# Set pair analysis for karst waterlogging risk assessment based on AHP and entropy weight

Jiajun Zeng and Guoru Huang

## ABSTRACT

Karst waterlogging is a natural disaster that occurs frequently and it adversely affects the social and economic development of affected areas. An analysis of the causes of karst waterlogging with respect to climate and topography can serve as a foundation for disaster assessment and prevention. In this study, a karst waterlogging risk assessment indexing system was established. The system was based on a comprehensive analysis of risk factors, including the severity of the disaster and the vulnerability of the affected area. DeBao County in GuangXi was used as the study area. Nine risk indicators were chosen as evaluation indexes and combination weights were divided into subjective and objective weights based on the analytic hierarchy process (AHP) and entropy weight. A geographical information system (GIS) was applied to help with the calculations of the set pair analysis (SPA) and to pile up the layers of the evaluation indicators. Then, a risk rating map was drawn using GIS techniques. The results showed that the high risk locations were mainly distributed in the center of DeBao County; therefore, the map could be used as reference for the prevention and management of karst waterlogging risk.

**Key words** | AHP, entropy weight, GIS, karst waterlogging, risk assessment, SPA

**Jiajun Zeng**
**Guoru Huang** (corresponding author)
School of Civil Engineering and Transportation,
South China University of Technology,
Guangzhou 510640,
China
E-mail: *huanggr@scut.edu.cn*

**Guoru Huang**
State Key Laboratory of Subtropical Building Science,
South China University of Technology,
Guangzhou 510640,
China

## INTRODUCTION

Karst waterlogging is a typical natural disaster that occurs in karst areas. These disasters impact large populations across southwest China every year where karst topography is located. Previously, the research on water quantity in karst areas mostly focused on drought (Keqiang *et al.* 2011) while karst waterlogging was rarely mentioned. As a result of ecological disruptions to the environment due to climate, karst waterlogging has begun to occur more frequently in areas of southwestern China, such as in Guangxi Province. This has aroused the attention of the government.

Karst terrain is a unique landscape that is produced by the dissolution of carbonate bedrock (Guo & Jiang 2010). The dissolution process leads to the development of caves, sinkholes, springs, and sinking streams, all of which are typical features of a karst system (Malagò *et al.* 2016). In the rainy season, the rainfall enters the Earth's subsurface and percolates down to groundwater networks, which consist of three types of spaces: pores, fissures, and caverns. Groundwater storage occurs in pores and fissures, and cavern conduits act as drains (Mimi & Assi 2009). In other words, the flow in an interconnected fissure network is concentrated in conduits that connect cavities and cave systems to create groundwater flow in a karst aquifer (Fiorillo 2009). These hydrological processes involve surface water, groundwater, and soil water with complicated interactions (Díez-Herrero *et al.* 2012). This results in difficulties in determining the causes of karst waterlogging. First, the direct external cause is torrential rainfall. When the volume of water exceeds the drainage capacity of the conduits in the area, the water will deepen in low-lying areas leading to

doi: 10.2166/nh.2017.265

waterlogging. When mud, soil, and trash are carried by the flow and deposited in sinkhole in-takes or conduits, which reduces cross-sectional flow and the discharge capacity of conduits, the potential for waterlogging is increased (Long *et al.* 2014). In addition, it should be noted that groundwater flow can be reversed from its normal flow direction and spill over the sinkhole while the water levels increase at the discharge points (Zhou 2007). Due to the interconnections of underground structures, the upstream rainfall water travels downstream through underground conduits. At this point, waterlogging may occur as the water pours out of the sinkhole.

A significant effort has been made to reduce the damage caused by karst waterlogging disasters, but these efforts have not completely addressed the issue, resulting in continued losses to life and property. This can be attributed to disasters' broad influence which last for a long time (Macdonald *et al.* 2014). As the karst areas are distributed in remote mountainous areas where the economy is underdeveloped with limited resources available to mitigate natural disasters, it is prudent for disaster management personnel to take into account the potential impact on people and property when considering the application of risk assessment methods (Zheng & Qi 2011). Analyzing the spatial distribution characteristics of karst waterlogging risk and evaluating the degree of risk are important for disaster warning, disaster evacuation, risk management, and decision-making (Jiang *et al.* 2009; Zou *et al.* 2013; Lai *et al.* 2015). In general, the risk of karst waterlogging is defined as the product of hazard and vulnerability. As a natural disaster, karst waterlogging is similar to flood disasters, and the methods associated with risk assessment of the latter can be used for the former.

To obtain accurate risk levels in risk assessment, the main difficulty is the multi-variable and non-linear relationship between indicators and risk levels. The analytic hierarchy process (AHP), the most widely used multi-criteria analysis method in risk assessment, has been applied to a wide range of scientific fields (Saaty 1990; Yang *et al.* 2013). Therefore, it is reasonable to use the AHP to analyze karst waterlogging risk and determine indicator weight. The method used to determine the attribute weights should be divided into subjective and objective weighting methods. Entropy weight was chosen as an objective way to measure the amount of useful information with the data provided while also avoiding human interference when deciding the weight factors (Zou *et al.* 2006; Jia *et al.* 2011). The set pair analysis (SPA) proposed by Zhao & Xuan (1996) is a theoretical method used to deal with uncertainty problems. The SPA was used to distinguish between uncertain and unknown indicator symptoms (Feng *et al.* 2014). A geographical information system (GIS) is a decision support system involving the integration of spatially referenced data in a problem solving environment (Chenini *et al.* 2010; Aurit *et al.* 2013). In GIS, database analytical tools and mathematical relationships among different layers can be combined into a decision support system, yielding a waterlogging risk zoning map (Kourgialas & Karatzas 2016).

At present there are not many studies on waterlogging of karst internationally, especially for karst waterlogging risk analysis. However, flood risk analysis has numerous existing classical and developed methods, and there are similar properties between flood risk and karst waterlogging risk. Combining the characteristics of karst areas we can use the analysis methods of flood risk for the risk assessment of karst waterlogging. In this study, the risk of karst waterlogging was considered as the product of hazard and vulnerability and a risk assessment index system was established. The weights of the disaster factors were combined with subjective and objective weights calculated using the AHP and entropy weights. Next, the SPA was established by employing GIS to generate an integrated zoning map for karst waterlogging to be used in risk management.

## STUDY AREA AND DATA

### Study area

DeBao County, located to the southwest of the Guangxi Zhuang Autonomous Region, is located at longitude 106°09′–106°59′, latitude 23°01′–23°39′ (Figure 1). DeBao County covers an area of 2,575 km$^2$. The terrain in the northwest is high while in the southeast it is low. The major landforms are karst topography in the county including peak-cluster depression area, peak forest-valley, etc. Low mountain terrain in the county comprises four terrain types: hilly, hills, plateaus, and plains. Mountainous areas account for 78.4% of the total area. The rainy season begins in late

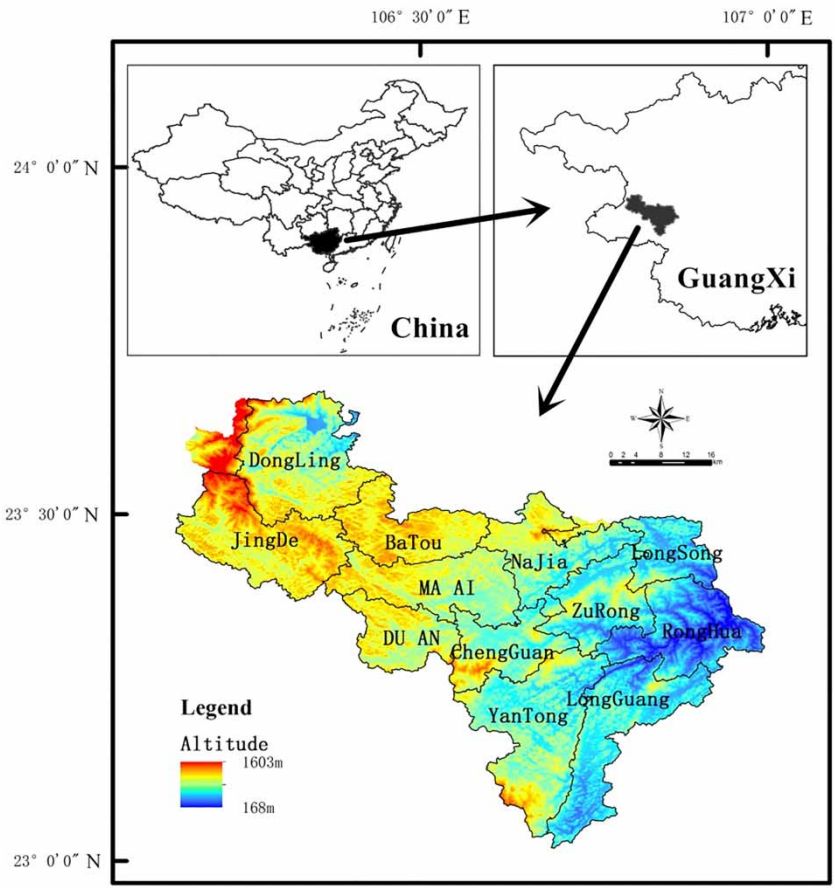


**Figure 1** | DeBao County map.

April and ends in early October, and the average annual rainfall is 1,463.7 mm.

DeBao County sits atop a typical karst area that experiences heavy rainfall over a short period, resulting in drainage holes and sinkholes becoming easily blocked by water erosion. Due to this, the county is considered high-risk with regards karst waterlogging. Karst waterlogging disasters have occurred for many years and have threatened the personal property of residents, as well as the growth of crops. According to the history of disaster records, the most serious karst waterlogging disaster occurred in 1998. It affected 80,700 people across 67 villages and up to six townships.

## Data

The selection of suitable risk indicators is the first step in karst waterlogging risk assessment. Different risk indicators from different study areas were selected based on the specific characteristics of each location. An index may show a high degree of impact for waterlogging risk in one area while it may not in another (Chaudhary *et al.* 2016). Seven hazard factor indicators and two vulnerability factors were selected to illustrate the influence factors of karst waterlogging (Figure 2(a)–2(i)). The factors were selected after taking field surveys into account, considering the actual conditions of karst waterlogging, and considering the relevant characteristics in the study area. These factors were presented in the form of grid maps in GIS with the scale of the raster set to 100 m × 100 m. These were used to determine the final karst waterlogging risk map.

The description of each factor in Figure 2 is as follows:

1. Maximum 3-day precipitation (M3DP, mm): Precipitation is the direct external cause of karst waterlogging disasters. M3DP was chosen since karst waterlogging

generally occurs with a long duration of rainfall. M3DP which is representative of the rainfall index was calculated based on daily precipitation data. These data were collected by the observational weather stations in DeBao County recorded for the years 1983–2013. These data were obtained from the Guangxi Water Resources Department for this study.

2. Digital elevation model (DEM, m): Waterlogging often occurs in low-lying areas, where rainfall easily flows from highlands. DEM was selected as an index to reflect the topographical conditions of the karst area. It was obtained from the Geospatial Data Cloud (http://www.gscloud.cn/). The SRTMDEM 90 m Original Resolution Elevation Data on the website was procured and the raster set to 100 m × 100 m in ArcGIS.
3. Slope (SL, °): This index reflects the degree of topographical change. The hoodoos and peak clusters, especially in karst areas, generally have sharp slopes that accelerate the convergence speed of rain. This results in a constant flood threat to lowland areas that aggravates soil erosion. The slope index was extracted from the DEM using GIS techniques.
4. Soil types (ST): Soil erosion conditions vary considerably under different ST. There are two types of soil losses in karst mountain areas: above-ground and underground soil losses. The soil conveyed by the current is deposited in the sinkholes or the groundwater pipes and these

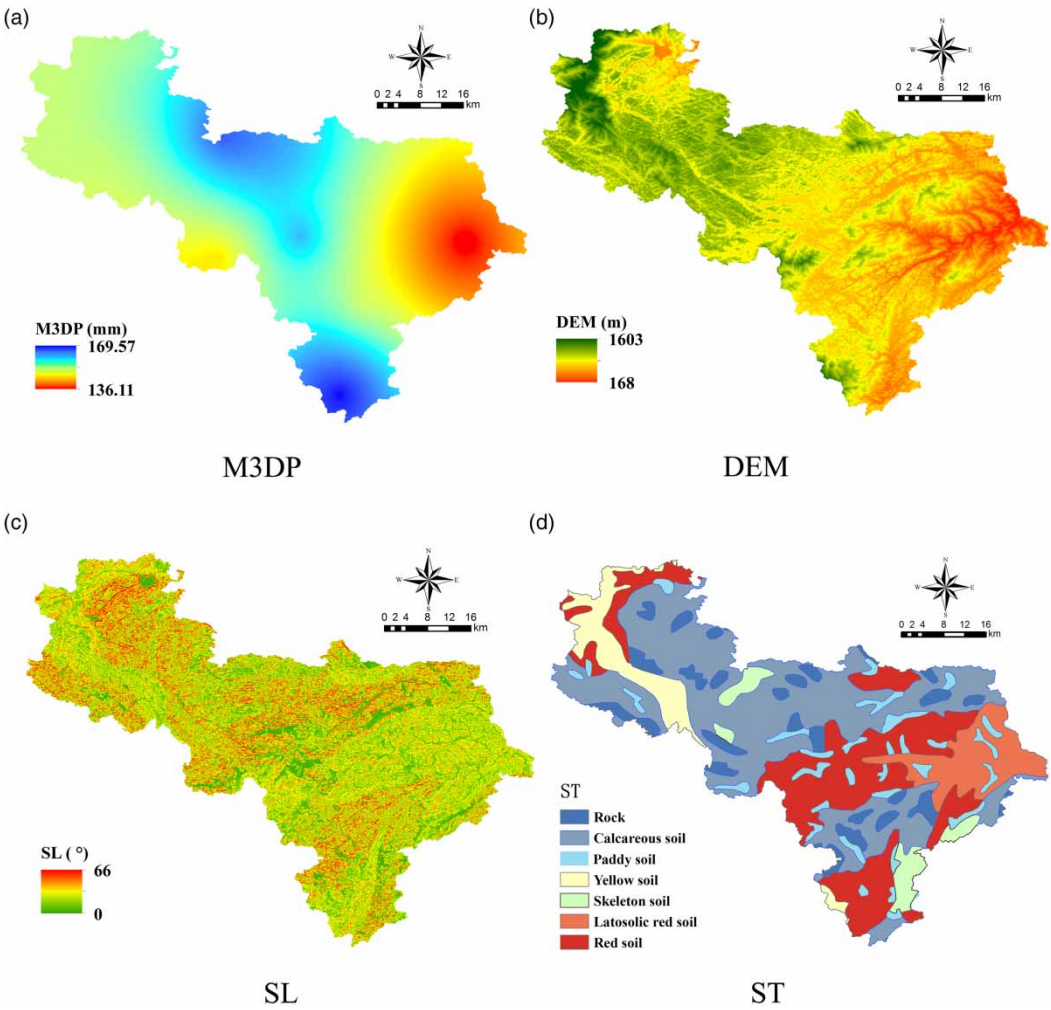


**Figure 2** | Characteristic distributions of evaluation indictors. *(Continued.)*

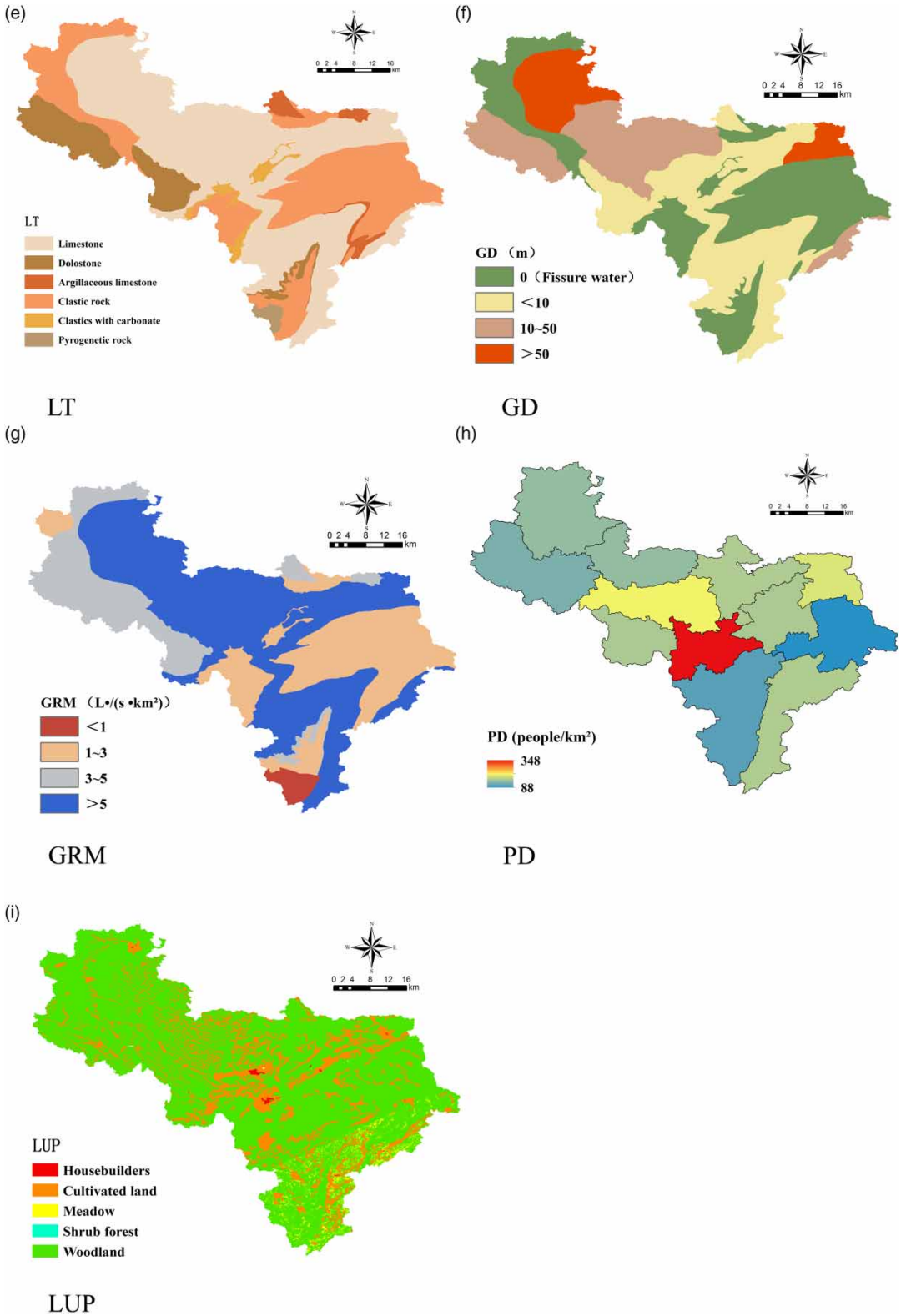


**Figure 2** | Continued.

blocked water channels may cause waterlogging. The data for this indicator were obtained from the Guangxi Water Resources Department.

5. Lithology (LT): This index reflects the degree of karst development. The higher the degree of karst development, the more prone the area is to waterlogging. These data were obtained from the Guangxi Geology Bureau.
6. Groundwater depth (GD, m): Shallow groundwater is not conducive to rainwater infiltration during a storm when it is more likely to gush from the sinkholes. In reality, we find that waterlogging rarely occurs in areas without groundwater. These data were obtained from the Guangxi Geology Bureau.
7. Groundwater runoff modulus (GRM, $L{\cdot}s^{-1}{\cdot}km^{-2}$): The GRM is the basic parameter used to evaluate the size of underground water in the form of underground runoff. A greater GRM leads to larger water contents in the area, which increases the probability of the occurrence of waterlogging. The data were obtained from the Guangxi Geology Bureau.
8. Population density (PD, people/$km^2$): This index reflects the population distribution in DeBao County in 2013. Densely populated areas experienced greater disaster losses than less densely populated areas. These data were obtained from the shared site of the Guangxi County Economic Net (China) (http://www.gxcounty.com/).
9. Land use pattern (LUP): Land used for housing and cultivation fall into the more vulnerable category of property loss when karst waterlogging occurs. The data were obtained from the Guangxi Water Resources Department.

## METHODOLOGY

### Analytic hierarchy process

The AHP is a multi-criteria decision-making method that provides a systematic approach for assessing and integrating the impacts of various factors (Saaty 1990). The AHP has been successfully applied to many fields, including engineering, evaluation, and management, because of its simplicity and flexibility (Chenini *et al.* 2010). Therefore, the AHP can be used to handle many risk-related problems, such as risk evaluation, risk management, and in particular hazard evaluation and analysis (Stefanidis & Stathis 2013). Carrying out a risk assessment of karst waterlogging based on AHP is therefore possible.

The AHP takes advantage of determining the relative importance of a set of factors and blending judgments on intangible qualitative criteria alongside tangible quantitative criteria (Jiang *et al.* 2009). The AHP combines the experience-based score of experts using pairwise comparisons of each factor to construct a judgment matrix. The most common methodology for performing comparisons is Saaty's (1990) comparative scale, which consists of the integers 1–9, and their reciprocal values (1/9, 1/8, 1/7, 1/6, 1/5, 1/4, 1/3, 1/2, 1, 2, 3, 4, 5, 6, 7, 8, 9). The minimum value, 1/9, represents the least relative influence, and the maximum value, 9, means that a factor is far more important than the others. The weights of each index are calculated with the appropriate mathematical model.

After a comprehensive analysis of the hazard factor, disaster-charged environment, and the disaster information from DeBao County, combined with a site visit and characteristics' analysis in a karst area, this study chose hazard and vulnerability as the karst waterlogging disaster risk evaluation index. A multi-level analysis model including goal, criterion, and sub-criterion levels was developed (Figure 3).

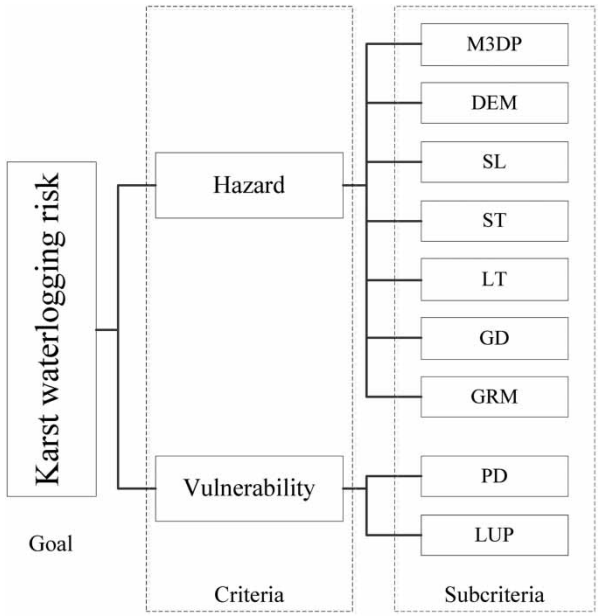


**Figure 3** | Hierarchical structure of the karst waterlogging risk index.

The hazard criteria of the karst waterlogging index illustrated the severity of the waterlogging disasters. Seven factors (M3DP, DEM, SL, ST, LT, GD, and GRM) were selected for the karst hazard waterlogging evaluation index. The judgment matrix of hazard (Matrix A) was constructed using index scoring assigned by experts. Meanwhile, the vulnerability criteria showed the possibility and severity of socio-economic losses caused by waterlogging. PD and LUP were chosen to evaluate the vulnerability of a karst waterlogging disaster area. A judgment matrix of vulnerability (Matrix B) was constructed using these two factors.

$$A = \begin{array}{c} \begin{matrix} \text{M3DP} & \text{DEM} & \text{SL} & \text{ST} & \text{LT} & \text{GD} & \text{GRM} \end{matrix} \\ \begin{bmatrix} 1 & 2 & 3 & 4 & 4 & 3 & 2 \\ 1/2 & 1 & 2 & 3 & 3 & 1 & 2 \\ 1/3 & 1/2 & 1 & 2 & 2 & 1/2 & 1 \\ 1/4 & 1/3 & 1/2 & 1 & 1 & 1/3 & 1/2 \\ 1/4 & 1/3 & 1/2 & 1 & 1 & 1/3 & 1/2 \\ 1/2 & 1 & 2 & 3 & 3 & 1 & 2 \\ 1/3 & 1/2 & 1 & 2 & 2 & 1/2 & 1 \end{bmatrix} \end{array} \begin{matrix} \text{M3DP} \\ \text{DEM} \\ \text{SL} \\ \text{ST} \\ \text{LT} \\ \text{GD} \\ \text{GRM} \end{matrix}$$

$$B = \begin{array}{c} \begin{matrix} \text{PD} & \text{LUP} \end{matrix} \\ \begin{bmatrix} 1 & 4 \\ 1/4 & 1 \end{bmatrix} \end{array} \begin{matrix} \text{PD} \\ \text{LUP} \end{matrix}$$

In order to improve the rationality and logic of the judgment matrix, it could be evaluated through a consistency ratio (*CR*). The *CR* must be <0.1 in order to be accepted. The *CR* was defined as follows (Saaty 1990):

$$CR = \frac{CI}{RI} \tag{1}$$

where *RI* is a random index representing the consistency of a randomly generated pairwise comparison matrix (Saaty 1990). *CI* represents the computing consistency index and is given as:

$$\text{CI} = \frac{\lambda_{max} - n}{n - 1} \tag{2}$$

where $\lambda_{max}$ represents the largest eigenvalue of the judgment matrix, and $n$ represents the size of the judgment matrix.

Considering that the dimensions of the evaluation index are different, non-dimensional standardization of the data is carried out using GIS, with which the layer index data are rasterized and reclassified according to the contribution of the data to the degree of risk.

## Entropy weight method

In information theory, entropy is used to measure the amount of disorder in a system. However, it could also be used to measure the amount of information carried in data. For a certain evaluation, if the difference between the indexes values increases, the entropy decreases. If the index also carried a relatively large amount of effective information, then the weight of the evaluation index value would be larger (Fagbote *et al.* 2014).

In the concrete implementation and application, according to each evaluation index value of the degree of differentiation, the entropy weight of each evaluation index is calculated (Zou *et al.* 2006). Then, the weights of the evaluation indexes were measured using the method of entropy weight in order to obtain objective assessment results. The calculation was accomplished in the following steps.

1. Normalization of the original evaluating matrix: Construct a judgment matrix, $X$, with $m$ evaluation objects and $n$ risk indices as $X = (x_{ij})_{n\times m}$; where $x_{ij}$ represent the $j$ grid of the $i$ risk indices:

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1m} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nm} \end{pmatrix} \tag{3}$$

Normalize this matrix to obtain matrix $R = (r_{ij})_{n\times m}$, where $r_{ij}$ is the proportion of the index value of the $j$-th evaluation item of the $i$-th evaluation index. Among these indicators, a larger attribute value of $x_{ij}$ relates to higher risk (Wang *et al.* 2009), thus:

$$r_{ij} = \frac{x_{ij} - x_{min}}{x_{max} - x_{min}} \tag{4}$$

while a larger attribute value of $x_{ij}$ relates to lower risk:

$$r_{ij} = \frac{x_{max} - x_{ij}}{x_{max} - x_{min}} \tag{5}$$

2. Definition of the entropy: Calculate the index's entropy value, $H_i$, with the following formula:

$$H_i = -k\sum_{j=1}^{m} f_{ij}\ln f_{ij} \tag{6}$$

in which, $f_{ij} = r_{ij}/\sum_{j=1}^{m} r_{ij}$; $k = 1/ln\ m$; $i = 1, 2 \ldots\ldots n$; $j = 1, 2 \ldots\ldots m$.

3. Definition of the weight of entropy: The weight of entropy of each index can be defined as:

$$u_i = \frac{1 - H_i}{n - \sum_{i=1}^{n} H_i} \tag{7}$$

where $0 \le u_i \le 1$, $\sum_{i=1}^{n} u_i = 1$.

## Set pair analysis

SPA is a theoretical method used to deal with the certainty and uncertainty of systematic analysis (Wang *et al.* 2017). The SPA constructs a set pair of two related sets of uncertain systems in order to analyze the identity, difference, and opposition of the set pair. The connection can be quantitatively characterized as:

$$\mu = \frac{S}{N} + \frac{F}{N}i + \frac{P}{N}j \tag{8}$$

where $\mu$ is the connection degree. $N$ is the total number of set characteristics, and $S$, $P$, and $F$ are the number of identical, contradictory, and discrepant characteristic terms, respectively. The connection degree of the set pair can be determined as identical, contradictory, or discrepant. $S$ represents the number of identical characteristics; $P$ is the number of contradictory characteristics; $F + N - S - P$ is the number of characteristics of these two sets that are neither identical nor contradictory (Wang *et al.* 2015). $i$ is the uncertain coefficient of the discrepancies and has an interval of $[-1,1]$. $j$ is the coefficient of the contraries and $j = -1$ as marked effect. Generally $a = S/N$, $b = F/N$, and $c = P/N$, with the formula then simplified to:

$$\mu = a + bi + cj \tag{9}$$

where $a$, $b$, and $c$ are connectivity components referred to as the identity degree, the discrepancy degree, and the contradictory degree. In addition, $a + b + c = 1$. $bi$ could be expanded into a multi-element connection degree ($k$ elements) and $\mu$ can be shown as:

$$\begin{aligned}\mu &= \sum_{r=1}^{m}\lambda_r\mu_r \\ &= \sum_{r=1}^{m}\lambda_r a_r + \sum_{r=1}^{m}\lambda_r b_{r,1}i_1 + \cdots \\ &\quad + \sum_{r=1}^{m}\lambda_r b_{r,k-2}i_{k-2} + \sum_{r=1}^{m}\lambda_r c_r j\end{aligned} \tag{10}$$

where $\lambda_r$ is the index combination weight presented in this paper.

When the values of the indicators are greater the risk is greater, for example, the risk is greater when the value of the M3DP is greater. The connection degree, $\mu$, can be defined as follows:

$$\mu_r = \begin{cases} 1 + 0i_1 + 0i_2 + \cdots + 0i_{k-2} + 0j & (x_r < s_1) \\ \dfrac{s_1 + s_2 - 2x_r}{s_2 - s_1} + \dfrac{2x_r - 2s_1}{s_2 - s_1}i_1 + \cdots + 0i_{k+2} + 0j & \left(s_1 \le x_r < \dfrac{s_1 + s_2}{2}\right) \\ 0 + \dfrac{s_1 + s_2 - 2x_r}{s_3 - s_1}i_1 + \dfrac{2x_r - s_2 - s_3}{s_3 - s_1}i_2 + \cdots 0i_{k-1} + 0j & \left(\dfrac{s_1 + s_2}{2} \le x_r < \dfrac{s_2 + s_3}{2}\right) \\ \vdots & \vdots \\ 0 + 0i_1 + \cdots + \dfrac{2s_{k-1} - 2x_r}{s_{k-1} - s_{k-2}}i_{k-2} + \dfrac{2x_r - s_{k-2} - s_{k-1}}{s_{k-1} - s_{k-2}}j & \left(\dfrac{s_{k-2} + s_{k-1}}{2} \le x_r < s_{k-1}\right) \\ 0 + 0i_1 + \cdots + 0i_{k-2} + j & x_r > s_{k-1}) \end{cases} \tag{11}$$

where $s_1, s_2, \cdots s_{k-1}$ are the extremes of the standard.

In contrast, when the indicator values are greater the risk is smaller, for example, the risk is lesser when the value of the DEM is higher, $\mu$ can be defined as follows:

$$\mu_r = \begin{cases} 1 + 0i_1 + 0i_2 + \cdots + 0i_{k-2} + 0j & (x_r \geq s_1) \\ \dfrac{2x_r - s_1 - s_2}{s_1 - s_2} + \dfrac{2s_1 - 2x_r}{s_1 - s_2} i_1 + \cdots + 0i_{k+2} + 0j & \left(\dfrac{s_1 + s_2}{2} \leq x_r < s_1\right) \\ 0 + \dfrac{2x_r - s_2 - s_3}{s_1 - s_3} i_1 + \dfrac{s_1 + s_2 - 2x_r}{s_1 - s_3} i_2 + \cdots 0i_{k-1} + 0j & \left(\dfrac{s_2 + s_3}{2} \leq x_r < \dfrac{s_1 + s_2}{2}\right) \\ \vdots & \vdots \\ 0 + 0i_1 + \cdots + \dfrac{2x_r - 2s_{k-1}}{s_{k-2} - s_{k-1}} i_{k-2} + \dfrac{s_{k-2} + s_{k-1} - 2x_r}{s_{k-2} - s_{k-1}} j & \left(s_{k-1} \leq x_r < \dfrac{s_{k-2} + s_{k-1}}{2}\right) \\ 0 + 0i_1 + \cdots + 0i_{k-2} + j & (x_r < s_{k-1}) \end{cases} \quad (12)$$

where $s_1, s_2, \cdots s_{k-1}$ are the extremes of the standard.

## RESULTS AND DISCUSSION

### Combination weight

The AHP is used for determining the subjective weights which are affected by the subjective judgment of the experts. The weights of the seven hazard factor indicators of hazard factors were calculated as 0.303(M3DP), 0.184(DEM), 0.105(SL), 0.060(ST), 0.060(LT), 0.184(GD), and 0.105(GRM). The judgment matrix A was established with a CR of 0.019, which satisfied the criteria of CR being less than 0.1. Similarly, the weights of vulnerability factors, PD and LUP, were determined to be 0.750 and 0.250, respectively. The CR of matrix B whose dimension was two-dimensional was bound to meet the requirements of consistency. Using previous risk assessments performed by experts as a reference, the karst waterlogging disaster risk assessment in DeBao County was assigned a hazard weight of 0.75 and a vulnerability weight of 0.25. Therefore, the weight of the nine indicators, determined by AHP, can be shown to be 0.2273(M3DP), 0.1378 (DEM), 0.0784(SL), 0.0452(ST), 0.0452(LT), 0.1378(GD), 0.0784(GRM), 0.1875(PD), and 0.0625(LUP).

The entropy weight method is used to determine objective weights, in order to reduce the subjective influence of the experts, and the information carried by the original data is fully excavated. However, the weight values may not be consistent with the actual importance of the corresponding indicators. In order to construct a judgment matrix, the ST, LT, and LUP layers need to be quantified in GIS (Aurit *et al.* 2013). The elements of the matrix are judged by the grid value. Through the calculation steps of entropy weight method, the entropy weight is determined using the raster calculator in GIS. The results of the entropy weight are as follows: 0.0013(M3DP), 0.0544(DEM), 0.2255 (SL), 0.1255(ST), 0.0749(LT), 0.2025(GD), 0.0643(GRM), 0.1279(PD), and 0.1237(LUP).

In this study, the subjective weights, $\omega_i$, and objective weights, $u_i$, are combined to obtain the combination weights, $\lambda_i$, which can be defined as follows:

$$\lambda_i = \alpha\omega_i + (1 - \alpha)u_i \quad (13)$$

in which a change in $\alpha$ changes the combination weights. According to previous risk assessments performed by experts (Liu *et al.* 2016), this coefficient $\alpha$ could be set in the range of 0.6–0.8. After examining the proportional change of the subjective and objective weights and combining with the field investigation, we set $\alpha = 0.725$ in this paper. Therefore, the combination weight results are shown in Table 1.

### Risk distribution analysis

In order to quantify the dimensionless indicators (ST, LT, LUP), the indicators are encoded based on the degree of danger they pose. This degree of danger is based on previous determinations made by experts (Table 2). As the

**Table 1** | Combination weights of the evaluation indicators

| Index | M3DP | DEM | SL | ST | LT | GD | GRM | PD | LUT |
|---|---|---|---|---|---|---|---|---|---|
| AHP | 0.2773 | 0.1378 | 0.0784 | 0.0452 | 0.0452 | 0.1378 | 0.0784 | 0.1875 | 0.0625 |
| Entropy weight | 0.0013 | 0.0544 | 0.2255 | 0.1255 | 0.0749 | 0.2025 | 0.0643 | 0.1279 | 0.1237 |
| Combination weight | 0.1652 | 0.1149 | 0.1189 | 0.0673 | 0.0534 | 0.1556 | 0.0745 | 0.1711 | 0.0793 |

**Table 2** | Degree of danger of the indicators

| ST | Code | LT | Code | LUP | Code |
|---|---|---|---|---|---|
| Rock | 1 | Pyrogenetic rock | 1 | Woodland | 1 |
| Yellow soil | 2 | Clastic with carbonate | 2 | Shrub forest | 2 |
| Latosolic red soil | 3 | Clastic rock | 3 | Meadow | 3 |
| Skeleton soil | 4 | Argillaceous limestone | 4 | Cultivated land | 4 |
| Paddy soil | 5 | Dolostone | 5 | Housebuilders | 5 |
| Red soil | 6 | Limestone | 6 | | |
| Calcareous soil | 7 | | | | |

degree of danger increases, the indicator is considered more dangerous. Each grid of layers represents a point of evaluation in ArcGIS (Wang *et al.* 2012) and each index was divided into five categories, using the method of natural breaks (Table 3). In the table, $s_1$ represented the threshold of the lowest risk, while $s_4$ represented one of the highest risk thresholds. For each layer, Equations (11) and (12) were used to calculate the connection degree using the raster calculator in ArcGIS. This resulted in the generation of individual grids for different layers with their own $a_r, b_{1,r}, b_{2,r}, b_{3,r}, c_r$. $r$ was the serial number of the layers.

The identify degree $a$ could be calculated as follows:

$$a = \sum_{r=1}^{9} a_r\lambda_r = a_1\lambda_1 + a_2\lambda_2 + \cdots + a_9\lambda_9 \quad (14)$$

where $\lambda_r$ was the individual weights of the indicators. According to the weights obtained above, weighting stacks the 9 evaluation indicators by using the raster calculator of the weighted aggregation method in GIS. Similarly, $b_1, b_2, b_3, c$ could be calculated as follows:

$$b_t = \sum_{r=1}^{9} b_{t,r}\lambda_r = b_{t,1}\lambda_1 + b_{t,2}\lambda_2 + \cdots + b_{t,r}\lambda_9, (t = 1, 2, 3) \quad (15)$$

$$c = \sum_{r=1}^{9} c_r\lambda_r = c_1\lambda_1 + c_2\lambda_2 + \cdots + c_9\lambda_9 \quad (16)$$

The confidence degree was set to 0.5. If $a_1 > 0.5$, the risk of this grid was the lowest; if $a_1 + b_1 > 0.5$, the risk was the lower; if $a_1 + b_1 + b_2 > 0.5$, the risk was the medium; if

**Table 3** | Standard grades and weights of the indicators

| Indicator | M3DP ($x_1$) | DEM ($x_2$) | SL ($x_3$) | ST ($x_4$) | LT ($x_5$) | GD ($x_6$) | GRM ($x_7$) | PD ($x_8$) | LUP ($x_9$) |
|---|---|---|---|---|---|---|---|---|---|
| $s_1$ | 144 | 1086 | 36 | 1 | 1 | 0 | 0 | 88 | 1 |
| $s_2$ | 150 | 894 | 26 | 2 | 2 | 1 | 1 | 119 | 2 |
| $s_3$ | 155 | 724 | 18 | 4 | 4 | 10 | 3 | 147 | 4 |
| $s_4$ | 160 | 537 | 10 | 6 | 5 | 50 | 5 | 183 | 5 |
| Weights | 0.1652 | 0.1149 | 0.1189 | 0.0673 | 0.0534 | 0.1556 | 0.0745 | 0.1711 | 0.0793 |

$a_1 + b_1 + b_2 + b_3 > 0.5$, the risk was the higher; if $a_1 + b_1 + b_2 + b_3 + c > 0.5$, the risk was the highest. By adding up these layers together, the flood risk assessment map was drawn, as shown in Figure 4.

According to Figure 4, the highest risk zones are concentrated in the central and southeast regions of DeBao County. The risk in MA AI and NaJia are particularly prominent, and parts of ChengGuan and LongGuang are located in high-risk zones. The higher-risk zones are mainly distributed in the north and southern regions of DeBao County, including BaTou, northern LongSang, western MA AI and DU AN, northern YanTong and LongGuang. The medium-risk zones are mainly located in the western region of DeBao County, which mainly includes JingDe and DongLing. The lower risk zones exist mainly in ZuRong, and the lowest risk zones are in RongHua.

Generally, the trend of karst waterlogging risk in DeBao County shows high risk in the center and low risk on both sides. According to the risk-zoning map, areas of the highest and higher risk account for 34.64% (Table 4) of the total risk area, have greater precipitation, and a high degree of karst development with abundant and relatively shallow underground water. Moreover, the PD in these areas is high. In contrast, the lowest risk zones are distributed in weak karstification areas with almost the least amount of rainfall (e.g., RongHua, where the density of population was small). It is suggested that karst waterlogging risk is mainly determined by natural conditions and social factors. In addition, the effect of topography on disaster risk should not be ignored. For example, NaJia and Long Sang do not experience high rainfall, but contain steep slopes and low elevation that is conducive to the convergence of precipitation in low-lying areas. Therefore, the risk of waterlogging in these areas should also receive attention.

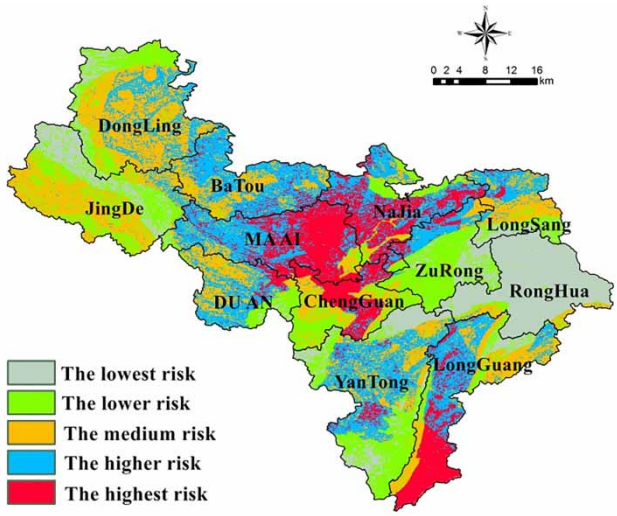


**Figure 4** | Karst waterlogging assessment map of DeBao County.

**Table 4** | Risk level percentage by area in DeBao County

| Risk level | The lowest risk | The lower risk | The medium risk | The higher risk | The highest risk |
|---|---|---|---|---|---|
| Percentage (%) | 15.33 | 19.31 | 24.64 | 26.42 | 14.30 |
| Areas ($km^2$) | 394.75 | 497.23 | 634.48 | 680.32 | 368.22 |

## DISCUSSION

According to the low-lying waterlogged area statistics surveyed in 2010, the results of risk assessment can be considered to be in line with actual situations. The statistics were provided by the Flood Control and Drought Relief Office of Debao County. The statistics show that DongLing, BaTou, MA AI, NaJia, and LongSang have experienced waterlogging disasters in history and the cumulative number of people affected is up to 6,074 people. The water depth of the affected area in NaJia is 0.5 to 1 m because the area is surrounded by mountains, and the disaster area of MA AI was affected by 1.5 to 2 m of water depth. The results of risk assessment are, therefore, in line with the survey results.

Although the direct external cause of risk is the rain, it could be seen from the risk-zoning map that the groundwater could be an important factor contributing to the risk of karst waterlogging. There is a strong possibility that the disaster occurred because the groundwater flows backwards from the sinkhole. In particular, this is the difference between karst waterlogging and a traditional flood disaster. It is suggested that studying the laws of groundwater variation could be helpful to researching the risk of karst

waterlogging. In addition, there are many research methods of risk analysis and system theory for risk analysis of flood disaster that could serve as a reference in karst waterlogging research.

## CONCLUSIONS

Karst waterlogging is a characteristic of karst terrain. Due to the influence of climate change and human activities, karst waterlogging has become more dangerous to people and the environment, and thus has attracted more attention. In this study, a synthetic assessment and analysis method was proposed involving nine risk indices, guided by the theory of natural disaster risk assessment. The conventional AHP was combined with entropy theory to calculate the weights of indicators that integrated subjective and objective weights. The SPA model was applied to depict connections between the indicators of karst waterlogging risk. Using a conditioned query and an overlay technique in GIS, a detailed risk-zoning map was obtained.

The evaluation results show that the disaster risk level was generally higher along the central axis of DeBao County, rather than in the eastern and western regions. Approximately 34.64% of the county is located in a high-risk region, including BaTou, MA AI, NaJia, ChengGuan, LongSang, LongGuang, and YanTong. MA AI and NaJia, especially, are located in a particularly prominent risk zone, and therefore those regions could serve as a demonstration area for the implementation of disaster prevention and control measures. The results would provide the government and local residents with an appropriate overview of karst waterlogging that would be beneficial to the practice of managing waterlogging risk. Transferring property of a high risk area reduces the loss caused by karst waterlogging disasters.

Karst waterlogging is a particular disaster in karst areas because both surface hydrological and groundwater processes interact with each other. In order to analyze karst waterlogging risk, hydrological and geological conditions need to be taken into account. Investigating the dynamic variation law of underground river discharge can help in karst waterlogging risk studies.

## ACKNOWLEDGEMENTS

This study was financially supported by the Science and Technology Project of Guangxi Water Resources Department (201530).

## REFERENCES

Aurit, M. D., Peterson, R. O. & Blanford, J. I. 2013 A GIS analysis of the relationship between sinkholes, dry-well complaints and groundwater pumping for frost-freeze protection of winter strawberry production in Florida. *PloS One* **8** (1), e53832.

Chaudhary, P., Chhetri, S. K., Joshi, K. M., Shrestha, B. M. & Kayastha, P. 2016 Application of an analytic hierarchy process (AHP) in the GIS interface for suitable fire site selection: a case study from Kathmandu Metropolitan City, Nepal. *Socio-Economic Planning Sciences* **53**, 60–71.

Chenini, I., Mammou, A. B. & El May, M. 2010 Groundwater recharge zone mapping using GIS-based multi-criteria analysis: a case study in Central Tunisia (Maknassy Basin). *Water Resources Management* **24** (5), 921–939.

Díez-Herrero, A., Ballesteros-Cánovas, J. A., Bodoque, J. M. & Ruiz-Villanueva, V. 2012 A new methodological protocol for the use of dendrogeomorphological data in flood risk analysis. *Hydrology Research* **44** (2), 234–247.

Fagbote, E. O., Olanipekun, E. O. & Uyi, H. S. 2014 Water quality index of the ground water of bitumen deposit impacted farm settlements using entropy weighted method. *International Journal of Environmental Science and Technology* **11** (1), 127–138.

Feng, L. H., Sang, G. S. & Hong, W. H. 2014 Statistical prediction of changes in water resources trends based on set pair analysis. *Water Resources Management* **28** (6), 1703–1711.

Fiorillo, F. 2009 Spring hydrographs as indicators of droughts in a karst environment. *Journal of Hydrology* **373** (3), 290–301.

Guo, F. & Jiang, G. 2010 Problems of flood and drought in a typical peak cluster depression karst area (SW China). In: *Advances in Research in Karst Media* (B. Andreo, F. Carrasco, J. Duran & J. LaMoreaux, eds). Springer, Berlin Heidelberg, pp. 107–113.

Jia, Z., Fan, Z. & Chen, S. 2011 Fuzzy comprehensive evaluation of peak load unit based on entropy weight. In: *Intelligent Computing and Information Science* (R. Chen, ed.). Springer-Verlag, Berlin Heidelberg, pp. 737–743.

Jiang, W., Deng, L., Chen, L., Wu, J. & Li, J. 2009 Risk assessment and validation of flood disaster based on fuzzy mathematics. *Progress in Natural Science* **19** (10), 1419–1425.

Keqiang, H., Jia, Y., Wang, F. & Lu, Y. 2011 Overview of karst geo-environments and karst water resources in north and south China. *Environmental Earth Sciences* **64** (7), 1865–1873.

Kourgialas, N. N. & Karatzas, G. P. 2016 A flood risk decision making approach for Mediterranean tree crops using GIS;

climate change effects and flood-tolerant species. *Environmental Science & Policy* **63**, 132–142.

Lai, C., Chen, X., Chen, X. & Wang, Z. 2015 A fuzzy comprehensive evaluation model for flood risk based on the combination weight of game theory. *Natural Hazards* **77** (2), 1243–1259.

Liu, S., Chan, F. T. & Ran, W. 2016 Decision making for the selection of cloud vendor: an improved approach under group decision-making with integrated weights and objective/subjective attributes. *Expert Systems with Applications* **55**, 37–47.

Long, D., Shen, Y., Sun, A., Hong, Y., Longuevergne, L., Yang, Y., Li, B. & Chen, L. 2014 Drought and flood monitoring for a large karst plateau in Southwest China using extended GRACE data. *Remote Sensing of Environment* **155**, 145–160.

Macdonald, A. M., Lapworth, D. J., Hughes, A. G., Auton, C. A., Maurice, L., Finlayson, A. & Gooddy, D. C. 2014 Groundwater, flooding and hydrological functioning in the Findhorn floodplain, Scotland. *Hydrology Research* **45** (6), 755–773.

Malagò, A., Efstathiou, D., Bouraoui, F., Nikolaidis, N. P., Franchini, M., Bidoglio, G. & Kritsotakis, M. 2016 Regional scale hydrologic modeling of a karst-dominant geomorphology: the case study of the Island of Crete. *Journal of Hydrology* **540**, 64–81.

Mimi, Z. A. & Assi, A. 2009 Intrinsic vulnerability, hazard and risk mapping for karst aquifers: a case study. *Journal of Hydrology* **364** (3), 298–310.

Saaty, T. L. 1990 How to make a decision: the analytic hierarchy process. *European Journal of Operational Research* **48** (1), 9–26.

Stefanidis, S. & Stathis, D. 2013 Assessment of flood hazard based on natural and anthropogenic factors using analytic hierarchy process (AHP). *Natural Hazards* **68** (2), 569–585.

Wang, W. S., Jin, J. L., Ding, J. & Li, Y. Q. 2009 A new approach to water resources system assessment – set pair analysis method. *Science China Technological Sciences* **52** (10), 3017–3023.

Wang, Z., Ma, H., Lai, C. & Song, H. 2012 Set pair analysis model based on GIS to evaluation for flood damage risk. *Procedia Engineering* **28** (8), 196–201.

Wang, T., Chen, J. S., Wang, T. & Wang, S. 2015 Entropy weight-set pair analysis based on tracer techniques for dam leakage investigation. *Natural Hazards* **76** (2), 747–767.

Wang, Y., Jing, H., Yu, L., Su, H. & Luo, N. 2017 Set pair analysis for risk assessment of water inrush in karst tunnels. *Bulletin of Engineering Geology & the Environment* **76** (3), 1199–1207. doi.10.1007/s10064-016-0918-y.

Yang, X. L., Ding, J. H. & Hou, H. 2013 Application of a triangular fuzzy AHP approach for flood risk evaluation and response measures analysis. *Natural Hazards* **68** (2), 657–674.

Zhao, K. Q. & Xuan, A. L. 1996 Set pair theory-a new theory method of non-define and its applications. *Systems Engineering* **14** (1), 18–23 (in Chinese).

Zheng, Z. & Qi, S. 2011 Potential flood hazard due to urban expansion in the karst mountainous region of North China. *Regional Environmental Change* **11** (3), 439–440.

Zhou, W. 2007 Drainage and flooding in karst terranes. *Environmental Geology* **51** (6), 963–973.

Zou, Z. H., Yi, Y. & Sun, J. N. 2006 Entropy method for determination of weight of evaluating indicators in fuzzy synthetic evaluation for water quality assessment. *Journal of Environmental Sciences* **18** (5), 1020–1023.

Zou, Q., Zhou, J., Zhou, C., Song, L. & Guo, J. 2013 Comprehensive flood risk assessment based on set pair analysis-variable fuzzy sets model and fuzzy AHP. *Stochastic Environmental Research and Risk Assessment* **27** (2), 525–546.

First received 16 November 2016; accepted in revised form 6 June 2017. Available online 21 August 2017