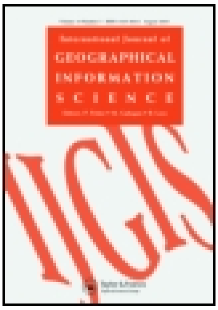

International Journal of Geographical Information Science

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tgis20

Uncertainty due to DEM error in landslide susceptibility mapping

Cheng-Zhi Qin [a] , Li-Li Bao [a b] , A-Xing Zhu [a b] , Rong-Xun Wang [b] & Xue-Mei Hu [a c]

[a] State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Anwai, Beijing, PR, China

[b] Department of Geography, University of Wisconsin-Madison, Madison, WI, USA

[c] College of Resources and Environment, University of Chinese Academy of Sciences, Beijing, PR, China

Version of record first published: 26 Mar 2013.

International Journal of Geographical Information Science, 2013
http://dx.doi.org/10.1080/13658816.2013.770515

Taylor & Francis
Taylor & Francis Group

Uncertainty due to DEM error in landslide susceptibility mapping

Cheng-Zhi Qin[a], Li-Li Bao[a,b], A-Xing Zhu[a,b]*, Rong-Xun Wang[b] and Xue-Mei Hu[a,c]

[a]*State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Anwai, Beijing, PR China;* [b]*Department of Geography, University of Wisconsin–Madison, Madison, WI, USA;* [c]*College of Resources and Environment, University of Chinese Academy of Sciences, Beijing, PR China*

(*Received 14 December 2011; final version received 16 January 2013*)

Terrain attributes such as slope gradient and slope shape, computed from a gridded digital elevation model (DEM), are important input data for landslide susceptibility mapping. Errors in DEM can cause uncertainty in terrain attributes and thus influence landslide susceptibility mapping. Monte Carlo simulations have been used in this article to compare uncertainties due to DEM error in two representative landslide susceptibility mapping approaches: a recently developed expert knowledge and fuzzy logic-based approach to landslide susceptibility mapping (*efLandslides*), and a logistic regression approach that is representative of multivariate statistical approaches to landslide susceptibility mapping. The study area is located in the middle and upper reaches of the Yangtze River, China, and includes two adjacent areas with similar environmental conditions – one for *efLandslides* model development (approximately 250 km^2) and the other for model extrapolation (approximately 4600 km^2). Sequential Gaussian simulation was used to simulate DEM error fields at 25-m resolution with different magnitudes and spatial autocorrelation levels. Nine sets of simulations were generated. Each set included 100 realizations derived from a DEM error field specified by possible combinations of three standard deviation values (1, 7.5, and 15 m) for error magnitude and three range values (0, 60, and 120 m) for spatial autocorrelation. The overall uncertainties of both *efLandslides* and the logistic regression approach attributable to each model-simulated DEM error were evaluated based on a map of standard deviations of landslide susceptibility realizations. The uncertainty assessment showed that the overall uncertainty in *efLandslides* was less sensitive to DEM error than that in the logistic regression approach and that the overall uncertainties in both *efLandslides* and the logistic regression approach for the model-extrapolation area were generally lower than in the model-development area used in this study. Boxplots were produced by associating an independent validation set of 205 observed landslides in the model-extrapolation area with the resulting landslide susceptibility realizations. These boxplots showed that for all simulations, *efLandslides* produced more reasonable results than logistic regression.

Keywords: DEM error; landslide susceptibility mapping; error propagation; Monte Carlo simulation; uncertainty

*Corresponding author. Email: axing@lreis.ac.cn; azhu@wisc.edu. Also affiliated with the Space Environment Prediction Center, Center for Space Science and Applied Research, Chinese Academy of Sciences, Beijing, PR China.

1. Introduction

As a major geological hazard landslides cause numerous human casualties and significant damage to natural ecosystems and human-built infrastructures (Chung *et al*. 1995, Dai and Lee 2002, Lee and Choi 2004, Guzzetti *et al*. 2005). To mitigate losses and damage from landslides, many landslide susceptibility studies have been carried out to map locations prone to landslides (Aleotti and Chowdhury 1999, Guzzetti *et al*. 1999, Dai *et al*. 2002, Ohlmacher and Davis 2003, Brenning 2005). Most approaches to landslide susceptibility mapping are data-driven approaches that have been developed based on multivariate statistical models (Dai *et al*. 2002). However, such models suffer from two critical deficiencies: (1) lack of reliability, stability, and portability (Carrara *et al*. 1991, van Westen *et al*. 1993); and (2) difficulty in representing complex nonlinear relationships between landslide susceptibility and predisposing factors (Zhu *et al*., under review).

Zhu *et al*. (2004), Wang (2008) and Zhu *et al*. (under review) have developed a new, expert knowledge and fuzzy logic-based approach to landslide susceptibility mapping (*efLandslides*), which combines GIS techniques and fuzzy logic. The basic idea of this approach is to find out the complex nonlinear relationships between landslide susceptibility and predisposing factors (e.g. geology, terrain) for a certain study area directly from local landslide experts. These relationships are then used to evaluate landslide susceptibility at various locations (or cells in a raster data model) across the study area. Compared with the statistical functions used in multivariate statistical approaches, the knowledge of experts formalized in fuzzy logic has been asserted to be more reliable, more portable, and more generally applicable (Zhu 1999, Zhu *et al*. 2004, Luger 2005).

Terrain attributes such as slope gradient and slope shape, which can be computed from a gridded digital elevation model (DEM), are one of the most important types of input data for landslide susceptibility mapping. These attributes have uncertainty due to elevation error, which is inevitable with DEM data (Fisher 1991, Holmes *et al*. 2000, Fisher and Tate 2006, Wechsler 2007, Wilson 2012). Therefore, DEM error will affect landslide susceptibility mapping in both the model-development and model-extrapolation stages. Little research has been done to date on comparison of the uncertainties due to DEM error in expert knowledge-based approaches and multivariate statistical approaches to landslide susceptibility mapping.

This article compares the uncertainties due to DEM error in an expert knowledge-based approach (such as *efLandslides*) and a widely used multivariate statistical approach (the logistic regression approach). Section 2 briefly introduces the *efLandslides* approach and the logistic regression approach. The uncertainty assessment method is described in Section 3. Section 4 presents the experimental design in a study area that consists of two parts, one for model development and the other for model extrapolation. Section 5 compares the uncertainty assessment results of *efLandslides* and the logistic regression approach in both areas. Conclusions are drawn in Section 6.

2. Approaches to landslide susceptibility mapping in this study

2.1. efLandslides: expert knowledge and fuzzy logic-based approach to landslide susceptibility mapping

efLandslides includes three general steps (Figure 1) (Zhu *et al*., under review): (1) knowledge extraction from local landslide experts, (2) characterization of predisposing factors, and (3) fuzzy inference (susceptibility mapping).

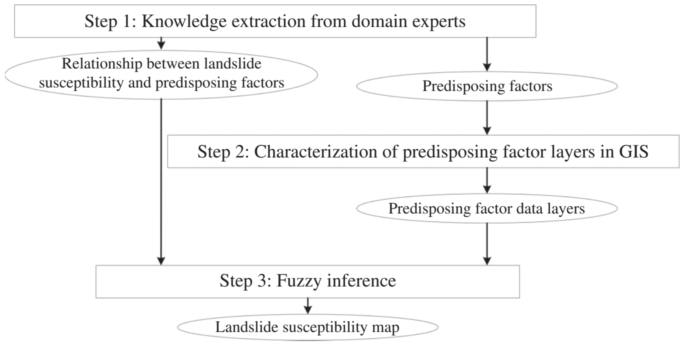

Figure 1. Framework of *efLandslides*.

2.1.1. Knowledge extraction from local landslide experts

Extracting empirical knowledge from local landslide experts is crucial and should be conducted by a proficient knowledge engineer following a well-defined knowledge acquisition process (Zhu 1999). *efLandslides* needs two types of knowledge obtained from local landslide experts: (1) knowledge of the predisposing factors affecting landslide susceptibility, and (2) knowledge of the relationships between these predisposing factors and landslide susceptibility. Thus the knowledge can be formalized in terms of a set of predisposing factors and a set of fuzzy membership functions describing the relationships between these predisposing factors and landslide susceptibility (Zhu *et al*. 2004).

2.1.2. Characterization of predisposing factors

Predisposing factors in *efLandslides* are those used in the formalized knowledge extracted from local landslide experts, which can be characterized using standard geographical information system (GIS) data-processing techniques or customized programs. Predisposing factors typically include variables derived through digital terrain analysis (e.g., slope gradient and slope aspect).

2.1.3. Fuzzy inference of landslide susceptibility

Landslide susceptibility at a given location can be inferred by evaluating individual landslide susceptibilities on the basis of predisposing factors or combinations of factors and then summarizing all these individual values at this location (Zhu *et al*. 2004). A landslide susceptibility classification map can also be derived from the landslide susceptibility map by using an expert-defined lookup table with landslide susceptibility values and levels (such as very low, low, high, and very high) (Wang 2008).

2.2. Logistic regression approach to landslide susceptibility mapping

Logistic regression belongs to the family of generalized linear models (Dobson 1990). It uses a linear combination of predictor factors (independent variables) to determine the probability of occurrence of an event (a dependent variable that has only two states). As one of the widely used multivariate statistical approaches to landslide susceptibility mapping (Atkinson and Massari 1998, Dai and Lee 2002, Ohlmacher and Davis 2003, Ayalew and Yamagishi 2005, Brenning 2005, Lee 2005, Bai *et al*. 2010), the logistic regression approach to landslide susceptibility mapping explains the relationships between landslide susceptibility (L) and a set of predisposing factors using the best fitting logistic regression model (Atkinson and Massari 1998):

$$\text{logit}\,(L) = \beta_0 + \sum_{i=1}^{n} \beta_i E_i,$$

where E_i is the ith input predisposing factor, β_i is the coefficient of E_i estimated from training data using the maximum-likelihood method, and $logit\,(\cdot)$ is the logistic link function:

$$\text{logit}\,(L) = \log\left[\frac{L}{1-L}\right].$$

Thus L can be rewritten as

$$L = \frac{1}{1+\mathrm{e}^{-Y}}, \text{where } Y = \beta_0 + \sum_{i=1}^{n} \beta_i E_i.$$

3. Method of uncertainty assessment

3.1. Basic idea

The Monte Carlo method (also known as stochastic simulation) has been widely used to analyze the propagation of uncertainty in a model when its functions are too complex for analytical approaches (Heuvelink 1998, Zhu 2005, Lindsay 2006). Based on the concept of realizations (Hammersley and Handscomb 1979), the method simulates a set of input realizations based on covariance among the inputs and the uncertainty associated with each input. For each input realization, the result for the model of interest is computed. The resulting realizations form a distribution. Parameters such as the mean and standard deviation (SD) of the distribution can be estimated from these realizations and can be used to assess both the expected results of the model and the uncertainty associated with the inputs (Heuvelink 1998, Zhu 2005).

The Monte Carlo method is a viable method for this study. The proposed application of the Monte Carlo method is shown in Figure 2. A DEM error field is first generated based on a user-defined DEM error model. Then a DEM realization is generated by adding the DEM error field to the original DEM. The realization of terrain attributes derived from the DEM realization can be used as input to each of the tested mapping approaches to produce a realization of landslide susceptibility. This process is repeated for a user-defined number of iterations to accomplish a so-called ‘simulation’ for assessing uncertainty in the mapping approach under test. This makes it possible for us to compare uncertainties due to DEM error in *efLandslides* and in logistic regression.

In the workflow of the Monte Carlo method, two key decisions must be made before application: (1) how to simulate the DEM error field, and (2) which evaluation aspects to apply to the resulting landslide susceptibility realizations.

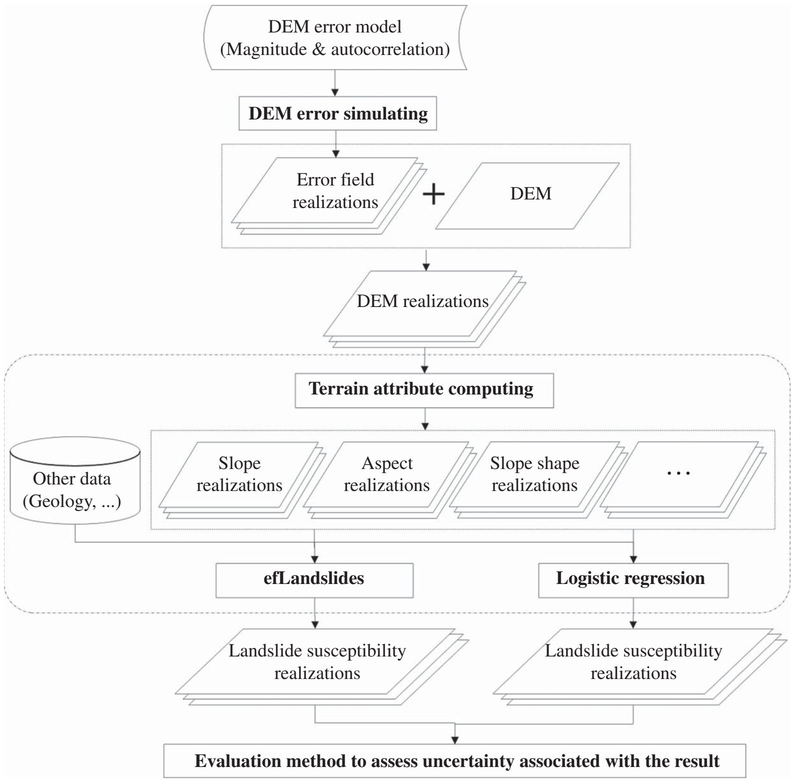

Figure 2. Workflow of Monte Carlo method for assessing uncertainties due to DEM error in *efLandslides* and logistic regression.

3.2. Sequential Gaussian simulation of the DEM error field

In this study, sequential Gaussian simulation (Goovaerts 1997) was used to simulate the DEM error field. This is a commonly used stochastic simulation method for generating DEM error fields with different magnitudes and degrees of spatial autocorrelation for assessing uncertainty due to DEM error in specific application models (e.g., Holmes *et al*. 2000, Temme *et al*. 2009). Based on the second-order stationarity assumption, each variable in the sequential Gaussian simulation is simulated sequentially according to a normal conditional cumulative distribution function, which is fully characterized by a simple kriging system. The conditioning data used in the kriging system consist of the original data and all previously simulated data within a neighborhood of the location being simulated. If the number of original data points is nonzero, the sequential Gaussian simulation is used as a conditional simulation (e.g., Holmes *et al*. 2000); otherwise, it is used as a nonconditional simulation (Goovaerts 1997). In this study, the sequential Gaussian simulation has been used as a nonconditional simulation because the objective is to assess the uncertainty in landslide mapping that occurs due to DEM error, not to characterize DEM quality. The magnitude and spatial autocorrelation of the stochastic field simulated by sequential Gaussian simulation are controlled by the SD and range, respectively, of the semi-variogram model.

3.3. Evaluation aspects

Uncertainty due to DEM error in the mapping approaches under test was quantitatively assessed based on the realizations of the landslide susceptibility map. A map of mean landslide susceptibility values was derived as the expected landslide susceptibility result under the simulated DEM error model. The SD of the susceptibility in each cell was calculated to assess the uncertainty due to DEM error in the tested mapping approaches. The larger the SD, the greater is the uncertainty. From the map of simulated SD values, the mean SD can be calculated to assess the overall uncertainty of the landslide susceptibility predictions due to the simulated DEM error model:

$$\text{Mean (SD)} = \frac{\sum_{i}\sum_{j}\text{SD}_{i,j}}{\text{count (cell)}}$$

where $\text{SD}_{i,j}$ is the SD of cell (i, j). Generally, the larger the mean SD, the greater is the overall uncertainty.

A boxplot is produced by associating an independent validation set of observed landslides in the study area with the realizations of the landslide susceptibility classification map from each simulation. This boxplot can be used to assess the uncertainty of the landslide occurrence density within each landslide susceptibility level derived by the mapping approaches under test.

4. Experimental design

4.1. Study area

The middle and upper reaches of the Yangtze River, China, where this study was conducted, are high landslide risk areas (Wu *et al*. 2001, Liu *et al*. 2004, Bai *et al*. 2010). The study area was divided into two parts with similar environmental conditions, separated by approximately 10 km. One part (the Kaixian County area) was used to develop the models for the mapping approaches, and the other (the Three Gorges area) was used for model extrapolation (Figure 3).

4.1.1. Model-development area in Kaixian County

The model-development area is located in Kaixian County of Chongqing Municipality. This area is approximately 250 km^2 in size and has generally high relief (Figure 4a). The average slope gradient is approximately 18° (Table 1). A DEM with a 25-m grid size was

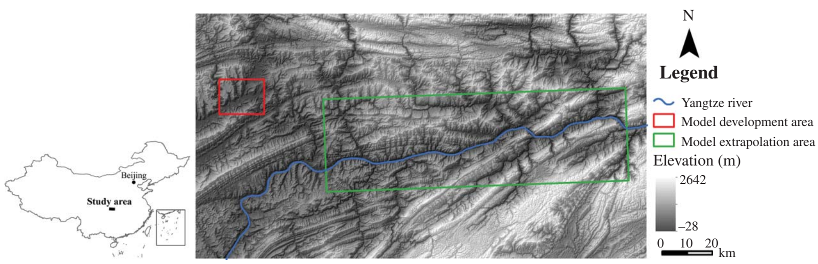

Figure 3. Location of the study area.

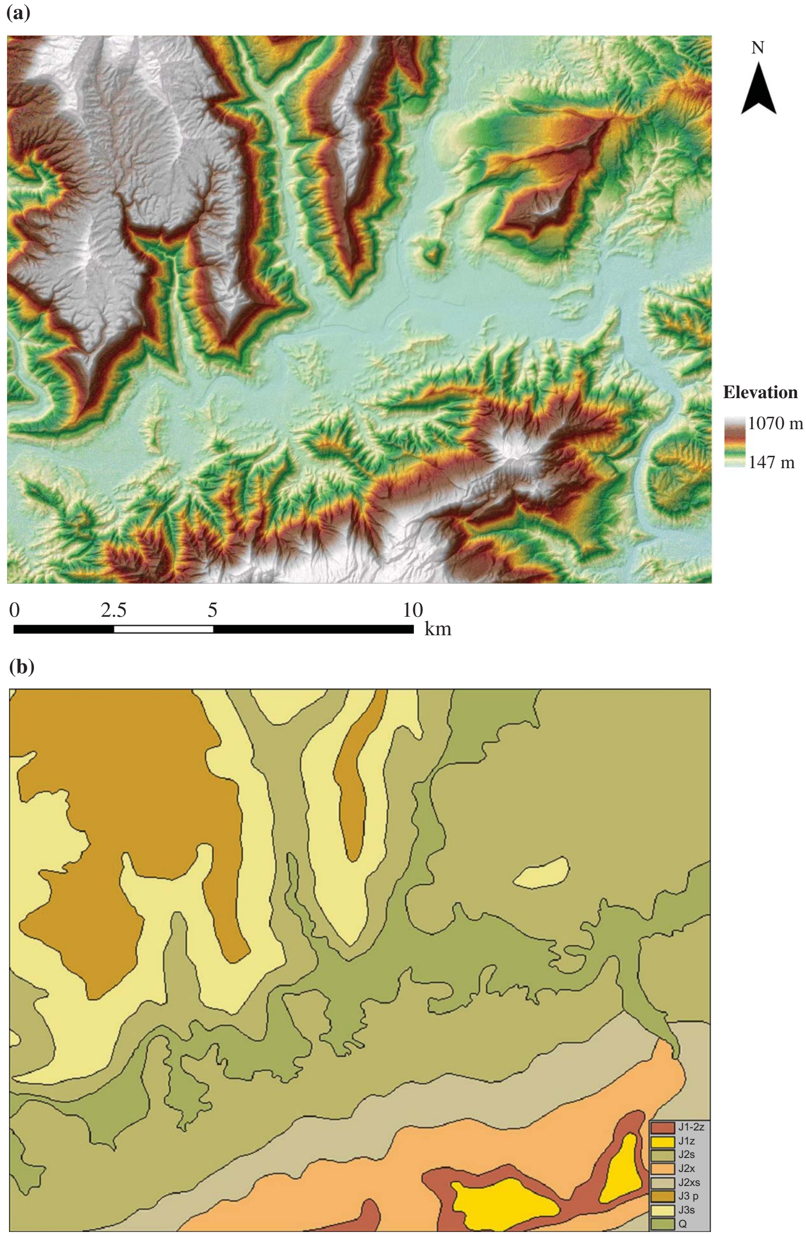

Figure 4. Map of the model-development area (Kaixian County): (a) topography; (b) lithology.

generated from a 1:10,000 scale topographic map, which was created in the 1960s (Wang 2008). The lithology in this area is of three major types: the lower to middle Jurassic system (including J2s, J2xs, J2x, J1-2z, and J1z), which is composed mainly of sandstone, siltstone, mudstone, and shale; the upper Jurassic system (including J3s and J3p), which is made up mainly of sandstone and siltstone; and the Quaternary system (Q), which is composed primarily of recent deposits in river valleys (Figure 4b).

Table 1. General statistics of the topography of the study area.

	Elevation (m)				Slope gradient (°)	
	Min	Max	Mean	SD	Mean	SD
Model-development area	147	1070	387.5	204.2	18.4	13.2
Model-extrapolation area	77	1961	686.9	362.9	23.3	12.2

4.1.2. Model-extrapolation area in the Three Gorges

The model-extrapolation area is located in the Three Gorges along the Yangtze River, in part of Chongqing Municipality (Figure 3, Table 1). This area is approximately 4600 km^2. The DEM was created from 1:50,000 scale topographic maps which were made in the 1960s (Wang 2008). The DEM grid size was also 25 m (Figure 5a).

The Three Gorges were formed by severe incision of massive limestone mountains of lower Paleozoic and Mesozoic age along narrow fault zones in response to Quaternary uplift (Liu *et al*. 2004). There are two major types of lithology in this area: (1) the

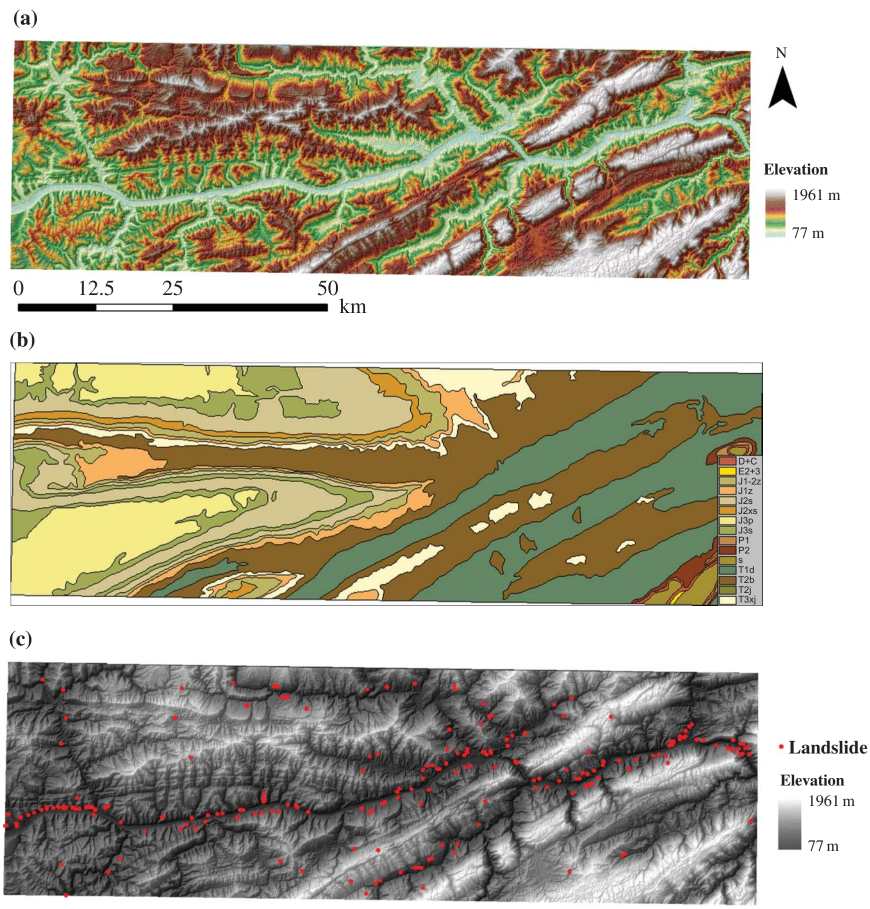

Figure 5. Map of the model-extrapolation area (Three Gorges area): (a) topography; (b) lithology; (c) location map of observed landslides.

Jurassic system (including J2s, J3p, J3s, J1z, J2xs, and J2x), consisting mainly of mudstone, sandstone, siltstone, shale, and coal; and (2) the Triassic system (including T2b, T1j, T3xj, and T1d), which is composed mainly of limestone, shale, claystone, dolomite, gypsum, sandstone, siltstone, and coal (Figure 5b).

In this area, 205 landslide occurrences compiled by Wang (2008) (Figure 5c) were used as the validation set to produce a boxplot for assessing the uncertainty of the landslide occurrence density within each landslide susceptibility level. The spatial distribution of observed landslides shows a linear pattern along the Yangtze River (see also Figure 3).

4.2. Modeling of expert knowledge-based landslide susceptibility mapping

4.2.1. Choosing and characterizing the predisposing factors

Based on the work reported in Zhu *et al*. (2004), Wang (2008) and Zhu *et al*. (under review), seven essential predisposing factors were chosen for *efLandslides* in the study area – three geologic attributes (lithology, stratum dip, and stratum strike) and four terrain attributes (slope gradient, slope aspect, slope height, and slope shape). Slope height (or the relative relief of a hill slope) was defined as the elevation difference between the highest elevation (at a divide or a local peak) and the lowest elevation (in a channel or pit) along a downslope profile of a hill slope (Giles and Franklin 1998, MacMillan *et al*. 2000). Slope shapes were classified into six general types by the local expert: (1) flat area; (2) concave hill slope; (3) upper concave, lower convex hill slope; (4) straight hill slope; (5) convex hill slope; and (6) upper convex, lower concave hill slope. Characterizations of these predisposing factors in the study area are summarized in Table 2.

4.2.2. Representing the relationship between landslide susceptibility and predisposing factors

Zhu *et al*. (2004, 2006) classified the lower to middle Jurassic system lithology type (including J2s, J2xs, J2x, J1-2z, and J1z) as the most susceptible type to landslides; the upper Jurassic system (including J3s and J3p) as a moderately susceptible type; and the Quaternary system (Q) as the least susceptible to landslide occurrence. Furthermore, knowledge from landslide experts on the relationships among the predisposing factors for

Table 2. Characterization of predisposing factors in the study area.

Predisposing factors		Characterization
Geology	Lithology	The geological data were created by digitizing and rasterizing a 1:200,000 geological map in the study area.
	Stratum dip	Stratum dip was generated using inverse distance weighted (IDW) interpolation.
	Stratum strike	Stratum strike was generated using nearest-neighbor interpolation.
Terrain	Slope gradient	Slope gradient and slope aspect were derived using Wood's (1996) algorithms with a neighborhood size of 200 m.
	Slope aspect	Slope height and slope shape were characterized by tracing the flow direction from ridge to valley (Giles and Franklin 1998, MacMillan *et al*. 2000, Wang 2008, Matsuura and Aniya 2012).
	Slope height	
	Slope shape	

this study area was encoded into four combinations of predisposing factors (Zhu *et al*. 2004, 2006): (1) the difference between the slope gradient and the stratum gradient, (2) slope gradient and lithology, (3) slope height and lithology, and (4) slope shape and lithology.

A landslide susceptibility value (in the range [0,1]) based on each individual combination of predisposing factors in the model-development area can be computed using the equations proposed in Zhu *et al*. (2006). The specifics of the equations are not included in this article and interested readers are referred to Zhu *et al*. (2006, under review) for details.

The landslide susceptibility maps for all individual predisposing factor combinations were then aggregated into a map of overall landslide susceptibility (with range [0, 4]). From each overall landslide susceptibility map a landslide susceptibility classification map can be derived by assigning the value ranges [0, 1), [1, 2), [2, 3), and [3, 4] to landslide susceptibility classes of very low, low, high, and very high, respectively.

For the model-extrapolation area, the knowledge of the relationships between landslide susceptibility and predisposing factors was transferred directly from the model-development area, with a slight adjustment because of the different lithology in the two areas (Wang 2008). In the model-extrapolation area, lithology types $D + \mathrm{C}$, J3s, S, T2b, and T3xj are the most susceptible types to landslides; lithology types J1z, J2s, J3p, and T2j are moderately susceptible types; and lithology types $\mathrm{E2} + 3$, $\mathrm{J1} - 2\mathrm{z}$, J2xs, P1, P2, and T1d are the least susceptible types (Zhu *et al*. 2004, Wang 2008, Zhu *et al*. under review).

4.3. Modeling of logistic regression model for landslide susceptibility mapping

The logistic regression model for the Kaixian study area using 21 locations where landslides have occurred since 1978 (as positive evidence), which were compiled by Wang (2008), and 21 randomly selected locations without landslide occurrences (as negative evidence) (Wang 2008). The predisposing factors used for the logistic regression model were the same as those used in *efLandslides*. The best fitting logistic regression model was

$$L = \frac{1}{1 + \mathrm{e}^{-(3.745\times\mathrm{Lithology}\,+\,0.497\,\times\,\mathrm{GradientDifference}\,-\,25.484\,\times\,\mathrm{SlopeGradient}\,+\,0.012\,\times\,\mathrm{SlopeHeight}\,-\,10.555)}},$$

where GradientDifference is the difference between the slope gradient and the stratum gradient.

4.4. Simulations of DEM error fields

Sequential Gaussian simulation was carried out using the GSLIB software (Deutsch and Journel 1998). The parameter-settings of DEM error fields in this study were based on the spatial characterizations of DEM error with similar data sources and terrain features in other DEM error-related studies (e.g. Oksanen and Sarjakoski 2005, Temme *et al*. 2009). To simulate DEM error fields of different magnitudes in the uncertainty assessment, three error magnitudes were used in the sequential Gaussian simulation by setting the error SD ($\mathrm{SD}_{\mathrm{err}}$) to 1, 7.5, and 15 m. These are considered optimistic, moderate, and pessimistic estimates, respectively, of DEM error magnitude. Three range parameter values (range $= 0$, 60, and 120 m) were also used in the sequential Gaussian simulation to simulate DEM error fields within a range from totally random or low spatial autocorrelation through high spatial autocorrelation when the spherical model was used to fit the semi-variogram and the nugget was set to zero. Thus, nine simulations (3 error magnitudes $\times$ 3 range values) were performed in both the model-development and the model-extrapolation areas. For each simulation the number of realizations was set to 100 because of the limited computing resources available for this study.

5. Results and discussion

5.1. Model-development area

Figure 6a shows the mean value of the SD map derived from the landslide susceptibility realizations from both *efLandslides* and logistic regression for each simulation in the model-development area. Uncertainty trends changing in opposite directions were observed under simulations with different DEM error magnitudes. For the simulations with the lowest DEM error magnitude ($SD_{err} = 1$ m), the overall uncertainties in both *efLandslides* and logistic regression are very close and slightly decrease with increasing spatial autocorrelation in the simulated DEM error field. By contrast, for simulations with greater DEM error magnitude ($SD_{err} = 7.5$ and 15 m), the overall uncertainties in both the *efLandslides* and logistic regression models increase with increasing spatial autocorrelation in the simulated DEM error field, although the overall uncertainty in *efLandslides* is clearly lower than in logistic regression.

The opposite changing uncertainty trends in *efLandslides* or logistic regression under simulations with different DEM error magnitudes may be attributed to the different characteristics of the terrain factor algorithms used. As shown in Table 2, the slope gradient and aspect were calculated using Wood's (1996) surface-fitting algorithm by specifying a larger neighborhood size (200 m). For a specific simulated DEM error magnitude, when the spatial autocorrelation of the simulated DEM error field increases from totally random (range = 0 m) to a higher level (e.g. range = 60 or 120 m), the fitted surface from the large neighborhood used for calculating slope gradient and aspect tends to be different from that fitted from the original DEM data. This means that the uncertainty in calculating slope gradient and aspect increases, as is apparent in the mean value of the SD map derived from the slope gradient realizations for each simulation in the model-development area (Figure 7a).

However, the uncertainty in calculating slope shape and height has a different trend than the uncertainty in calculating slope gradient and aspect. Because slope shape and height were calculated by tracing the flow direction from ridge to valley (MacMillan *et al*. 2000, Wang 2008), the impact of DEM error on the calculation of slope shape and height would not increase with spatial autocorrelation of simulated DEM error. The mean value of the SD map derived from the slope height realizations for each simulation in the model-development area shows that the uncertainty in calculating slope height remains stable or decreases with increasing spatial autocorrelation of simulated DEM error (Figure 7b). The larger the simulated DEM error magnitude, the faster the uncertainty in calculating

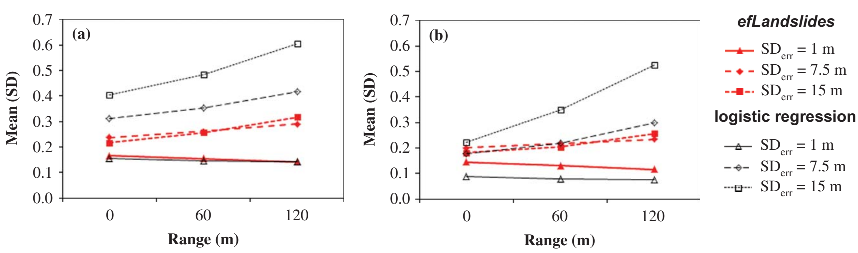

Figure 6. Mean SD derived from the landslide susceptibility realizations of each simulation using both *efLandslides* and logistic regression: (a) the model-development area; (b) the model-extrapolation area.

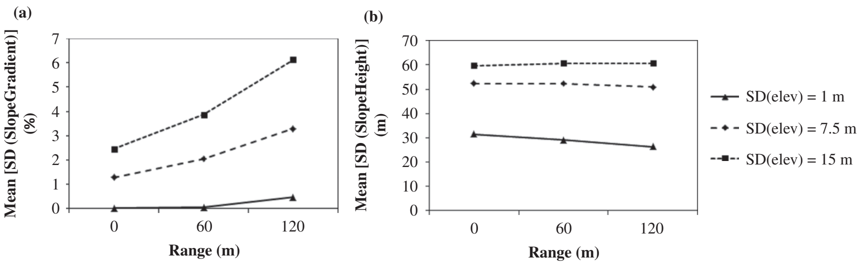

Figure 7. Mean SD derived from the terrain attribute realizations of each simulation in the model-development area: (a) slope gradient; (b) slope height.

slope height decreases with increasing spatial autocorrelation of simulated DEM error (Figure 7b).

Therefore, the authors suppose that when the high-accuracy DEM ($SD_{err} = 1$ m) was simulated for *efLandslides* and for logistic regression, the uncertainty associated with slope shape and height contributed more to the overall uncertainty than did the characterization of slope gradient and aspect. This resulted in an overall uncertainty decrease with increasing spatial autocorrelation in the simulated DEM error field. By contrast, when a lower-accuracy DEM ($SD_{err} = 7.5$ or 15 m) was used for *efLandslides* or for logistic regression, the uncertainty associated with slope gradient and aspect contributed more to the overall uncertainty than did the characterization of slope shape and height. This resulted in an overall uncertainty increase with increasing spatial autocorrelation of the simulated DEM error field.

Figure 6a also shows that for each tested range the overall uncertainty in *efLandslides* first increased and then remained stable with further increases in simulated DEM error magnitude. By contrast, the overall uncertainty in logistic regression increased monotonically with the simulated DEM error magnitude. The change in overall uncertainty in *efLandslides* was much less than the change in overall uncertainty in logistic regression, which means that the overall uncertainty in *efLandslides* is less sensitive to DEM error magnitude than the overall uncertainty in logistic regression. This phenomenon might be attributed to the data-driven nature of multivariate statistical models, which makes the performance of logistic regression rely more heavily on DEM data quality than does *efLandslides* (Wang 2008, Zhu *et al*. under review).

5.2. Model-extrapolation area

Figure 6b shows mean SD values derived from the landslide susceptibility realizations of *efLandslides* and logistic regression for each simulation in the model-extrapolation area. Under the simulations with the greatest DEM error magnitude ($SD_{err} = 15$ m), the overall uncertainty in *efLandslides* is clearly lower than in logistic regression, although the case with the lowest DEM error magnitude ($SD_{err} = 1$ m) shows the opposite situation. This means that in the model-extrapolation area, the overall uncertainty in *efLandslides* is less sensitive to DEM error magnitude than the overall uncertainty in logistic regression.

A comparison between Figure 6a and b shows that the basic trend of mean SD values from *efLandslides* and logistic regression simulations in the model-extrapolation area is similar to that in the model-development area. Furthermore, overall uncertainties in both *efLandslides* and logistic regression in the model-extrapolation area are generally lower than those in the model-development area, which means that both *efLandslides* and logistic regression proved to be portable in this study from the perspective of overall uncertainty.

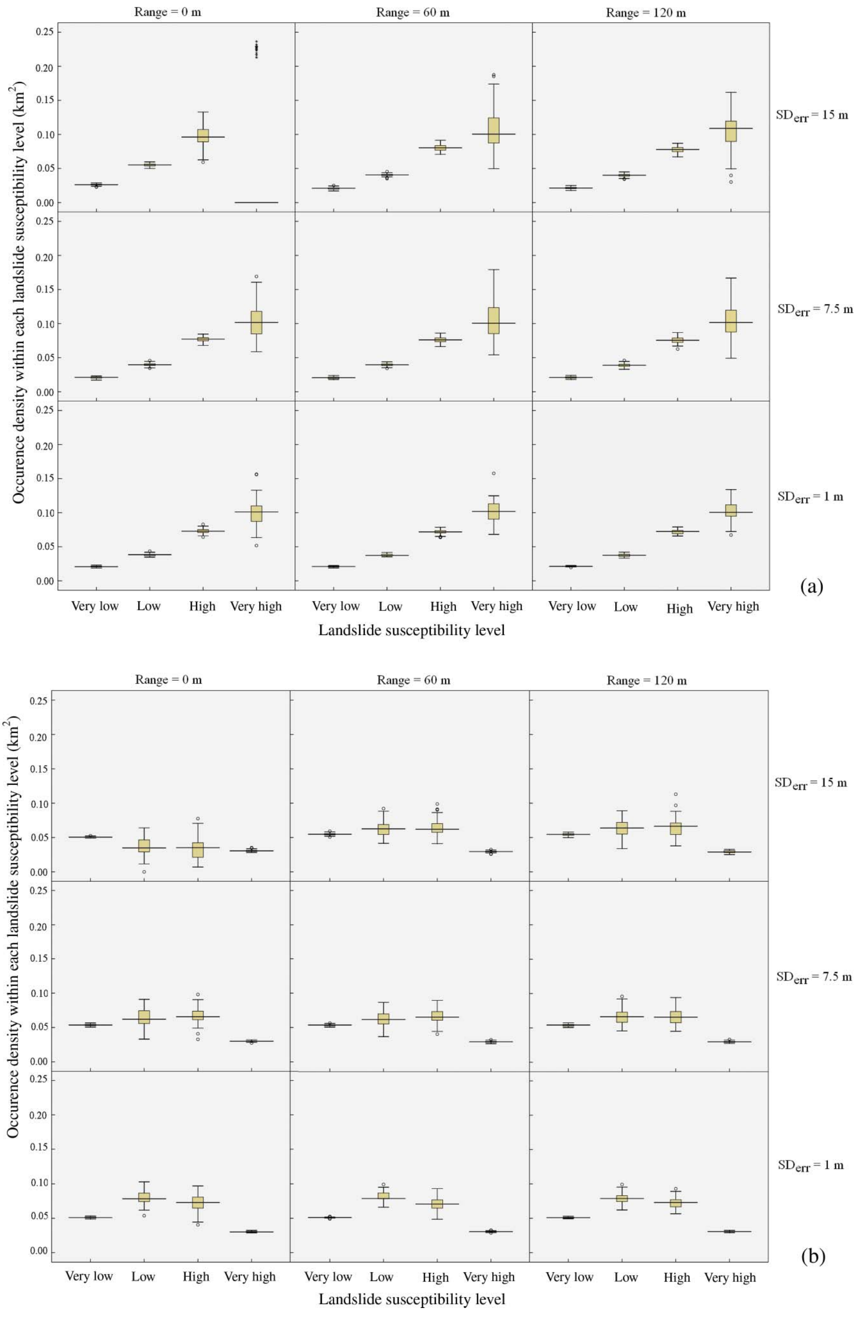

Figure 8. Boxplots of the landslide occurrence density within each landslide susceptibility level in the model-extrapolation area: (a) *efLandslides*; (b) logistic regression (the circular marks in the figure represent outliers).

Figure 8a and b shows boxplots of the landslide occurrence density within each landslide susceptibility level for both *efLandslides* and logistic regression in the model-extrapolation area. For *efLandslides* under all simulations, not only landslide occurrence density but also the uncertainties in landslide occurrence density increase with landslide susceptibility level and landslide occurrence density. The only exception is the case of 'very high' landslide susceptibility level from the simulation with the greatest magnitude of error and a totally random DEM error field, i.e., $SD_{err} = 15$ m and range = 0 m (Figure 8a). It is reasonable for a landslide susceptibility mapping approach to behave in this way.

By contrast, logistic regression showed different performance. As shown in Figure 8b, for logistic regression under all simulations the uncertainties in landslide occurrence density within 'low' and 'high' landslide susceptibility levels are higher than those within 'very low' and 'very high' landslide susceptibility levels. Within 'low' and 'high' landslide susceptibility levels under all simulations, the uncertainties in the landslide occurrence density from logistic regression are higher than those from *efLandslides*. Within 'very high' landslide susceptibility levels under all simulations, the opposite situation occurs, that is, the uncertainties in the landslide occurrence density from logistic regression are much lower than those from *efLandslides*. However, this situation is accompanied by an unreasonable result from logistic regression, in which the mean landslide occurrence density within 'very high' landslide susceptibility level is lower than those within other landslide susceptibility levels.

The difference between the uncertainties in landslide occurrence density from *efLandslides* and logistic regression may be attributed to the different areas of landslide susceptibility level predicted by these two approaches to landslide susceptibility mapping. As shown in Figure 9, the uncertainties in the areas of landslide susceptibility level predicted by both *efLandslides* and logistic regression are very low under all simulations. However, there is clear difference between the predicted areas of landslide susceptibility level by *efLandslides* and by logistic regression. For *efLandslides* under all simulations, the minimum predicted areas of landslide susceptibility level are always the cases of 'very high' landslide susceptibility level (Figure 9a). Therefore, it is reasonable for *efLandslides* that the uncertainties in landslide occurrence density attain their highest values for the cases of 'very high' landslide susceptibility level (see Figure 8a). For logistic regression under all simulations, the predicted areas of 'low' and 'high' landslide susceptibility levels are always much smaller than those of 'very low' and 'very high' landslide susceptibility levels (Figure 9b). Therefore, it is understandable that for logistic regression, the uncertainties in landslide occurrence density within 'low' and 'high' landslide susceptibility levels are always higher than those within 'very low' and 'very high' landslide susceptibility levels, as shown in Figure 8b.

6. Conclusions

The main conclusions from this study using a nonconditional stochastic simulation method include the following.

(1) In the model-development area, at greater DEM error magnitude ($SD_{err} = 7.5$ and 15 m), the overall uncertainty in *efLandslides* was clearly lower than in logistic regression. Moreover, the overall uncertainties in both *efLandslides* and logistic regression increased with increase in spatial autocorrelation of the simulated DEM error field. By contrast, under the simulations with the lowest DEM error magnitude ($SD_{err} = 1$ m), the overall uncertainties in both *efLandslides* and logistic regression were very similar and were found to decrease slightly with increasing spatial autocorrelation in the simulated DEM error field.

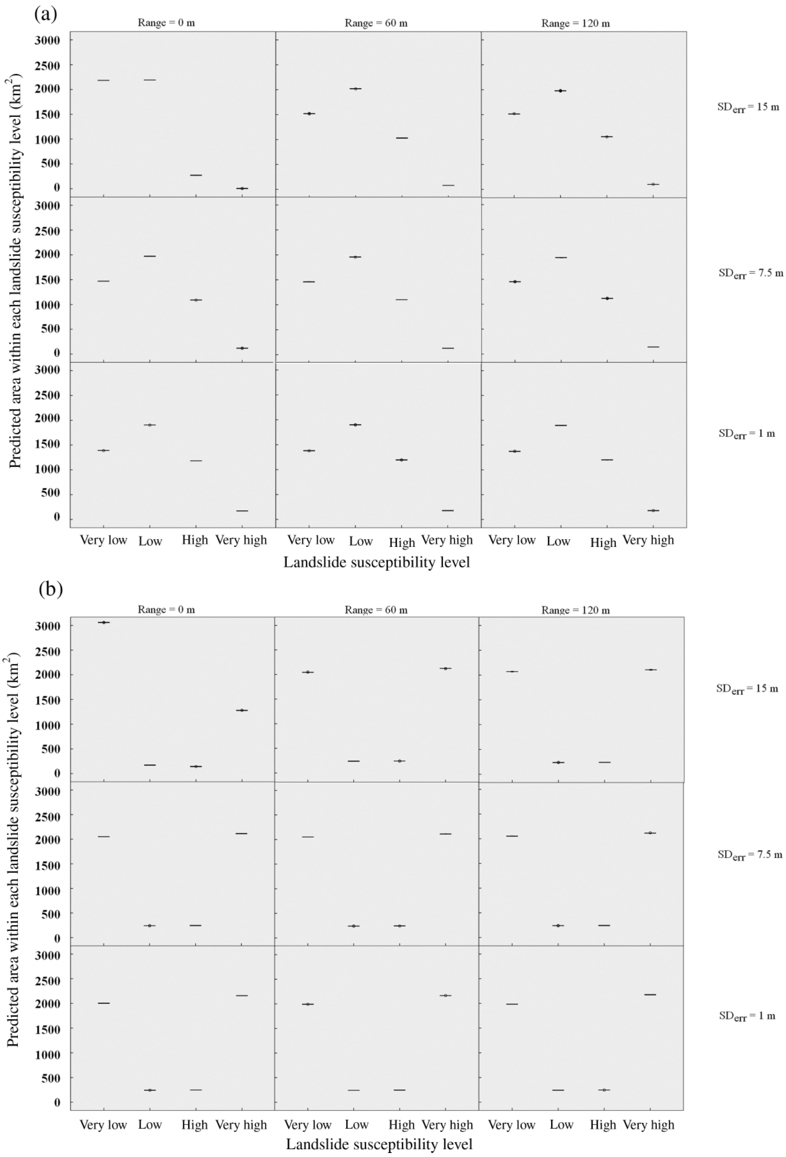

Figure 9. Boxplots of the area of each landslide susceptibility level predicted by *efLandslides* and logistic regression in the model-extrapolation area: (a) *efLandslides*; (b) logistic regression.

(2) Also in the model-development area, the change in the overall uncertainty in *efLandslides* with different DEM error magnitudes was much less than the change in the overall uncertainty in logistic regression under these circumstances. This means that the overall uncertainty in *efLandslides* is less sensitive to DEM error magnitude than the overall uncertainty in logistic regression.

(3) In the model-extrapolation area, the overall uncertainty in *efLandslides* was less sensitive to DEM error magnitude than the overall uncertainty in logistic regression. For *efLandslides*, the uncertainty in landslide occurrence density increases with landslide susceptibility level and landslide occurrence density. The uncertainties in the predicted areas of landslide susceptibility level by both *efLandslides*

and logistic regression are very low under all simulations. The performance of *efLandslides* was more reasonable than that of logistic regression in this part of the study.

Acknowledgements

This study was supported by the National Natural Science Foundation of China (No. 40971235); the National High-Tech Research and Development Program of China (No. 2011AA120302); and the Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences (No. 2011RC203). We thank Dr. Ian S. Evans and other anonymous reviewers for their constructive comments on an earlier version of this article.

References

Aleotti, P. and Chowdhury, R., 1999. Landslide hazard assessment: summary review and new perspectives. *Bulletin of Engineering Geology and the Environment*, 58, 21–44.

Atkinson, P.M. and Massari, R., 1998. Generalised linear modelling of susceptibility to landsliding in the central Apennines, Italy. *Computers and Geosciences*, 24, 373–385.

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

Bai, S.-B., *et al*., 2010. GIS-based logistic regression for landslide susceptibility mapping of the Zhongxian segment in the Three Gorges area, China. *Geomorphology*, 115, 23–31.

Brenning, A., 2005. Spatial prediction models for landslide hazards: review, comparison and evaluation. *Natural Hazards and Earth System Sciences*, 5, 853–862.

Carrara, A., *et al*., 1991. GIS techniques and statistical models in evaluating landslide hazard. *Earth Surface Processes and Landforms*, 16, 427–445.

Chung, C.F., Fabbri, A.G., and van Westen, C.J., 1995. Multivariate regression analysis for landslide hazard zonation. *In*: A. Carrara and F. Guzzetti, eds. *Geographical information systems in assessing natural hazards*. Dordrecht, The Netherlands: Kluwer Academic Publishers, 107–133.

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

Dai, F.C., Lee, C.F., and Ngai, Y.Y., 2002. Landslide risk assessment and management: an overview. *Engineering Geology*, 64, 65–87.

Deutsch, C.V. and Journel, A.G., 1998. *GSLIB: Geostatistical software library and user's guide*. 2nd ed. New York: Oxford University Press.

Dobson, A.J., 1990. *An introduction to generalized linear models*. London: Chapman and Hall.

Fisher, P., 1991. First experiments in viewshed uncertainty: the accuracy of the viewshed area. *Photogrammetric Engineering and Remote Sensing*, 57 (10), 1321–1327.

Fisher, P. and Tate, N.J., 2006. Causes and consequences of error in digital elevation models. *Progress in Physical Geography*, 30 (4), 467–489.

Giles, P.T. and Franklin, S.E., 1998. An automated approach to the classification of the slope units using digital data. *Geomorphology*, 21, 251–264.

Goovaerts, P., 1997. *Geostatistics for natural resources evaluation*. New York: Oxford University Press.

Guzzetti, F., *et al*., 1999. Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study, Central Italy. *Geomorphology*, 31, 181–216.

Guzzetti, F., *et al*., 2005. Probabilistic landslide hazard assessment at the basin scale. *Geomorphology*, 72, 272–299.

Hammersley, J.M. and Handscomb, D.C., 1979. *Monte Carlo methods*. London: Chapman and Hall.

Heuvelink, G.B.M., 1998. *Error propagation in environmental modelling with GIS*. London: Taylor & Francis.

Holmes, K.W., Chadwick, O.A., and Kyriakidis, P.C., 2000. Error in a USGS 30-meter digital elevation model and its impact on terrain modeling. *Journal of Hydrology*, 233 (1–4), 154–173.

Lee, S., 2005. Application of logistic regression model and its validation for landslide susceptibility mapping using GIS and remote sensing data. *International Journal of Remote Sensing*, 26 (7), 1477–1491.

Lee, S. and Choi, J., 2004. Landslide susceptibility mapping using GIS and the weight-of-evidence model. *International Journal of Geographical Information Science*, 18 (8), 789–814.

Lindsay, J.B., 2006. Sensitivity of channel mapping techniques to uncertainty in digital elevation data. *International Journal of Geographical Information Science*, 20 (6), 669–692.

Liu, J.G., *et al*., 2004. Landslide hazard assessment in the Three Gorges area of the Yangtze river using ASTER imagery: Zigui-Badong. *Geomorphology*, 61, 171–187.

Luger, G.F., 2005. *Artificial intelligence: structures and strategies for complex problem solving*. New York: Addison-Wesley.

MacMillan, R.A., *et al*., 2000. A generic procedure for automatically segmenting landforms into landform elements using DEMs, heuristic rules and fuzzy logic. *Fuzzy Sets and Systems*, 113, 81–109.

Matsuura, T. and Aniya, M., 2012. Automated segmentation of hillslope profiles across ridges and valleys using a digital elevation model. *Geomorphology*, 177–178, 167–177.

Ohlmacher, G.C. and Davis, J.C., 2003. Using multiple logistic regression and GIS technology to predict landslide hazard in northeast Kansas, USA. *Engineering Geology*, 69, 331–343.

Oksanen, J. and Sarjakoski, T., 2005. Error propagation of DEM-based surface derivatives. *Computers & Geosciences*, 31 (8), 1015–1027.

Temme, A.J.A.M., *et al*., 2009. Geostatistical simulation and error propagation in geomorphometry. *In*: T. Hengl. and H.I. Reyter, eds. *Geomorphometry: concepts, software, applications*. Amsterdam: Elsevier, 121–140.

Wang, R.X., 2008. *An expert knowledge-based approach to landslide susceptibility mapping using GIS and fuzzy logic*. Thesis (PhD). University of Wisconsin.

Wechsler, S.P., 2007. Uncertainties associated with digital elevation models for hydrologic applications: a review. *Hydrology and Earth System Sciences*, 11, 1481–1500.

Wilson, J.P., 2012. Digital terrain modeling. *Geomorphology*, 137, 107–121.

Wood, J.D., 1996. *The geomorphological characterisation of digital elevation models*. Thesis (PhD), University of Leicester, UK, http://www.soi.city.ac.uk/~jwo/phd.

Wu, S., *et al*., 2001. Zonation of the landslide hazards in the forereservoir region of the Three Gorges Project on the Yangtze River. *Engineering Geology*, 59, 51–58.

Zhu, A.X., *et al*., under review. An expert knowledge-based approach to landslide susceptibility mapping. *Geomorphology*.

Zhu, A.X., 1999. A personal construct-based knowledge acquisition process for natural resource mapping. *International Journal of Geographical Information Science*, 13 (2), 119–141.

Zhu, A.X., *et al*., 2004. Mapping landslide susceptibility in the Three Gorges area, China using GIS, expert knowledge and fuzzy logic. *In*: *GIS and Remote Sensing in Hydrology, Water Resources and Environment, IAHS Publication 289 (IAHS Red Book)*, Wallingford, UK: International Association of Hydrological Sciences, 385–391.

Zhu, A.X., 2005. Research issues on uncertainty in geographic data and GIS-based analysis. *In*: E.L. Usery and B.M. Robert, eds. *A research agenda for geographic information science*. Boca Raton, Florida: CRC Press, 197–223.

Zhu, A.X., *et al*., 2006. A landslide susceptibility mapping approach using fuzzy logic under GIS. *Progress in Geography*, 25 (4), 1–12 (in Chinese with English abstract).

van Westen, C.J., *et al*., 1993. *GISSIZ: Training package for geographic information systems in slope instability zonation, ITC Publication No. 15*, Enschede, The Netherlands: ITC.