

Landslide susceptibility mapping based on global and local logistic regression models in Three Gorges Reservoir area, China

Miao Zhang¹ · Xuelian Cao¹ · Ling Peng² · Ruiqing Niu¹

Received: 27 January 2016 / Accepted: 19 May 2016 / Published online: 2 June 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract This paper investigates the spatial stationarity of the relationship between landslide susceptibility and associated factors in Three Gorges Reservoir area, a landslide-rich area in China. Two logistic regression (LR) models have been used: A global LR (LR) assumes that the regression coefficients remain constant over the whole region, whereas a geographically weighted LR (GWLR) allows the regression coefficients to differ at the local scale. In LR model, lithology seems to have positive influence on the location of landslides, as it has a positive regression coefficient (0.005), while the other factors all have negative effects on landslide susceptibility as they all have negative coefficients. However in GWLR model, lithology does not always keep positive influence, as its coefficients range from -0.533 to 0.695 . These results indicate a degree of spatial variation in the relationship between landslide susceptibility and the influencing factors in the study area. Furthermore, six evaluation criteria, based on the fit and complexity of the models, were used to compare the two approaches: deviance, corrected Akaike's information criterion (AICc), local percent deviance explained (pdev), receiver operating characteristic curve (ROC), Bayesian information criterion (BIC), and residual

Moran's I. The results suggest that GWLR model provides potential advantages in landslide susceptibility mapping and sheds new light on the spatial non-stationarity of the relationship between landslide susceptibility and its influencing factors.

Keywords Spatial non-stationarity · Landslide susceptibility mapping (LSM) · Logistic regression (LR) · Geographically weighted logistic regression (GWLR) · Three Gorges Reservoir area

Introduction

Landslides are characterized the most natural hazards after earthquakes, which continue to cause human and financial loss. In Sep 2014, Lifengyuan Hydropower Station (originally called Dalingshan) in China was completely destroyed by landslides. It is the first time hydropower station has been destroyed by landslides in Three Gorges Reservoir area. Landslide susceptibility mapping (LSM) is a solution to understanding and predicting hazards to mitigate their consequences (Feizizadeh and Blaschke 2011).

For decades, a variety of methods for LSM have been proposed, most of which are GIS-based and are related to qualitative, quantitative, or hybrid approaches (Fig. 1). Some of methodology overviews list as follows: Varnes 1984; Van Westen 1994; Leroi 1996; Aleotti and Chowdhury 1999; Guzzetti et al. 1999; Van Westen et al. 2003; Brenning 2005; Glade and Crozier 2005; and Budimir et al. 2015.

Landslides are complex processes, mainly because of the role of many factors, including geology, geomorphology, hydrology, and the role of people. If we try to quantify the relationships between influencing factors and landslide

This article is part of a Topical Collection in Environmental Earth Sciences on "Environmental Research of the Three Gorges Reservoir", guest edited by Binghui Zheng, Shengrui Wang, Yanwen Qin, Stefan Norra, and Xiafu Liu.

✉ Xuelian Cao
hustcxl@163.com

¹ Institute of Geophysics and Geomatics, China University of Geosciences, 430074 Wuhan, People's Republic of China

² China Institute of Geo-Environment Monitoring, 100081 Beijing, People's Republic of China

occurrence, a question may be asked: whether the influence of factors on landslide occurrence is homogeneous everywhere? That is to say, are relationships between landslide location and influencing factors stationary? It is necessary to note that the literature focusing on the spatial non-stationarity in landslide susceptibility assessment is very limited (Zhou et al. 2002; Erener 2009; Erener and Düzgün 2010; Atkinson and Massari 2011; Chalkias et al. 2011, 2014; Feuillet et al. 2014). Unfortunately, most studies have considered the relationships between predictors and landslide occurrence as fixed effects (Feuillet et al. 2014).

The methods which ignore spatial dependence or autocorrelation characteristics of data in susceptibility assessment, like logistic regression (LR), can be called as global models (Feuillet et al. 2014). Relatively, the methods which consider spatial variability in the effect of influencing parameters at the local scale, like geographically weighted

logistic regression (GWLR), can be called as local models. The global model (LR) makes the regression coefficients of landslide predisposing factors remain constant throughout the whole study area, whereas the local model (GWLR) allows the regression coefficients different everywhere. The concept of geographical weighting was proposed in 1996 and initiated using geographically weighted regression (GWR) to capture spatial non-stationarity (Brunsdon et al. 1996; 1998; Fotheringham et al. 2002). GWLR is then developed to explore the relations between riverbank erosion and geomorphological controls (Atkinson et al. 2003). Erener (2009) first used GWLR in LSM in her PhD thesis, in the case of Bartın Kumluca watershed in Turkey; then, she compared LR, spatial regression (SR), and GWLR in the case of More and Romsdal, Norway (Erener and Düzgün 2010); Chalkias et al. showed the differences of landslide susceptibility estimations between GWLR and LR in South Greece (Chalkias et al. 2011; 2014). Feuillet et al. (2014) focused on investigating the spatial non-stationarity in the relationships between paraglacial variables and landslide locations in northern Iceland.

In this paper, LR and GWLR models are implemented for LSM in Qinggan River basin, a small part of the Three Gorges Reservoir area in China. And six evaluation criteria are used to compare the two approaches.

Study region

Three Gorges lies in the mountains separating Sichuan Basin and Jiangnan Plain, and along the middle reaches of the Yangtze River. The study area, located ~50 km west of the Three Gorges Dam, covers a surface area of 46.8 km² and lies between latitudes from 30.56'13" N to 31.0'37" N and longitudes from 110.33'46" E to

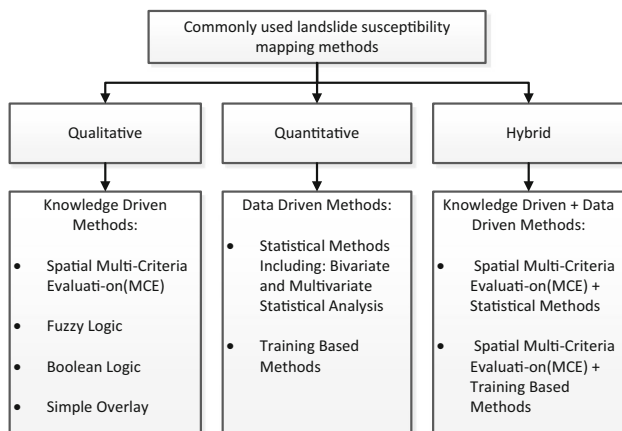


Fig. 1 A schematic illustration of commonly used LSM methods (by Sabokbar et al. 2014)

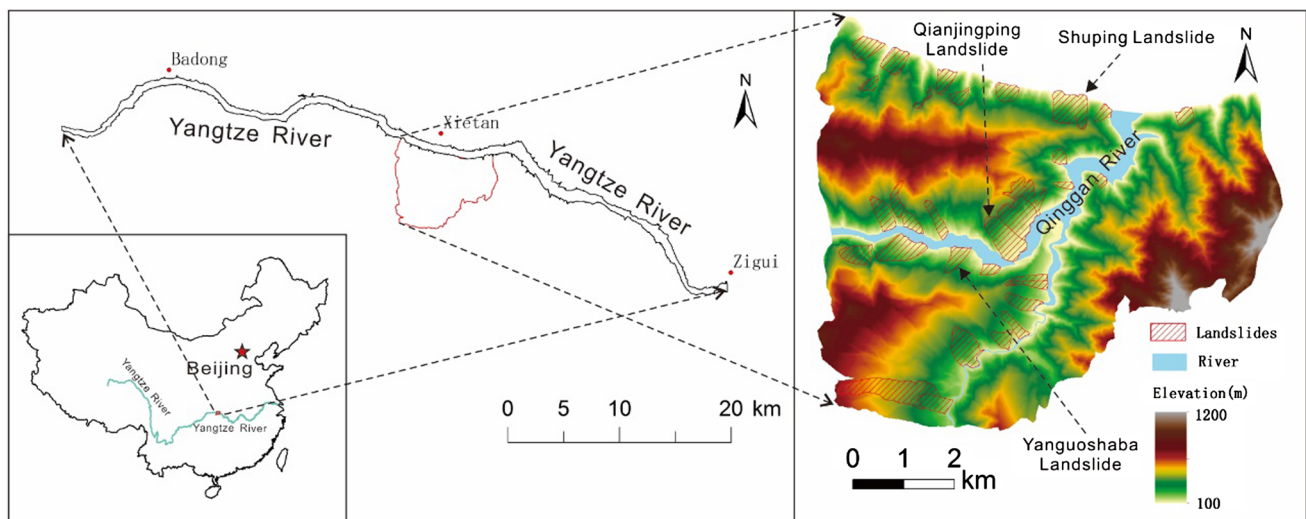


Fig. 2 Study area in Three Gorges Reservoir area with landslide events

110.39°/36"E (Fig. 2). The minimum elevation in the area is 140 m, and the maximum elevation is 1210 m. The terrain consists of a series of limestone ridges and gorges, with intergorge valleys underlain primarily by interbedded mudstones, shale, and thinly bedded limestone. Landslides tend to occur in areas underlain by failure-prone rock units are exposed in the intergorge valleys (Fourniadis et al. 2007a, 2007b). Although the slope gradients range from 0° to 78°, the majority are between 20° and 30°. Steep slopes have developed in areas underlain by easily erodible, soft materials, and landslides are common in these areas. The average annual precipitation is 1100 mm. The rainfall is generally concentrated in the spring and summer, and the summer average can be as high as 200–300 mm per month (Peng et al. 2014, 2015).

The study area is located in the southern tip of Zigui syncline axis. The strata exposed along the northwest to the southeast are (from the old to the new): Triassic Badong Formation (T_{2b}), Triassic Jiuligang Formation (T_{2j}), Jurassic Tonglinyuan Formation (J_{1t}), Jurassic Qianfoyan Formation (J_{2q}), Jurassic lower Shaximiao Formation (J_{2x}), and Jurassic upper Shaximiao Formation (J_{2s}).

A landslide inventory map surveyed by the Three Gorges Headquarters was used to obtain landslide locations (Fig. 2). The mapped landslides cover a total area of 5.46 km², representing 11.1 % of the study region. Landslides mostly distribute along the waters. The global Moran's I is 0.617 with *P* value 0.001, which reveals obvious spatial agglomeration. Examples of large and the most destructive landslides include Shuping landslide, Qianjiangping landslide and Yanguosha landslide (Fig. 2).

Data

Influencing factors of landslides

Nine variables were selected for the LSM study area: elevation, slope, modified normalized difference water index (MNDWI, Pelletier et al. 1997), normalized difference vegetation index (NDVI), distance-to-stream (dis_stream), distance-to-fault (dis_fault), distance-to-road (dis_road), lithology, and bedding structure. These variables were based on previous studies of landslide susceptibility in Three Gorges Reservoir area (Liu et al. 2004, 2009; Fourniadis et al. 2007a, 2007b; Bai et al. 2010; Peng et al. 2014; Bi et al. 2014; Talaei 2014).

Elevation seems to have no direct relation with landslide. However, water development and human activities are closely associated with elevation. Then as elevation reduces, the probability of the surface material disturbance increases. Through the histogram statistics and analysis,

there is no landslide when elevation is more than 700 m. Therefore, negative correlation may exist between elevation and landslides.

Slope is thought to be important factor affecting the stability of landslides. In the study area, most landslides distribute between 20° to 30°, and there is no landslide distribution after more than 50°.

It is essential to group lithology properly. We divided lithology into three groups: mudstone, shale and Quaternary deposits; sandstones and thinly bedded limestones; and limestones and massive sandstones.

The bedding structure is a continuous raster layer representing the angular relationship between topography and strata attitude. The relationship can be characterized by the product of the bed dip direction and angle, slope angle, and aspect (Meentemeyer and Moody 2000; Peng et al. 2014). The classification for bedding structure is shown in Fig. 3. Then, these data were used to generate a bedding structure map (Fig. 4).

Elevation, slope, dis_stream, dis_fault, and dis_road were derived from 1:10,000-scale digital topographic maps and 1:10,000-scale digital geological maps surveyed by the Three Gorges Headquarters. MNDWI and NDVI were calculated from Landsat imagery. Lithology was obtained from 1:50,000 geo-map. Bedding structure was obtained from 1:10,000 topo-map and 1:50,000 geo-map.

By the analysis of multicollinearity, dis_stream showed a high correlation with elevation, and MNDWI showed high correlation with NDVI. As a result, dis_stream and NDVI were excluded from the further calculations. In the study area, the main triggering factor for landsliding is the high amount of precipitation (mainly consists of rainfall). However, the regression analysis does not include precipitation because rainfall is relatively uniform throughout the whole study area.

Slope unit

In this study, slope unit was chosen as the LSM unit, a partition of the landscape based on the surface hydrologic analysis. Partition of a region into subbasins or slope units is obtained from high-quality DEM'S and hydrological regions between drainage and divide lines (Carrara 1988; Carrara et al. 1991). Depending on the type of instability to be investigated (deep-seated vs. shallow slides or complex slides vs. debris flows) the mapping unit may correspond either to the subbasin or to the main slope unit (right/left side of the subbasin) (Guzzetti et al. 1999). In nature, there exists a clear physical relationship between landsliding and the fundamental morphological elements of a hilly or mountainous region, namely drainage and divide lines. Therefore, the slope unit-based mapping unit has more representative power for the landslide phenomena (Erener 2009).

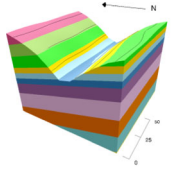
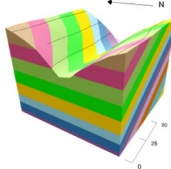
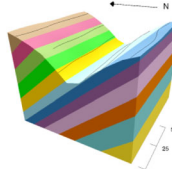
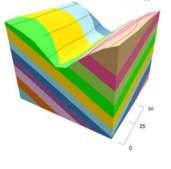
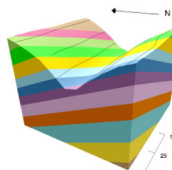
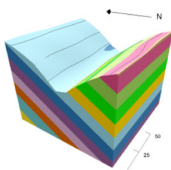
3D sketch	Definition	Type	3D sketch	Definition	Type
	$ \alpha - \beta \in [0^\circ, 30^\circ)$ or $ \alpha - \beta \in [330^\circ, 360^\circ)$, $\gamma > 10^\circ$ and $\delta > \gamma$	over-dip slope		$ \alpha - \beta \in [60^\circ, 120^\circ)$ or $ \alpha - \beta \in [240^\circ, 300^\circ)$	transverse slope
	$ \alpha - \beta \in [0^\circ, 30^\circ)$ or $ \alpha - \beta \in [330^\circ, 360^\circ)$, $\gamma > 10^\circ$ and $\delta < \gamma$	under-dip slope		$ \alpha - \beta \in [120^\circ, 150^\circ)$ or $ \alpha - \beta \in [210^\circ, 240^\circ)$	anaclinal-oblique slope
	$ \alpha - \beta \in [30^\circ, 60^\circ)$ or $ \alpha - \beta \in [300^\circ, 330^\circ)$	dip-oblique slope		$ \alpha - \beta \in [150^\circ, 210^\circ)$	anaclinal slope

Fig. 3 Classification of the bedding structure. α : slope aspect; β : bed dip direction; γ : bed dip angle; and δ : slope angle (by Peng et al. 2014)

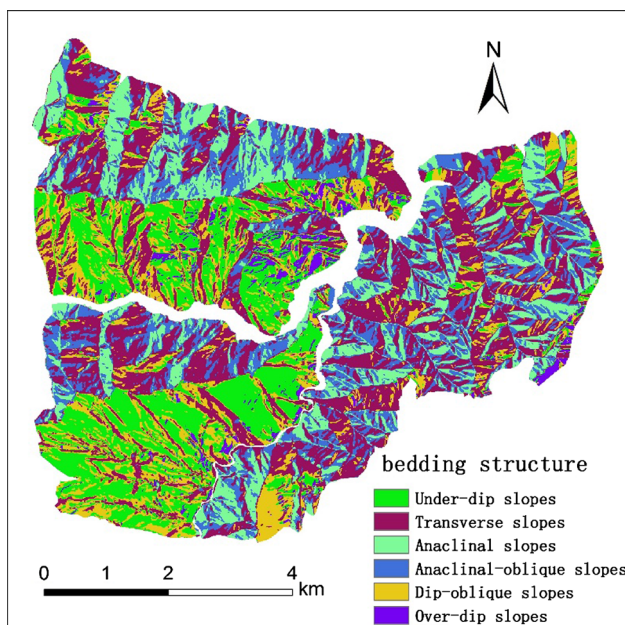


Fig. 4 Bedding structure map for the whole region

Methods

Two logistic regression models were developed for quantitative mapping of landslide susceptibility: LR was used to establish the relationship between landslide influencing factors and landslide occurrence at a global scale, while GWLR was used to investigate the relationship at a local scale. Figure 5 is a schematic representation of the methodology.

Global logistic model (LR)

LR is a multivariate analysis model for predicting the presence (or the absence) of a phenomenon, based on the values of predictor variables (Lee 2005). It has at least two advantages over traditional multivariate linear regression (Atkinson et al. 2003). First, LR allows the independent variables that are categorical or continuous, or any combination of both types. Second, residuals do not need to be normally distributed about their mean and no assumptions are made about their error distributions. Thus, most of the limitations of traditional regression are removed (Atkinson et al. 2003). Using a LR model, the relationship between landslide occurrence (Y) and landslide influencing factors (X_1, X_2, \dots, X_n) is established as:

$$Y = \log \text{it} (p_i) = \ln\left(\frac{p_i}{1 - p_i}\right) = \beta_0 + \sum_{k=1}^p \beta_k * X_k \quad (1)$$

where p_i is the probability of Y occurring at location i, $p_i / (1 - p_i)$ is the “odds ratio” or likelihood ratio, β_0 is the intercept, and $\beta_1, \beta_2, \dots, \beta_p$ are the regression coefficients. If a coefficient is positive, its transformed log value will be greater than one, meaning the event is more likely to occur. If a coefficient is negative, its transformed log value will be less than one and the odds of the event occurring decrease (Ayalew and Yamagishi 2005).

In the application of LR model for LSM, some scientists have created layers of binary values for each class of an influencing parameter (Guzzetti et al. 1999; Lee and Min

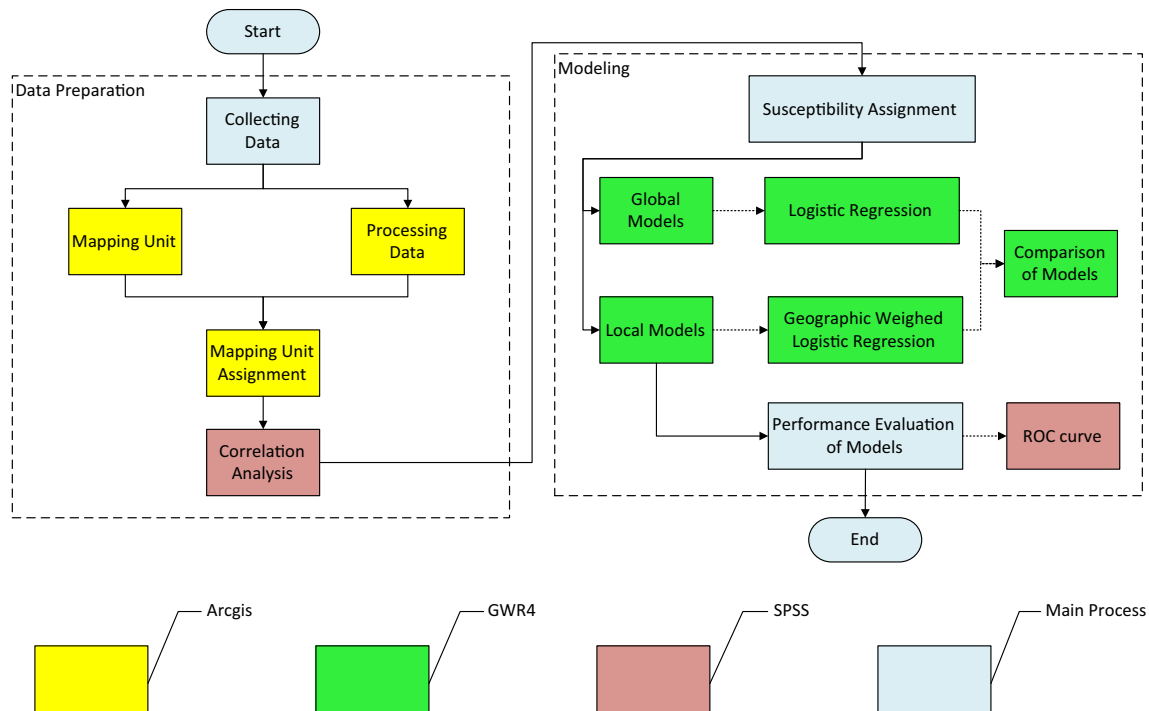


Fig. 5 Schematic representation of the methodology

2001; Dai et al. 2001; Dai and Lee 2002; Ohlmacher and Davis 2003). This might lead to a great number of independent variables. If many variables are included, the regression equation will be very long and it may even introduce numerical problems. One solution is to arrange classes of all parameters according to their corresponding landslide densities (Ayalew and Yamagishi 2005). Another alternative method is to transform the categorical variables to numeric variables by using landslide densities (Yesilnacar and Topal 2005; Zhu and Huang 2006; Bai et al. 2010).

In this study, we implemented the two logistics models by using the continuous data standardized to range from 0 to 1 and using landslide density to transform the categorical variables to numeric variables. Landslide density is calculated as (Carrara 1992):

$$\text{Landslide density} = \frac{(B_i/A_i)}{\sum_i (B_i/A_i)} \quad (2)$$

where A_i is the area of the i th type of the influencing variable and B_i is the landslide area in i th type of the influencing variable.

Local logistic model (GWLRL)

GWLRL is a geographically weighted version of the above traditional LR model. It is first developed to explore the relations between riverbank erosion and geomorphological controls (Atkinson et al. 2003). Then, it has been applied in

the spatial simulation of regional land use patterns (Liao et al. 2010), the exploration of spatial non-stationarity of fisheries survey data (Windle et al. 2010), other studies in geosciences (Erener and Düzgün 2010; Chalkias et al. 2011; 2014; Cossart 2013; Wu and Zhang 2013; Feuillet et al. 2014; Zini et al. 2015), and in human geography (Windle et al. 2010; Saefuddin et al. 2012; Yang and Matthews 2012).

A key step in the development of GWLRL is the choice of a spatial weighting function for estimating local parameters. Specifically, GWLRL uses a distance-based weighting scheme since it is assumed that observations near point i have more influence on the estimation of the parameters than observations located farther from i (Feuillet et al. 2014). Then, the weighted least square estimates of β_i are:

$$\hat{\beta}_i = (x^T w_i x)^{-1} x^T w_i y \quad (3)$$

where w_i is n by n weighting matrix, whose off-diagonal elements are zero and diagonal elements are the geographical weighting:

$$w_i = \begin{pmatrix} w_{i1} & & & \\ & w_{i2} & & \\ & & \ddots & \\ & & & w_{in} \end{pmatrix} \quad (4)$$

The choice of w_i depends on the selection of kernel function, which may be in the form of fixed (i.e., fixed

Table 1 Summary statistics for LR and GWLR regression coefficients

Model	Evaluating indicator	Intercept	Elevation	MNDWI	Slope	dis_fault	dis_road	Bedding structure	Lithology
LR	Estimate	-1.903	-1.523	-0.593	-0.401	-0.217	-0.253	-0.101	0.005
	Stand error	0.085	0.116	0.073	0.069	0.073	0.083	0.062	0.070
	Z(Est/SE)	-22.379	-13.138	-8.149	-5.786	-2.970	-3.048	-1.620	0.075
	Exp(Est)	0.149	0.218	0.553	0.669	0.805	0.777	0.904	1.005
GWLR	Min	-4.647	-2.460	-1.304	-1.506	-1.120	-3.454	-0.305	-0.533
	Max	-1.633	-0.870	-0.182	0.135	1.451	0.065	0.199	0.695
	Lwr quartile	-2.784	-1.754	-0.811	-0.443	-0.430	-1.043	-0.199	-0.236
	Median	-2.209	-1.669	-0.646	-0.320	-0.246	-0.631	-0.145	-0.032
	Upr quartile	-1.797	-1.475	-0.515	-0.131	-0.134	-0.254	-0.029	0.123
	Stand error	0.710	0.296	0.208	0.314	0.395	0.615	0.113	0.245
	Exp(Median)	0.110	0.189	0.524	0.726	0.782	0.532	0.865	0.969

bandwidth) or adaptive kernels (i.e., varying bandwidths). We choose bi-square function (Fotheringham et al. 2002):

$$w_{ij} = \begin{cases} [1 - (d_{ij}/b)]^2 & d_{ij} \leq b \\ 0 & d_{ij} > b \end{cases} \quad (5)$$

where b is the bandwidth and d_{ij} is the spatial distance between location i and j . Points close to location i are highly weighted, and the weighting reduces as d_{ij} increases. The choice of bandwidth size is influent and is determined by minimizing the corrected Akaike's information criterion (AICc), which is based on the log likelihood of the model (Johnson and Omland 2004). In this study, the bandwidth that minimizes the AICc is 805 m in the GWLR model with 7 variables. But we choose a fixed bandwidth (1600 m) in order to compare with other GWLR models.

All the regression calculations were performed using GWR4.0 freeware, while interpolation and mapping were computed with ArcGIS 10.2, and spatial statistics was done in GeoDa.

Model comparison criteria

Six evaluation criteria, considering both fitting and complexity of the models, are used to compare LR and GWLR. BIC/MDL is appropriate for arguing the degree of complexity of the process to be analyzed. Lower value of AICc indicates a more efficient model. The model error is given by deviance, and the lower the value is, the better the fitting is. Local percent deviance explained (pdev) is another goodness-of-fit measure, also known as a type of pseudo-R square. The higher the value is, the better the model fits to the data. The receiver operating characteristic curves (ROC) are also given. The area under the ROC curve (AUC) is used to evaluate the accuracy of

model. It is normally above 0.5 (random discrimination) and not higher than 1 (perfect separation of the classes) (Swets 1988). The global Moran's I is used as a measure on the spatial distribution of the residual error. The closer the index is to 0, the more random the spatial distribution is.

Results and discussion

Model comparison results

The statistics of model parameters are summarized in Table 1. In LR model, the regression coefficients are constant over the whole region. But in GWLR model, the regression coefficients are different for each slope unit (Table 1 and Fig. 6).

From Table 1, bedding structure and lithology seem to be not significant, owing to the small Z scores. But in the study area, they are very important factors. Then, should we remove these two factors? For the sake of answer, two experiments, one with all the factors except lithology, the other with all the factors except bedding structure and lithology, have been compared to see whether the factors could be removed. The comparison results are given in Table 2, and the ROC are shown in Fig. 7.

Table 2 shows that, as factor removed, the model complexity is decreased slightly (BIC/MDL decreased gradually), both for LR and GWLR model. However, there is almost no change in LR results except BIC/MDL, whether the goodness of fit, or the spatial autocorrelation of residuals. But GWLR model behaves different. When lithology factor being removed, the goodness of fit of GWLR significantly reduced, with less randomness of residual distribution. Details about the six evaluation criteria see the last section.

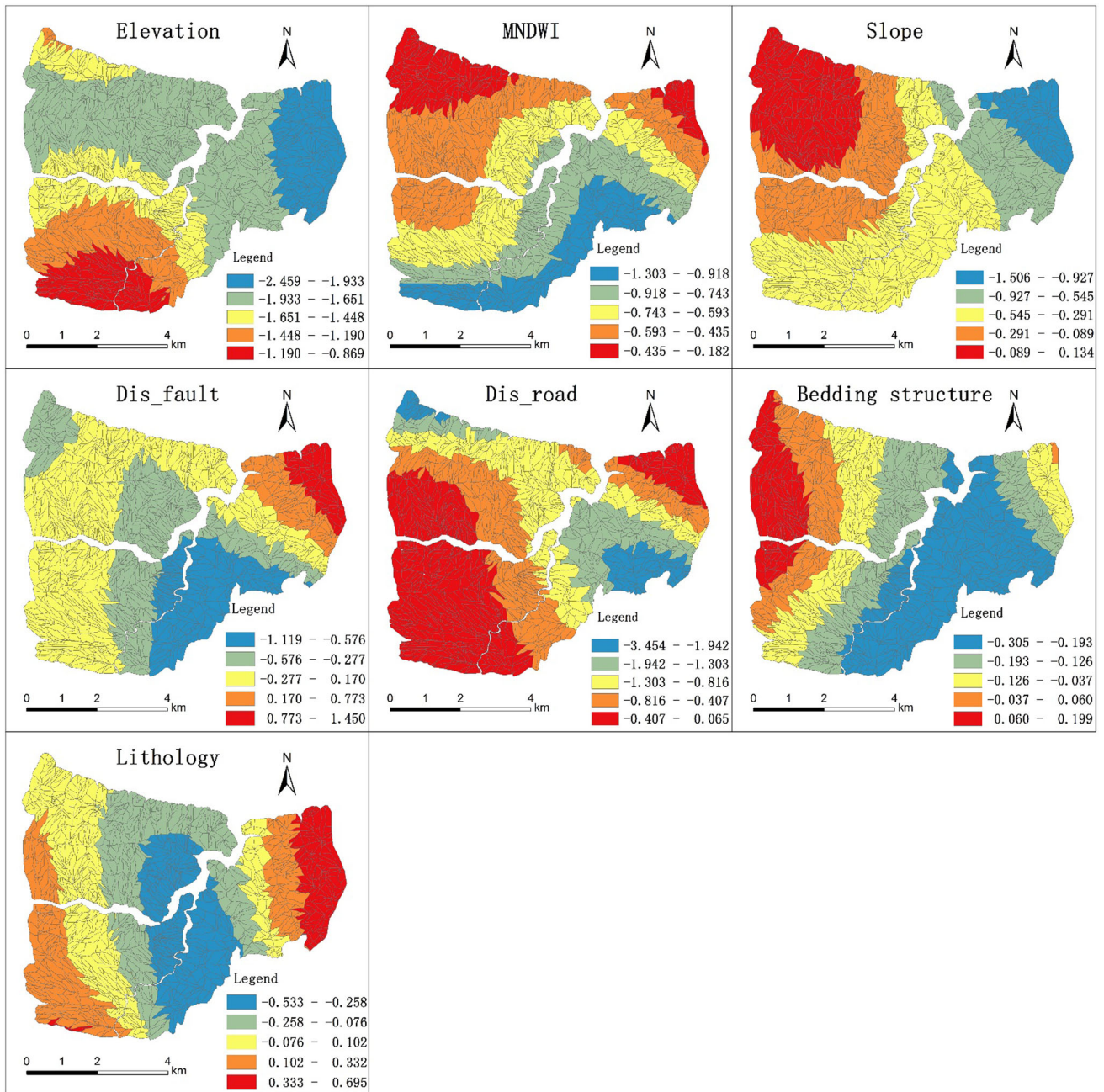


Fig. 6 Spatial variation of coefficients values from GWLR calculation

Table 2 Comparison diagnostics between LR and GWLR models with different number of factors

Factors number	Model	BIC/MDL	AICc	Deviance	pdev	AUC	Residual Moran's I
7	LR	1768.435	1723.434	1707.364	0.203	0.806	0.481
	GWLR	1742.372	1592.510	1538.325	0.282	0.848	0.436
6	LR	1760.807	1721.425	1707.370	0.203	0.806	0.481
	GWLR	1726.842	1652.472	1625.822	0.241	0.826	0.464
5	LR	1755.815	1722.052	1710.012	0.202	0.806	0.482
	GWLR	1716.088	1652.134	1629.245	0.239	0.825	0.466

The comparative results told us two things. On the one hand, it is likely to weed out some important factors by significance tests; on the other hand, GWLR model could better reflect the importance of lithology factor and protect important factors being removed. Therefore, we decided not to remove the last two factors.

Next, let us compare LR and GWLR models. Again, from Table 1, we can easily find that LR lithology appears to have a little positive influence on landslide occurrence,

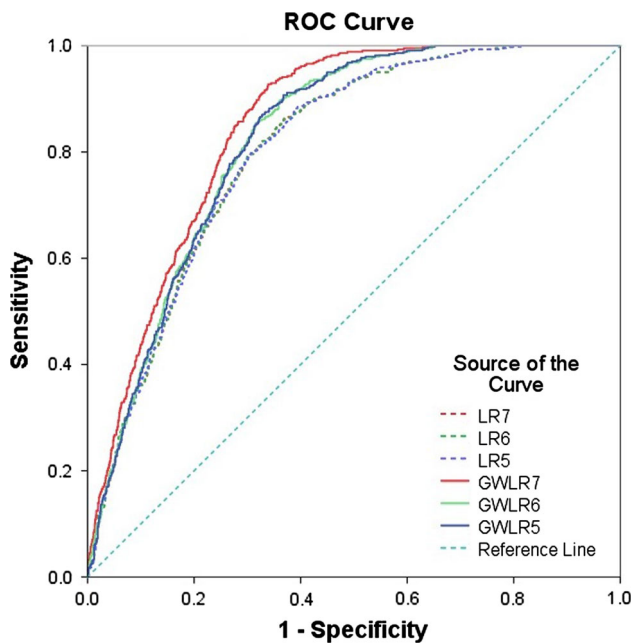


Fig. 7 ROC of LR and GWLR models with different number of factors

as it has a small positive regression coefficient (0.005), and the other factors have negative effects in landslide formation as they all have negative coefficients. However, the situation changes in GWLR model; lithology does not always keep positive influence at the local scale, as its coefficients range from -0.533 to 0.695 , negative in the northern and middle parts of the study region and positive in the eastern and western other parts (Fig. 6). And the other factors do not always have negative effects in landslide formation, except for elevation and MNDWI. The hypothesis that elevation is negatively related to landslides has been confirmed by both LR and GWLR. However, GWLR results indicate a degree of spatial variation in the relationship between landslide susceptibility and the influencing factors in the study area, which seems more reasonable than LR. For example, through the slope histogram statistics before the regression analysis, we know that slope is impossible always negatively related to landslides, but for LR slope has a negative effect (-0.401) over the whole study area.

Nevertheless, GWLR has well maintained the general trend with LR, for the comparability between the median of the regression coefficients in GWLR and the corresponding coefficients in LR (Table 1). And for both the two models, elevation shows the strongest correlation with landslide, as it has a maximum absolute value, next are MNDWI and slope. On the contrary, bedding structure and lithology have little influence on landslide occurrence (Table 1).

Furthermore, it is easy to find from Table 2 that GWLR model has less complexity than LR, as it has smaller BIC/MDL values. GWLR also shows a better fit of the statistics,

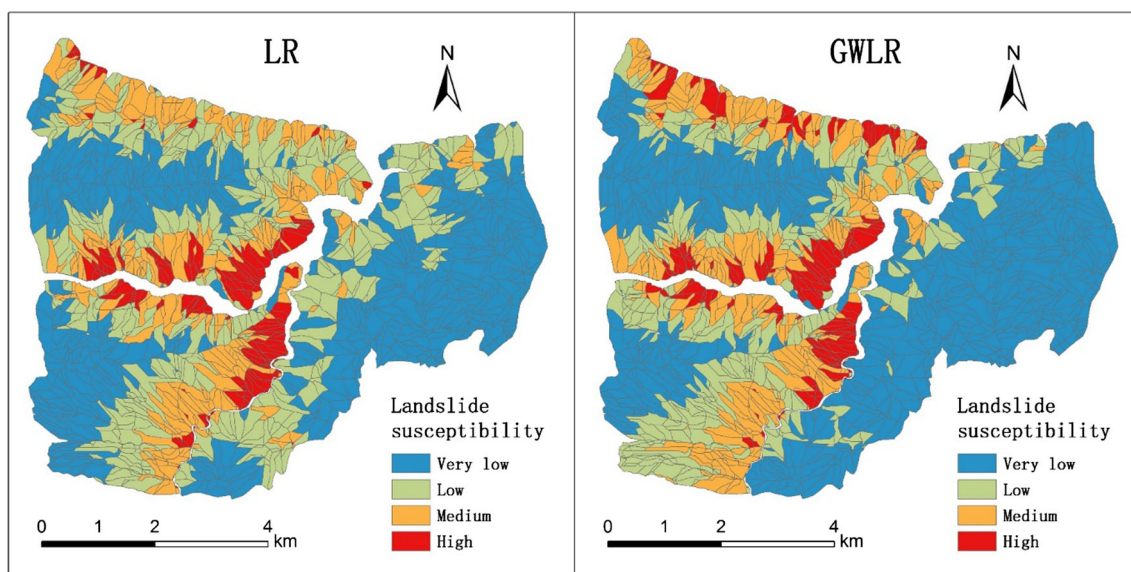


Fig. 8 Landslide susceptibility maps produced by LR and GWLR with 7 variables

Table 3 Area statistics for each susceptibility zones

Susceptibility	LR		GWLR	
	% of study area	% of inventoried landslides	% of study area	% of inventoried landslides
Very low	44.65	5.29	57.36	2.20
Low	31.42	24.18	20.69	22.41
Medium	17.43	38.97	15.29	40.65
High	6.49	31.56	6.66	34.75

with lower values of AICc, deviance, and higher values of pdev and AUC. Moreover, GWLR gives a more randomly distributed spatial pattern, as it has lower residual Moran’s I.

Landslide susceptibility mapping

Landslide susceptibility maps were created after obtaining the two regression models (Fig. 8). The probability of landslide occurrence, i.e., a proxy for the landslide susceptibility index values has been obtained from the regression calculation. For LR, the probability ranges from 0.0003 to 0.826, while for GWLR it ranges from 0 to 0.834. The natural breaks algorithm in ArcGIS was used to divide the probability maps into four susceptibility zones: very low, low, medium, and high. The breakpoints are: 0.114, 0.273, and 0.470 for LR, and 0.119, 0.291, and 0.477 for GWLR, respectively.

Looking at the two maps (Fig. 8), there are places where differences are subtle and there are also areas with dissimilarities. In both of the prediction models, the eastern part of the region is less susceptible to landslides, compared to the middle and southwestern parts of the regions which are highly susceptible to landslides. It should be taken into consideration that the more the distance from the effective landslide events the more the intrinsic uncertainty is to the interpolated points. Figure 8 shows that the eastern area is more uniformly depicted in GWLR technique than in LR since there are no real landslides. In contrast, the northern part is more susceptible for GWLR model, which seems more acceptable, for its adjacent to the Yangtze River.

Percentage statistics given in Table 3 show the differences between the two susceptibility maps quantitatively. It is found that 70 % of all the landslides which occurred in the study region in the past lie in the high and medium susceptibility zones in LR, comparing to 75 % that exist in the high and medium susceptibility zones in GWLR. 5 % of the past landslides lie in the very low susceptibility zone in LR, comparing to 2 % that exist in the very low susceptibility zone in GWLR. As it can be seen from the findings, the reliability of the GWLR model is relatively higher.

Conclusion

This paper implemented two logistic regression models (LR and GWLR) for LSM in the Three Gorges Reservoir area, China. The summary statistics for regression coefficients (Table 1) and the comparative results summarized (Table 2, 3) show that GWLR has less complexity and higher goodness of fit than LR. It could better reflect the importance of lithology factor and protect important factors being removed. Moreover, GWLR gives a more randomly distributed spatial pattern, as it has lower Moran’s I value of the residual error. The results reveal that GWLR provides potential advantages in LSM and sheds new light on the spatial non-stationarity of the relationship between landslide susceptibility and its influencing factors. It is worth noting that GWLR parameters are specific to the area under study, but the general methodology is “general,” since it is reproducible to any geographical context.

Acknowledgments The study is supported by the National High-tech R&D Program of China (863 Program, Grant 2012AA121303). The authors are grateful to the Three Gorges Headquarters for providing data and material, and also thank Ministry of Science and Technology of the People’s Republic of China for the above sponsored funding. Finally, the authors thank reviewers for the valuable suggestions.

References

Aleotti P, Chowdhury R (1999) Landslide hazard assessment: summary review and new perspectives. *Bull Eng Geol Environ* 58:21–44. doi:10.1007/s100640050066

Atkinson PM, Massari R (2011) Auto-logistic modelling of susceptibility to landsliding in the Central Apennines, Italy. *Geomorphology* 130(1–2):55–64

Atkinson PM, German SE, Sear DA, Clark MJ (2003) Exploring the relations between riverbank erosion and geomorphological controls using geographically weighted logistic regression. *Geographical Anal* 35(1):58–82

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

Bai SB, Wang J, Lu 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 Geomorphol* 115:23–31

Bi R, Schleier M, Rohn J, Ehret D, Xiang W (2014) Landslide susceptibility analysis based on ArcGIS and artificial neural

- network for a large catchment in Three Gorges region. *China Environ Earth Sci* 72(6):1925–1938
- Brenning A (2005) Spatial prediction models for landslide hazards: review, comparison and evaluation. *Nat Hazards Earth Syst Sci* 5(6):853–862
- Brunsdon C, Fotheringham AS, Charlton ME (1996) Geographically weighted regression: a method for exploring spatial nonstationarity. *Geographical Anal* 28(4):281–298
- Brunsdon C, Fotheringham AS, Charlton ME (1998) Geographically weighted regression. *J Royal Stat Soc: Series D (The Statistician)* 47(3):431–443
- Budimir MEA, Atkinson PM, Lewis HG (2015) A systematic review of landslide probability mapping using logistic regression. *Landslides* 2015:1–18
- Carrara A (1988) Drainage and divide networks derived from high-fidelity digital terrain models. In: CF Chung (ed), *Quantitative analysis of mineral and energy resources*. Reidel, Dordrecht, pp 581–597, NATO-ASI Series
- Carrara A (1992) Uncertainty in evaluating landslide hazard and risk. *Adv Nat Technol Hazards Res* 2:172–183
- Carrara A, Cardinali M, Detti R, Guzzetti F, Pasqui V, Reichenbach P (1991) GIS Techniques and statistical models in evaluating landslide hazard. *Earth Surf Proc Land* 16:427–445
- Chalkias C, Kalogirou S, Ferentinou M (2011) Global and local statistical modeling for landslide susceptibility assessment. A comparative analysis[C]//Proceedings of the 17th European Colloquium on Quantitative and Theoretical Geography. Harokopio University, Athens, Greece, pp 2–5
- Chalkias C, Kalogirou S, Ferentinou M (2014) Landslide susceptibility, Peloponnese Peninsula in South Greece. *J Maps* 10(2):211–222
- Cossart E (2013) Influence of local vs regional settings on glaciation pattern in the French Alps. *Geografia Fisica e Dinamica Quaternaria* 36(1):39–52
- Dai FC, Lee CF (2002) Landslide characteristics and slope instability modeling using GIS Lantau Island, Hong Kong. *Geomorphology* 42:213–238
- Dai FC, Lee CF, Li J, Xu ZW (2001) Assessment of landslide susceptibility on the natural terrain of Lantau Island. *HongKong Environ Geol* 40:381–391
- Erener A (2009) An approach for landslide risk assessment by using geographic information systems (GIS) and remote sensing (RS). Ph.D. Thesis, Middle East Technical University, Turkey
- Erener A, Düzgün HSB (2010) Improvement of statistical landslide susceptibility mapping by using spatial and global regression methods in the case of More and Romsdal (Norway). *Landslides* 7(1):55–68
- Feizizadeh B, Blaschke T (2011) Landslide risk assessment based on GIS multi-criteria evaluation: a case study in Bostan-Abad County. *Iran J Earth Sci Eng* 1:66–71
- Feuillet T, Coquin J, Mercier D, Cossart E, Decaulne A, Jónsson HP (2014) Focusing on the spatial non-stationarity of landslide predisposing factors in northern Iceland Do paraglacial factors vary over space? *Prog Phys Geogr* 38(3):354–377
- Fotheringham AS, Brunsdon C, Charlton M (2002) Geographically weighted regression: the analysis of spatially varying relationships. John Wiley & Sons, Chichester
- Fourniadis IG, Liu JG, Mason PJ (2007a) Landslide hazard assessment in the Three Gorges area, China, using ASTER imagery: Wushan-Badong. *Geomorphology* 84:126–144
- Fourniadis IG, Liu JG, Mason PJ (2007b) Regional assessment of landslide impact in the Three Gorges area, China, using ASTER data: Wushan-Zigui. *Landslides*. 4(3):267–278. doi:10.1007/s10346-007-0080-5
- Glade T, Crozier MJ (2005) A review of scale dependency in landslide hazard and risk analysis. In: Glade T, Anderson MG, Crozier MJ (eds) *Landslide risk assessment*. John Wiley, New York, pp 75–138
- 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* 31:181–216
- Johnson J, Omland KS (2004) Model selection in ecology and evolution. *Trends Ecol Evol* 19(2):101–108
- Lee S (2005) Application of logistic regression model and its validation for landslide susceptibility mapping using GIS and remote sensing data. *Int J Remote Sens* 26(7):1477–1491
- Lee S, Min K (2001) Statistical analysis of landslide susceptibility at Yongin. *Korea Environ Geol* 40(9):1095–1113
- Leroi E (1996) Landslide hazard-Risk maps at different scales: objectives, tools and developments. *Landslides* 1:35–51
- Liao Q, Li M, Chen Z, Shao Y, Yang K (2010) Spatial simulation of regional land use patterns based on GWR and CLUE-S model. In *Geoinformatics, 2010 18th International Conference on IEEE, China*, pp 1–6
- Liu JG, Mason PJ, Clerici N, Chen S, Davis A, Miao F, Deng H, Liang L (2004) Landslide hazard assessment in the Three Gorges area of the Yangtze river using ASTER imagery: Zigui-Badong. *Geomorphology* 61:171–187
- Liu CZ, Liu YH, Wen MS, Li TF, Lian JF, Qin SW (2009) Geo-Hazard Initiation and Assessment in the Three Gorges Reservoir. In: Wang FW, Li TL (eds) *Landslide disaster mitigation in Three Gorges Reservoir, China*. Springer Verlag, Berlin Heidelberg, pp 3–40
- Meentemeyer RK, Moody A (2000) Automated mapping of conformity between topographic and geological surfaces. *Comput Geosci-UK* 26(7):815–829
- Ohlmacher GC, Davis JC (2003) Using multiple logistic regression and GIS technology to predict landslide hazard in northeast Kansas. *USA Eng Geol* 69(3):331–343
- Pelletier J, Malamud B, Blodgett T, Turcotte D (1997) Scale-invariance of soil moisture variability and its implications for the frequency-size distribution of landslides. *Eng Geol* 48:255–268
- Peng L, Xu S, Hou J, Peng J (2015) Quantitative risk analysis for landslides: the case of the Three Gorges area, China. *Landslides* 12(5):943–960
- 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
- Sabokbar HF, Roodposhti MS, Tazik E (2014) Landslide susceptibility mapping using geographically-weighted principal component analysis. *Geomorphology* 226:15–24
- Saefuddin A, Setiabudi NA, Fitrianto A (2012) On comparison between logistic regression and geographically weighted logistic regression: with application to Indonesian poverty data. *World Appl Sci J* 19(2):205–210
- Swets JA (1988) Measuring the accuracy of diagnostic systems. *Science* 240(4857):1285–1293
- Talaei R (2014) Landslide susceptibility zonation mapping using logistic regression and its validation in Hashtchin Region, northwest of Iran. *J Geol Soc India* 84(1):68–86
- Van Westen CJ (1994) GIS in landslide hazard zonation: a review, with examples from the Andes of Colombia. In: Heywood I, Price M (eds) *Mountain environments & geographic information systems*, London, pp 135–166
- Van Westen CJ, Rengers N, Soeters R (2003) Use of geomorphological information in indirect land-slide susceptibility assessment. *Nat Hazards* 30:399–419
- Varnes DJ (1984) *Landslide hazard zonation: a review of principles and practice*. UNESCO, Natural Hazards, No.3
- Windle MJS, Rose GA, Devillers R, Fortin MJ (2010) Exploring spatial non-stationarity of fisheries survey data using

- geographically weighted regression (GWR): an example from the Northwest Atlantic. *ICES J Mar Sci* 67:145–154. doi:10.1093/icesjms/fsp224
- Wu W, Zhang L (2013) Comparison of spatial and non-spatial logistic regression models for modeling the occurrence of cloud cover in north-eastern Puerto Rico. *Appl Geogr* 37:52–62
- Yang TC, Matthews SA (2012) Understanding the non-stationary associations between distrust of the health care system, health conditions, and self-rated health in the elderly: a geographically weighted regression approach. *Health & Place* 18(3):576–585
- Yesilnacar E, Topal T (2005) Landslide susceptibility mapping: a comparison of logistic regression and neural networks methods in a medium scale study, Hendek region (Turkey). *Eng Geol* 79(3):251–266
- Zhou CH, Lee CF, Li J, Xu ZW (2002) On the spatial relationship between landslides and causative factors on Lantau Island, Hong Kong. *Geomorphology* 43(3):197–207
- Zhu L, Huang JF (2006) GIS-based logistic regression method for landslide susceptibility mapping in regional scale. *J Zhejiang University Sci A* 7(12):2007–2017
- Zini A, Grauso S, Verrubbi V, Falconi L, Leoni G, Puglisi C (2015) The RUSLE erosion index as a proxy indicator for debris flow susceptibility. *Landslides* 12:847–859