

Assessment of Geological Hazards in Ningde Based on Hybrid Intelligent Algorithm

Shiliang Zhang,^{1*} Yuxia Wang,² and Tingcheng Chang¹

¹Department of Computer Science, Ningde Normal University, Fujian 352100, China

²Institute of Remote Sensing and Geographic Information Systems, Peking University, Beijing 100087, China

(Received August 21, 2017; accepted November 14, 2017)

Keywords: geological hazard, neural network, hazards assessment, WNN

In recent years, geological hazards have occasionally occurred throughout the world, and have caused immense damage to roads and lives in places where landslides have occurred. Thus, it is of great importance to coordinate the work of regional hazard prevention and reduction. Under these circumstances, hybrid intelligent algorithm (HIA) combined with genetic algorithm (GA) and wavelet neural network (WNN) is proposed with the geological risk assessment analysis in our study. The HIA integrated both the geographic information system (GIS) technology and the artificial neural network model. In the HIA, GA is adopted to initialize the network connection weights and thresholds of WNN. Moreover, in simulations, measurement-obtained data of geological hazards were collected by a geological environment monitoring station and statistic data which were extracted from the map and statistics text data from Ningde City in eastern China. The proposed HIA provides us with increased accuracy compared with established methods using traditional back propagation (BP) neural networks. Our result is of great importance for regional geological hazard prevention, land resources rational development, and proper protection of the geological environment.

1. Introduction

Geological disasters always affect the development of human society and economic progress.⁽¹⁾ Landslides are a serious natural disaster next only to earthquakes and floods, and are 50–90% of the total number of geological disasters each year.⁽²⁾ Thus, landslide recognition is very important for disaster prevention, monitoring, and other applications.⁽³⁾ In recent years, theories and methods have been developed in this field to predict and analyze hazards more accurately, which involve earthquakes, landslides, mudslides, and other geohazards. The major findings of these geological hazard recognition (GHR) analyses and assessments have become the principal guidelines of hazard mitigation and rescue implementation.⁽⁴⁾ At present, GHR assessments in most countries have adopted the following methods: expert system,^(5,6) multivariate statistical method,⁽⁷⁾ geostatistics,⁽⁸⁾ time series analysis,^(9,10) fuzzy comprehensive evaluation method,⁽¹¹⁾ grey system theory,^(12,13) analytic hierarchy process,⁽¹⁴⁾ gravity and magnetic methods,⁽¹⁵⁾ back

*Corresponding author: e-mail: shiliangzh@163.com
<http://dx.doi.org/10.18494/SAM.2018.1770>

propagation (BP) neural network,⁽¹⁶⁾ and a few other approaches. However, these methods still have some imperfections. A novel approach known as hybrid intelligent algorithm (HIA) is proposed in this paper. On the basis of the advantages of the wavelet neural network (WNN) algorithm's solution to nonlinear problems, the genetic algorithm (GA) is employed to calculate initial connection weights and thresholds. Thus, HIA can avoid falling into a local optimal solution and improve its convergence rate and obtain more accurate results. In the second section we introduce GA, WNN, and the HIA based on the two mentioned methods; in the third section we introduce the data sources and pretreatment; in the fourth section we show the establishment of the experiments and the analysis results based on HIA method, and finally in the fifth section we summarize this paper.

2. Materials and Methods

2.1 GA

GA is a stochastic search algorithm based on the theory of natural evolution,⁽¹⁷⁾ including the process of selection, crossover, and mutation. When GA is used for solving optimization problems, a set of possible solutions called populations are usually obtained. Each possible solution is treated as a chromosome or an individual in the population. Multiple individuals together form a solution space set, or population space. Randomly selected individuals in the population space are used as the initial solution of the target optimization problem, which will be either kept or eliminated according to their degree of adaptation in the population, i.e., whether they are in compliance with the principle of survival of the fittest in the evolutionary process. The good individuals will be retained until the next generation while the bad ones will be eliminated. After crossover and mutation procedures, all the outstanding individuals will form the next population generation. After continuous iterative optimization, the global optimal solution of the target problem is finally obtained.

2.2 WNN

WNN is a combination of wavelet analysis theory and neural network theory,⁽¹⁸⁾ which takes the wavelet base function as the transfer function of the hidden layer node, and the neural network with the error of the signal forward propagation.

In Fig. 1, x_1, x_2, \dots, x_M are the input parameters of the WNN, and y_1, y_2, \dots, y_N are the predictive outputs of the WNN. w_{ij} is the connection weights between the input layer and the hidden layer. When the input signal sequence is x_i ($i = 1, 2, \dots, M$), the output of the hidden layer can be expressed as

$$h(j) = h_j \left(\frac{\sum_{i=1}^M w_{ij} x_i - b_j}{a_j} \right), \quad j = 1, 2, \dots, N, \quad (1)$$

where $h(j)$ is the output of the j th hidden layer node, w_{ij} is the connection weight between the input layer and the hidden layer, a_j is the scaling factor of the wavelet basis function h_j , and b_j

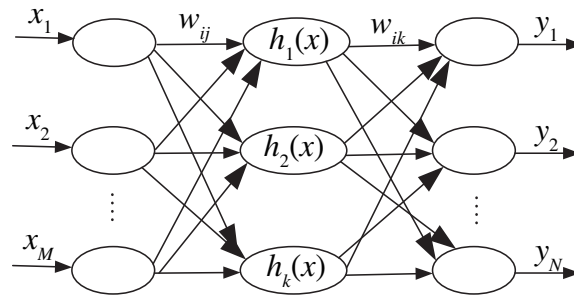


Fig. 1. Wavelet neural network topology.

is the translation factor h_j . In this paper, the Morlet wavelet basis function is employed as the transfer function of hidden layer nodes, namely

$$y = \cos(1.75x)e^{-x^2/2}. \quad (2)$$

The output of the WNN can be expressed as

$$y(k) = \sum_{i=1}^l w_{ik}h(i), \quad j = 1, 2, \dots, N, \quad (3)$$

where w_{ik} denotes the hidden layer to output layer weights, $h(i)$ is the output of the i th hidden layer nodes, l represents the number of hidden layer nodes and N is the number of output layer nodes.

2.3 Optimization of WNN parameters with GA

The weight correction method of traditional WNN is similar to the weight optimization method of the BP neural network, i.e., the error back propagation method. However, this method is prone to have slow convergence and local optimum drawbacks, thereby resulting in low prediction accuracy. In this paper, we use the GA to optimize the initial parameters of the WNN. In the search process, the GA can maintain the diversity of population, and more likely find the global optimal solution. The parameters to be optimized include the translation factor b_j of the WNN, the scaling factor a_j , and the connection weights w_{ij} . In the population $X_i = \{x_{i1}, x_{i2}, \dots, x_{iT}\}$, individual x_{ij} is expressed as $x_{ij} = (a_1, a_2, \dots, a_k, b_1, b_2, \dots, b_k, w_1, w_2, \dots, w_k)$. Each individual represents a possible solution of the solution space. We optimize the steps as follows:

Step 1: determination of WNN structure: including the input layer nodes, the hidden layer nodes, and the output layer nodes, etc.;

Step 2: population initialization: including population size, the individual initialization, target error, etc.;

- Step 3: population classification: according to the fuzzy clustering method, the initial population is classified into k clustering centers c_k ;
- Step 4: similar individual selection: the degree of similarity between individuals in the same class is determined in accordance with Eq. (4). Only one individual with the highest degree of similarity is retained and genetically passed to the next generation while the rest of the individuals are removed;
- Step 5: the individuals in the same population are selected according to the ranking of the fitness values;
- Step 6: determine whether the population size meets the requirements. If yes, parallel genetic operations such as crossover and mutation between individuals from different classes are performed; otherwise, return to step 5;
- Step 7: determine whether or not the set target error has been achieved, if not, return to step 3;
- Step 8: online learning: the optimal parameters b_j , a_j , and w_{ij} are substituted into the WNN to conduct network training;
- Step 9: network testing: the trained WNN is used to forecast the traffic flow, and the prediction results are analyzed.

3. Data Sources and Pretreatment

3.1 Study area

Ningde City is formed mostly of hills, with many mountains and a few plains. It has a hot and humid climate with heavy rainfall. Moreover, it has a complex geological structure, intense rock erosion, significant changes in topography and widespread construction of slope cutting. With the influence of lithology, geological structure, rainfalls, human activities, and many other factors, geological hazards are prone to occur in Ningde. According to a 2010 geological disaster investigation, there are 147 potential risks for geological disasters, including 64 landslides, 39 geological collapses, nine mudslides, and five risky slopes and ground subsidences. The types of geological disasters include landslides, avalanches, unstable slopes, and mudslides, with landslides and collapses accounting for about 90% of the total disasters. In our study, the research area is divided into 5957 cells; each cell has a size of $30 \times 30 \text{ m}^2$ (Fig. 2).

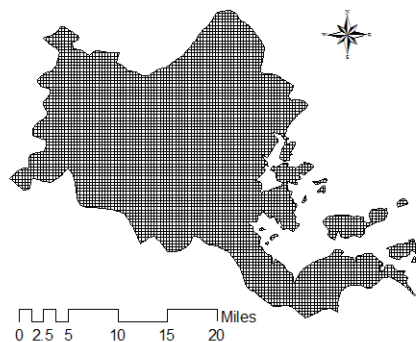


Fig. 2. Fishnet of research area.

3.2 Data sources and pretreatment

Landslide debris flow is a product of the interactions of many factors. Through a lot of research and field investigations in the past, in this paper, we selected several kinds of influential factors: soil vegetation, lithology, elevation, slope, distance from the river, distance from the road, land use, human activities, rainfall, basic physical parameters of soil, deformation, and other factors as the hazard assessment factors of the research area, and used these influential factors to generate the risk evaluation factor data. Moreover the evaluation factors are divided into two categories: (1) statistic data factors, including geological structure, topography, natural factors, and human activities; and (2) measurement-obtained data factors, including rainfall, basic physical parameters of soil, and deformation. More detailed factors are shown in Table 1, including rainfall, channel flow velocity, soil water content, groundwater level, soil pore water pressure, soil pressure, deep displacement, and surface displacement of soil.

3.2.1 Measurement-obtained data extraction

According to the historical disaster distribution points, 147 monitoring points of geological hazards in the research area are established to collect measurement-obtained data, including rainfall, basic physical parameters of soil, and deformation. Because of the location, traffic and power supply of the monitoring point, we decided to use wireless intelligent sensors to monitor the geological hazards dynamically. From these wireless sensors to data terminals, we set up an integration data center to receive data from data terminals, and the 3G and general packet radio service (GPRS) communication networks together formed the research area of the geological hazard monitoring network (Fig. 3). Terminal intelligent equipment includes water level meter, displacement meter, rainfall meter, water temperature meter, soil moisture content meter, groundwater level meter, and so forth. The collected data are stored and transmitted through wireless GPRS to the server for further processing.

Table 1
Acquisition of measurement-obtained geological hazard factors.

| Monitoring project | Items | Methods/tools | Transmission |
|-----------------------------------|--------------------------|--------------------------------|---------------------------|
| Rainfall | Rainfall intensity | Automatic rainfall station | GPRS |
| | Channel flow | Flow meter | GPRS |
| Basic physical parameters of soil | Moisture rate | Soil moisture sensor | GPRS |
| | Pore pressure | Ohmmeter | GPRS |
| | Matrix suction | Soil water potential sensor | GPRS |
| | Groundwater level | Groundwater level apparatus | GPRS |
| | Soil pressure | Soil pressure sensor componets | GPRS |
| Deformation | Internal deformation | Fixed inclinometer | GPRS |
| | Surface deformation | Crack displacement instrumet | GPRS |
| Images | Video monitoring station | PTZ camera | GPRS Automatic storage |

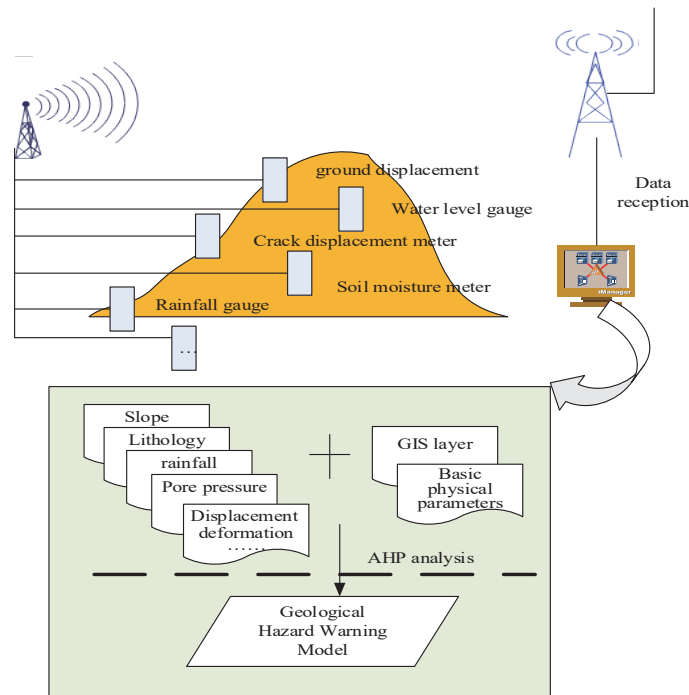


Fig. 3. (Color online) Design topology of the system.

The sensor was selected and calibrated properly, then it was buried at a specific point. For example, a surface displacement meter should be installed at a point where the surface is cracked or might crack in the future (Fig. 4).

As shown in Fig. 5, the fixed inclinometer (also called a slanting probe) is the real-time displacement monitoring instrument for the measurement of subgrade, slope, and other underground deep deformations.

3.2.2 Statistic data extraction

Geographic information system (GIS) technology can extract spatial data quickly and accurately, and establish the spatial data for evaluation factors. In this study, the main statistic data source for the map data and text statistics are based on the ArcGIS platform to extract the spatial data. First of all, raster data and text data should be converted into vector data, because GIS is based on grid cell computing. Therefore, after the completion of each evaluation factor layer extraction, it is necessary to rasterize the factor layer and the geologic hazard distribution map, and then perform the operation of the grid layer. The specific extraction processes are as follows: (1) The study area 1:50000 geological map, land use status map in 2010 and 1:50000 topographic map are registered and projected, and then ArcGIS software is used to transfer the geological disaster survey text data into the same coordinate system. After the interactive data, the following vector data are obtained: lithology map, linear structure diagram, river water system map, historical disaster point data distribution map, land use status map, and

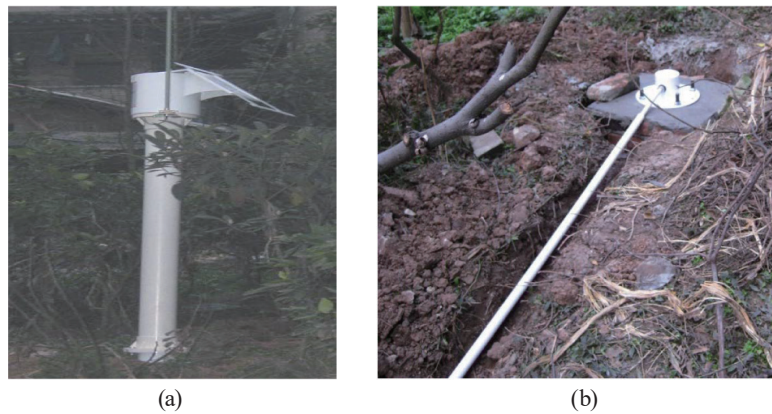


Fig. 4. (Color online) Installation method of (a) displacement meter and (b) soil moisture meter.

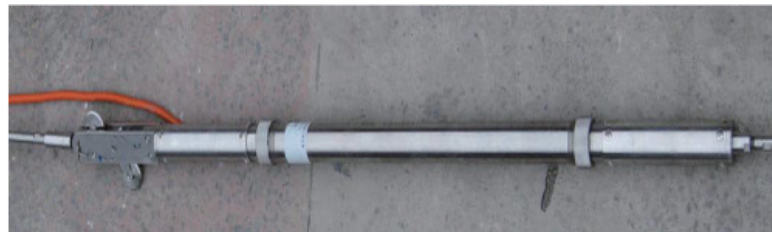


Fig. 5. (Color online) Fixed inclinometer.

topographic map. (2) Using the ground statistic module in ArcGIS, the linear structure isoline map, the contour map of river network density and the contour map of historical geologic disaster distribution density are generated. (3) According to the evaluation criteria, the evaluation factors are graded, and each factor layer is converted into the raster data of $30 \times 30 \text{ m}^2$, and the factor evaluation layer is generated (Fig. 6).

3.3 Data pretreatment

In this paper, we determined the weight of the statistic data factors with the analytic hierarchy process (AHP) algorithm in Ref. 19 to generate four factors as network input data.⁽¹⁹⁾ The measurement-obtained data factors use the sum method to generate three factors as network input data. The research area is divided into 5957 cells. Among them, 147 cells of measurement-obtained data are obtained from the 147 monitoring points, and the remainder are obtained by interpolation.

However, owing to the different dimensions and units of each factor, the predicted results will be impacted. In the process of network learning, the activation function of neurons is a bounded function, and the effects of different dimensions and units of each factor shall be eliminated to prevent some neurons from reaching the saturation state, while the larger input shall be located in the region with a large gradient of neuron activation function. Therefore,

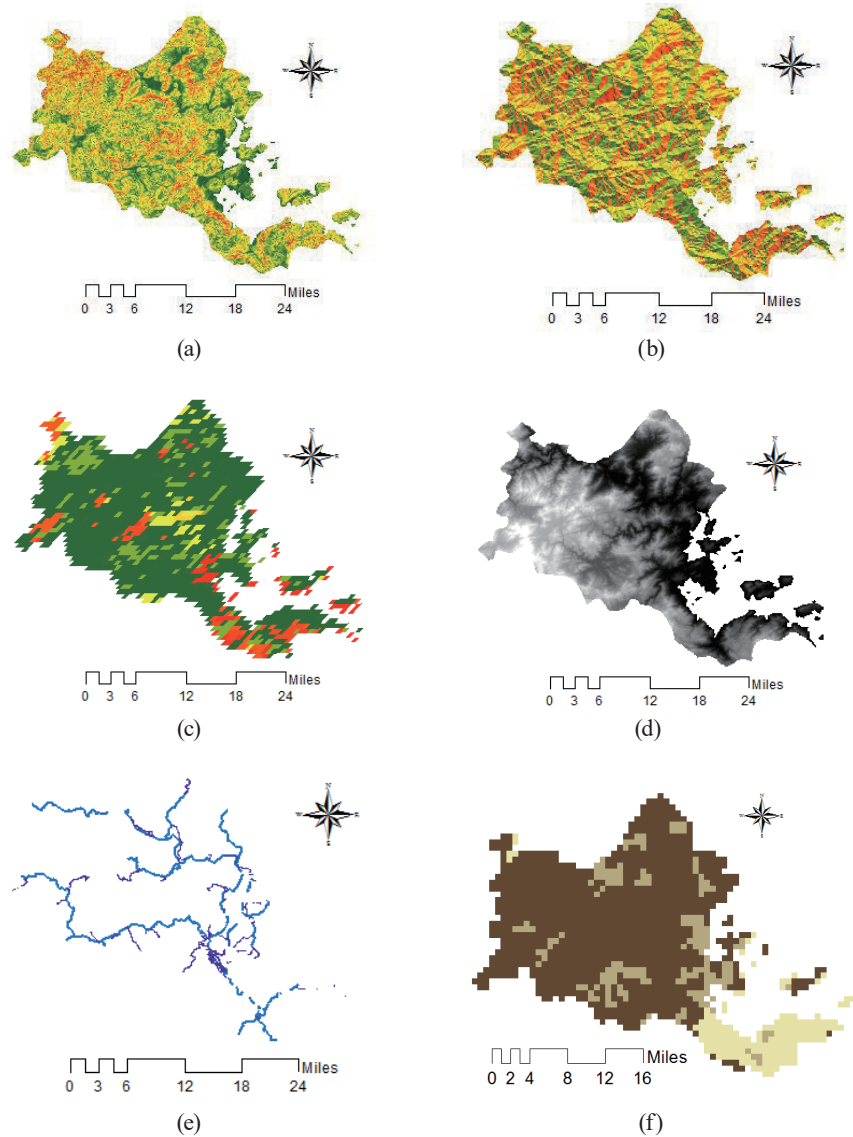


Fig. 6. (Color online) Distribution of some attributes. (a) Distribution of slope, (b) distribution of aspect, (c) distribution of vegetation cover, (d) distribution of DEM, (e) distribution of road, and (f) distribution of soil type.

before training and forecasting, the input vectors shall be normalized by Eq. (4) and the raw data shall be processed as data between 0.1 and 0.9:

$$x'_i = 0.1 + (0.9 - 0.1) \times \frac{x_i - x_{imin}}{x_{imax} - x_{imin}}. \quad (4)$$

Here, x_i represents the x'_i value of the input neurons before pretreatment, and represents the value of the input neurons after pretreatment; x_{imax} denotes the maximum value of each neuron i in the input units, and x_{imin} denotes the minimum value of each neuron i in the input units.

4. Analysis with HIA

Firstly, 110 cells of data are selected randomly from 147 cells of samples as the training samples and the remaining 37 cells of data are taken as test samples. All the data are normalized because of their different dimensions.

On the basis of the measurement-obtained data and statistic data, we design a WNN with three-layer structures, including the input layer, hidden layer and output layer. The Morlet mother wavelet function is used as the transfer function of the hidden layer nodes; the WNN structure used in this paper is 7-8-4. Specifically, the seven input layer nodes represent four kinds of statistic data and three kinds of measurement-obtained data; the eight hidden layer nodes and the output layer are set for four neurons y_1 – y_4 , which respectively correspond to geological hazard risks in the four levels, namely, high-risk zones, medium-risk zones, low-risk zones, and no-risk zones. In order to facilitate the calculation, the encoding process uses four-digit encoding, which is composed of 0 and 1, namely, high-risk, medium-risk, low-risk, and no-risk zones, which correspond to 1000, 0100, 0010, and 0001, respectively. The initial parameters of WNN are the optimal values optimized with GA. The GA considers the initial parameters of WNN as population, where the population size is 100, the number of evolution generation required for termination is 100, the crossover probability is 0.8, and the mutation probability is 0.1.

To better illustrate the effectiveness of the HIA, we build three kinds of prediction model and compare them with the parameters of the mean absolute error (MAE) and the mean relative error (MRE). The three models assessed geological hazard based on BP, WNN, and HIA. In the experiment, Table 2 shows the results of MAE and MRE, which are obtained from these different models.

It can be seen from Table 2 that the MAE and MRE obtained with HIA are lower than those obtained from BP and WNN. Thus, the HIA model has achieved satisfactory results, which are closer to the field data than those obtained with the other two models. By the same method, we applied the trained network of HIA to obtain the output layer values of the other 5810 cells and they are shown in Fig. 7.

According to the weighted comprehensive evaluation model, a Ningde geological hazard zoning diagram is compiled based on ArcGIS (Fig. 7). By statistical analysis, the area of the high-risk zone is 19% of the total area, and the medium-risk zones for geological disasters are relatively widely distributed, and the area is 37% of the total area. The medium-risk zones and the high-risk zones are mainly located on the tectonic belt of the Zhejiang–Fujian uplift area, which is the transition of the Alpine region and the coastal area of medium and low mountainous areas. Widespread destruction of the mountain accumulated a large number of loose deposits. Some mountains have cracks of up to hundreds of meters, which may induce new geological hazards. These zones show the characteristics of distribution along the high and medium-high risk zones. They are mainly distributed in mountain areas with steeper terrain, with some of them distributed along the secondary fault zones. On the other hand, the low and medium-low risk zones are distributed most widely, and the area is 44% of the total area. They are mainly distributed in medium sloped areas, vegetation covered lands, and rocky areas.

Table 2
Comparison of training and testing with MAE and MRE in different models.

| Model | Training | | Testing | |
|-------|----------|--------|---------|--------|
| | MAE | MRE | MAE | MRE |
| BP | 21.007 | 0.3068 | 22.256 | 0.3766 |
| WNN | 16.028 | 0.2893 | 17.112 | 0.3321 |
| HIA | 12.436 | 0.2133 | 12.823 | 0.2491 |

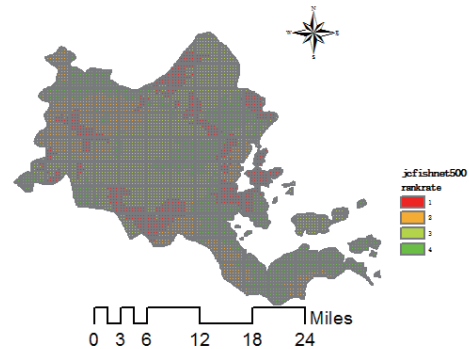


Fig. 7. (Color online) Sensitivity levels of HIA.

The zones more prone to geological disasters are mainly located in the southeast of the Jiaonan-Jiaobei King Temple-Zhangwan-Shangtang-Xiatang-Midwest Hordong area, Chixi, Yangzhong, outside of Hubei basin and Huangjiadun-Hulanli, the northeast corner of the Linyangtou-Wushan. Its area is about 285 km². The zones prone to geological disasters are mainly distributed in the middle east of Jiaocheng District and Sandu area of 456 km². The zones less prone to geological hazards are mainly distributed in the northwest and southeast of Jiaocheng District. Its area is about 550 km². Geological disasters do not easily occur in the west, sparsely populated middle, low mountainous areas, and low plains. Its area is about 190 km², accounting for about 13% of the land area.

5. Conclusions

To overcome the slow convergence speed and local optimal problems of the WNN, we proposed the HIA model, which uses GA to optimize the initial parameters of WNN. The proposed model overcomes the unidirectional optimization, relatively low predication accuracy and other shortcomings of the gradient correction method, which uses traditional WNN to optimize the initial parameters. Our experimental results show that the prediction accuracy of HIA is significantly higher than that of simple BP and WNN. Therefore, it is feasible to apply HIA to geological hazard risk prediction.

A hybrid model of prediction for geological hazard is developed. Within the model, the prediction was performed using different geological and geophysical methods. The advantages of each method are combined, while the disadvantages are avoided. Our model is characterized by the emphasis on geological analysis, and the points of geological hazards are the key points for prediction. As shown in these case studies, this model is an effective method of predicting geological hazards. Because the application of the prediction model requires professional experience of geology and geophysics, efforts should be made to develop an expert system in the future.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 41271386), the Science and Technology Project of Education Department of Fujian Province of China (No. JAT170652), and School Innovation Team of Ningde Normal University (2017T05). The data set is provided by Data Center for Resources and Environmental Sciences, Chinese Academy of Sciences (<http://www.resdc.cn>). The authors would like to appreciate the valuable comments and suggestions from the editors and reviewers.

References

- 1 J. S. M. Fowze, D. T. Bergado, S. Soralump, P. Voottipreux, and M. Dechasakulsom: *Geotext. Geomembr.* **30** (2012) 50.
- 2 R. M. Iverson, D. L. George, K. Allstadt, M. E. Reid, and B. D. Collins: *Earth Planet. Sci. Lett.* **412** (2015) 197.
- 3 B. Punsly: *CATENA*. **96** (2012) 28.
- 4 R. Hadji, A. E. Boumazbeur, Y. Limani, and M. Baghem: *Quat. Int.* **302** (2013) 224.
- 5 M. V. D. Eeckhaut and J. Hervás: *Geomorphology* **139** (2012) 545.
- 6 F. Gutiérrez, R. Linares, C. Roqué, and M. Zarroca: *Geomorphology* **241** (2015) 224.
- 7 H. Hong, B. Pradhan, C. Xu, and D. T. Bui: *CATENA* **133** (2015) 266.
- 8 F. Lucà, D. D'Ambrosio, and G. Robustelli: *Comput. Geosci.* **67** (2014) 163.
- 9 Z. Wang and S. Zhengcha: *Electron. J. Geotech. Eng.* **18** (2013) 5481.
- 10 A. K. Mishra, S. Deep, and A. Choudhary: *EJRS* **18** (2015) 181.
- 11 G. Akkaya, B. Turanoğlu, and S. Öztaş: *Expert Syst. Appl.* **42** (2015) 9565.
- 12 P. Chaudhary, S. K. Chhetri, K. M. Joshi, B. M. Shrestha, and P. Kayastha: *Socio-Economic Planning Sci.* **53** (2016) 60.
- 13 A. Dehnavi, I. N. Aghdam, and B. Pradhan: *CATENA* **135** (2015) 122.
- 14 C. Y. Lu, C. L. Tang, and Y. C. Chan: *Eng. Geol.* **183** (2014) 14.
- 15 N. V. Kuznetsov, G. A. Leonov, M. V. Yuldashev, and R. V. Yuldashev: *Commun. Nonlinear Sci. Numer. Simul.* **51** (2017) 39.
- 16 S. L. Zhang: *IJASM* **2** (2015) 207.
- 17 N. S. Bajestani, A. V. Kamyad, E. N. Esfahani, and A. Zare: *Biocybern. Biomed. Eng.* **37** (2017) 281.
- 18 M. J. Alizadeh and M. R. Kavianpour: *Mar. Pollut. Bull.* **98** (2015) 171.
- 19 T. Euler, S. Haverkamp, T. Schubert, and T. Baden: *Nat. Rev. Neurosci.* **15** (2014) 507.