Available online at www.sciencedirect.com

ScienceDirect

Progress in Natural Science 19 (2009) 1419–1425

Progress in Natural Science

www.elsevier.com/locate/pnsc

Short communication

Risk assessment and validation of flood disaster based on fuzzy mathematics

Weiguo Jiang [a,b], Lei Deng [a,b], Luyao Chen [c], Jianjun Wu [a,b,*], Jing Li [a,b]

[a] Key Laboratory of Environment Change and Natural Disaster, MOE, Beijing Normal University, Beijing 100875, China
[b] Academy of Disaster Reduction and Emergency Management, MOCA&MOE, Beijing Normal University, Beijing 100875, China
[c] Department of Geography, Texas A&M University, College Station, TX 77843-3147, USA

Received 6 October 2008; received in revised form 24 November 2008; accepted 14 December 2008

Abstract

Floods often take place around rivers and plains, which indicates a higher risk of flooding in these areas. This paper adopts fuzzy comprehensive assessment (FCA), simple fuzzy classification (SFC), and the fuzzy similarity method (FSM) to assess flood disaster risk in Kelantan, Malaysia. Validation data, such as the flooded area, paddy area, urban area, residential area, and refuges, were overlaid to validate and analyze the accuracy of flood disaster risk. The results show that (1) 70–75% of flooded areas lie within the higher and highest risk zones, which shows an effective assessment accuracy; (2) paddy, built-up, and residential areas concentrated in the higher and highest risk zones are more likely to be destroyed by flood disasters; (3) 200–225 refuges in the higher and highest risk zones account for around 50% of all refuges, which means that more refuges should be built in the higher and highest risk zones to meet the accommodation requirement; (4) three methods proved to be feasible and effective in evaluating flood disaster risk, among which FCA is more suitable for the study area than the two other methods.

Keywords: Flood disaster; Risk assessment; Risk validation; Fuzzy mathematics; Malaysia

1. Introduction

Flood disasters are among the world's most frequent and damaging types of disaster [1]. Flood hazard, risk, and disasters are the products of an interaction between environmental and social processes [2]. Risk is defined as the expected losses (of lives, persons injured, property damaged, and economic activity disrupted) due to a particular hazard for a given area and reference period. Based on mathematical calculations, risk is the product of hazard and vulnerability [3,4]. The assessment of flood disaster risk is a synthetic evaluation and analysis of several factors. These factors are the stability of the disaster-breeding setting, the risk of the disaster-inducing environment, and the vulnerability of the hazards-bearing body [5–7]. Flood risk assessment has widely been used in flood insurance, floodplains management, flood disaster evacuation, disaster warning, disaster evaluation, flood influence evaluation, and improving the public's flood risk awareness. It is also an important scientific basis for flood disaster risk management and decision-making. Floods often take place around rivers and plains, which indicates a higher risk in these areas. Flood disaster risk is essentially a three-dimensional concept related to non-profitability, uncertainty, and complexity. The geographic, remote sensing, and statistical information used to present these hazard factors have multiplicity, complexity, uncertainty, inaccuracy, and diversity of assessment methods, which make risk assessment and validation of flood disasters a worldwide problem in the field of natural science and technology [8,9]. Given the differences in flood disaster risk assessment indexes, applica-

* Corresponding author. Tel.: +86 10 58802923; fax: +86 10 58806178.
E-mail address: wjj@ires.cn (J. Wu).

1002-0071/$ - see front matter
doi:10.1016/j.pnsc.2008.12.010

tion theories, and application levels, numerous risk assessment methods have been developed. Common methods include the mathematical statistic analysis method [10–12], the uncertainty method [9,13], the decision-making analysis method [14–16], and the system dynamic method [17]. The fuzzy mathematical method, a type of uncertainty method, has an advantage in the complex uncertainty problem-solving and analysis used in flood disaster risk assessment. Many scholars use the fuzzy mathematical method to study flood disaster forecasting and risk assessment [9,18,19].

The contents, methods, and techniques of flood disaster evaluation have been studied all over the world, but few studies on the validation and analysis of flood disasters with multi-data have found ideal assessment methods. This is because there are fewer studies on coastal flood disaster risk. In response to this apparent lack of research, this paper adopts fuzzy comprehensive assessment (FCA), simple fuzzy classification (SFC), and the fuzzy similarity method (FSM) to assess flood disaster risk to coastal zones by overlaying validation data such as the flooded area, paddy area, urban area, residential area, and refuges. In addition, this study aims to validate and analyze the accuracy of flood disaster risk, and to find an optimal method of coastal flood disaster risk assessment.

2. Study area and data

2.1. Study area

Seven counties of Malaysia: Kota Bharu, Tumpat, Bachok, Pasir Mas, Pasir Puteh, Tanah Merah, and Machang were selected as the study area. The study area, situated in the northeast of the Kelantan Delta in Malaysia (bordering Thailand in the northwest and next to the South China Sea in the northeast), is the main flood disaster area (Fig. 1). Due to the influence of strong northeast monsoons, flood disasters happen regularly from October to the following January. The area, threatened by flooding, is about 50% of the whole study area. The population threatened by flooding is about 60% of all the population in the study area. Flooding also results in huge economic losses every year [20]. In the year 2000, within the study area of 3240 km^2, the population was about 1,080,000. The gross domestic product (GDP) output was about 8,935,000 USD. Geographically, the area is high in the south and low in the north. Cocoa, palm oil, hevea rubber, and paddy rice are the main crops in the study area.

2.2. Study data

2.2.1. Assessment data

Flood disaster risk is determined by three ingredients comprising the stability of the disaster-breeding setting, the risk of the disaster-inducing environment, and the vulnerability of the hazards-bearing body [5–7]. According to the actual situation in Malaysia, three-day maximum rainfall data and rainstorm times were selected as the assessment indexes of the disaster-breeding factor. Elevation standard deviation (reciprocal is taken), drainage density, and vegetation coverage (reciprocal is taken) were selected as assessment indexes of the disaster-inducing factor. The total population in the unit area, the population of the young and the old in the unit area, the GDP output in the unit area, the number of motorbikes owned per household (the reciprocal is taken), televisions owned per household (the reciprocal is taken), and the proportion of paddy area were selected as assessment indexes of vulnerability [9,14–16]. Through standardized processing, 11 data layers were converted into grid data with a cell size of 100×100 m in the float data type (Fig. 2). There was a positive correlation between the numerical values of the

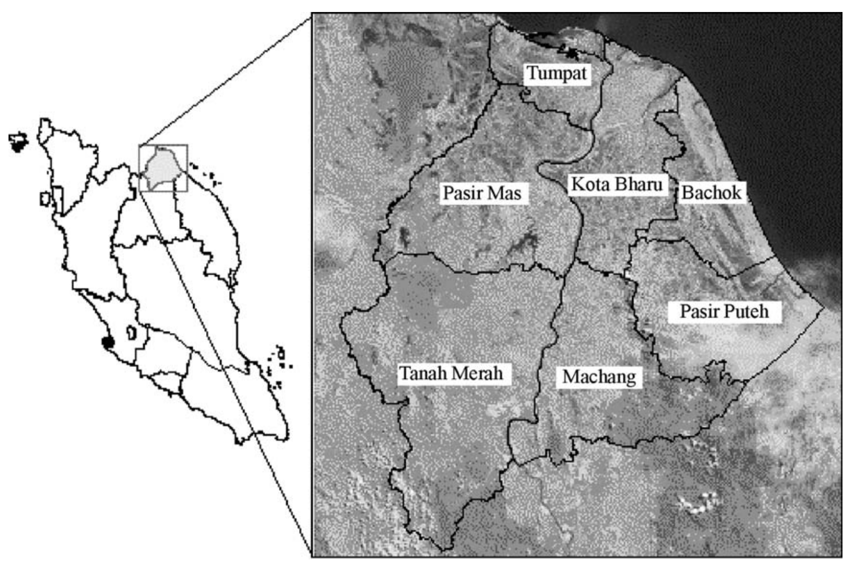

Fig. 1. Location of study area.

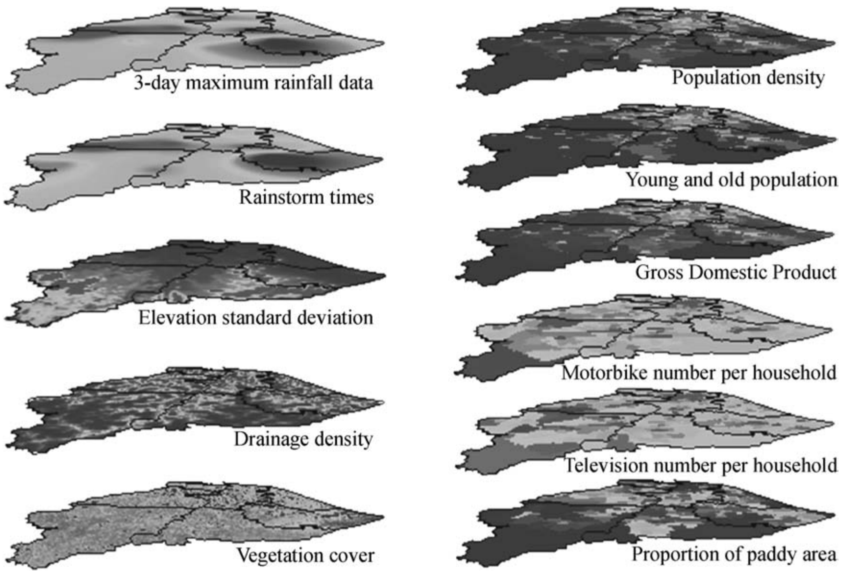

Fig. 2. Flood risk assessment dataset.

index data layers and the risk of flood disaster. This means that a higher value of index data coincides with higher flood disaster risk.

2.2.2. Validation data

Because there is no risk map or division standard for flood disaster in Malaysia, the flooded area, land use data (including the paddy and urban area), residential area, and refuges are selected as validation data in this paper (Fig. 3). Flooded area data were extracted from RADARSAT images with a resolution of 15 m for the years 1998, 2003, and 2004. Land use data were extracted from LANDSAT images with a resolution of 15 m for the year 2000 through the object-oriented classification method. Residential area data were acquired from SPOT images with a resolution of 2.5 m for the year 2004 through visual interpretation. According to the refuge data, there were 424 disaster shelters in the study area, which accommodate 94,685 people. This covers 70% of the population.

3. Methodology

This study adopts fuzzy synthetic evaluation (FSE) to assess flood disaster risk. FSE divides data into several categories according to predetermined quality criteria. This eliminates the possible fuzziness. FSE then synthesizes and evaluates several individual components of a process as a whole [21].

The flood disaster risk is divided into five classes: lowest risk zone, lower risk zone, medium risk zone, higher risk zone, and highest risk zone. Grade interval values can be obtained through the statistics of standard deviation and the average value in each raster layer. The minimum value of the standard deviation difference and average is regarded as the grade interval value (Δ) of the raster data fuzzy subset [9,14]. Each thematic raster data is graded equally according to the same interval value; thus, each has 5 interval points: D_1, D_2, D_3, D_4, and D_5 (Table 1).

Through the piecewise linear function (descending semi-trapezoid, ascending semi-trapezoid, and triangle) in fuzzy mathematics (Fig. 4), the membership function of each grade can be determined and the assessment indexes can be processed by the fuzzy subset classification method. u_{ij} is the membership level of index i and grade j,

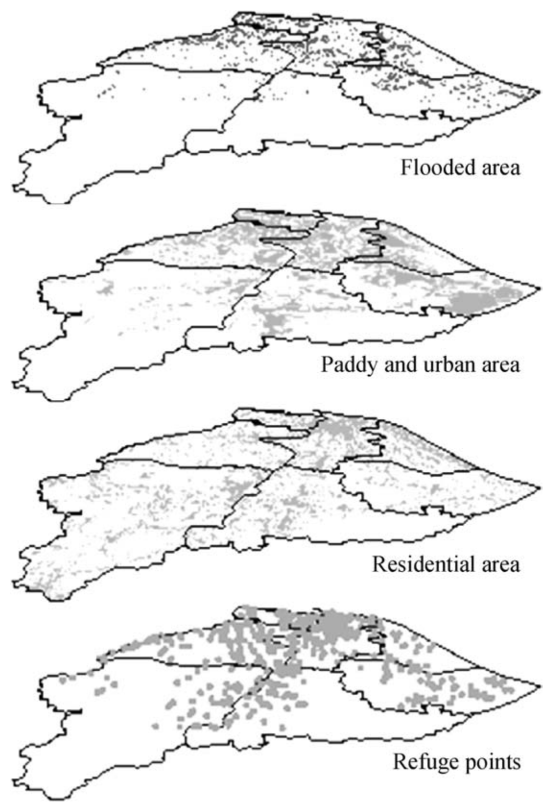

Fig. 3. Flood risk validation dataset.

Table 1
The interval value of every factor.

Index type	Δ	D_1	D_2	D_3	D_4	D_5
Maximum 3-day precipitation (mm)	60	130	190	250	310	370
Times of the rainstorm (times/year)	21	20	41	61	81	101
Vegetation coverage (the reciprocal is taken)	0.14	1.23	1.37	1.51	1.65	1.79
Drainage density (m/m^2)	0.0012	0.0006	0.0018	0.0030	0.0042	0.0054
Elevation standard deviation (the reciprocal is taken) (m)	0.028	0.014	0.042	0.07	0.098	0.126
Population in the unit area (population/10^5 m^2)	2.75	1.38	4.13	6.88	9.63	12.38
Old/young population in the unit area (population/10^5 m^2)	1.09	0.054	1.144	2.234	3.324	4.414
Motorbikes owned per household (the reciprocal is taken)	0.03	0.07	0.10	0.13	0.17	0.20
Televisions owned per household (the reciprocal is taken)	0.22	0.51	0.73	0.95	1.17	1.39
Proportion of the paddy area (%)	23	11	33	56	79	100
Gross domestic product (USD)	2316	1158	3474	5790	8106	10422

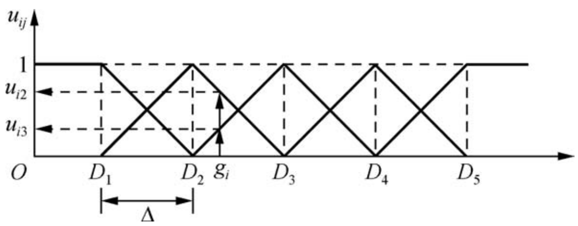

Fig. 4. Fuzzy set of membership functions.

$i = 1,2, \ldots,11$; $j = 1,2,3,4,5$. g_i represents the raster data of index i. In Fig. 4, g_i has two membership levels: u_{i2} and u_{i3}.

The value of the fuzzy membership function of each factor related to the five assessment levels can be calculated by a set of formulas as follows (Eqs. (1)–(5)):

$$u_{i1}(g_i) = \begin{cases} 1 & 0 \le x \le D_1 \\ \frac{D_2-g_i}{D_2-D_1} & D_1 < x < D_2 \\ 0 & x \ge D_2 \end{cases} \tag{1}$$

$$u_{i2}(g_i) = \begin{cases} 0 & g_i \le D_1 or g_i \ge D_3 \\ \frac{g_i-D_1}{D_2-D_1} & D_1 < g_i < D_2 \\ 1 & g_i = D_2 \\ \frac{D_3-g_i}{D_3-D_2} & D_2 < g_i < D_3 \end{cases} \tag{2}$$

$$u_{i3}(g_i) = \begin{cases} 0 & g_i \le D_2 or g_i \ge D_4 \\ \frac{g_i-D_2}{D_3-D_2} & D_2 < g_i < D_3 \\ 1 & g_i = D_3 \\ \frac{D_4-g_i}{D_4-D_3} & D_3 < g_i < D_4 \end{cases} \tag{3}$$

$$u_{i4}(g_i) = \begin{cases} 0 & g_i \le D_3 or g_i \ge D_5 \\ \frac{g_i-D_3}{D_4-D_3} & D_3 < g_i < D_4 \\ 1 & g_i = D_4 \\ \frac{D_5-g_i}{D_5-D_4} & D_4 < g_i < D_5 \end{cases} \tag{4}$$

$$u_{i5}(g_i) = \begin{cases} 0 & g_i \le D_4 \\ \frac{g_i-D_4}{D_5-D_4} & D_4 < g_i < D_5 \\ 1 & g_i \ge D_5 \end{cases} \tag{5}$$

The evaluation matrix, $\boldsymbol{R}$, is generated by the membership values and corresponding flood disaster risk parameters.

$$\boldsymbol{R} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{15} \\ r_{21} & r_{22} & \cdots & r_{25} \\ \vdots & \vdots & & \vdots \\ r_{i1} & r_{i2} & \cdots & r_{i5} \end{bmatrix} \tag{6}$$

$r_{ij} = u_j(g_i), i = 1, 2, \cdots, 11, j = 1, 2, \cdots, 5.$

As the risk assessment of flood disaster is a complicated problem with multi-levels, this paper applies raster data layers as a risk index of flood disaster, and thus establishes a hierarchical system of flood disaster risk assessment indexes based on the analytic hierarchy process (AHP). Also the weights of the 11 factors are reasonably determined which separately are 0.3598, 0.1798, 0.0485, 0.1603, 0.0882, 0.0429, 0.0214, 0.0643, 0.0150, 0.0075, and 0.0123.

Once the weights, membership functions, and the single-factor matrix $\boldsymbol{R}$ are determined, three different fuzzy synthetic assessment methods (fuzzy comprehensive assessment, simple fuzzy classification, and the fuzzy similarity method) are selected [22–24]. After comparison and validation of the three methods, a proper fuzzy integrated risk assessment of flood disaster can be retrieved.

In fuzzy comprehensive assessment (FCA), the fuzzy weight vector $\overline{\boldsymbol{W}}$ is multiplied by the fuzzy relation matrix to get the resulting vector $\overline{\boldsymbol{B}}_{\boldsymbol{j}}$.

$$\overline{\boldsymbol{B}}_{\boldsymbol{j}} = \overline{\boldsymbol{W}} \times \boldsymbol{R} = [\boldsymbol{w_1}, \boldsymbol{w_2}, \cdots, \boldsymbol{w_i}] \times \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1j} \\ r_{21} & r_{22} & \cdots & r_{2j} \\ \vdots & \vdots & & \vdots \\ r_{i1} & r_{i2} & \cdots & r_{ij} \end{bmatrix} \tag{7}$$

In simple fuzzy classification (SFC), the resulting vector $\overline{\boldsymbol{B}}_{\boldsymbol{j}}$ can be calculated by weight w_i and r_{ij}.

$$\overline{\boldsymbol{B}}_{\boldsymbol{j}} = \sqrt{\frac{\sum_{1}^{i}(\boldsymbol{w_i} \times \boldsymbol{r_{ij}})^2}{\sum_{1}^{j}\sum_{1}^{i}(\boldsymbol{w_i} \times \boldsymbol{r_{ij}})^2}} \tag{8}$$

In the fuzzy similarity method (FSM), matrix $\boldsymbol{R}$ is transformed to matrix $\boldsymbol{R}'$ through standard deviation transformation, and then the resulting vector $\overline{\boldsymbol{B}}_{\boldsymbol{j}}$ can be achieved through the max–min method.

$$r' = \begin{bmatrix} r'_{11} & r'_{12} & \cdots & r'_{1j} \\ r'_{21} & r'_{22} & \cdots & r'_{2j} \\ \vdots & \vdots & & \vdots \\ r'_{i1} & r'_{i2} & \cdots & r'_{ij} \end{bmatrix} \quad (9)$$

$$\overline{B}_j = \frac{\sum_I^i \min[r_{ij}(x), r'_{ij}]}{\sum_I^i \max[r_{ij}(x), r'_{ij}]} \quad (10)$$

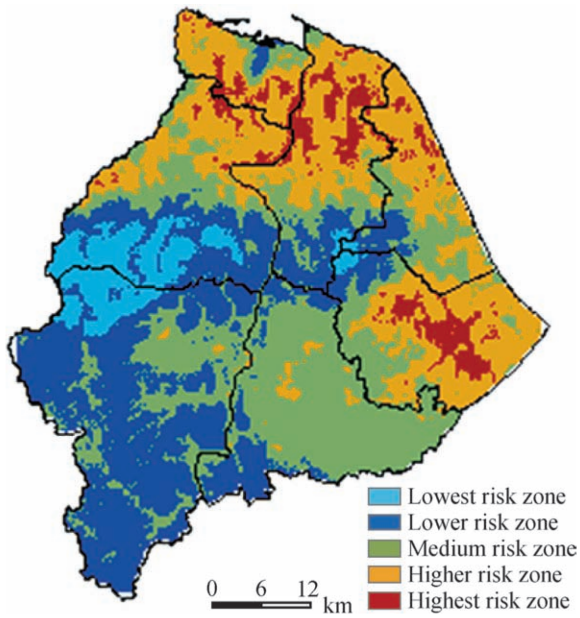

Fig. 6. Integrated risk maps of flood disaster based on SFC.

4. Results analysis

4.1. Results evaluation

Three integrated flood disaster risk maps were obtained through FCA, SFC, and FSM (Figs. 5–7). Each map includes five zones: lowest risk, lower risk, medium risk, higher risk, and highest risk. The spatial distribution of the integrated risk of flood disaster decreases from northeast to southwest. The region of the highest and higher risk zones lie in the middle and northern parts of Kota Bharu County, part of Tumpat County, the eastern and northern parts of Bachok County, the northern part of Pasir Mas County, and the middle and southwest parts of Pasir Puteh County. The higher and highest risk zones are located in the coastal areas or areas along rivers. These areas are characterized by plains, with a high density of agriculture, economic activity, and population. The lowest and lower risk zones are in the central and southwest of the study area. These areas have high mountains and less population.

From Table 2, the bias among the three methods is relatively small, with a maximum bias of 3.52%. In contrast, however, the highest and higher risk zones of FCA are greater in area than that of SFC and FSM. Contrarily, the areas of the lower and lowest risk zones of FCA are smaller than that of SFC and FSM. In the distribution pattern of risk zones based on the FCA method, the highest risk zone accounts for 6.91%, the higher risk zone 26.69%, the medium risk zone 33.95%, the lower risk zone 26.71%, and the lowest risk zone 6.24%. The variation of percentage among risk zones is inconspicuous in the three risk maps (Table 2).

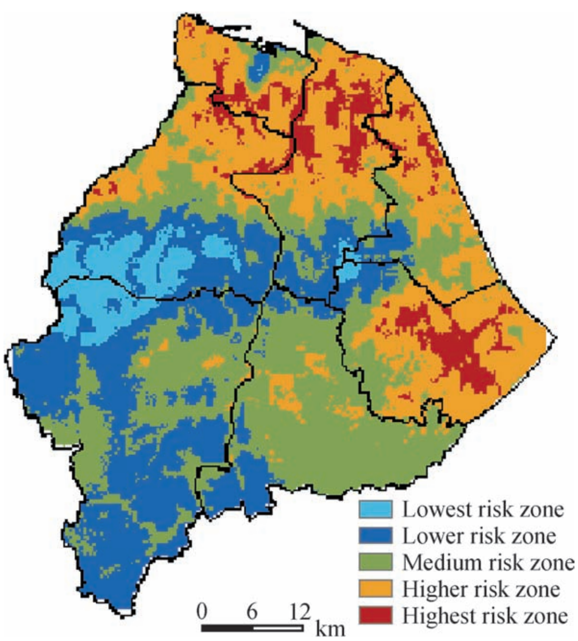

Fig. 5. Integrated risk maps of flood disaster based on FCA.

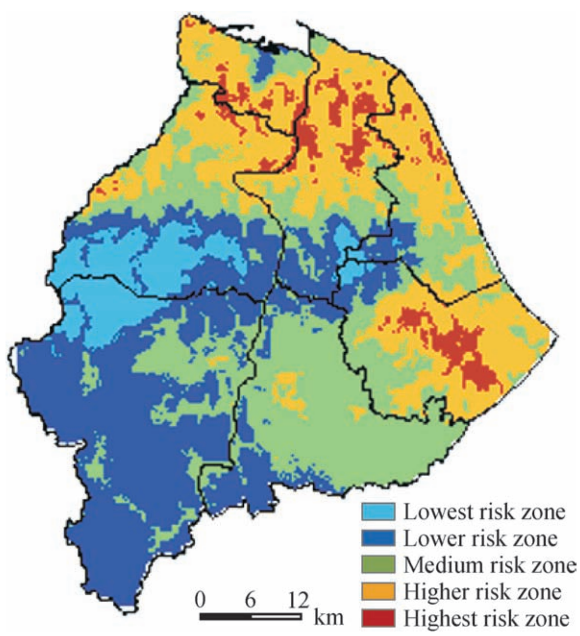

Fig. 7. Integrated risk maps of flood disaster based on FSM.

Table 2
Percentage of risk zones using three methods.

Method	Highest risk zone	Higher risk zone	Medium risk zone	Lower risk zone	Lowest risk zone
FCA	6.91	26.69	33.95	26.21	6.24
SFC	5.66	25.46	34.60	27.46	6.82
FSM	4.55	25.65	32.84	29.73	7.23

4.2. Risk validation

Validation is to judge the risk assessment accuracy of flood disasters by using other data to validate the reliability of the higher risk and highest risk zones. It is generally acknowledged that (i) the flooded area lies within the higher and highest risk zones; (ii) there should be more refuges in the higher and highest risk zones; (iii) in the higher and highest risk zones, paddy, built-up, and residential areas are densely located. These data are overlaid with integrated risk maps of flood disaster. The percentage of validation data is calculated in the higher and highest risk zones. Based on the results, the integrated risk assessment accuracy of flood disasters can be validated, and the optimum method can be determined.

In Table 3, results of the three assessments are pretty similar: the difference between risk levels of flooded areas is = <4.54%; the difference between risk levels of the paddy, built-up, and residential areas is = <3.84%; the difference between risk levels of refuges is less than 17. According to the validation results of flood disasters, 70–75% of flooded areas belong to the higher risk and highest risk zones. This shows high assessment accuracy. According to other validation results, 58–61% of paddy areas, 64–67% of built-up areas, and 52–56% of residential areas belong to the higher and highest risk zones. This shows that paddy, built-up, and residential areas are concentrated in the higher and highest risk zones and are more likely to be destroyed by flood disasters. There are 200–225 refuges, accounting for about 50%, located in the higher and highest risk zones. These refuges can only accommodate about 60,000 people. This means that more refuges should be built within the area to meet the accommodation requirements. Because validation results are similar and there is little bias between different methods, the three methods are feasible and effective to evaluate flood disaster risk. According to the amount of risk level of flooded areas, the precision of FCA is higher than the other two methods; therefore, FCA is more suitable for the study area than the two other methods.

Table 3
Validation results of the three assessments.

Assessment	Flooded area (%)	Paddy area (%)	Urban area (%)	Residential area (%)	Refuges (No.)
FCA	75.43	60.14	67.10	55.87	224
SFC	72.15	61.06	65.64	53.50	212
FSM	70.89	58.96	64.12	52.03	207

5. Conclusion

The Kelantan delta is a fertile coastal plain, but it faces strong northeast monsoons; as a result, flood disasters occur regularly in the coastal delta resulting in higher flood hazard and flood risk than other coastal zones of Malaysia.

Either flood disaster risk itself or risk level classification has fuzziness and uncertainty. Thus fuzzy mathematics can effectively be applied to analyze uncertainty problems. Fuzzy mathematics, one of the widely used methods in multi-index synthetic evaluation, has widely been applied in natural disaster studies.

Risk assessment and validation of flood disasters are difficult problems in the natural disaster field. Due to the lack of risk maps and division standards of flood disasters in Malaysia, this study uses data of flooded areas, paddy area, urban area, residential area, and refuges as validated data. This paper validates and compares the results of three assessment methods: FCA, SFC, and FSM, to find an optimal assessment according to the proportion of validation data in the higher and highest risk zones. From the validation results of the flooded areas, 70–75% of flooded areas belong to the higher and highest risk zones. This shows high validation accuracy. Paddy, built-up, and, residential areas concentrated in the higher and highest risk zones are more likely to be destroyed by flood disasters. The establishment of refuges is not proper because refuges in the higher and highest risk zones only account for 50% of all refuges and can only accommodate 60,000 people. Because the refuges cannot meet the requirements of the rescue, more refuges should be made.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant Nos. 40701172 and 40671122) and the International Program for Cooperation in Science and Technology (2007DFA20640). All the data for this study were provided by the Malaysian Centre for Remote Sensing (MACRES). The authors express their sincere gratitude to the Director of MACRES and the MACRES Airborne Remote Sensing (MARS) project manager for their continuous guidance and support in this study. The authors also thank all MARS project team members, MACRES research officers, Technology Park Malaysia, and the staff of Cilix Corp. Sdn. Bhd. for their precious comments and valuable assistance.

References

[1] International Federation of Red Cross and Red Crescent Societies. World disaster report. Oxford: Oxford University Press; 1998.

[2] Parker DJ. Introduction to floods and flood management. Floods, vol. 1. London: Routledge; 2000, p. 3–39.

[3] United Nations Department of Humanitarian Affairs (UNDHA). Internationally agreed glossary of basic terms related to disaster management. Geneva; 1992.
[4] World Meteorological Organization (WMO). Comprehensive risk assessment for natural hazards. Geneva: WMO/TD No 955, Switzerland; 1999.
[5] Xu YP, Li LG, Cai GM. Flood risk map system in medium and small basins in China. Sci Geogr Sin 2004;24(4):452–7, [in Chinese].
[6] Hu BQ, Jiang MX, Jun SL, et al. Application of map and file information visualization system to comprehensive division of natural disasters taking the Changjiang valley as an example. Chin Geogr Sci 2001;11(4):326–35.
[7] Huang CF. Theory and application of natural disaster risk assessment. Beijing: Science; 2004, [in Chinese].
[8] Wang XZ, Shi WZ, Wang SL. Process of fuzzy spatial information. Wuhan: Wuhan University Press; 2003, [in Chinese].
[9] Jiang WG, Li J, Li ZW, et al. Fuzzy assessment of the population risk of flood disaster. J Hunan Univ [Nat Sci] 2008;35(9):84–7, [in Chinese].
[10] Pandey GR, Nguyen VTV. A comparative study of regression based methods in regional flood frequency analysis. J Hydrol 1999;225:92–101.
[11] Jiang WG, Li J, Wang L. Compositive analysis of global flood disaster from 1950 to 2004. J Beijing Norm Univ [Nat Sci] 2006;42(5):530–3, [in Chinese].
[12] Hundecha Y, Bardossy A, Theisen HW. Development of a fuzzy logic based rainfall–runoff model. Hydrol Sci J 2001;46(3):363–77.
[13] Li X, Wang XY, Li YL. Forecasting flood disasters in the Chaohu lake basin based on Grey-Markov theory. J China Hydrol 2006;26(4):43–6, [in Chinese].
[14] Zhou CH, Wan Q, Huang SF. A GIS-based approach to flood risk zonation. Acta Geogr Sin 2000;55(1):15–24, [in Chinese].
[15] Gao JX, Pan YZ, Liu HY. Assessment on regional vulnerability to flood. Res Environ Sci 2004;17(6):31–4, [in Chinese].
[16] Tang C, Zhu J. A GIS based regional torrent risk zonation. Acta Geogr Sin 2005;60(1):87–94, [in Chinese].
[17] Shi PJ. Theory on disaster science and disaster dynamics. J Nat Disaster 2002;11(3):1–9, [in Chinese].
[18] Mao DH, Wang LH. Diagnosis and assessment on vulnerability of the urban flood-waterlogged disaster in human province. Resour Environ Yangtze Basin 2002;11(1):89–93, [in Chinese].
[19] Nayak PC, Sudheer KP, Ramasastri KS. Fuzzy computing based rainfall – runoff model for real time flood forecasting. Hydrol Process 2005;19:955–68.
[20] Department of Irrigation and Drainage (DID). National register of river basins. Final report, vol. 2.8. Updating of condition of flooding in Malaysia: State report for Kelantan; 2003.
[21] Lu RS, Lo SL, Hu JY. Analysis of reservoir water quality using fuzzy synthetic evaluation. Stochastic Environ Res Risk Assess 1999;13:327–36.
[22] Chang NB, Chen HW, Ning SK. Identification of river water quality using the fuzzy synthetic evaluation approach. J Environ Manage 1999;63(3):293–305.
[23] Haiyan W. Assessment and prediction of overall environmental quality of Zhuzhou city, Hunan Province, China. J Environ Manage 2002;66(3):329–40.
[24] Guleda OE, Ibrahim D, Halil H. Assessment of urban air quality in Istanbul using fuzzy synthetic evaluation. Atmos Environ 2004;38:3809–15.