen@lzu.edu.cn

[1] Key Laboratory of West China's Environmental System (Ministry of Education), College of Earth and Environmental Sciences, Lanzhou University, Lanzhou 730000, People's Republic of China

[2] Gansu Environmental Geology and Geohazards Engineering Research Centre, Lanzhou University, Lanzhou 730000, People's Republic of China

[3] Sina Mobile Technology Co., Ltd, Beijing 100080, People's Republic of China

several large landslide events have occurred in the region. The most recent destructive example was the Gansu mudslide, which took place on August 8, 2010. This rainfall-triggered fast-moving debris flow thrust through the urban area, destroying all building along the flow, then rushed into the Bailong River and formed a barrier lake. The debris flow destroyed a large area and claimed at least 1287 lives (Yu et al. 2010; Bai et al. 2013; Cui et al. 2013). Therefore, mapping areas susceptible to landslides is essential for effective land-use planning, disaster management and hazard mitigation in this area.

The landslide susceptibility (LS) mapping technique can be grouped into four categories: landslide inventories, heuristic, statistical and deterministic. Overviews of these categories and their advantages and disadvantages can be found in many publications (Carrara et al. 1995; Soeters and van Westen 1996; Guzzetti et al. 1999; Chacón et al. 2006; Bai et al. 2010; Yilmaz 2010). Within these techniques, statistical methods have been widely applied to determine the landslide susceptibility zones. The statistical methods relying on the basic assumption that environmental factors led to slope failure in the past are susceptible to landslides in the future (Varnes 1984) and are considered to be more suitable for LS mapping over large and complex terrains (Van Westen et al. 2006; Thiery et al. 2007). The bivariate weight of evidence (WoE) was the most favorable bivariate model that could be used to assess the relationship between different categories of each factor and landslide occurrence. This robust and flexible mapping approach has been adopted by numerous landslide studies (Van Westen et al. 2003; Süzen and Doyuran 2004; Thiery et al. 2007; Regmi et al. 2010; Kayastha et al. 2012; Ozdemir and Altural 2013). Among the most popular and widely used statistical method is the logistic regression (LR) as its simple form to solve the complex problem (e.g., landslide occurrence), this method has been applied for many LS zonation at local and regional scale (Dai and Lee 2002; Lee 2004; Ayalew and Yamagishi 2005; Greco et al. 2007; Bai et al. 2010; Kavzoglu et al. 2014). The LR determines the weight of landslide causal factors based on the relative contribution of each in the presence or absence of landslides within a defined land unit. More recent advances in knowledge-based techniques including fuzzy logic, artificial neural networks, decision trees, and support vector machines have been also applied for LS mapping. Specifically, support vector machine (SVM) was a recently developed machine learning (non-parametric technique) tool, which is based on statistical theory and has enhanced ability to solve the nonlinear problem. It is a particularly suitable means of solving the highly complicated relationship between a landslide and its conditioning factors and was preferred for landslide susceptibility mapping

(Yao et al. 2008; Yilmaz 2010; Xu et al. 2012; Pourghasemi et al. 2013; Pradhan 2013; Peng et al. 2014).

Different statistical models provide various results because the predictions of the model are largely determined by the model applied even if the datasets are the same. In such a case, in order to obtain reliable result (for example, results with better prediction accuracy and more realistic which are more suitable for the practical application) in a given region, a comparative study of LS mapping using different methods is necessary and can be highly significant. In the scientific literature, some studies compare the prediction and interpretation capabilities of different methods and techniques for LS assessment (Yalcin 2008; Nandi and Shakoor 2010; Yalcin et al. 2011; Akgun 2012; Mohammady et al. 2012; Schicker and Moon 2012; Hong et al. 2016; Paulín et al. 2016; Youssef et al. 2016). However, LS models include uncertainties and limitations in determining the spatial and temporal risks inherit in areas sensitive to geohazards; in addition, the results of the LS models have seldom been compared with or evaluated by the regional surface deformation information which can be obtained by remote sensing technique such as Interferometric Synthetic Aperture Radar (InSAR) (Kincal et al. 2010). It is necessary to apply more efficient techniques to gain more accurate results to evaluate the model result; this will support a more effective management of geohazards and enable safer land planning.

The main objective of this study is to obtain a reliable LS map in Zhouqu to Wudu segment along Bailong River basin, a region that is seriously affected by landslides. For this purpose, landslide susceptibility maps were prepared by three different statistical methods including WoE, LR and SVM. The results of these models will then be evaluated by the deformation signals measured by the Small Baseline Subset InSAR (SBAS-InSAR) and Receiver Operating Characteristic curve (ROC).

Study area

The Bailong River catchment was the secondary tributary of the Yangtze River. The reach from Zhouqu to Wudu with an area of 5739 km^2 in southern Gansu (Fig. 1), being the region most susceptible to landslide, was selected for analysis. The study area is located in the middle south of the west wing of Qinling orogeny. The dominant landform is alpine canyons, with elevation range from 835 to 4577 m asl. A wide variety of lithological strata with complex structure are present, including materials from the Silurian, Devonian, Carboniferous, Permian, Triassic, Mesozoic Era and Quaternary ages (Fig. 2). In particular, the Silurian
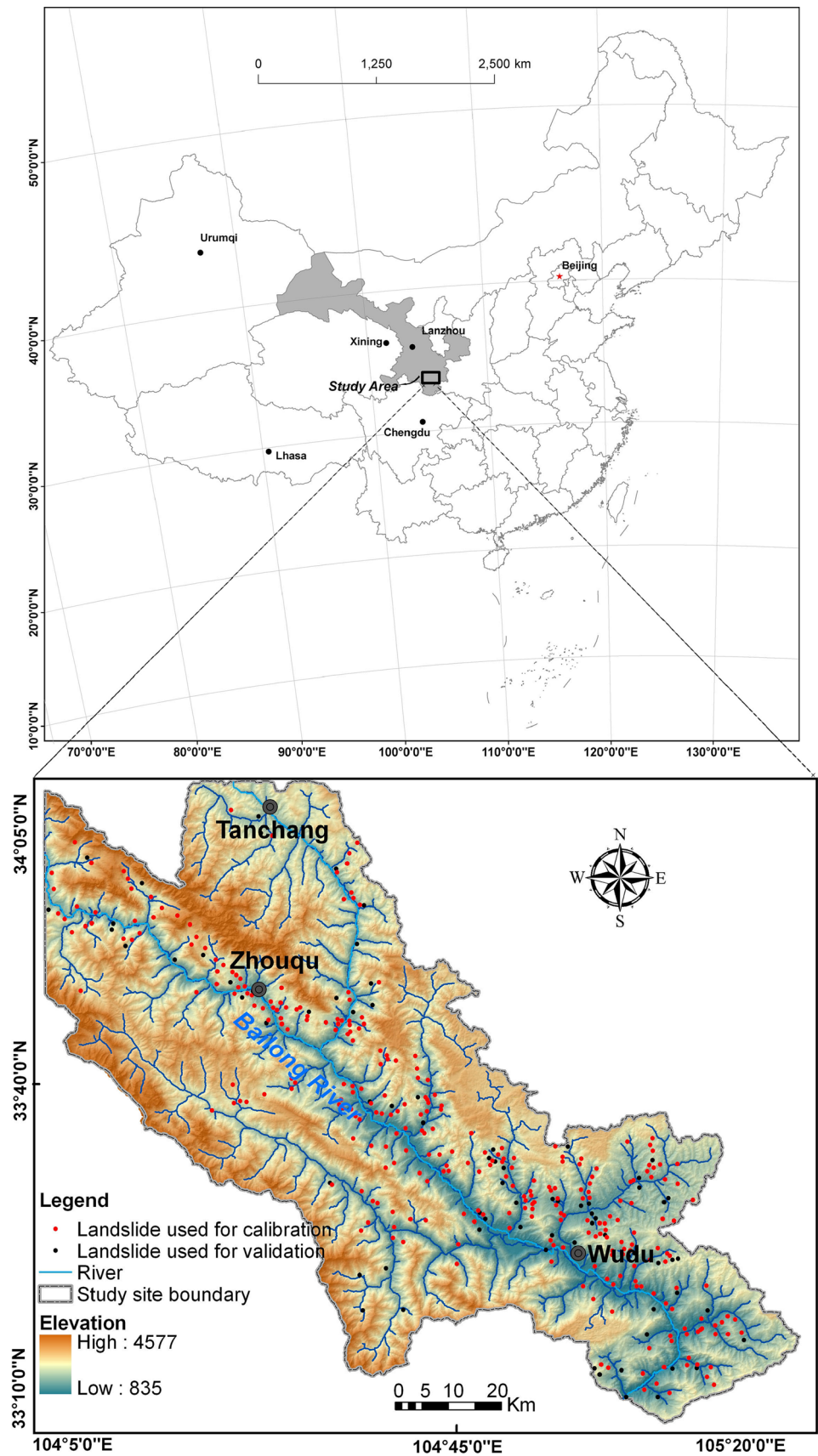
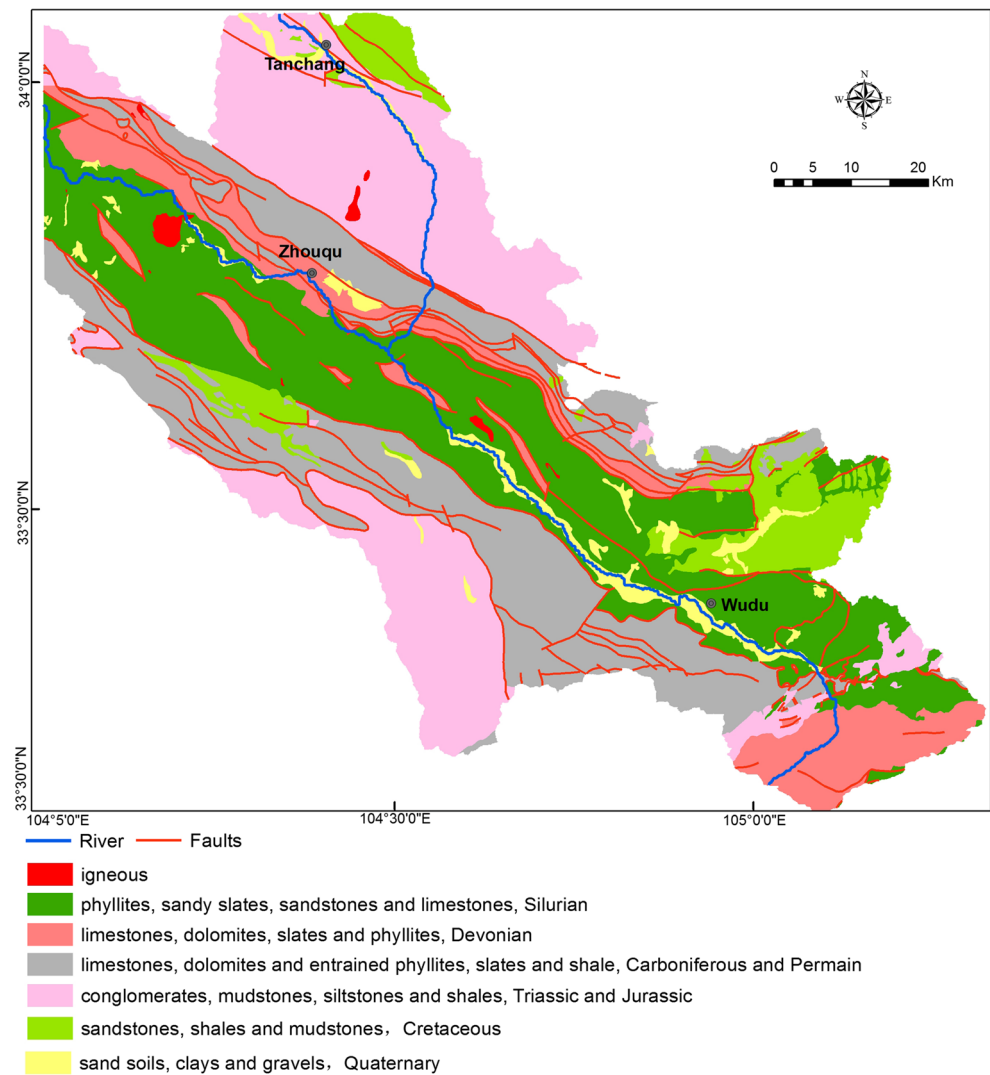**Fig. 1** Location of the study area

Fig. 2 Geological map of the study area

Legend:

— River   — Faults

- igneous
- phyllites, sandy slates, sandstones and limestones, Silurian
- limestones, dolomites, slates and phyllites, Devonian
- limestones, dolomites and entrained phyllites, slates and shale, Carboniferous and Permain
- conglomerates, mudstones, siltstones and shales, Triassic and Jurassic
- sandstones, shales and mudstones，Cretaceous
- sand soils, clays and gravels，Quaternary

phyllite, slate and schist are well known for their high susceptibility to landslide. The study area was strongly controlled by Qinghai–Tibet tectonic belt and Wudu arc-shaped structure and affected by the tectonic uplift of the Qinghai–Tibet plateau (Derbyshire et al. 2000). This segment lies within a seismically active belt, in which the historical Tianshui earthquake in 1654 (Ms = 8.0) and Wudu earthquake in 1879 (Ms = 8.0) were located and in addition, the 2008 Wenchuan earthquake that triggered a great number of landslides and caused movement of abundant loose debris as well as weakening the stability of the slopes in the study area (Xu et al. 2013). Study area receives average annual rainfall ranges from 500 to 900 mm, and 80% of the total rainfall concentrated in the period from June to September (Bai et al. 2013). Human activities are mainly concentrated along the banks of the Bailong River or valley floors, such as Tanchang, Zhouqu

and Wudu (the county town). The main land cover types consist of cultivated land, grass and forest lands.

Data acquisition

Landslide inventory map

In this research, the landslides were detected in the study area by interpretation of SPOT 5 images (scanned on 24th March 2009 with 2.5-m resolution), analysis of available data and extensive field surveys (Chen et al. 2014). The locations of the individual landslides were drawn on at a scale of 1:50,000. Most of the landslides are deep-seated rotational or translational and are generally large scale [according to Varnes' (1984) classification criteria; Fig. 3], with minimum and maximum areas of 0.12 and 1.75 km^2,
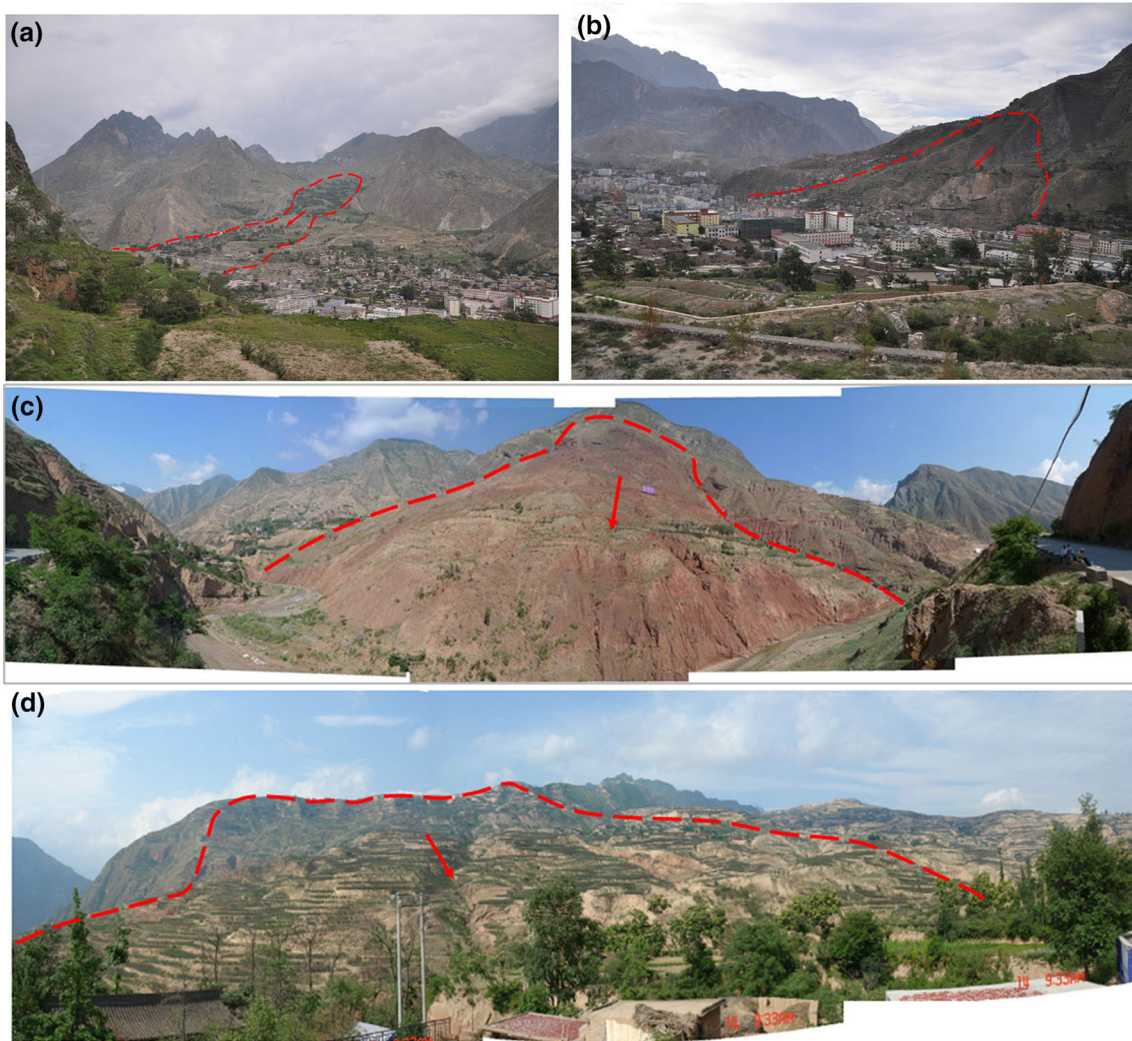
Fig. 3 Typical large deep-seated landslide with potential risk in the study area: **a** suoertou landslide, **b** nanshan landslide, **c** hongtupo landslide, **d** sanjiadi landslide

respectively. In the study area, 375 landslides were delineated in the inventory map (Fig. 1). The inventory map shows only deep-seated landslides and does not include small, shallow landslides, because the former exhibit high risk in the area and are therefore of major concerns. Shallow landslides are more frequent compared to the deep-seated and require high-resolution imagery to acquire their inventories.

Landslide conditioning factors

Seventeen conditioning factors related to landslide occurrence were considered: elevation, slope angle, slope aspect, slope curvature, roughness, lithology, distance from fault, fault density, peak ground acceleration—*PGA*, distance from drainage, drainage density, topographic wetness index—*TWI*, specific catchment area—*SCA,* rainfall, settlement density, land cover and normalized difference vegetation index—*NDVI*. These variables were selected based on the previous study of LS mapping in Bailong River Basin (Bai et al. 2012; Chen et al. 2014).

Topography

Topography controls the spatial variation of soil moisture and the groundwater flow, which plays a significant role in landslide occurrence. (Dai and Lee 2002).

Topographic factors including elevation, slope, aspect, slope curvature and roughness were extracted or calculated based on a digital elevation model (DEM) with 30 m × 30 m resolution. The DEM was created using 1:50,000 scale topographic maps (Fig. 4a–c).
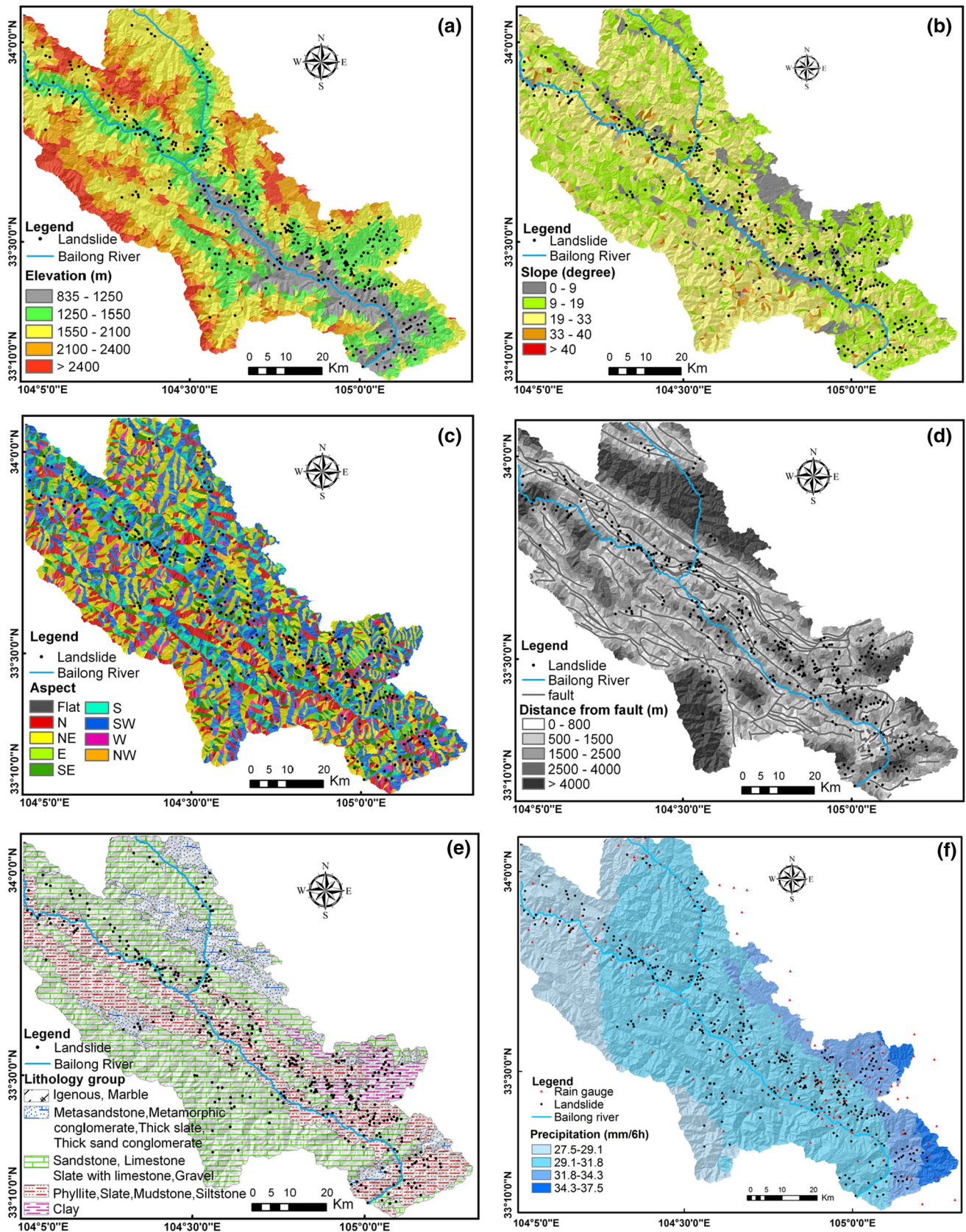
Fig. 4 Landslide and PCA selected factors applying zonal statistics in ArcGIS **a** elevation, **b** slope, **c** aspect, **d** distance from fault, **e** lithology group, **f** precipitation, **g** distance from drainage, **h** land-use type, **i** settlement density
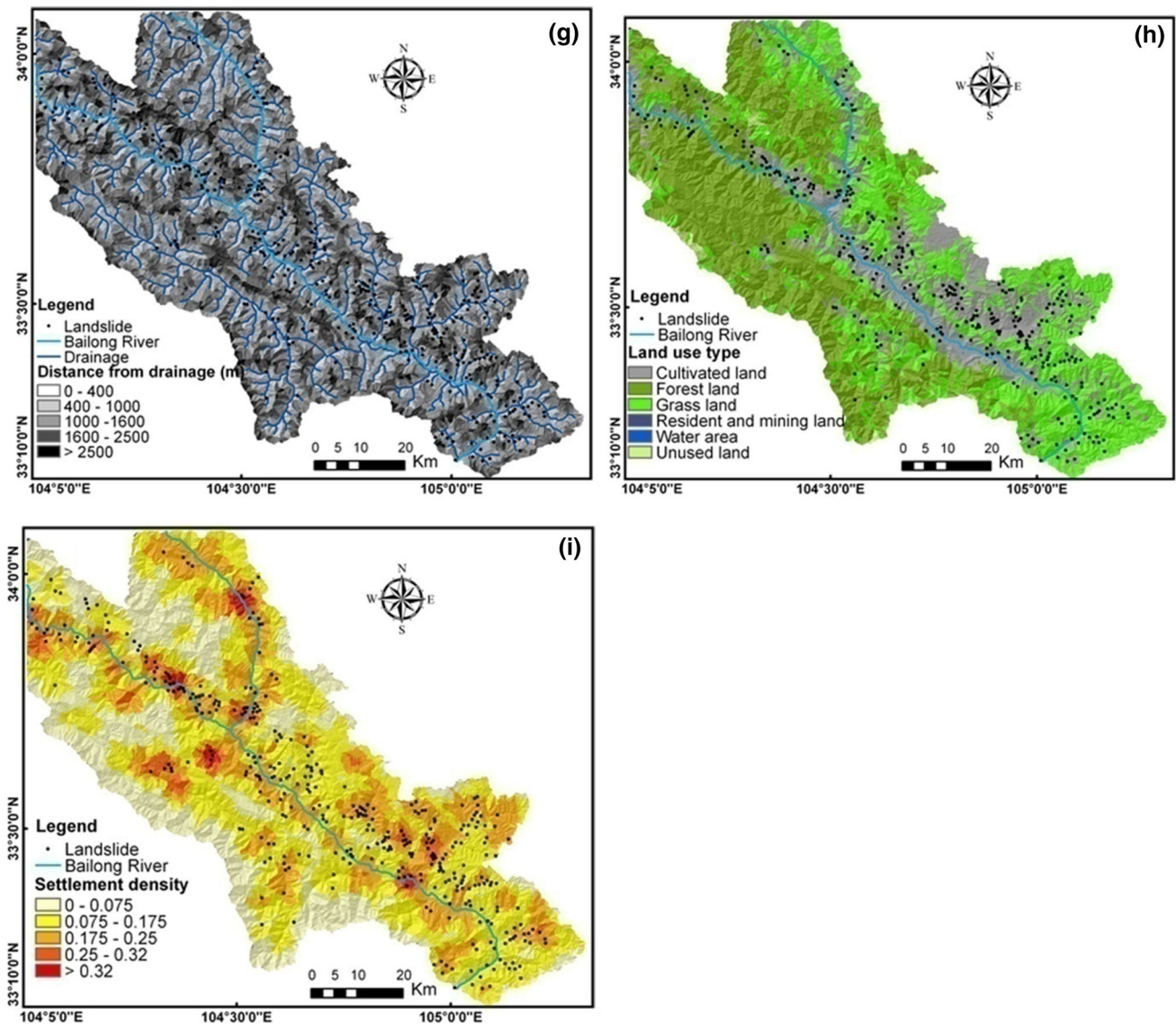
Fig. 4 continued

Geology

Tectonic structure influences the spatial distribution of landslide in this area (Chen et al. 2014). The distance from faults was calculated based on the 1:200,000 geological map, which is the Euclidean distance from the fault (Fig. 4d). The fault density was also calculated in ArcGIS. PGA map derived from the National Seismological Bureau (Lu et al. 2010; Bai et al. 2012) was applied to assess the correlation between landslide and earthquake acceleration on the ground.

Lithology plays a significant role in the distribution of landslides and is associated with the properties of slope-forming materials such as rock mass strength and structure. Geological map of scale 1:200,000 has been divided into five categories (Fig. 4e) according to different lithologies present in the study area (Chen et al. 2006).

Hydrology

Rainfall data from 224 rainfall gauges surrounding the study area were available in this research (Fig. 4f). The rain gauge in the higher elevation are lacked as the stations are mainly concentrated in the river valley. Multi-temporal rainfall data at interval of 6 h were interpolated using spatial Kriging within ArcGIS.

Distance from drainage is the Euclidean distance from the stream network; drainage line density was calculated using ArcGIS 9.3 (Fig. 4g). The TWI and SCA were also calculated based on the topographic data.

Human activities

Land cover is widely considered as an important factor for landslides because it is correlated with their hydrological and mechanical effects. Land-use-type data were processed with supervised classification in ENVI software from Landsat TM5 imagery with a 30-m resolution, which was then verified by field survey (Fig. 4h).

Human activities concentrated in or nearby the towns such as excavation and ramp loading can cause slope instability. Settlement density was considered as another indicator of human activities. Based on the residential data, the settlement density was then computed using ArcGIS (Fig. 4i).

Ecological

The NDVI is a measure of surface reflectance and gives a quantitative estimate of the vegetation growth and biomass and was calculated based on Landsat TM5 image data, which reflects the relation between the vegetation condition and landslides.

Modeling strategy

In this study, LS maps were generated using GIS-based WoE, LR and SVM methods. In order to obtain reliable results, 80% of the landslides were randomly selected to train the model, and an equivalent number of slope units situated outside of the landslides were randomly selected in the areas without landslides.

Selection of major and independent parameters

In order to select the significant and independent factors, principal component analysis (PCA) in combination with Chi-square test was adopted. As a consequence, 17 conditioning factors related to landslide occurrence were applied for PCA analysis. Before the Chi-square test was conducted, the major factors contribute to landslide occurrence had to be classified into different classes, and the maximum likelihood ratio method proposed by Chung and Fabbri (2003) and Chung (2006) was used.

LS mapping using three models

Weight of evidence (WoE)

WoE is a bivariate method using the statistical approach, known as log-linear form of the Bayesian probability model, to estimate the relative importance of evidence (Bonham-Carter 1994).

The WoE approach was employed to calculate the weight of a certain category of a factor map related to the landslide occurrence, expressed as follows:

$$W_i^+ = \log_e\left(\frac{P\{F|L\}}{P\{F|\overline{L}\}}\right) \tag{1}$$

$$W_i^- = \log_e\left(\frac{P\{\overline{F}|L\}}{P\{\overline{F}|\overline{L}\}}\right) \tag{2}$$

$$W_i = W_i^+ - W_i^- \tag{3}$$

where P is the landslide probability, F is the presence of landslide conditioning factor, $\overline{F}$ is the absence of landslide conditioning factor, L is the presence of landslide and $\overline{L}$ is the absence of landslide. W_i^+ and W_i^- indicate that the causative factor is present (positive correlation) and absent (negative correlation), respectively. C (contrast, W_i) is determined by the difference between W_i^+ and W_i^-.

The continuous variables were classified into different categories using the maximum likelihood ratio method proposed by Chung and Fabbri (2003). W_i^+, W_i^- and C for each class of the factors were calculated based on Eqs. (1)–(3) (Table 1).

As the WoE method is restricted by the independence of the input variables, the conditional independence of the major influence factors that contribute to landslide was further examined through pair-wise comparison using Chi-square statistics. Two steps, including conversion of all factors causing landslides into binary mode and preparation of the 2×2 contingency table for all possible pairs of the primary causative factors, should be conducted when the Chi-square test is executed at the 99% significance level and 1 degree of freedom. If the χ^2 value in the contingency table is below 6.63, the pair of dichotomous predictor patterns is independent (Oh and Lee 2010). The studentized value of C referred as the ratio of C to its standard deviation $S(C)$ need to be computed before dichotomous transform (Bonham-Carter 1994). On the basis of the inflection point in the $C/S(C)$ graph, classes which are more susceptibility to landslide were assigned to the class given the value W^+ with the sort have maximum $C/S(C)$. Conversely, W^- with the same rating was assigned to the class less sensitive to landslide occurrence.

Logistic regression model (LR)

The LR is one of the most widely applied multivariate methods in LS mapping as it involves a multivariate regression between a dependent variable and several independent variables, and can yield less biased results (Hosmer and Lemeshow 1989; Atkinson and Massari 1998). In the case of LS mapping, logistic regression can

Table 1 Likelihood ratio classified factor categorical and WoE calculated weight for each class

Parameter	Class	W^+	W^-	W_i	$N(W_i)$
Elevation	835–1250 m	1.43	−0.21	1.64	1.00
	1250–1550 m	0.76	−0.42	1.18	0.91
	1550–2100 m	−0.41	0.18	−0.59	0.55
	2100–2400 m	−2.15	0.20	−2.35	0.19
	>2400 m	−3.11	0.16	−3.27	0.00
Slope	0°–9°	0.03	0.00	0.03	0.79
	9°–20°	0.19	−0.14	0.33	1.00
	20°–31°	−0.16	0.12	−0.29	0.57
	31°–39°	−0.34	0.01	−0.34	0.53
	>39°	−1.10	0.01	−1.11	0.00
Aspect	Flat	−0.69	0.00	−0.70	0.00
	0–45	−0.33	0.06	−0.39	0.24
	45–90	−0.15	0.05	−0.20	0.38
	90–135	0.07	0.00	0.08	0.60
	135–180	0.23	−0.03	0.26	0.74
	180–225	0.20	−0.03	0.22	0.71
	225–270	0.14	−0.04	0.18	0.68
	270–315	0.58	−0.02	0.60	1.00
	315–360	−0.29	0.01	−0.30	0.31
Distance from fault	0–800 m	0.28	−0.09	0.37	1.00
	800–1500 m	0.03	−0.01	0.05	0.81
	1500–2400 m	0.13	−0.03	0.17	0.88
	2400–4500 m	0.08	−0.01	0.09	0.84
	>4500 m	−1.23	0.13	−1.36	0.00
Lithology*	A				
	B	−1.10	0.09	−1.19	0.00
	C	−0.42	0.46	−0.88	0.13
	D	0.79	−0.33	1.11	1.00
	E	0.80	−0.08	0.88	0.90
Rainfall	27.5–29.1 mm	−1.06	0.19	−1.25	0.33
	29.1–31.8 mm	0.16	−0.37	0.53	0.94
	31.8–34.3 mm	0.61	−0.08	0.70	1.00
	34.3–37.5 mm	−2.20	0.03	−2.22	0.00
Distance from drainage	0–400 m	−1.10	0.06	−1.15	0.28
	400–1000 m	0.57	−0.29	0.86	1.00
	1000–1600 m	0.27	−0.13	0.41	0.84
	1600–2500 m	−0.59	0.14	−0.73	0.43
	>2500 m	−1.79	0.13	−1.92	0.00
Land-use	Cultivated land	0.92	−0.54	1.46	1.00
	Forest land	−1.29	0.47	−1.76	0.00
	Grass land	0.04	−0.02	0.06	0.57
	Resident land				
	Water area land				
	Unused land				
Settlement density	0–0.075	−2.58	0.34	−2.92	0.00
	0.075–0.175	0.14	−0.11	0.25	0.67
	0.175–0.25	0.49	−0.22	0.71	0.76
	0.25–0.32	0.62	−0.04	0.66	0.75
	>0.32	1.79	−0.03	1.83	1.00

Lithology* (A: igneous, marble; B: metastone, metamorphic conglomerate, thick layer slate, thick layer sand conglomerate; C: sandstone, medium limestone, slate with limestone, gravel; D: phyllite, slate, mudstone, thin layer limestone, siltstone, E: clay)

determine the best-fit model to describe the relation between the dependent variable (presence or absence of a landslide) and a set of independent parameters, such as slope angle, lithology and distance to drainage (Ayalew and Yamagishi 2005). Furthermore, the regression coefficient determined in the logistic regression can be interpreted as a measure of the relative importance of the independent variables. The logistic model representing the maximum likelihood regression model can be expressed in its simplest form as:

$$P = \frac{1}{1 + e^{-z}} \tag{4}$$

where P is the probability of an occurrence (landslide) that varies from 0 to 1; z is defined as the following equation:

$$z = b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_n x_n \tag{5}$$

where b_0 is a constant, n is the number of independent variables, x_i ($i = 1, 2, 3, \ldots, n$) represents the value of the independent variable, and b_i ($i = 1, 2, 3, \ldots, n$) is the slope coefficient of the model.

Support machine learning (SVM)

The SVM is a machine learning method which was firstly proposed by Vapink (1995). It is based on the statistical approach in order to find an optimal hyper-plane for separating two classes (Kavzoglu et al. 2014). A more detailed SVM algorithm for landslide assessment has recently been depicted (Yao et al. 2008; Marjanović et al. 2011; Xu et al. 2012); it can be summarized as follows:

Consider a set of linear separable training vectors x_i ($i = 1, 2, \ldots, n$) consisting of two classes, denoted as $y_i = \pm 1$. The goal of the SVM is to search for an n-dimensional hyper-plane differentiating the two classes by the maximum gap. Mathematically, it is expressed as:

$$\frac{1}{2} \|w\|^2 \tag{6}$$

subject to the following constraints:

$$y_i((w \cdot x_i) + b) \geq 1 \tag{7}$$

where $\|w\|$ is the norm of the normal of the hyper-plane, b is a scalar base, and $(\cdot)$ denotes the scalar product operation. Introducing the Lagrangian multiplier, the cost function can be defined as:

$$L = \frac{1}{2} \|w\|^2 - \sum_{i=1}^{n} \lambda_i (y_i((w \cdot x_i) + b) - 1) \tag{8}$$

where λ_i is the Lagrangian multiplier. The solution can be achieved by dual minimizing of Eq. (8) with respect to w and b through the standard procedures; the detailed

discussions can be found in Vapink (1995) and Tax and Duin (2002).

For application to a complicated non-separable problem, the Slack variable ξ_i can be introduced to modify the following limitation:

$$y_i((w \cdot x_i) + b) \geq 1 - \xi_i \tag{9}$$

And modification of Eq. (6) is as follows:

$$L = \frac{1}{2} \|w\|^2 - \frac{1}{vn} \sum_{i=1}^{n} \xi_i \tag{10}$$

where $v(0, 1)$ is introduced to account for misclassification. In addition, Vapink (1995) brought in a kernel function $K(x_i, x_j)$ to account for a nonlinear decision boundary (Yao et al. 2008).

LS class rating

In this study, we applied the maximum likelihood ratio classification method (Chung and Fabbri 2003) to divide the LS values. The likelihood ratio classification method was then compared with the automated "natural break" methods in order to verify the robustness of this method. The latter method is conventional and widely used for classification of the susceptibility map (Schicker and Moon 2012). Relative landslide density (RLD) is derived from the ratio of percentages of total landslide area in each susceptibility category to the total area in the class and gives an indication of the goodness of fit. (Santacana et al. 2003; Arora et al. 2004)

$$RLD = (n_i/N_i)/\sum(n_i/N_i) \tag{11}$$

where n_i is the sum of the landslide area with susceptibility level i, and N_i is the total area of the susceptibility class i. By deriving RLD for each class, comparative charts for the two classifications method can be plotted.

Accuracy assessment

Future landslides provide a much better means of assessing the performance of the models. As there was no temporal information on the landslide dataset in this research, the spatiotemporal surface deformation which could reflect the activity of landslide to a certain extent was firstly applied to validate the three models. InSAR can be used to detect ground deformation by calculating the phase differences in complex SAR images acquired in similar geometric conditions, but at two different epochs. The SBAS-InSAR allows reliable deformation measurements to be obtained by implementing an easy combination of the SAR interferograms generated from an appropriate selection of SAR data pairs characterized by a small spatial and temporal

baseline (Tizzani et al. 2007). The StaMPS (Stanford Method for Persistent Scatterers) package has proven its ability to study landslide dynamics even in a densely vegetated environment. In this study, the Delft Object-Oriented Radar Interferometric Software (DORIS) and StaMPS package were used to process 55 ENVISAT images from descending track 018 and 290 collected between November 2003 and September 2010, and the detailed data processing of SBAS-InSAR in the study area could be found at Zhang et al. (2016). Due to the decorrelation caused by layover and shadows in mountainous regions, we only selected areas of less rugged terrain in populated valleys as our SBAS-InSAR calculated areas (Zhang et al. 2016). The SBAS-derived mean velocity map of the Bailong River Basin is used for further model validation.

In order to access the models in a more quantitative way, landslides, other than those from the calibrating dataset with a total of 75 landslides and equivalent slope units, were randomly selected from the non-landslide locations and prepared for model verification. The receiver operating characteristic (ROC) curve which shows goodness of fit was also applied to assess the performance of the three LS methods.

Results

Principal and conditional independent factors

As shown in Table 2, nine factors including elevation, slope gradient, aspect, lithology, distance from fault, rainfall, distance from drainage, land-use and settlement density were identified as the main factors contributing to landslide occurrence and accounting for 79.3% of the total variance (Fig. 4). A multicollinearity diagnosis was further implemented to test the correlation for each of the major factors; the variance inflation factors (VIF) and tolerance (TOL) are two important indexes for multicollinearity diagnosis. According to Allison (2001), the variables with VIF > 2 and TOL < 0.4 were identified as multicollinearity with other factors; as a result (Table 3), nine major factors were found to be "independent" of each other.

Table 4 summarizes the result of the Chi-square test, the χ^2 values higher than 6.64 reflecting a stronger association with other factors. Of the 9 key influencing factors, only six factors (consisting of elevation, slope, aspect, distance from fault, lithology and settlement density) were found to be extremely independent of each other; these six factors were utilized to construct the three LS mapping models in this research. The outcome shows that the influencing factors including elevation, rainfall

Table 2 Total variance explained of the nine selected variables

Component	% of variance	Cumulative %
Elevation	19.325	19.325
Slope	12.992	32.317
Distance to fault	9.637	41.954
Distance to drainage	8.640	50.594
Settlement density	7.147	57.740
Lithology	6.578	64.318
Aspect	5.450	69.768
Precipitation	4.953	74.721
Land-use	4.628	79.349

Table 3 Multicollinearity diagnosis indexes for principle variables

Independent variables	TOL	VIF
Elevation	0.482	2.073
Slope	0.931	1.074
Distance to fault	0.953	1.050
Distance to drainage	0.853	1.172
Settlement density	0.588	1.700
Lithology	0.821	1.219
Aspect	0.994	1.006
Precipitation	0.864	1.157
Land-use	0.703	1.423

Table 4 Pairwise Chi-square statistics of the nine principal factors

	EL	SL	AS	DF	DD	PR	LI	LA	SD
EL		1.32	0.81	6.96	**12.45**	**11.6**	4.37	**22.56**	1.05
SL			1.56	1.67	4.68	0.15	0.91	0.04	0.55
AS				0.14	**8.11**	0.01	0.02	0.69	0.08
DF					2.72	0.58	2.35	0.01	4.9
DD						0.03	0.33	1.23	1.03
PR							2.26	**25.29**	0.21
LI								**14.5**	1.46
LA									0
SD									

Chi-square values in bold indicate the conditional dependence

EL elevation, *SL* slope, *AS* aspect, *DF* distance from fault, *DD* distance from drainage, *PR* precipitation, *LI* lithology, *LA* land-use type, *SD* settlement density

and land-use that were identified as independent of each other based on the VIF and TOL judgment criteria when linear regression was undertaken; in contrast, the three factors are closely correlated within the Chi-square test. It could be concluded that when the conditional independent of the factors was required, the Chi-square test was more effective to detect the correlation of the input factors in this research.
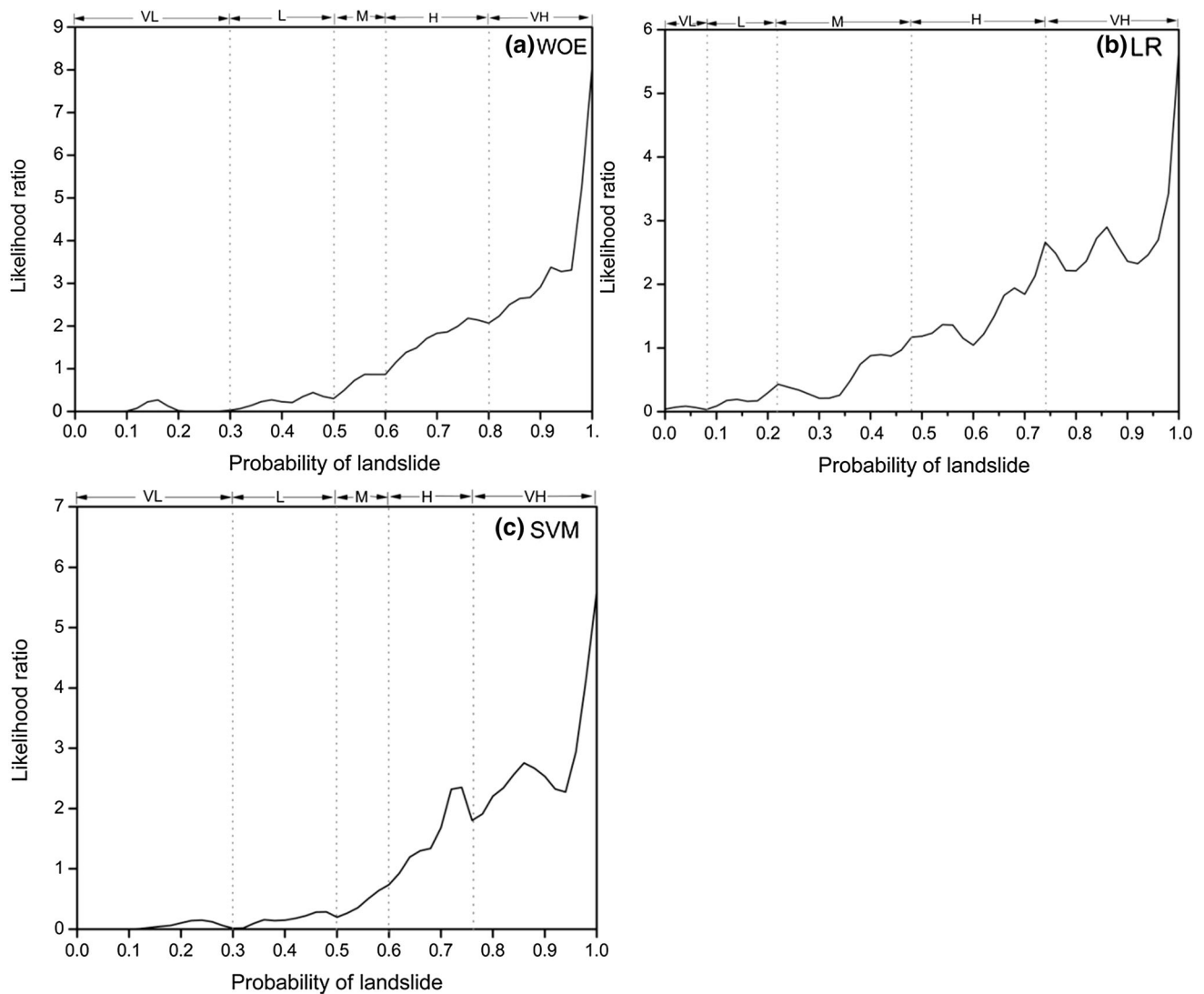
Fig. 5 Empirical likelihood ratio classification of the three models **a** WoE, **b** LR and **c** SVM, (*VL* very low, *L* low, *M* moderate, *H* high, *VH* very high)

LS class rating and LS maps

Figure 5 shows the likelihood ratio curve of landslide probability with uptrend across the whole range; this is consistent with the real condition that higher probability corresponds to larger likelihood ratio. Their cutoff points were determined according to their characteristic inflection point (Table 5). Susceptibility maps can be classified into the following categories: very low susceptibility, low susceptibility, moderate susceptibility, high susceptibility and very high susceptibility. On the basis of the charts shown in Fig. 6, it is clear that in both the likelihood ratio and natural breaks classifications considered here, the likelihood ratios are apparently superior for the LR as this method produces the highest densities in both the high and very high classes and the lower densities from the very low to

Table 5 Landslide probability range of different LS class for the three models

Model	LS class				
	Very low	Low	Moderate	High	Very high
WOE	0–0.30	0.3–0.50	0.50–0.60	0.60–0.80	0.80–1.00
LR	0–0.08	0.08–0.20	0.20–0.48	0.48–0.74	0.74–1.00
SVM	0–0.30	0.30–0.50	0.50–0.60	0.60–0.77	0.77–1.00

the moderate class. In the case of the WoE and SVM, the likelihood ratios and the natural breaks are very similar relative for these two models. Both classified methods in the WoE and SVM show low density in the very low and low susceptibility classes, and peak relative density in the very high class in both likelihood ratio and natural breaks.
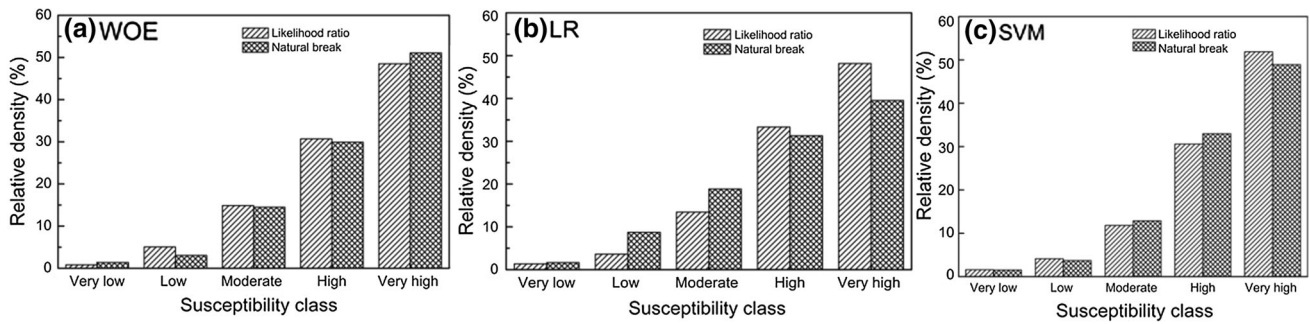
Fig. 6 Comparison of the relative density for the three models applied using likelihood ratio and natural breaks classified methods **a** WoE, **b** LR and **c** SVM
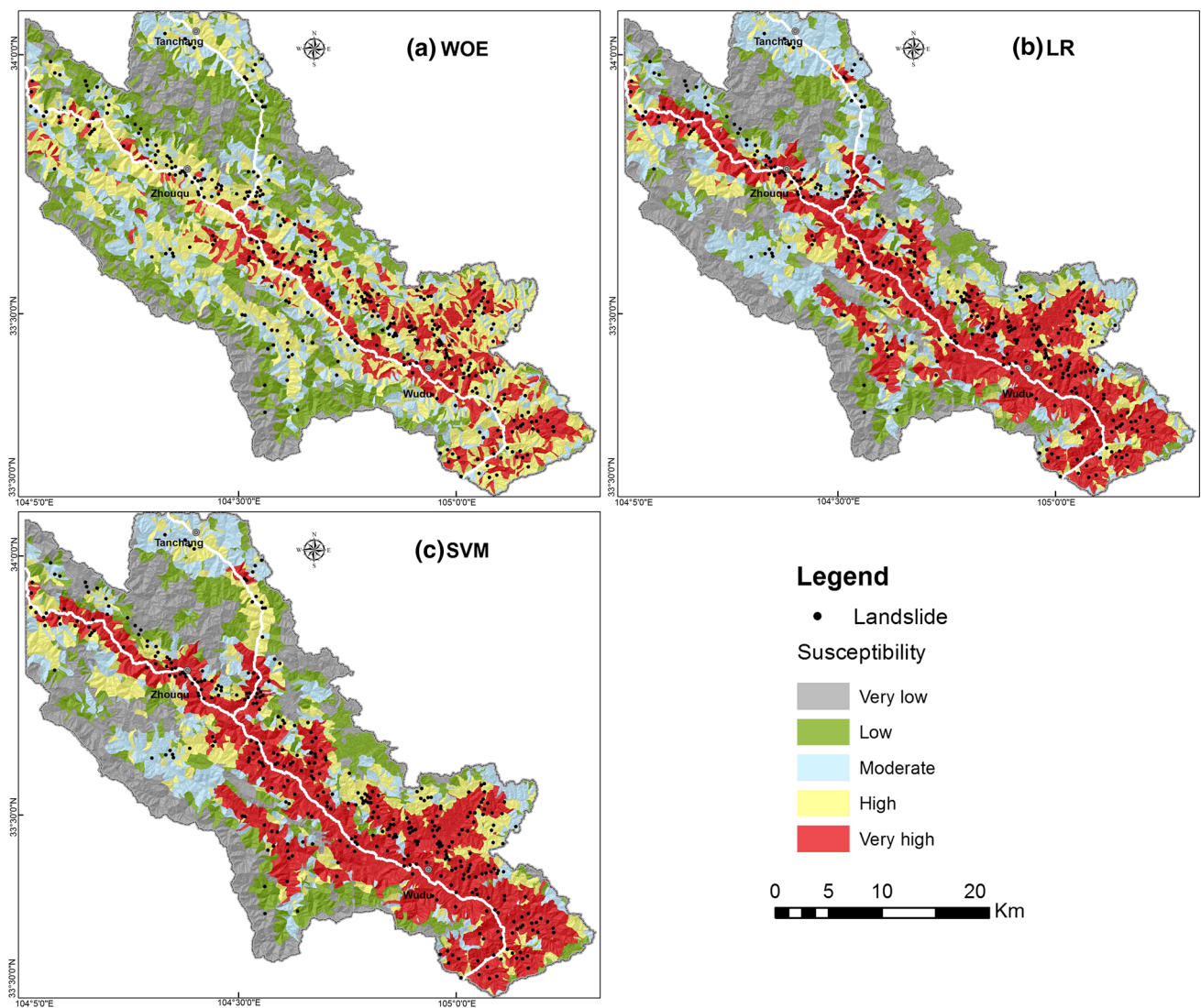

Fig. 7 Landslide susceptibility mapping generated by **a** WoE, **b** LR and **c** SVM

These results indicate that the likelihood ratio method can also result in robust classification in all three models, providing a categorization that most sensitively reflects the true distribution of landslides in the region.

The LS maps produced by the three models (WoE, LR and SVM) are presented in Fig. 7a–c, while Table 6 presents the numeric results.

Table 6 LS class distribution of the three models

LS class	Model		
	WOE	LR	SVM
Very low (%)	13.46	24.20	23.48
Low (%)	13.07	14.86	16.50
Moderate (%)	25.38	19.52	14.15
High (%)	17.97	14.58	15.97
Very high (%)	30.12	26.84	29.90

Accuracy evaluation

To evaluate the three LS models, the deformation information derived from SBAS-InSAR was overlaid onto the susceptibility maps. Since the majority of the basin has been a rural environment with its consequent vegetation cover, high slope inclination, and unsuitable orientation, most of the region has no radar coverage (Colombo et al. 2006) and single descending acquisition data therefore lead to an absence of information on movement on NE-facing slopes (Meisina et al. 2008; Zhang et al. 2016). The mean velocity maps of the Bailong River Basin were constructed (Fig. 8a–c). The velocity was extracted along the line of sight (LOS) of the satellite, which is 23° on average from the vertical. Reliability of the high precise deformation derived from the SBAS-InSAR in this study area has been proved by field survey (Zhang et al. 2016). In Fig. 8a–c, red (−25 to −8 mm/year) and orange (−8 to −2 mm/year) points indicate subsidence, green points (−2 to 2 mm/year) represent stable, and light blue (2–7 mm/year) and blue (7–22 mm/year) points imply uplift. It should be noted that the density of SBAS-InSAR points does not indicate the magnitude of deformation. The substantial clusters of
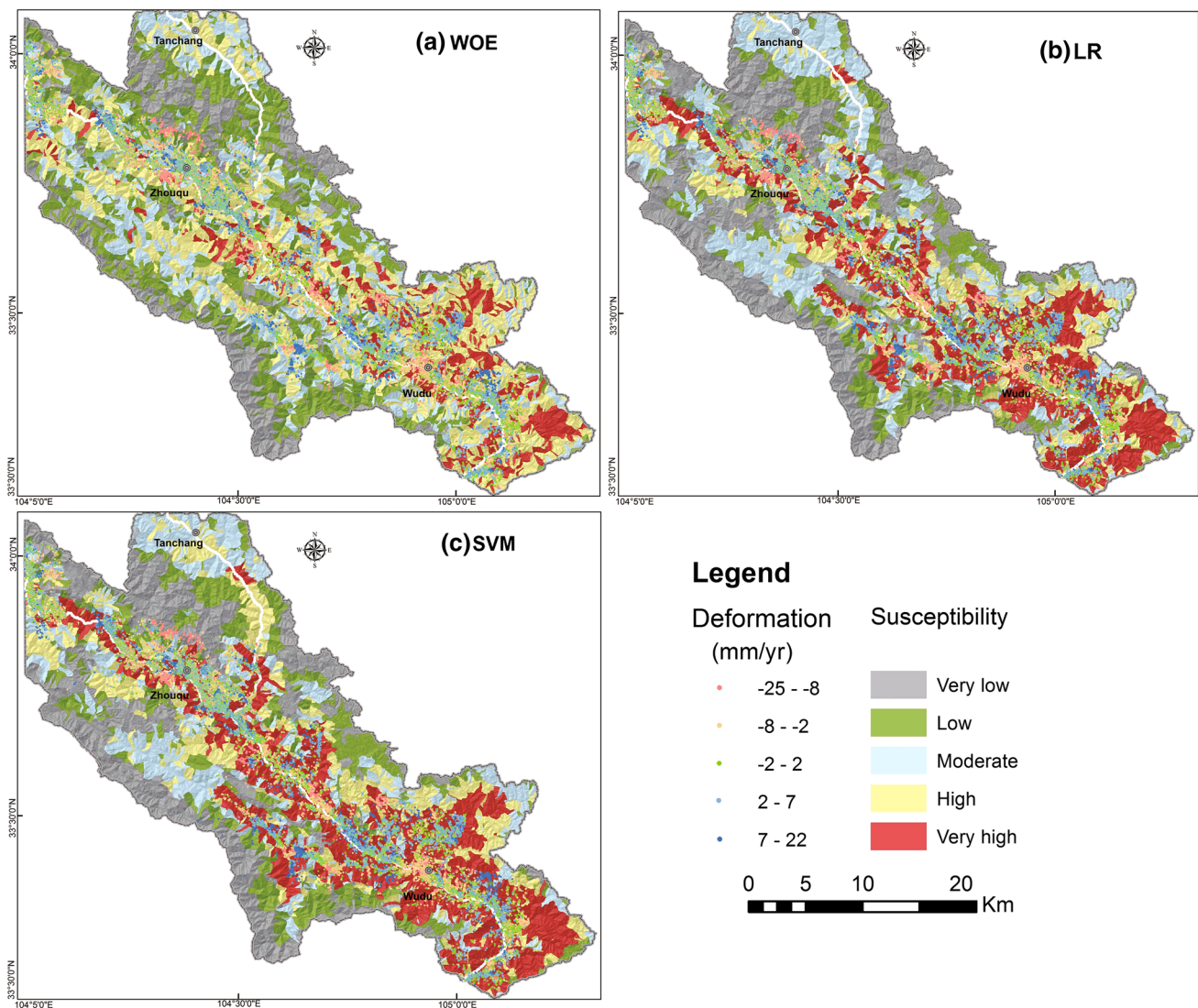

Fig. 8 Landslide susceptibility maps overlaid with SBAS-InSAR derived deformation. **a** WoE, **b** LR, **c** SVM

Table 7 Percentages of SDFP points versus landslide susceptibility classes for WoE, LR and SVM models

Deformation (mm/year)	Susceptibility				
	Very high (%)	High (%)	Moderate (%)	Low (%)	Very low (%)
WoE					
−25 to −8 (subsidence)	3.56	2.91	0.39	0.56	0.01
−8 to −2 (subsidence)	13.04	9.31	0.86	0.56	0.00
−2 to 2 (stable)	15.79	14.42	1.19	0.48	0.01
2–7 (uplift)	13.46	14.17	1.42	0.27	0.00
7–22 (uplift)	3.07	3.81	0.67	0.06	0.00
LR					
−25 to −8 (subsidence)	5.31	0.95	0.45	0.21	0.51
−8 to −2 (subsidence)	18.76	2.69	1.52	0.39	0.38
−2 to 2 (stable)	25.77	3.48	1.88	0.47	0.28
2–7 (uplift)	23.51	2.77	2.47	0.40	0.16
7–22 (uplift)	5.90	0.82	0.76	0.12	0.04
SVM					
−25 to −8 (subsidence)	5.80	0.89	0.06	0.26	0.42
−8 to −2 (subsidence)	19.86	2.92	0.13	0.65	0.20
−2 to 2 (stable)	27.09	3.48	0.20	1.03	0.10
2–7 (uplift)	24.66	3.64	0.20	0.76	0.05
7–22 (uplift)	6.11	1.29	0.08	0.10	0.02

SBAS-InSAR points mostly correspond to more populated areas in the river valley (e.g. towns of Zhouqu, Wudu) which possibly exhibit high radar backscatters from man-made buildings, and also the slowly decorrelating filtered phase (SDFP) points were not evenly distributed. It can be seen in Fig. 8a–c and Table 7 that there is a good correspondence of SDFP points to the LS models, 63.31, 60.71 and 65.16% of the deformation points (both subsidence and uplift) are located in areas classified by the LS models as high and very high susceptibility to landslide, for the WoE, LR and SVM, respectively, and the stable points correspond to the low and very low LS, for the WoE, LR and SVM were 0.49, 0.76 and 1.13%, respectively. It could be concluded that the SVM classified more instability slope as high LS level, and also more stable slope as low LS level. As a consequence, it indicates that the result of SVM is more reasonable and has a good engineering application value. Nevertheless, we also should believe that it is not appropriate to consider each deformation point in Fig. 8a–c to be a landslide or mass movement.

Determination of the accuracy of different models applied in this study was also achieved by plotting ROC curves. The area under the ROC curve (AUC) can be used to assess the quantitative prediction accuracy. For the calibration set, the AUC values for the WoE, LR and SVM were found to be 0.797, 0.840 and 0.844, respectively (Fig. 9a), which indicates that the LR and SVM are models with a better training capability. With respect to the prediction skill, the three models show a similar tendency with regard to the calibration set. The AUC of the WoE is 0.777, while the AUCs of the LR and SVM methods are 0.812 and 0.830, respectively (Fig. 9b). Inspection of Fig. 9 clearly shows the similar performance of the LR and SVM. On this basis, it can be concluded that estimations drawn from the LR and SVM are relatively good in determining LS in the study area, whereas the WoE model is a relative poor estimator. The SVM model is with better prediction ability and is much more robust statistically. And this is consistent with the result evaluated by high precise deformation derived from SBAS-InSAR.

Discussion

The reliability of LS maps depends mostly on the quality of available data and on the model applied (Ayalew and Yamagishi 2005; Yilmaz 2010; Pourghasemi et al. 2013). As can be seen in Fig. 7, several landslides in Northwestern Zhouqu were classified into low to moderate LS level, and these landslides were not highlighted by an apparently increased LS level in any of the three models. This may be caused by the limited quality of input data: as shown in Fig. 4d, most of the landslides in Northwestern Zhouqu are located in limestone areas, and this stratum was classified as relative hard rock; however, the detailed rock property such as weathering rate and joint distribution could not be referenced in the current applied geological map. In addition, rain gauges in the Northwestern Zhouqu are mainly
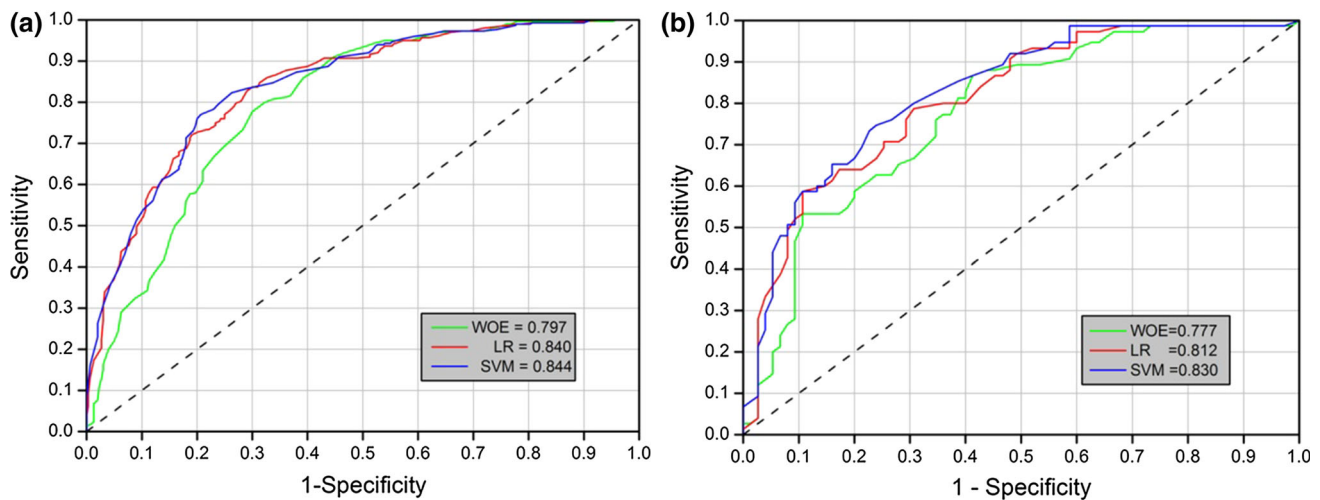
Fig. 9 ROC curve evaluation of the model fit and predictive capability of the three models. **a** Calibration set, **b** validation set

concentrated in the river valley, with no rain gauges emplaced at higher elevation areas (Fig. 4f). This may also influence the LS result to some extent. Therefore, better quality input data would be needed to solve this problem.

Model evaluation is a significant process for LS mapping, which determines whether the LS map could be applied for land planning and hazard mitigation; in addition, it could provide reliable basis on model comparison. Traditional evaluation mostly used the contemporaneous landslide or "future" landslide which were not applied for calibration to test the model performance; besides, these dataset could not reveal the real activity state of landslide. In this study, the surface deformation derived from the SBAS-InSAR was applied to evaluate the three LS models. The surface deformation can also reveal the dynamic deformation processes of landslides, and these activity states of landslides are with more concern to the decision maker. As seen from Fig. 8a–c and Table 7, for the three LS models, most of the deformation points correspond well to the high and very high LS levels. And it could also challenge the traditional view that the knowledge-based SVM was prone to overfitting (Brenning 2005; Pradhan 2013), however, in this study, more stable points (1.13%) corresponding to low and very low LS levels for the SVM when compared with the WoE and LR. This could provide significant support that the SVM is effective in predicting the slope units with deformation as high or very high LS level and also classifying the stable slope units as low and very low LS level.

However, there still are some limitations by applying the SBAS-InSAR for evaluating the model accuracy in this research. Firstly, most of the SDFP points were concentrated in the populated valley; in addition, these SDFP points also cover other deformation types including potential landslides, the movement of debris and

subsidence (Zhang et al. 2016). Although some researches have made use of SAR interferometry for LS assessment or updating the risk map (Singh et al. 2005; Lu et al. 2014), how to efficiently integrate LS models and InSAR technique to obtain more reliable LS maps is still a problem. A potential solution might be to seek a strategy that could demonstrate the temporal deformation information, either with highly deformed history or with long-term stability, of landslides in the LS map. This could possibly be achieved by analyzing the relationship between the ground deformation which is obtained by time series analysis of PS-InSAR and SBAS-InSAR, and environmental factors with statistical models.

The lack of interpretability regarding the contribution of individual variables when applied to statistical methods for LS mapping has always been criticized. In the case of the three models used here, the importance of the different classes of the influencing factors related to slope instability could be highlighted in the WoE model (Table 1), as this method using an categorization to transform the nominal variables to numeric variable, the weighted value of each class could be obtained; the spatial distribution of LS obtained by the WoE models was mainly influenced by elevations from 835 to 1150 m, slope ranges between 9° and 20°, aspect between 270° and 315°, the 0–800 m class of proximity to faults, lithology with phyllite, slate, mudstone, thin limestone and siltstone groups, and settlement more than 0.3, as relatively high weights assigned to these classes. In addition, the most important contributing factors could also be identified from the WoE model; for example, the settlement density, elevation and lithology were found to be having the higher maximum contrast weight (Table 1). With regard to the LR, each factor can be assigned a weight; there is an element which cannot be attributed to any one factor but to the group as a whole

which is represented by the model intercept in LR (Schicker and Moon 2012). Therefore, relative weights of the each influence factor can be derived from the model (Table 3). The coefficient for the elevation, settlement density and distance from faults was relatively higher in the LR model, indicating that these three factors dominated the results issued by the LR model. For the SVM, it is not an interpret model just like a black box which cannot provide relative weights for the variables.

Conclusion

In this study, three statistical models, including the WoE, LR and SVM pertain to the bivariate statistical, multivariate statistical and knowledge-based statistical methods, respectively, are compared on LS mapping in a study of the Zhouqu to Wudu segment in Bailong River Basin, China. Six major independent explanatory variables, i.e., elevation, slope, aspect, distance from fault, lithology and settlement density, related to landslide occurrence were selected after implementation of the PCA and Chi-square test. The accuracy of the models was evaluated by the SBAS-InSAR derived high precise deformation and ROC. Results show that the three LS maps agree well with our SBAS-InSAR-derived deformation maps, most of the deformation points (subsidence or uplift) are correspond to the high and very high LS level, which are 63.31, 60.71 and 65.16% for the WoE, LR and SVM, respectively, particularly more points of stable correspond to the low and very low LS level for the SVM. The AUCs values for the prediction of the WoE, LR and SVM are 0.777, 0.812 and 0.830, respectively. The accuracy evaluation provides a strong support of SVM with better performance on LS mapping in the study area.

With regard to their operational and interpretation capability, the WoE is a simple bivariate model, whereas the LR and SVM are much more complicated when applied to LS mapping. The WoE is effective in determining the contribution of different class rates in the controlling factors, which derives the class of factors consisting of the elevation from 835 to 1150 m, slope range from 9° to 20°, aspect between 270° and 315°, the class 0–800 m of proximity to fault, lithology with phyllite, slate, mudstone, thin limestone and siltstone group, and settlement density more than 0.3 most contributed to landsliding in the study area; with regard to the LR, the relative weights in each factor as a group can be derived reveal that factors including elevation, settlement density and distance from fault played major roles in landslide occurrence and distribution. Finally, the SVM is not an interpretable model; rather, it is like a black box in which the internal processing steps are difficult to follow.

To sum up, the SVM with better prediction ability was the most reasonable model for LS mapping of the study area, whereas the WoE and LR have better interpretation capability could be applied to enhance our understanding between the landslide occurrence and the landslide conditioning factors. Thus, the SVM LS maps could be used as the preliminary basis by decision makers, planners, and engineers to avoid and/or minimize the damage and losses caused by existing and future landslides. However, site-specific studies need to be undertaken so as to complement the assessment and when the results of these models are taken into practical engineering applications.

References

Akgun A (2012) A comparison of landslide susceptibility maps produced by logistic regression, multi-criteria decision, and likelihood ratio methods: a case study at İzmir, Turkey. Landslides 9(1):93–106

Allison PD (2001) Logistic regression using SAS System: theory and application. Wiley Interscience, New York

Arora M, Das Gupta A, Gupta R (2004) An artificial neural network approach for landslide hazard zonation in the Bhagirathi (Ganga) Valley, Himalayas. Int J Remote Sens 25(3):559–572

Atkinson P, Massari R (1998) Generalised linear modelling of susceptibility to landsliding in the central Apennines, Italy. Comput Geosci 24(4):373–385

Ayalew L, Yamagishi H (2005) The application of GIS-based logistic regression for landslide susceptibility mapping in the Kakuda–Yahiko Mountains, Central Japan. Geomorphology 65(1):15–31

Bai SB, Wang J, Lü GN, Zhou PG, Hou SS, Xu SN (2010) GIS-based logistic regression for landslide susceptibility mapping of the Zhongxian segment in the Three Gorges area, China. Geomorphology 115(1):23–31

Bai S, Wang J, Zhang Z, Cheng C (2012) Combined landslide susceptibility mapping after Wenchuan earthquake at the Zhouqu segment in the Bailongjiang Basin, China. CATENA 99(4):18–25

Bai S, Wang J, Thiebes B, Cheng C, Yang Y (2013) Analysis of the relationship of landslide occurrence with rainfall: a case study of wudu county, china. Arab J Geosci 7(4):1277–1285

Bonham-Carter GF (1994) Geographic information systems for geoscientists: modelling with GIS, 13. Computer methods in the geosciences. Pergamon Press, New York, p 414

Brenning A (2005) Spatial prediction models for landslide hazards: review, comparison and evaluation. Nat Hazards Earth Syst Sci 5(6):853–862

Carrara A, Cardinali M, Guzzetti F, Reichenbach P (1995) GIS technology in mapping landslide hazard. In: Carrara A, Guzzetti F (eds) Geographical information systems in assessing natural hazards. Kluwer, Dordrecht, pp 135–176

Chacón J, Irigaray C, Fernández T, Hamdouni RE (2006) Engineering geology maps: landslides and geographical information systems. Bull Eng Geol Envron 65(4):341–411

Chen WW, Zhao ZF, Liu G, Liang SY (2006) The engineering geological problems study of Gansu section of Lanzhou–Haikou highway. Lanzhou University Press, Lanzhou (in Chinese)

Chen G, Meng X, Tan L, Zhang F, Qiao L (2014) Comparison and combination of different models for optimal landslide susceptibility zonation. Q J Eng GeolHydrogeol 47(4):283–306

Chung CJ (2006) Using likelihood ratio functions for modeling the conditional probability of occurrence of future landslides for risk assessment. Comput Geosci 32(8):1052–1068

Chung CJF, Fabbri AG (2003) Validation of spatial prediction models for landslide hazard mapping. Nat Hazards 30(3):451–472

Colombo A, Mallen L, Pispico R, Giannico C, Bianchi M and Savio G (2006) Mappatura regionale delle aree monitorabili mediante l'uso della tecnica PS. Proceedings of National Conference ASITA, Bolzano, Italy, pp 14–17

Cui P, Zhou GG, Zhu X, Zhang J (2013) Scale amplification of natural debris flows caused by cascading landslide dam failures. Geomorphology 182(427):173–189

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

Derbyshire E, Meng X, Dijkstra TA (2000) Landslides in the thick loess terrain of North-West China. Wiley, Chichester, pp 1–256

Greco R, Sorriso-Valvo M, Catalano E (2007) Logistic regression analysis in the evaluation of mass movements susceptibility: the Aspromonte case study, Calabria, Italy. Eng Geol 89(1):47–66

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):181–216

Hong H, Pourghasemi HR, Pourtaghi ZS (2016) Landslide susceptibility assessment in lianhua county (china): a comparison between a random forest data mining technique and bivariate and multivariate statistical models. Geomorphology 259:105–118

Hosmer DW, Lemeshow S (1989) Applied regression analysis. Wiley, New York, p 307

Kavzoglu T, Sahin EK, Colkesen I (2014) Landslide susceptibility mapping using GIS-based multi-criteria decision analysis, support vector machines, and logistic regression. Landslides 11(3):425–439

Kayastha P, Dhital MR, De Smedt F (2012) Landslide susceptibility mapping using the weight of evidence method in the Tinau watershed, Nepal. Nat Hazards 63(2):479–498

Keefer DK, Larsen MC (2007) Assessing landslide hazards. Science(Washington) 316(5828):1136–1138

Kincal C, Singleton A, Li Z, Drummond J, Hoey T, Muller J, Qu W, Zeng Q, Zhang J, Du P (2010) Mass movement susceptibility mapping using satellite optical imagery compared with insar monitoring: Zigui county, three gorges region, China. Dragon-2 Symposium

Lee S (2004) Application of likelihood ratio and logistic regression models to landslide susceptibility mapping using GIS. Environ Manag 34(2):223–232

Lu D, Cui J, Li X, Lian W (2010) Ground motion attenuation of M S8. 0 Wenchuan earthquake. Earthq Sci 23(1):95–100

Lu P, Catani F, Tofani V, Casagli N (2014) Quantitative hazard and risk assessment for slow-moving landslides from persistent scatterer interferometry. Landslides 11(4):685–696

Marjanović M, Kovačević M, Bajat B, Voženílek V (2011) Landslide susceptibility assessment using SVM machine learning algorithm. Eng Geol 123(3):225–234

Meisina C, Zucca F, Notti D, Colombo A, Cucchi A, Savio G, Giannico C, Bianchi M (2008) Geological interpretation of PSInSAR data at regional scale. Sensors 8(11):7469–7492

Mohammady M, Pourghasemi HR, Pradhan B (2012) Landslide susceptibility mapping at Golestan Province, Iran: a comparison between frequency ratio, Dempster–Shafer, and weights-of-evidence models. J Asian Earth Sci 61(15):221–236

Nandi A, Shakoor A (2010) A GIS-based landslide susceptibility evaluation using bivariate and multivariate statistical analyses. Eng Geol 110(1):11–20

Oh HJ, Lee S (2010) Assessment of ground subsidence using GIS and the weights-of-evidence model. Eng Geol 115(1):36–48

Ozdemir A, Altural T (2013) A comparative study of frequency ratio, weights of evidence and logistic regression methods for landslide susceptibility mapping: Sultan Mountains, SW Turkey. J Asian Earth Sci 64(5):180–197

Paulín GL, Pouget S, Bursik M, Quesada FA, Contreras T (2016) Comparing landslide susceptibility models in the río el estado watershed on the sw flank of pico de orizaba volcano, mexico. Nat Hazards 80(1):127–139

Peng L, Niu R, Huang B, Wu X, Zhao Y, Ye R (2014) Landslide susceptibility mapping based on rough set theory and support vector machines: a case of the Three Gorges area, China. Geomorphology 204(1):287–301

Pourghasemi HR, Jirandeh AG, Pradhan B, Xu C, Gokceoglu C (2013) Landslide susceptibility mapping using support vector machine and GIS at the Golestan Province, Iran. J Earth Syst Sci 122(2):349–369

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(2):350–365

Regmi NR, Giardino JR, Vitek JD (2010) Modeling susceptibility to landslides using the weight of evidence approach: western Colorado, UA. Geomorphology 115(1):172–187

Santacana N, Baeza B, Corominas J, De Paz A, Marturiá J (2003) A GIS-based multivariate statistical analysis for shallow landslide susceptibility mapping in La Pobla de Lillet Area (Eastern Pyrenees, Spain). Nat Hazards 30(3):281–295

Schicker R, Moon V (2012) Comparison of bivariate and multivariate statistical approaches in landslide susceptibility mapping at a regional scale. Geomorphology 161–162(7):40–57

Singh LP, Van Westen C, Ray PC, Pasquali P (2005) Accuracy assessment of insar derived input maps for landslide susceptibility analysis: a case study from the swiss alps. Landslides 2(3):221–228

Soeters R, van Westen CJ (1996) Slope instability recognition, analysis and zonation. In: Turner AK, Schuster RL (eds) Landslides, investigation and mitigation. Transportation Research Board, National Research Council, Special Report 247. National Academy Press, Washington, pp 129–177

Süzen ML, Doyuran V (2004) A comparison of the GIS based landslide susceptibility assessment methods: multivariate versus bivariate. Environ Geol 45(5):665–679

Tax DM, Duin RP (2002) Uniform object generation for optimizing one-class classifiers. J Mach Learn Res 2(2):155–173

Thiery Y, Malet J-P, Sterlacchini S, Puissant A, Maquaire O (2007) Landslide susceptibility assessment by bivariate methods at large scales: application to a complex mountainous environment. Geomorphology 92(1):38–59

Tizzani P, Berardino P, Casu F, Euillades P, Manzo M, Ricciardi GP, Zeni G, Lanari R (2007) Surface deformation of long valley

caldera and mono basin, california, investigated with the SBAS-InSAR approach. Remote Sens Environ 108(3):277–289

Van Westen C, Rengers N, Soeters R (2003) Use of geomorphological information in indirect landslide susceptibility assessment. Nat Hazards 30(3):399–419

Van Westen C, Van Asch TW, Soeters R (2006) Landslide hazard and risk zonation—why is it still so difficult? Bull Eng Geol Env 65(2):167–184

Vapink VN (1995) The nature of statistical learning theory. Springer, New York

Varnes D (1984) International Association of Engineering Geology Commission on Landslides and Other Mass Movements on Slopes. Landslide Hazard Zonation: A Review of Principles and Practice: Int Assoc Eng Geol, UNESCO Natural Hazards Series, No. 3. P 63

Xu C, Dai F, Xu X, Lee YH (2012) GIS-based support vector machine modeling of earthquake-triggered landslide susceptibility in the Jianjiang River watershed, China. Geomorphology 145–146(2):70–80

Xu C, Xu X, Yao Q, Wang Y (2013) GIS-based bivariate statistical modelling for earthquake-triggered landslides susceptibility mapping related to the 2008 Wenchuan earthquake, China. Q J Eng Geol Hydrogeol 46(2):221–236

Yalcin A (2008) GIS-based landslide susceptibility mapping using analytical hierarchy process and bivariate statistics in Ardesen (Turkey): comparisons of results and confirmations. CATENA 72(1):1–12

Yalcin A, Reis S, Aydinoglu A, Yomralioglu T (2011) A GIS-based comparative study of frequency ratio, analytical hierarchy process, bivariate statistics and logistics regression methods for landslide susceptibility mapping in Trabzon, NE Turkey. CATENA 85(3):274–287

Yao X, Tham L, Dai F (2008) Landslide susceptibility mapping based on support vector machine a case study on natural slopes of Hong Kong, China. Geomorphology 101(4):572–582

Yilmaz I (2010) Comparison of landslide susceptibility mapping methodologies for Koyulhisar, Turkey: conditional probability, logistic regression, artificial neural networks, and support vector machine. Environ Earth Sci 61(4):821–836

Youssef AM, Pourghasemi HR, Pourtaghi ZS, Al-Katheeri MM (2016) Landslide susceptibility mapping using random forest, boosted regression tree, classification and regression tree, and general linear models and comparison of their performance at wadi tayyah basin, asir region, saudi arabia. Landslides 13(5):839–856

Yu B, Yang Y, Su Y, Huang W, Wang G (2010) Research on the giant debris flow hazards in Zhouqu County, Gansu Province on August 7, 2010. J Eng Geol 18(4):437–444 (in Chinese)

Zhang Y, Meng X, Chen G, Qiao L, Zeng R, Chang J (2016) Detection of geohazards in the bailong river basin using synthetic aperture radar interferometry. Landslides 13(5):1273–1284