Nature Environment and Pollution Technology
An International Quarterly Scientific Journal
p-ISSN: 0972-6268
e-ISSN: 2395-3454
Vol. 18 No. 5 (Special Issue) pp. 1647-1654 2019
Original Research Paper Open Access

Study on Mathematical Modelling of Geological Hazard Assessment

Bin Zhou*, Dan He*† and Rodeano Roslee**
*Hunan Institute of Technology, Hunan 421002, China
**Faculty of Science and Natural resources, University Malaysia Sabah, Malaysia
†Corresponding author: Dan He

Nat. Env. & Poll. Tech.
Website: www.neptjournal.com
Received: 19-08-2019
Accepted: 12-10-2019

Key Words:
Geological disaster
Spatial database
Data model
Mathematical model
Zhijin county

ABSTRACT

Zhijin county is one of the areas with fragile geological environment and frequent natural disasters in China. In recent years, with the intensification of human activities, the geological environment has been continuously damaged, causing geological disasters of different scales and serious life threats and property losses to people. Therefore, the purpose of introducing this mathematical model is to make the data more intelligent and professional by combining the data of geological disasters in mathematics. Based on the geological disaster monitoring, warning and decision support system of Guizhou province, this paper takes Zhijin county as the research object and uses Oracle database to manage the data by building and designing data structure and data model. The results show that geological hazard assessment can not only provide reliable data for analysis and research, but also provide reference value for similar database construction: (1) to construct spatial data model to realize effective correlation between spatial data and business data. (2) spatial entity objects have the characteristics of multi-dimension and polysemy, and "relational-object" mathematical modelling is adopted.

INTRODUCTION

Guizhou is a plateau slope zone, east to the department with more low mountain hills, has a strong karst terrain, there always erosion along the river in the direction of cutting, surfaces are broken seriously, residues of mountains, deep canyons, topographical features visible at any time. Guizhou has a unique karst geological environment, strong karst action; it has widely distributed variety of karst landforms, a variety of karst engineering geological problems and karst geological disasters. Because of these special geological features and geological structure, it has become a national key geological disaster management area. Zhijin county is located in Bijie area of Guizhou province, in the middle of the plateau, where the Liuchong river in the north source meets the Yanchi river in the upper reaches of the Wujiang river system in the south source mountain. Due to the relatively poor geographical conditions, geological structure and topography and other conditions are quite complex, located in a very fragile environment. With the rapid development of economy, the lack of attention to the environment in the process of economic construction and resource development, has led to frequent geological disasters and accidents. The scale and scope of disaster have been continuously expanded, causing great losses and inconvenience to local people and great losses to national property. The research object of this paper is Zhijin county, Guizhou province, where the disaster area is relatively wide and the frequency of geological disasters is high, and the people are seriously threatened. According to the survey, as of December 2009, the county has a total of 211 potential geological disasters, including 3 debris flows, 43 collapses, 80 ground cracks and 81 landslides. According to preliminary statistics, by 2009, 6 people had died, over 3,000 houses had been destroyed, over 4,000 mu of farmland had been destroyed, and the direct economic loss reached 98 million yuan. In addition, there are many potential disaster spots in the region, including houses, schools, bridges, roads, farmland and so on. Threatened economic losses of up to 200 million yuan.

This paper takes the key scientific research project of Guizhou province "decision-making support system and geological disaster monitoring and warning" as the background, and takes "the construction of geological disaster database of Zhijin county" as the theme, discusses the construction process and application of geological disaster database, hoping to provide reference for the prediction and prevention of geological disasters in the future.

EARLIER RESEARCH

Mathematical modelling is the process of how to apply mathematics to real life (Sivakumar & Ghosh 2017). Mathematical modelling technology provides a basic evaluation mechanism for people to solve problems on the basis of collecting information and mining rules (Ismail-Zadeh et al. 2017). ArcGIS

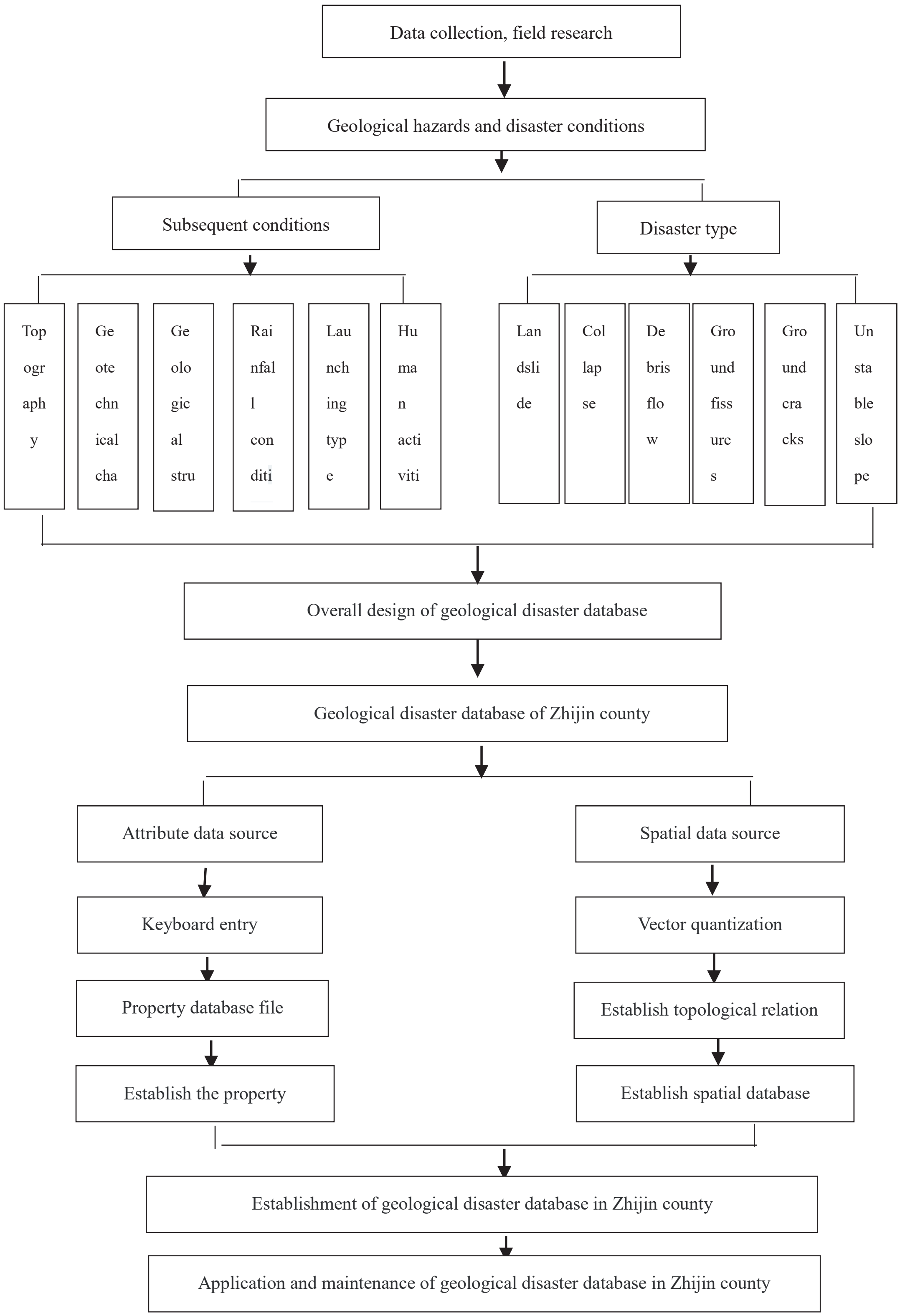

Fig. 1: Technical route.

spatial analysis and MATLAB mathematical modelling functions to build GIS-based GA-BP model (that is, genetic algorithm optimized BP neural network) to conduct nonlinear quantitative evaluation of the geological ecological environment in the research area (Tao et al. 2017). Specifically, a previous study adopted construction operation results and efficiency ratio to construct a scientific index system, combined subjective and objective evaluation methods, and carried out mathematical modelling and simulation calculation according to index weight and characteristic preference (Yepes et al. 2018). A study established a variable weight model based on incentive punishment and calculated the variable weight of each factor. On this basis, the EEGPs partition mathematical model is established (Echavarribravo et al. 2017). Karamouz (2017) used the performance evaluation model (k-pam) of Korea Atomic Energy Research Institute (KAERI) to carry out a new risk-based safety assessment method for the geological disposal system of nuclear waste. A recent study combined the weighting method with the conventional cloud model and proposed a new water inrush evaluation method. In the evaluation of geological hazards, it is necessary to use artificial intelligence to accurately evaluate severe geological hazards (Manea et al. 2017). A research proposed a risk assessment method for geological disasters. By calculating the threat degree of geological hazards and the vulnerability of the subject affected by the hazards, the risk of geological hazards in counties and regions was evaluated by qualitative comprehensive evaluation (Filatov et al. 2019). Specifically, the evaluation index system is constructed and each index is quantitatively divided into four grades (Furlan et al. 2018). In addition, the discipline of geostatistics provides a variety of quantitative methods for the establishment of random models that conform to the measured data and spatial correlation (Xiao et al. 2018).

MATERIALS AND METHODS

With the rapid development of science and technology and the increasing amount of data, how to store and manage these data well becomes a question. Data types are also diverse, not only numbers, pictures, graphics and code. So, the application needs to have a high compatibility, different data types are stored and effectively combined, can be in a large number of data efficient, fast query their needs. Due to the large and diverse geological disaster data, how to process and manage this batch of data quickly and efficiently has become the primary problem of how to make use of geological disaster data. According to the national standards, industry standards and the actual application requirements, complete database design and construction of Zhijin county geological disaster, the data to make use of the Oracle database management, is a good way to manage database of Zhijin county geological disaster, to better manage the geological disaster data, by constructing and designing the data structure, data model to manage the geological disaster data, technical route as shown in Fig. 1.

RESULTS AND DISCUSSION

Data model design: Database is a place to store data. By formalizing data to describe the logical structure and operation of data, it is more convenient for users to process spatial data. The database model designer describes and expresses views through the Moaishu7 data logical relationships. In general, the data model consists of four parts: data structure, operation set and integrity rule set, as shown in Fig. 2:

The development of data model mainly goes through three stages: hierarchical model, mesh model and relational model. Moreover, the establishment of data model needs a high clear concept, mathematical basis and independent data as the foundation. Therefore, for the general database it is difficult to extend, cannot do too much operation, also cannot provide a good index mechanism between the ground and space data. The integration of spatial data was also very difficult and the performance was poor, so people started to try it with object orientation.

But, at present the object oriented database is not mature enough, cannot compare with general theoretical basis for solid data model, but it has some common data types and features, such as it has a certain scalability, it also has a certain advantage in data query, so the GIS software integration data also begin ripping into objects to relational databases, such as MapInfo Spatial Ware, ESRI SDE, etc.

***Conceptual model design*:** In other words, the conceptual data model, called information model, is developed and expanded based on the entity-relationship (E-R) theory. This model is mainly used for the conceptual design of data modelling.

When people do conceptual design for data modelling, the first step is to transform the real world into a conceptual world, and the second step is to transform the conceptual world into a machine world. It can also be understood that objective things can be analysed as Entity first. For the so-called CDM model, it does not need specific requirements for the computer system and DBNS system. The physical data model is the data model that CDM transforms into that a DBMS can support on the computer, namely CDM (Fig. 3).

CDM is a set of strictly defined model elements, through which the static and dynamic characteristics and constraints of the system can be accurately described. CDM mainly consists of three parts: data operation, data structure and integrity constraints.

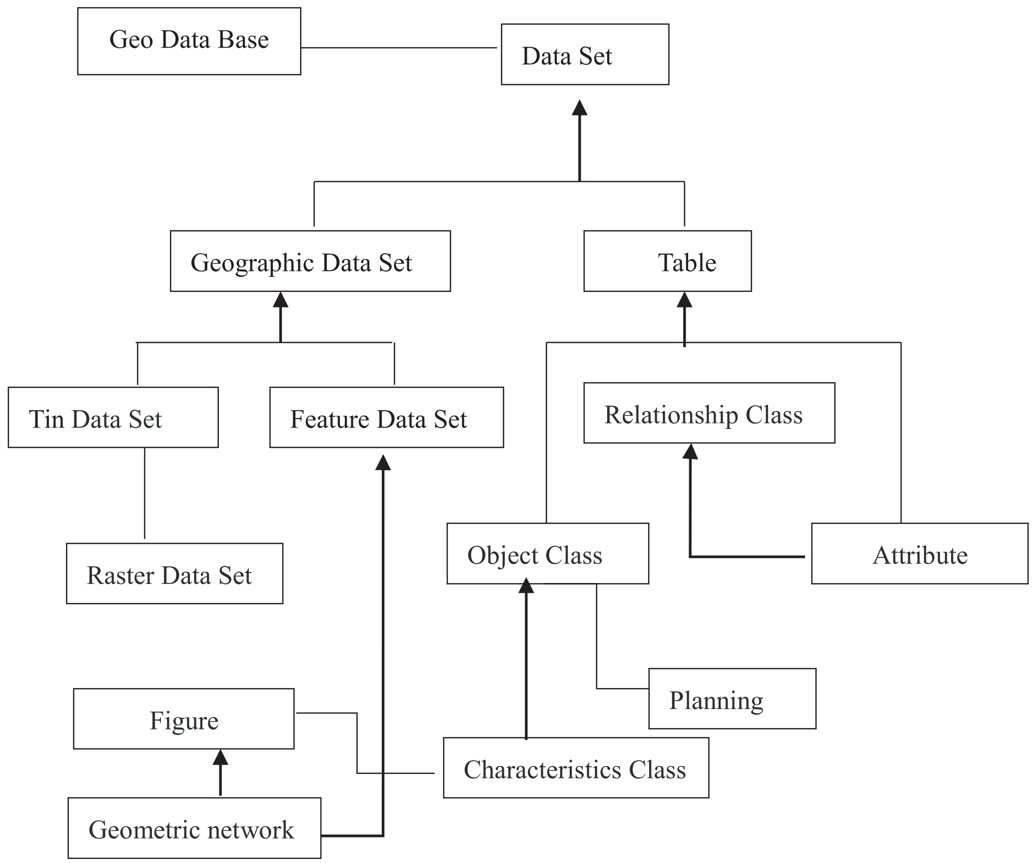

Fig. 2: Geo Data Base model.

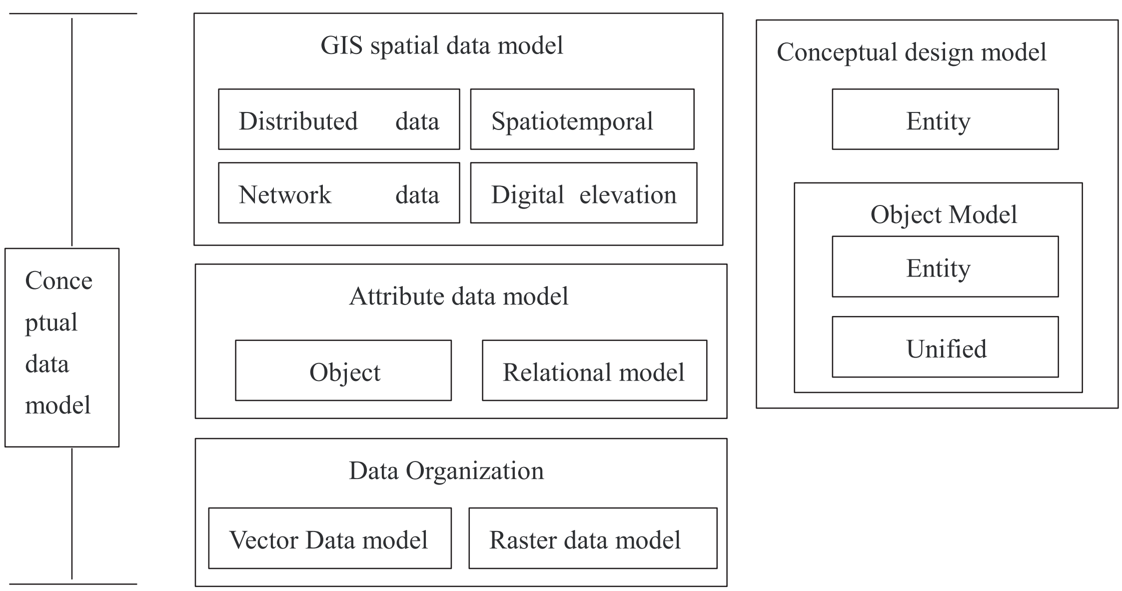

Fig. 3: Conceptual data model diagram.

1. According to the expression form of data structure, it is divided into data entity and data attribute.
2. Data operation expression is made by recording operations in the data entity, including deletion, insertion, query, modification, etc.
3. Data integrity constraints are used for complete constraint expression and referential integrity constraints between different data.

***Logical model design*:** Logical data model is transformed from conceptual model, but it can describe the logical structure of geographical data representation in GIS and the content of database, which is the middle level of abstract data. Common logical data models include network data model, relational data model and hierarchical data model. The database design model is mainly determined by the management system. The geological disaster database of Zhijin county is managed by oracle database. Therefore, the logical model design is selected for this database design.

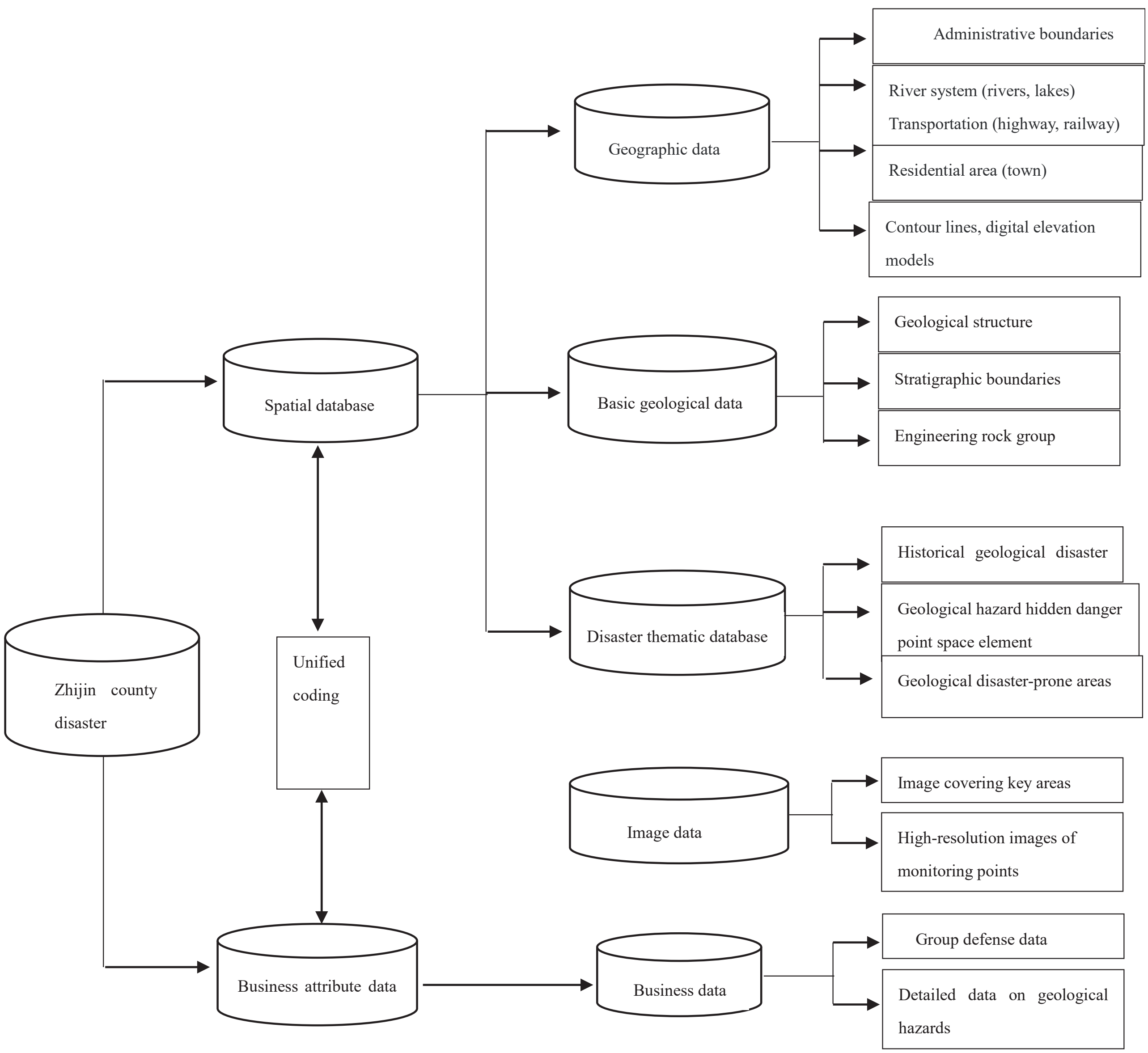

Fig. 4: General logical structure Design of Geological Hazard Database in Zijin County.

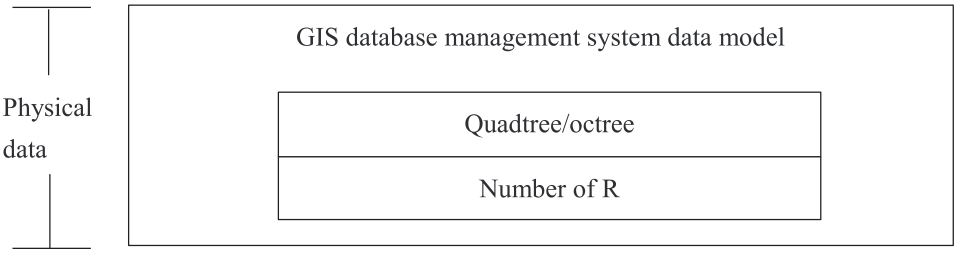

Fig. 5: Physical data model diagram.

***Physical model design*:** Physical model is an abstract data model, including spatial storage and acquisition methods of spatial data as well as its physical organization and database information storage structure. File index is one of the commonly used access methods, file index general methods are R tree, B tree, quadtree, etc. Only by understanding the relationship between entities in the real world can the established data model accurately reflect the relationship in the real world. Among the conceptual models, the E-R model is one of the most powerful tools of E, the connection between E-entities and R-entities (E), so it is also called the entity-connection model. Using the E-R model, the E-R graph can express the entities in the real world and their relations directly. Entity, attribute and relation are the three basic components of E-R graph. Without information storage structure and access paths, and questions about the computer case, E-R model to describe the real world is better, the E-R model as a conceptual model, data cannot be considered in the process of implementation in DBMS, and the DBMS (Database Management System, Database Management System) has no contact, so this kind of data model more reflect the result, because the data model is more close to the real world, as shown in Fig. 5.

Mathematical Model Design: A causative factor for the occurrence of geological disasters has many factors, the complexity of the inducing factors and the influence of the overlap between each other, it is difficult to forecast evaluation of the disaster. Research on disaster prediction method, analysis of the existing prediction models and other qualitative and quantitative forecasting methods, this paper focuses on the subject of Zhijin county geological disaster to study the effect of influencing factors of geological disasters.

***Mathematical Principles of Modelling*:**

(1) The principle of principal component analysis describes the original m variables by establishing n new variables ($m \le n$), which need to meet two requirements: m variables are unrelated to each other, and m variables retain as much information as possible from n variables. This method is called the principal component method. Through principal component analysis, on the one hand, establish the relationship between the factors and on the other hand, the degree of influence of factors.

$X = (X_1, X_2, X_3 \ldots X_n)^T$ is an n-dimensional random variable, whose second-order matrix exists. To use $Y_1, Y_2, Y_3 \ldots.. Y_m$ to replace the original $X_1, X_2, X_3 \ldots X_n$ variables, then $Y_1, Y_2, Y_3 \ldots.. Y_m$ should be a linear combination of $X_1, X_2, X_3 \ldots X_n$, if $Y_1, Y_2, Y_3 \ldots.. Y_m$: Y_1 is the largest difference between any shape $Y = I^T X$, and Y_k is the largest difference between any shape $Y = I^T X Y_1, Y_2, Y_3$. Y_{k-1} is unrelated and has the largest variance (k = 2..., m), Y_i is the i principal component of X (i =1,2...m). Set X as n random variables, the mean $E_X = 0$, coordinated differential matrix $\Sigma = \text{cov}(X)$, Σ n of order is: the main characteristics of the $\lambda_1 \ge \lambda_2 \ge \ldots\ldots \lambda_n \ge 0$, l_1, $l_2 \ldots\ldots l_n$. If ln is the corresponding unit eigenvector, then the ith principal component of X is:

$$Y_i = l_i X \qquad i = 1, 2 \ldots.., n \qquad \ldots(1)$$

Correlation coefficient between principal component Y_k and original variable X_i:

$$P(Y_k, X_i) = \sqrt{\lambda_k} t_{ik} / \sqrt{\sigma_{ii}} \qquad \ldots(2)$$

Where t_{ik} is the ith component of the KTH eigenvector; To $\lambda_k / \Sigma\lambda_i$ for $Y_1, Y_2, \ldots\ldots, Y_m$'s cumulative contribution rate; the number of selected variables is determined by the cumulative contribution rate, and the cumulative contribution rate is determined. The cumulative contribution rate reflects the amount of information of the variable X extracted by the principal component, which is generally not less than 0.85, known as the principle of 0.85. According to the principal component calculation results, m comprehensive variables $T_1, T_2, \ldots, T_m$, the calculation formula:

$$T_p = \Sigma\alpha_{ip} X_i \times W_p \qquad \ldots(3)$$

Where, W_p is the contribution rate of the p-th principal component, X_i is the factor quantization value of the i-th influencing factor of sample X, and α_{ip} is the correlation coefficient of the i-th influencing factor of the p-th principal component.

(2) Markov Distance Test

Let's say m (m > 2) population G_1, G_2, …G_m, mean codifference matrices are μ_1, μ_2, …μ_m and Σ_1, Σ_2, … Σ_m and its value to calculate unbiased estimator known or according to the sample. A sample X is discriminated, and the distances from X to m populations are calculated respectively

$$D^2 (X, G_i) = (X - \mu_i)^T \Sigma^{-1} (X - \mu_i) \quad i = 1, 2 ... m \qquad ...(4)$$

Let G_i's total $R_i = \{X: D^2 (X, G_j) \leq \min D^2 (X, G_j)\}$, and the discriminant rule is: $X \in G_i$, when X falls in R_i (i = 1, 2 ... m)

***Establishment of the model*:** Based on principal component analysis and discriminant analysis, a regional disaster prediction model is established. The modelling steps are as follows:

1. Determine the influencing factors of disasters and determine the influencing factors according to the mechanism of the influencing factors;
2. Taking the influencing factors as random variables and the factor factors as the specific values of variables, the quantitative criteria of influencing factors were established according to the role degree of the factor factors by random assignment;
3. Collect disaster data of the study area, analyse the factors affecting the factors, quantify the published disaster samples according to the quantitative criteria, and obtain the quantitative matrix of the disasters, denoted as A_1;
4. Conduct principal component analysis onA_1, extract principal components F_1, F_2, ...F_m (m ≤ n, n is the number of influencing factors), and calculate factor load matrix A_2 and the contribution rate of each principal component W_1,W_2,.....W_m;
5. According to the formula (3), the specific expressions of the comprehensive variables are calculated to obtain the simplified data matrix $A_3(T_\varepsilon)$;
6. To simplify the A_3 for disaster similar data matrix eigenvalue μ and Σ unbiased estimator û;
7. The existing geological disaster points are judged back, and the discriminant function adopts the Markov distance, takes the threshold D_k and estimates the error rate, divides the D_k, and links it with the hazard degree;
8. For any area to be forecasted, the geological conditions are analysed to determine the factor states of influencing factors. The factor factors are quantified by quantitative criteria. According to the load matrix A_2 of disaster factors and the corresponding principal component contribution rate, the comprehensive variable vector $H = (T_1, T_2, ... T_P)$, and calculate the Markov distance between the vector and the disaster occurrence class according to the following formula

$$DH = (H - \hat{u})^T \Sigma^{-1} (H - \hat{u}) \qquad ...(5)$$

CONCLUSION

In this paper, the related standards of geological disaster database in Zhijin county are studied, and the standards of geological disaster database are established. The characteristics of basic geological disaster data are analysed. The design and construction of the basic geological disaster database are completed. This paper studies the optimization technology of geological disaster data access based on cache.

The integrated processing of geological disaster spatial data and attribute data is realized through database research and development, the correlation between thematic data is enhanced, the effective management, maintenance and efficient utilization of geological disaster information are realized, and effective data service and technical support are provided for various functional systems of the monitoring and warning platform.

The main achievements of this paper are as follows:

1. To study the data standards of geological hazards and establish the data standards of the basic database of geological hazards in Zhijin county according to the actual situation of the basic data of geological hazards in Zhijin county.
2. The mathematical model was constructed to extract the principal component, which was the comprehensive variable of each factor, and the relationship between the influencing factors and the principal component was reflected by the load of each factor, and the mathematical expression between the principal component and the variables of each influencing factor was determined.

REFERENCES

Echavarribravo, V., Paterson, L., Aspray, T. J., Porter, J. S., Winson, M. K. and Hartl, M. 2017. Natural marine bacteria as model organisms for the hazard-assessment of consumer products containing silver nanoparticles. Marine Environmental Research, 130: S0141113617303653.

Filatov, D. M. and Lyubushin, A. A. 2019. Precursory analysis of GPS time series for seismic hazard assessment. Pure and Applied Geophysics, pp. 22-1.

Furlan, E., Torresan, S., Critto, A. and Marcomini, A. 2018. Spatially explicit risk approach for multi-hazard assessment and management in marine environment: the case study of the Adriatic sea. Science of the Total Environment, 618: 1008-1023.

Ismail-Zadeh, A., Soloviev, A., Sokolov, V., Vorobieva, I., Müller, B. and Schilling, F. 2017. Quantitative modelling of the lithosphere dynamics, earthquakes and seismic hazard. Tectonophysics, S0040195117301415.

Karamouz, M., Razmi, A., Nazif, S. and Zahmatkesh, Z. 2017. Integration of inland and coastal storms for flood hazard assessment using a distributed hydrologic model. Environmental Earth Sciences, 76(11): 395.

Manea, E.F., Michel, C., Hobiger, M., Fäh, D., Cioflan, C.O. and Radulian, M. 2017. Analysis of the seismic wavefield in the Moesian platform (Bucharest area) for hazard assessment purposes. Geophysical Journal International, 210(3): 1609-1622.

Sivakumar, R. and Ghosh, S. 2017. Earthquake hazard assessment through geospatial model and development of Eahaas to tool for visualization: an integrated geological and geoinformatics approach. Environmental Earth Sciences, 76(12): 442.

Tao, W., Liu, J., Shi, J. and Wu, S. 2017. The influence of DEM resolution on seismic landslide hazard assessment based upon the Newmark displacement method: a case study in the loess area of Tianshui, China. Environmental Earth Sciences, 76(17): 604.

Xiao, Y., Yi, S. and Tang, Z. 2018. A spatially explicit multi-criteria analysis method on solving spatial heterogeneity problems for flood hazard assessment. Water Resources Management, 32(10): 1-19.

Yepes, H., Audin, L., Alvarado, A., Beauval, C., Aguilar, J. and Font, Y. et al. 2018. A new view for the geodynamics of Ecuador: implication in seismogenic source definition and seismic hazard assessment. Tectonics, 35(5): 1249-1279.