

# Landslide Susceptibility Analysis Based on Data Field

Xianmin Wang

Institute of Geophysics and Geomatics, China University of Geosciences, Wuhan, China  
Email: xianminwang@163.com

Ruiqing Niu

Institute of Geophysics and Geomatics, China University of Geosciences, Wuhan, China  
Email: rqniu@163.com

**Abstract**—The Three Gorges are the areas in which the geological disasters are very serious. There often happen great landslide disasters, which brings tremendous threat to normal running of the Three Gorge Dam and the properties and lives of the residents in the reservoir. So landslide susceptibility analysis is an important task of prevention and cure of landslides in the Three Gorges. In this paper, landslide susceptibility analysis in the Three Gorges is studied based on spatial data mining. ETM+ image, 1:50000 geological graph and 1:10000 relief map are adopted as the data origins to produce the factors closely related to landslide transmutation, including slope structure, engineering rock group, slope level, fluctuation influence of reservoir water and land exploration. A spatial data mining method is proposed which is suitable for landslide susceptibility analysis. Firstly data field method is adopted to synthetically analyze the spatial distribution of landslides and the key factors influencing landslide transmutation and extract the potential centers. Secondly cloud model method is adopted to describe the concept represented by each potential center, and the synthesized cloud method elevates the concepts to produce the high-level concepts. Finally clustering analysis is made according to the membership degree of each data point to each high-level concept, and realizes landslide susceptibility analysis in the Three Gorges. The experimental results have shown that the method proposed in this paper obtains a good prediction result, which is priori to the ones of the other 3 methods (IsoData, K-Means and Parallelepiped). So the method can well realize landslide susceptibility analysis in the Three Gorges.

**Index Terms**—Landslide, susceptibility, data field, geology

## I. INTRODUCTION

Landslide is one of the ten primary nature disasters in the world and is a serious geological disaster just inferior to earthquake and floodwater. The Three Gorges are the areas in which the geological disasters are frequent and landslide is the most primary disaster in the Three Gorges. According to the statistic data, in the Three Gorges there exist 2490 landslides and there happen more than 200 landslides per year. So landslide susceptibility analysis has become a very important task of prevention and cure of landslides in the Three Gorges.

At present the key techniques and methods for landslide prediction are both hot research topics and challenges in the landslide study area. In the past 30 years, national and international researchers have conducted many studies on landslide susceptibility analysis [1-4]. Ying [5-6] has made deep and systemic research on spatial forecast and stability partition of landslide and slope and proposed such models as Information Analysis Model, Multi-factor Regress Model, Clustering Analysis Model and Judgement Analysis Model. William [7] used Light Detection and Ranging (LIDAR) data to visually map landslides, headscarps, and denuded slopes in Seattle, Washington. Four times more landslides were mapped than by previous efforts that used aerial photographs. Many individual landslides were adopted to create the mapped landforms. He also analyzed the landform's relative susceptibility to future landsliding according to the spatial densities of historical landslides. Zhu [8] applied double logistic regression in landslide susceptibility map. The first logistic regression equation showed that elevation, proximity to road, river and residential area were main factors triggering landslide occurrence in the study area. The prediction accuracy of the first landslide susceptibility map was showed to be 80%. Some non-landslide areas are incorrectly divided into high and medium landslide susceptibility zone. A second logistic regression was applied in high landslide susceptibility zone to improve the status. The new logistic regression equation indicated that only areas with unstable engineering and geological conditions were prone to landslide during large scale engineering activity. The experimental results showed that double logistic regression analysis improved the non-landslide prediction accuracy. Havenith [9] analyzed the influence of geological and morphological factors upon landslide occurrence on a regional scale based on a digital data set including landslides triggered in 1992 and several older landslides as well as various types of digital elevation models (DEMs), ASTER image data, and geological and active fault maps. By statistical methods these data were combined to compute landslide susceptibility (LS) maps. Günther [10] evaluated both the structurally-controlled

failure susceptibility of the fractured cretaceous chalk rocks and the topographically-controlled shallow landslide susceptibility of the overlying glacial sediments for the Jasmund cliff area on Rügen Island, Germany, and produced a landslide susceptibility map for the area. The shallow landslide susceptibility analysis involved a physically-based slope stability evaluation utilizing material strength and hydraulic conductivity data, and a bivariate landslide susceptibility analysis exploiting landslide inventory data and thematic information on ground conditioning factors. Sarkar [11] used an integrated remote sensing and GIS approach to conduct landslide susceptibility mapping in the study area – a part of the Darjeeling Himalaya. IRS satellite data, topographic maps, field data, and other informative maps were used as the inputs to identify the important terrain factors and generate the corresponding thematic data layers. The resulting landslide susceptibility map delineates the area into different zones of four relative susceptibility classes: high, moderate, low, and very low. Guinau [12] studied a methodology for landslide susceptibility assessment focusing on the developing countries. He chose a region in NW Nicaragua as the study area, which was among the most severely hit during the Mitch event. He produced a high-resolution inventory landslide map at 1 : 10000 scale and selected the relevant instability factors from a Terrain Units Map. He developed an analysis of failure-zones and terrain factors in an attempt to classify the land into zones according to the propensity to landslides triggered by heavy rainfalls.

However the traditional studies take little consideration on the uncertain and nonlinear action of a landslide system, lack in the mining and extraction of various complicated information of a landslide system, often need much manual intervention and possess poor intelligence. Landslide data come from various origins, and along with the improvement of data collection and accumulation means, the large amount of landslide data don't get adequate mining and utilization at all. The introducing of data mining provides a new idea to solve the above problems. Data mining can mine the potential rules and knowledge from a large amount of multi-theme landslide data to make landslide susceptibility analysis and provide decision support with the prevention and cure of the landslide disasters. And landslide data possess the characters of space, diversity and large amount. There exist both discrete and continuous data, and even exist absent or faulty data. So in the present study it is a difficulty how to synthetically consider the landslide spatial distribution and the key factors influencing landslide transmutation and how to elevate the concept levels. Therefore a data mining algorithm suitable for landslide prediction must be studied. At present, the research on introducing data mining into landslide prediction is rarely few. Yao [13] applied SVM in landslide susceptibility. He selected the study area from the natural terrain of Hong Kong and used slope aspect, elevation, slope angle, profile curvature of slope, lithology, vegetation cover and topographic wetness index (TWI) as the environmental parameters which influence the landslide transmutation.

One-class SVM, two-class SVM and logistic regression (LR) were used to make landslide susceptibility. The results showed that two-class SVM obtained better prediction result than logistic regression and one-class SVM. Caniani [14] applied the techniques derived from artificial intelligence (Artificial Neural Network - ANN) in landslide susceptibility in the area of Potenza. He adopted the parameters of slope aspect, topographical index, elevation, slope gradient, topographical shape, land use and lithology to make landslide susceptibility. Chu [15] proposed an approach to assess landslide susceptibility spatially, which integrated decision tree model and spatial cluster statistic. The method used tree graphs to explicitly represent the relationships of landslides and instability factors and adopted the local Getis-Ord statistics to cluster cells with high landslide probability. Then the method classified the analytic result from the local Getis-Ord statistics to establish a map of landslide susceptibility zones.

In this paper focusing on the Three Gorges, spatial data mining is introduced to make landslide susceptibility analysis. ETM+ images with the re-sampling resolution 20m, 1:50000 geological map and 1:10000 terrain map were adopted to produce the key factors closely related to landslide transmutation, including slope structure, engineering rock group, slope level, fluctuation influence of reservoir water and land exploration. In this paper a spatial data mining method suitable for landslide prediction is proposed. Firstly the method adopts data field to synthetically analyze the spatial distribution of landslides and the key factors influencing landslide transmutation. Secondly it mines the potential centers and describes each concept represented by each potential center based on cloud models. Thirdly it adopts the synthesized clouds to elevate the concept and knowledge level and produces the high-level concepts. Finally it makes clustering analysis according to the membership of each data point to each high-level concept and realizes landslide spatial prediction.

## II. DATA FIELD

Wang [16] studied a spatial data mining method based on cloud model and data field. Here we apply the method in landslide susceptibility analysis in the Three Gorges.

The key of a clustering method lies in the choice of the initial clustering center [16-17]. Data field [17] can impersonally and rationally describe the influence among the spatial entities. The potential of data field is the fusion of spatial information and attribute information and well reflects the importance of each spatial object in the whole region. The potential centers reflect the positions of the gravity centers of spatial objects. So the characters of data groups in different levels can be reflected by the radiation factor of the data field and the initial clustering center can be exactly found.

### A. Fusion of space and attribute information

The nature and severity of landslide disasters are related to many factors such as geological structure, stratum, lithology, terrain and physiognomy, rainfall and

human activities [18]. In this paper, the factors closely related to landslides in the Three Gorges are analyzed and classified into 4 classes: (1) texture, (2) spectra, (3) geology, geomorphology and environment, and (4) human activity. The texture factors are chosen as the 7 textural values of the GLCM (Gray level co-occurrence matrix) of TM4, which are mean, contrast, entropy, variance, homogeneity, dissimilarity, and angle second moment. The spectral factors are chosen as the 4<sup>th</sup>, 3<sup>th</sup>, and 2<sup>th</sup> bands of the ETM+ images. The geological, geomorphology and environmental factors are chosen as slope structure, reservoir water fluctuation, engineering rock group and slope level. Slope structure is classified into 5 classes: inverse slope, reverse slope, dip slope, forward slope, and lateral slope. Reservoir water fluctuation is classified into 4 classes: fluctuating region, strongly influenced region, moderately influenced region and poorly influenced region. Engineering rock group is classified into 3 classes: alternately soft and hard stratum, hard rock, and soft rock. Slope level is classified into 4 classes: steep slope, medium slope, gently inclined slope, and gentle slope. The human activity factor is chosen as land utilization, which is classified into 5 classes: water, vegetation, residential area, bare rock and bare soil.

Because the unit and amount level of each factor differs much, the factor values must be processed for scalar [16].

$$a'_i = \frac{a_i - a_{\min}}{a_{\max} - a_{\min}} \times 100 \quad (1)$$

$a_{\max}$  and  $a_{\min}$  are respectively the maximum and minimum values of the factor  $a_i$ , and  $a'_i$  is the scalar value without the unit. After procession each scalar value is between 0 and 100, so they can be compared between each other.

The spatial distribution of data points and the scalar values of the factors can be fused by the weights [16].

$$fusion = w_1 * x + w_2 * y + \sum_{k=3}^n w_k * a'_i \quad (2)$$

In Formula (2)  $\sum_{i=1}^n w_i = 1$ ,  $(x, y)$  is the spatial coordinate of a data point, and  $fusion$  is the attribute value after fusion.

### B. Potential transform of data field

In the data space, each data radiates its energy into the space to form the data field. The potential function of any point  $x$  in the data field is defined as the summation of the influence from all the data points. For the given  $n$  data points, the potential function of the point  $x$  based on the distance  $D = \{d_1, d_2, \dots, d_n\}$  is defined as [16, 17]:

$$F(x) = \sum \mu_i \exp\left(-\frac{(d(x, d_i))^2}{2\delta^2}\right) \quad (3)$$

In Formula (3)  $\mu_i$  is the value of the point,  $d(x, d_i)$  is the distance between the point  $d_i$  and  $x$ , and  $\delta$  is the radiation factor. The total potential  $F$  of the point  $x$  is the

summation of its potentials. The potential function is related to the position or distance and can be piled up, so in a data space, each data object contributes to the potential of any point, and the contribution is inversely to the square of the distance between them.

The introducing of data field can well solve the problem of the segmentation between the spatial information and the attribution information, and provide the relationship between them.

### C. Establishment of initial clustering center

The initial clustering centers can be established according to the position of the potential centers. The potential centers reflect the characters of the objects on one or more than one attribute values, and all the potential centers in the potential field constitute a character space of spatial objects. The definition of the potential center is as follows [17]:

$$F_{center} \geq F(i, j) \quad (4)$$

In Formula (4)  $F_{center}$  is the value of the potential center,  $F(i, j)$  is the potential value of one of the 8 directions around the potential center, and  $(i, j)$  is the coordinate of the point.

Formula (4) has illuminated that the potential center is the point of the local maximum potential value and represents the position of the gravity center of a class, so the potential centers are just the initial clustering centers.

## III. CLOUD MODEL

### A. Normal cloud model

Cloud model [19-20] can integrate the fuzziness and randomness of a linguistic term in a unified way. Prof. Deyi Li defined the concept of Compatibility Cloud as follows [19]: Let  $U$  be the set  $U = \{x\}$ , as the university of discourse, and  $T$  a linguistic term associated with  $U$ . As the membership degree of  $x$  in  $U$  to the linguistic term  $T$ ,  $\mu_r(x)$  is a random number with a stable tendency, the distribution of which in the university of discourse  $U$  is called Compatibility Cloud, shortly Cloud.

The values of  $\mu_r(x)$  are in  $[0, 1]$ . A compatibility cloud is a mapping from the university of discourse  $U$  to the unit interval  $[0, 1]$ , namely

$$\mu_r(x) : U \rightarrow [0, 1] \quad (5)$$

A cloud illuminates the qualitative connotation of a linguistic atom with three digital characteristics: expected value  $Ex$ , entropy  $En$  and deviation  $He$  [19]. The expected value  $Ex$  of a compatibility cloud corresponds to the gravity center of the cloud. The expected value  $Ex$  in the universe of discourse  $U$  is completely consistent with the linguistic atom. The entropy  $En$  of a linguistic atom reflects the fuzziness of the concept within the universe of discourse  $U$  and illuminates the amount of the elements in the universe of discourse which are subjective to the linguistic atom. The bigger is the entropy, the more elements are subjective to the linguistic atom, and the fuzzier is the concept. The deviation  $He$  is the entropy of  $En$ , which reflects the discretization of the

cloud drops. The bigger is the deviation, the more discrete is the cloud drops, the more random is the membership degree, and the bigger is the thickness of the cloud.

The three digital characteristics of a cloud model integrate the fuzziness and randomness of a linguistic term in a unified way, establish the mapping between qualities and quantities, and is the basis of concept elevation and knowledge representation.

The normal compatibility clouds [19-20] are most useful in representing linguistic atoms because normal distributions have been supported by results in every branch of both social sciences and natural sciences. The MEC of a normal compatibility cloud is defined as follows [19, 21]:

$$MEC_A(x) = \exp[-(x - Ex)^2 / (2En^2)]. \quad (6)$$

Each qualitative concept  $C_i$  ( $i=1, 2, \dots, m$ ) represented by each potential center can be described by a normal cloud. The concrete steps are as follows [19-21]:

(1) Produce the 3 digital characteristics of a normal cloud according to the samples of Concept  $C_i$ .

① Establishing the expected value  $Ex$  of Concept  $C_i$  according to the attribute value of the potential center;

② Calculating the entropy  $En$  of Concept  $C_i$  as

$En = \sqrt{\frac{\pi}{2}} * \frac{1}{n} \sum_{j=1}^n |x_j - Ex|$ , in which  $x_j$  is a sample of Concept  $C_i$ ;

③ Calculating the deviation  $He$  of Concept  $C_i$  as

$He = \sqrt{s^2 - En^2}$ , in which  $s^2 = \frac{1}{n-1} \sum_{j=1}^n (x_j - \bar{X})^2$  and

is the sample variance, and  $\bar{X}$  is the sample mean. In practical, the thickness  $He$  of a compatibility cloud can also be established according to the need.

(2) Produce cloud drops according to the above 3 digital characteristics.

① Producing normal random numbers  $c_i$  with the expected value  $Ex$  and standard deviation  $En$ , namely  $c_i = G(Ex, En)$ ;

② Producing normal random numbers  $e_i$  with the expected value  $En$  and standard deviation  $He$ , namely  $e_i = G(En, He)$ ;

③ Calculating  $\mu_i = \exp[-\frac{(x_i - Ex)^2}{2e_i^2}]$ , and  $(c_i, \mu_i$

) is a cloud drop. According to the above steps, a normal cloud composed of any cloud drops can be produced.

### B. Synthetical cloud

According to the character of human thought, the concept and knowledge level should be elevated to produce a high-level concept according to the structure of the concept level tree [16, 19, 21]. In this paper the synthetical cloud is adopted to optimize the concept.

Suppose there be two atom clouds  $C_1(Ex_1, En_1, He_1)$  and  $C_2(Ex_2, En_2, He_2)$ , and  $Ex_1 \leq Ex_2$ , then the synthetical cloud  $C_3(Ex_3, En_3, He_3)$  is defined as [16, 19, 21]:

$$Ex_3 = (Ex_1 + Ex_2) / 2 + (En_2 - En_1) / 4.$$

$$En_3 = (En_1 + En_2) / 2 + (Ex_2 - Ex_1) / 4.$$

$$He_3 = \max(He_1, He_2). \quad (7)$$

Two low-level concepts can be synthesized by the synthetical cloud method to produce a high-level concept, and each synthetical cloud corresponds to a clustering center.

For the data point  $x_j$  in the attribute space, calculate its membership degree  $\mu_i$  to each synthetical cloud  $C_i$ , make it belong to the synthetical cloud with the maximum membership degree and realize the clustering of the data points.

## IV. LANDSLIDE SUSCEPTIBILITY ANALYSIS IN THE THREE GORGES

In this paper a part of Badong County in the Three Gorges is chosen as the study area in which the landslide disaster is quite frequent. Badong County is primarily composed of the stratum of First Section Badong Group  $T_2b^1$ , Second Section Badong Group  $T_2b^2$ , Third Section Badong Group  $T_2b^3$ , Forth Section Badong Group  $T_2b^4$ , First section Jialingjiang Group  $T_{1j}^1$ , Second section Jialingjiang Group  $T_{1j}^2$ , Third section Jialingjiang Group  $T_{1j}^3$ , Forth section Jialingjiang Group  $T_{1j}^4$  and so on, and the geological structure is quite complicated. In the study area there distribute more than 30 landslides, such as Huangtupo Landslide, Hongshiliang Landslide and Zhujiadian Landslide. ETM+ image with the re-sampling resolution 20m is adopted, which is shown in Fig. 1. The image piled up with the landslide distribution graph is shown in Fig. 2.

The key factors closely related to landslide transmutation include land utilization, slope structure, reservoir water fluctuation, engineering rock group and slope level, which are respectively shown in Fig. 3 – Fig. 7 (with the re-sampling resolution 30m).

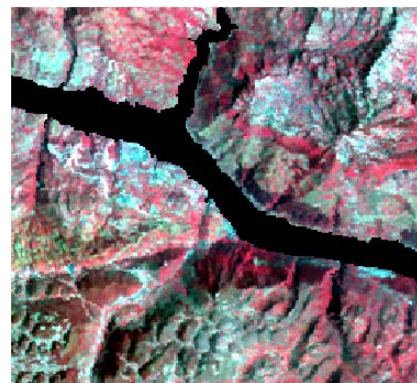


Figure 1. ETM+ image of 432 bands in a part of Badong County.

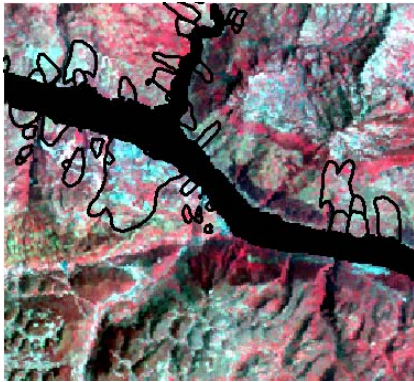


Figure 2. ETM+ image piled up with landslide distribution graph

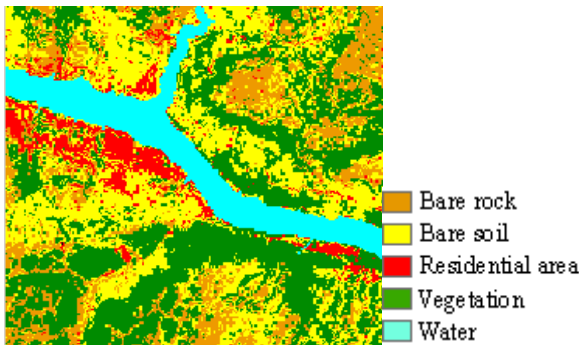


Figure 3. Land utilization.

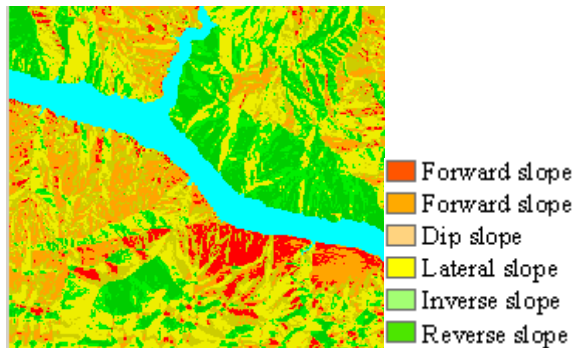


Figure 4. Slope structure.

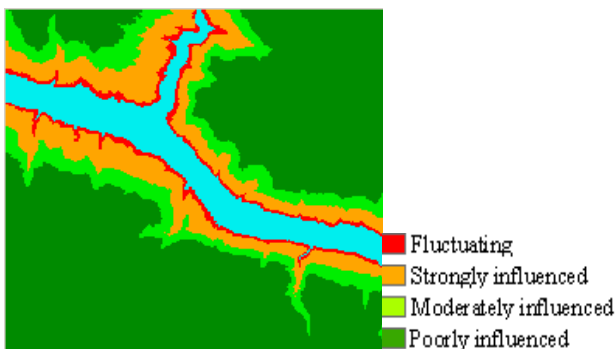


Figure 5. Reservoir water fluctuation.

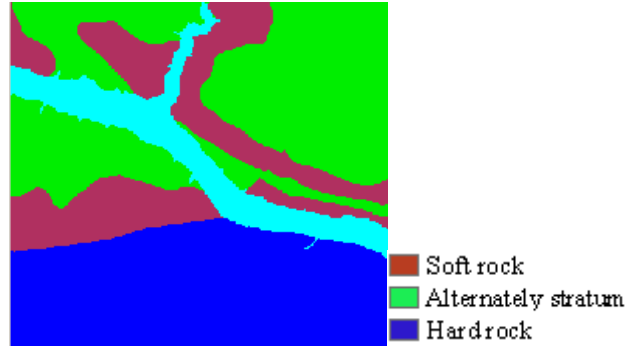


Figure 6. Engineering rock group.

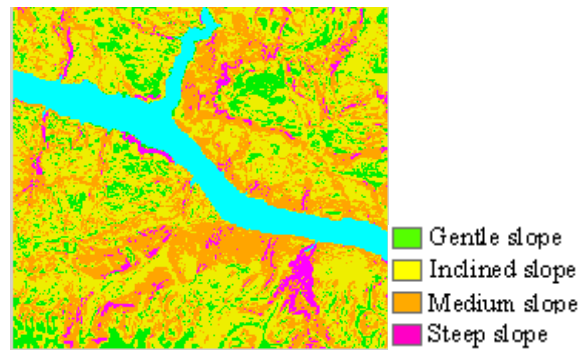


Figure 7. Slope level.

The stability of landslides can be classified into 4 circumstances: dangerous, unstable, basically stable and stable. The potential function is produced by 2650 sample points and is shown in Fig. 8. The potential centers were chosen from the data points which possess the local maximum potential function values and of which the cumulated potential function values  $F > 2750$ . Then 16 potential centers were obtained, which is shown in Fig. 9. By analyzing the characters of the sample points, 5 potential centers were added. The concept represented by each potential center is described by a cloud model. By synthetical cloud method, the concept level is optimized and elevated to produce 12 synthetical clouds, namely 12 clustering centers, which are shown in Fig. 10. In this paper the 3 methods of IsoData, K-Means and Parallelepiped are also adopted to make landslide susceptibility analysis and compared with the method of cloud model and data field, the prediction results are shown in Fig. 11-14. The experimental results have shown that the methods of IsoData and K-Means cannot well recognize the dangerous and unstable regions. The method of Parallelepiped cannot well distinguish the unstable, basically stable and stable regions. It misclassifies most regions into unstable ones. And the method proposed in this paper can distinguish the dangerous, unstable, basically stable and stable regions and the prediction result is prior to the other 3 ones.

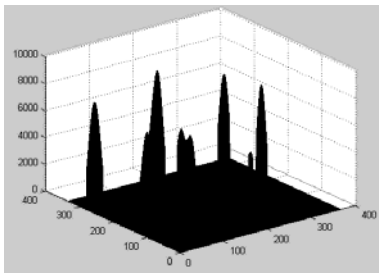


Figure 8. Distribution graph of potential function.

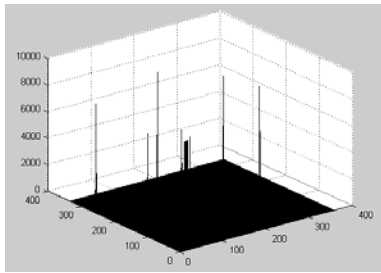


Figure 9. Distribution graph of potential centers.

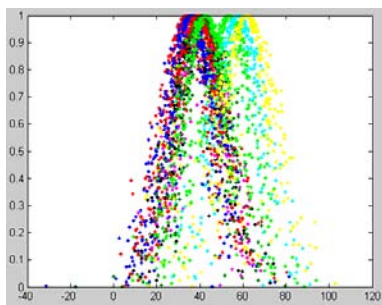


Figure 10. Synthetical cloud graph.

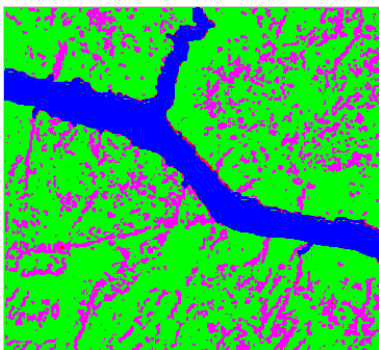


Figure 11. Forecast result of IsoData method

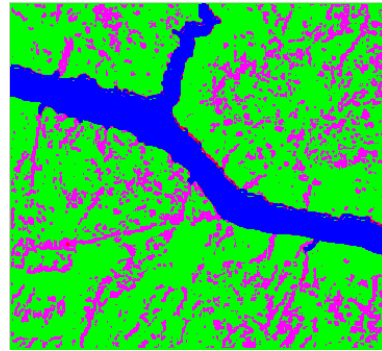


Figure 12. Forecast result of K—Means method.

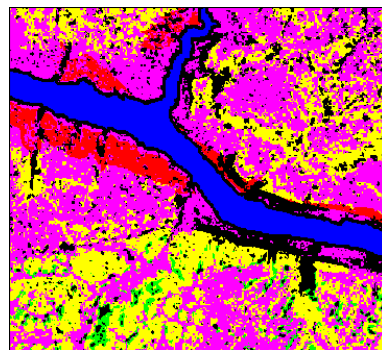


Figure 13. Forecast result of Parallelepiped.

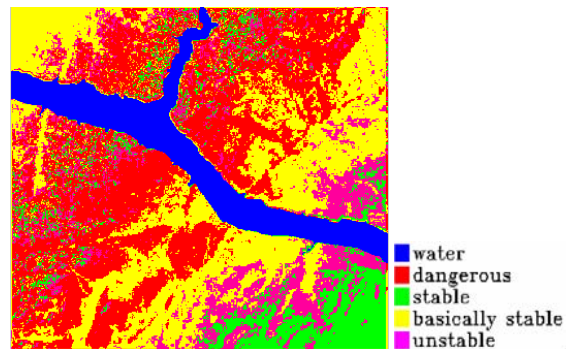


Figure 14. Forecast result of cloud model and data field

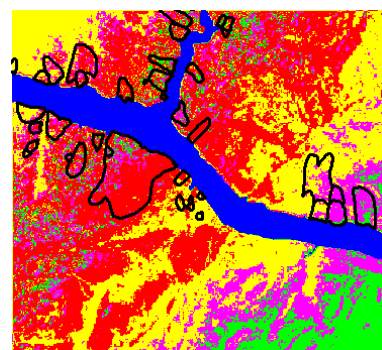


Figure 15. Result piled up with disaster distribution.

The landslide susceptibility analysis result obtained by the method proposed in this paper is piled up with the landslide distribution graph, which is shown in Fig. 15. The piled image has shown that most of the landslides lie in the dangerous and unstable regions, so the method proposed in this paper has obtained good prediction result. The method can synthetically consider the spatial distribution of landslides and the key factors closely related to landslide transmutation, realize the optimization and elevation of the concepts and provide a new idea with landslide intelligent prediction.

### V. CONCLUSION

The Three Gorges are the areas in which the geological environment is very poor and the serious landslide disasters happen frequently. That brings tremendous threat to normal run of the Three Gorge Dam, normal

sluice of the reservoir, and the properties and lives of the residents in the reservoir. So it is a very important task of prevention and cure of landslides in the Three Gorges to make landslide susceptibility analysis.

The traditional studies take little consideration on the uncertain and nonlinear action of a landslide system, lack in the mining and extraction of various complicated information of a landslide system, often need much manual intervention and possess poor intelligence. And the large amount of landslide data doesn't get adequate mining and utilization at all. The introducing of data mining provides a new idea to solve the above problems. However at present, the research on introducing data mining into landslide prediction is rarely few.

In this paper spatial data mining is introduced to make landslide susceptibility analysis and a suitable spatial data mining method for landslide susceptibility is proposed. ETM+ images with the re-sampling resolution 20m, 1:50000 geological map and 1:10000 terrain map were adopted as the data origins to produce the key factors closely related to landslide transmutation, including slope structure, engineering rock group, slope level, fluctuation influence of reservoir water and land exploration. Firstly the method adopts data field to synthetically analyze the spatial distribution of landslides and the key factors closely related to landslide transmutation, and fuse the spatial positions and various attribute values to mine the potential centers. Secondly it based on cloud models describes each concepts represented by each potential center, adopts the synthesized clouds to optimize the concept and elevate the knowledge level, and produces the high-level concepts. Each high-level concept represents a clustering center. Finally the method makes clustering analysis according to the membership of each data point in attribution space to each high-level concept and realizes landslide susceptibility analysis. The prediction result of the method proposed in this paper is compared with the ones of the other 3 methods (IsoData, K-Means and Parallelepiped). The experimental results have shown that the method proposed in this paper can distinguish the dangerous, unstable, basically stable and stable regions and most of the landslides lie in the dangerous and unstable regions. The prediction result is prior to the other 3 ones. Therefore the method proposed in this paper is a suitable data mining method for landslide susceptibility analysis. It can synthetically consider the landslide spatial distribution and the key factors closely related to landslide transmutation, realize the optimization of the concepts and the elevation of knowledge, and provide a new idea with landslide susceptibility analysis.

#### ACKNOWLEDGMENT

The research is funded by National Science Item (40902099), National 863 plan (2007AA12Z160) and Excellent Youthful Teacher Science Fund of China University of Geosciences (CUGQNL0813).

#### REFERENCES

- [1] J. K. Ghosh, D. Bhattacharya. Knowledge-Based Landslide Susceptibility Zonation System. *Journal of Computing in Civil Engineering*, vol. 24, 2010, pp. 325-334.
- [2] S. Lee. Landslide Susceptibility Mapping Using An Artificial Neural Network in the Gangneung Area, Korea. *International Journal of Remote Sensing*, vol. 28, 2007, pp. 4763-4783.
- [3] J. R. Minder, G. H. Roe, D. R. Montgomery. Spatial Patterns of Rainfall and Shallow Landslide Susceptibility. *Water Resources Research*, vol. 45, 2009, pp. w04419-w04430.
- [4] T. Fernández, C. Irigaray, R. El Hamdouni, J. Chacón. Methodology for Landslide Susceptibility Mapping by Means of A GIS. Application to the Contraviesa Area (Granada, Spain). *Natural Hazards*, vol. 30, no. 3, 2003, pp. 297-308.
- [5] Kunlong, Yin; Tongzhen, Yan.. Landslide Forecast and Related Models. *Chinese Journal of Rock Mechanics and Engineering*, 1996, 15, 1-8.
- [6] K. L. Ying. Mechanism and Dynamic Simulation of Landslide by Precipitation. *Geological Science and Technology Information*, vol. 22, no. 1, 2002, pp. 75-78
- [7] W. H. Schulz. Landslide Susceptibility Revealed by LIDAR Imagery and Historical Records, Seattle, Washington. *Engineering Geology*, vol. 89, 2007, pp. 67-87.
- [8] L. Zhu, J. F. Huang. GIS-based Logistic Regression Method for Landslide Susceptibility Mapping in Regional Scale. *Journal of Zhejiang University Science A*, vol. 7, no. 12, 2006, pp. 2007-2017.
- [9] H. B. Havenith, A. Storm, F. Caceres, E. Pirard. Analysis of Landslide Susceptibility in the Suusamy Region, Tien Shan: Statistical and Geotechnical Approach. *Landslides*, vol. 3, no. 1, 2006, pp. 85-96.
- [10] A. Günther, C. Thiel. Combined Rock Slope Stability and Shallow Landslide Susceptibility Assessment of the Jasmund Cliff Area (Rügen Island, Germany). *Natural Hazards and Earth System Sciences*, vol. 9, 2009, pp. 687-698.
- [11] S. Sarkar, D. P. Kanungo. An Integrated Approach for Landslide Susceptibility Mapping Using Remote Sensing And GIS. *Photogrammetric Engineering And Remote Sensing*, vol. 70, no. 5, 2002, pp. 617-625.
- [12] M. Guinau, R. Palla's, J. M. Vilaplana. A Feasible Methodology for Landslide Susceptibility Assessment in Developing Countries: A Case-study of NW Nicaragua after Hurricane Mitch. *Engineering Geology*, vol. 80, 2005, pp. 316-327.
- [13] X. Yao, L. G. Tham, F. C. Dai. Landslide Susceptibility Mapping Based on Support Vector Machine: A Case Study on Natural Slopes of Hong Kong, China. *Geomorphology*, vol. 101, 2008, pp. 572-582.
- [14] D. Caniani, S. Pascale, F. Sdao, A. Sole. Neural Networks and Landslide Susceptibility: A Case Study of the Urban Area of Potenza. *Natural Hazards*, vol. 45, no. 1, 2008, pp. 55-72.
- [15] C. M. Chu, B. W. Tsai, K. T. Chang. Integrating Decision Tree and Spatial Cluster Analysis for Landslide Susceptibility Zonation. *World Academy of Science, Engineering and Technology*, vol. 59, 2009, pp. 479-483.
- [16] H. J. Wang, Y. Deng. Spatial clustering method based on cloud model and data field. *Lecture Notes in Computer Science*, vol. 4683, 2007, pp. 420-427.

- [17] H. J. Wang, Y. Deng, L. Wang. A c-means algorithm based on data field. *Geomatics and Information Science of Wuhan University*, vol. 34, no. 5, 2009, pp. 626-629.
- [18] X. M. Wang, R. Q. Niu. Spatial forecast of landslides in Three Gorges based on spatial data mining. *Sensors*, vol. 9, no. 3, 2009, pp. 2035-2061.
- [19] D. Y. Li, K. C. Di, D. R. Li, X. M. Shi. Mining association rules with linguistic cloud models. *Journal of Software*, vol. 11, 2000, pp. 143-158.
- [20] F. R. Meng, C. J. Song, Z. P. Zheng, S. X. Xia. Mining the association rules of coal mine security monitoring data based on the cloud theory. *Journal of Chinese Computer Systems*, vol. 29, no. 9, 2008, pp. 1622-1626.
- [21] K. C. Di, D. Y. Li, D. R. Li. Cloud theory and its application in spatial data mining and knowledge discovery. *Journal of Image and Graphics*, vol. 4, no. 11, 1999, pp. 930-935.

**Xianmin Wang** was born in Fuzhou, China in 1978. Wang received a BS degree and a DS degree from Wuhan University respectively in 2001 and 2005.

She is now working as a teacher in China University of Geosciences. Her research has been supported by the National Science Foundation and National 863 Planning Foundation. She is currently interested in spatial data mining, geological disaster detection and forecast.

**Ruiqing Niu** was born in Nanyang, China in 1969. Niu received a DS degree in the field of earth exploration and information technique from China University of Geosciences in 2005.

He is now an associate professor in China University of Geosciences. His research has been supported by the National Science Foundation and National 863 Planning Foundation. He is currently interested in earth exploration, remote sensing geology and geological disaster detection and forecast.

Dr. Niu is the superintendent of Institution of Earth Spatial Information and the master of Department of Earth Information Science and Technique in China University of Geosciences, China.