

The Practical Application of Fuzzy Information Analysis in Flood Forecasting

Lihua Feng¹ and Weihu Hong²

¹ Department of Geography, Zhejiang Normal University, Jinhua 321004, China

² Department of Mathematics, Clayton College & State University, GA 30260, USA
fenglh@zjnu.cn

Abstract. The relationship between the peak stage and peak discharge is influenced by many factors in the flood system. Therefore, different peak discharges may occur under the same peak stage, while the same peak discharge may also occur under the different peak stages. If the peak stage set with similar peak discharges is taken as the fuzzy subset in the stage universe, then the membership function of these fuzzy subsets can be hypothesized to manifest a normal distribution graph. According to a_k and b_k , the mid-value of the universe element of the peak stage can be substituted into the normal distribution graph, and the fuzzy relational matrix can be obtained. Thus, the peak discharge can be calculated according to the peak stage using the fuzzy deduction theory. The relationship between the peak stage and peak discharge as $Q = f(H)$ has an important impact on the determination of the peak discharge during the high-water level period in the flood forecast. In this paper, the fuzzy information analysis method is used to forecast the peak discharge, with the result in accordance with the actual event. This method can be seen as a new and effective method of flood prediction and forecasting.

Keywords: fuzzy information analysis; universe; fuzzy relational matrix; flood forecast.

1 Introduction

Flooding is a common occurrence in China, and one that causes serious damage and harm. It is estimated that of the total economical loss caused by all kinds of disasters, 40% was due to flooding. Floods occur frequently, affect large areas of the community, and hence constitute a huge threat to human life and property. Flood damage has been aggravated by the rapid progress of the economy in recent years [1].

However, flood systems are complicated systems with strong indeterminacy [2-3]. Therefore, a method of recognition of flood systems using the fuzzy information analysis method [4] is devised and discussed in this paper.

2 Method of Fuzzy Information Analysis

Selection of a system model with higher reliability is difficult. A simple method is to directly construct a fuzzy relational matrix R without any transition through language variables and message net blocks [5-6].

Peak stage H and peak discharge Q are the most important parameters for the formation of flood damage. Unfortunately, at present the exact relationship between the two parameters as

$$Q = f(H) \tag{1}$$

cannot be found. In fact, there exists a fuzzy relationship between the two parameters. Therefore, the universe ranges of the peak stage and peak discharge need to be determined first. For the purpose of flood prevention, the range of the peak stage can be selected as m classes, and therefore the universe of the peak stage should be $U_H = \{H_1, H_2, \dots, H_m\}$. The range of the peak discharge can also be selected as n classes and the corresponding universe of the peak discharge should be $U_Q = \{Q_1, Q_2, \dots, Q_n\}$.

The method of determining the fuzzy relation through hypothesizing of the distribution form can only be performed upon analysis of the physical process. The relationship between the peak stage and peak discharge has been influenced by many factors in the flood system. If the peak stage set with similar peak discharges is taken as the fuzzy subset in the stage universe, then the membership function $\mu_k(H)$ of these fuzzy subsets can be hypothesized to manifest a normal distribution graph as follows:

$$\mu_k(H) = \exp \left[- \left(\frac{H - a_k}{b_k} \right)^2 \right] \tag{2}$$

In the equation, k means the class of the peak discharge, with a_k and b_k as constants that can be obtained from the historical flood samples. Suppose the number of flood samples with the class of the peak discharge of k is N , then a_k and b_k can be calculated according to the following two equations:

$$a_k = \frac{1}{N} \sum_{j=1}^N H_j \tag{3}$$

$$b_k^2 = \frac{1}{N} \sum_{j=1}^N (H_j - a_k)^2 \tag{4}$$

In the equation H_j means the value of the peak stage of the j th sample.

According to a_k and b_k , the mid-value of the universe element H_i of the peak stage can be substituted into the equation (2), and the fuzzy relational matrix R can be obtained. If the mid-value of H_i is \bar{H}_i , then the element of R should be $r_{ik} = \mu_k(\bar{H}_i)$.

After determination of the fuzzy relational matrix R , the peak discharge Q can be calculated according to the peak stage H using the fuzzy deduction theory as follows:

$$Q = H \circ R \tag{5}$$

The operator “ \circ ” means the combination operation in the equation.

As to the first forecast model, the relative fuzzy expression can be represented as $H = 1/H_i$ using the concept concerning language variables in fuzzy mathematics. Thus, the following expression can be constructed:

$$Q = (0, \dots, 0, 1, 0, \dots, 0) \circ \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ r_{21} & r_{22} & \dots & r_{2n} \\ \vdots & \vdots & & \vdots \\ r_{m1} & r_{m2} & \dots & r_{mn} \end{bmatrix}$$

$$= \frac{r_{i1}}{Q_1} + \frac{r_{i2}}{Q_2} + \dots + \frac{r_{in}}{Q_n} \tag{6}$$

The above equation represents a fuzzy subset on the universe of the peak discharge. It is the result of the fuzzy recognition.

3 Application Illustration

The relationship between the peak stage and peak discharge as $Q = f(H)$ has an important impact on the determination of the peak discharge during the high-water level period in the flood forecast. Unfortunately, this relationship remains undetermined owing to the complexity of the factors that influence the flood. Therefore, different peak discharges may occur under the same peak stage, while the same peak discharge may also occur under the different peak stages. Table 1 lists data of the annual maximum peak stage H and annual maximum peak discharge Q surveyed in the Hankou hydrometric station of the Changjiang River of China. From the table it can be seen that the peak stage values H are about the same in 1948 and 1949 at the Hankou hydrometric station. Conversely, the two corresponding peak discharge values Q are different. In addition, the peak discharge values Q are about the same in 1935 and 1952, whilst the two corresponding peak stage values H are very different. Furthermore, a situation where the peak discharge value Q became even larger at a small peak stage value H was also observed.

Table 1. Data of the annual maximum peak stage H and annual maximum peak discharge Q in the Hankou hydrometric station of Changjiang River

Year	H (m)	Q (m ³ /s)	Year	H (m)	Q (m ³ /s)
1948	27.03	56000	1935	27.58	59300
1949	27.12	52700	1952	26.60	59500

In fact, a definite relationship can only exist within a certain range between the annual maximum peak stage H and annual maximum peak discharge Q , owing to the action of the factors that influence flooding. Therefore, flood data was selected for the

present discussion in a range from 1865 to 1970 at the Hankou hydrometric station of the Changjiang River. This area was scarcely influenced by human activity and so is less likely to be affected by extra factors, although the station was short of flood data for the period 1944-1945. The annual maximum peak stage H in the range was divided into nine classes. From this the universe of H should be $U_H = \{H_1, H_2, \dots, H_9\}$ as shown in Table 3. The annual maximum peak discharge Q in the range was also divided into ten classes. From these classes, the universe of Q should be $U_Q = \{Q_1, Q_2, \dots, Q_{10}\}$ as shown in Table 2. According to the calculation from equations (3) and (4), the values of a_k and b_k can be obtained (as listed in Table 2). As a result of the scarcity of the sample number of Q_8, Q_9, Q_{10} , several constants have been determined as $b_8^2 = b_9^2 = b_{10}^2 = 0.2$, which are just the average

Table 2. The range of the annual maximum peak discharge and its calculated results

U_Q	Range of peak discharge (m^3/s)	Sample number	a_k	b_k^2
Q_1	29000-34000	2	22.05	0.2550
Q_2	34000-39000	9	23.34	0.0498
Q_3	39000-44000	16	24.36	0.0634
Q_4	44000-49000	33	25.12	0.3665
Q_5	49000-54000	24	26.13	0.1747
Q_6	54000-59000	11	26.63	0.1683
Q_7	59000-64000	7	27.21	0.3147
Q_8	64000-69000	1	27.36	0.2000
Q_9	69000-74000	0	28.55	0.2000
Q_{10}	74000-79000	1	29.73	0.2000

Table 3. The fuzzy relational matrix between the annual maximum peak stage H and annual maximum peak discharge Q

U_H	$H(m)$	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	Q_7	Q_8	Q_9	Q_{10}
H_1	21-22	0.31	0	0	0	0	0	0	0	0	0
H_2	22-23	0.44	0	0	0	0	0	0	0	0	0
H_3	23-24	0	0.58	0	0	0	0	0	0	0	0
H_4	24-25	0	0	0.74	0.34	0	0	0	0	0	0
H_5	25-26	0	0	0	0.67	0.10	0	0	0	0	0
H_6	26-27	0	0	0	0	0.45	0.90	0.20	0.02	0	0
H_7	27-28	0	0	0	0	0	0.01	0.76	0.90	0	0
H_8	28-29	0	0	0	0	0	0	0	0	0.98	0
H_9	29-30	0	0	0	0	0	0	0	0	0.01	0.76

values of $b_1^2 \sim b_7^2$. Afterwards, the mid-value \bar{H}_i of the universe of the annual maximum peak stage was substituted into the equation (2). From this the fuzzy relational matrix R between H and Q can be obtained, as shown in Table 3.

With the help of the fuzzy relational matrix R , the annual maximum peak discharge Q can be obtained according to equation (5). The fit rate of the calculated result for the flood data from 1865 to 1970 in the Hankou hydrometric station is satisfactory at a rate of 73%.

Thus, the annual maximum peak discharge in the Hankou hydrometric station can be predicted according to the present fuzzy relational matrix R . For example, the annual maximum peak stage in 1971 in the Hankou hydrometric station is 24.21m. At this time the value of H can satisfy the relation as $H = 1/24.21 = 1/H_4$. Therefore, fuzzy deduction can be performed according to equation (5) and the result can be represented as follows:

$$Q = \frac{0.74}{Q_3} = \frac{0.34}{Q_4}$$

Table 4. The forecast of the annual maximum peak discharge Q in the Hankou hydrometric station and its fit conditions

Year	$H(m)$	U_H	U_Q	Range of peak discharge (m^3/s)	Q (m^3/s)	Fit Conditions
1971	24.21	H_4	Q_3	39000-44000	43700	√
1972	22.15	H_2	Q_1	29000-34000	36400	×
1973	26.85	H_6	Q_6	54000-59000	54300	√
1974	26.19	H_6	Q_6	54000-59000	54900	√
1975	25.00	H_4	Q_3	39000-44000	43800	√
1976	26.50	H_6	Q_6	54000-59000	58400	√

If taking the value of Q with the highest probability as the forecasted value, then the annual maximum peak discharge in 1971 can be forecasted in a range between $39000m^3/s$ and $44000m^3/s$. In fact, the actual value of the annual maximum peak discharge in 1971 is $Q=43700m^3/s$, an exact forecasted value. The annual maximum peak discharges from 1971 to 1976 in the Hankou hydrometric station were also forecasted, as listed in Table 4. The forecasted values within those five years are in accordance with the actual situations with the exception of 1972.

4 Conclusion

The relationship between the peak stage and peak discharge is hard to determine due to the complexity of the factors which can influence flooding. In fact, there exists a fuzzy relationship between the two terms. Therefore, the fuzzy information analysis method can be used to study the exact relationship between the above two terms. The

method can also be used to forecast the corresponding peak stage or peak discharge and precipitation runoff between the hydrometric stations in the upper and lower reaches of the same river. Although the result of the flood forecast is always within a certain range, the forecast can satisfy the demands of most situations. In conclusion, flood forecasting using the fuzzy information analysis technology is effective.

Acknowledgments. This work was supported by Zhejiang Provincial Science and Technology Foundation of China (No. 2006C23066).

References

1. Feng, L.H., Chen, L.R.: Three Large Floods Along the Yangtze River in the 20th Century. *Journal of Natural Disasters* 1 (2001) 8-11
2. Fan, Z.P., Ma, J., Zhang, Q.: An Approach to Multiple Attribute Decision Making Based on Fuzzy Preference Information on Alternatives. *Fuzzy Sets and Systems* 1 (2002) 101-106
3. Cordon, O., Moya, F., Zarco, C.: A New Evolutionary Algorithm Combining Simulated Annealing and Genetic Programming for Relevance Feedback in Fuzzy Information Retrieval Systems. *Soft Computing* 5 (2002) 308-319
4. Huang, C.H., Wang, J.D.: *Analysis of Fuzzy Information and Its Application*. Beijing Normal University Publisher (1992) 177-182
5. Nils, H., Peter, S.: Precautionary Saving and Fuzzy Information. *Economics Letters* 1 (2001) 107-114
6. Isabelle, B., Thierry, G.: Representation and Fusion of Heterogeneous Fuzzy Information in the 3D Space for Model-Based Structural Recognition--Application to 3D Brain Imaging. *Artificial Intelligence* 1-2 (2003) 141-175