Flood Forecast Based on Fuzzy Information Analysis

Lihua Feng[1] Xingcai Zhang[2]

[1]Department of Geography, Zhejiang Normal University, Jinhua 321004, China. E-mail: fenglh@zjnu.cn
[2]Institute of Remote Sensing, Zhejiang Normal University, Jinhua 321004, China

Abstract

Flood system is a complicated system. The relationship between peak stage and peak discharge still remains undetermined owing to the complexity of flood influencing factors. There exists a fuzzy relation between the peak stage and peak discharge. Therefore, fuzzy information analysis method was used to forecast the peak discharge in the present paper and the result is in accordance with the actual situation. The method may be an effective way for flood prediction.

Keywords: Fuzzy information analysis; Universe; Fuzzy relational matrix; Flood forecast

1. Introduction

Flooding in China is a common natural disaster, 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 floods. They occur frequently, affect large areas of the community, hence constituting a huge threat to human life and property. Flood damage becomes even more aggravated with the rapid progress of the economy in recent years[1].

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

2. Fuzzy information analysis method

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 variable and message net block[5-6].

Peak stage *H* and peak discharge *Q* are the most important parameters for the forming of flood damage. Unfortunately, at present the exact relationship between the two parameters as $Q = f(H)$ cannot be found. In fact, there exists a fuzzy relation between the two parameters. Therefore, the universe ranges of the peak stage and peak discharge first need to be determined. 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 $U_H = \{H_1, H_2, \cdots, H_m\}$. The range of the peak discharge can also be selected as *n* classes and the corresponding universe of the peak discharge should be $U_Q = \{Q_1, Q_2, \cdots, Q_n\}$.

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

$$\mu_k(H) = \exp\left[-\left(\frac{H - a_k}{b_k}\right)^2\right] \quad (1)$$

In the equation, *k* means the class of the peak discharge, a_k and b_k are constants and can be obtained from the historical flood samples. Suppose the number of the 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:

$$a_k = \frac{1}{N}\sum_{j=1}^{N} H_j \quad (2)$$

$$b_k^2 = \frac{1}{N}\sum_{j=1}^{N} (H_j - a_k)^2 \quad (3)$$

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 (1), then the fuzzy relational matrix *R* can be obtained. Suppose the mid-value of H_i is $\overline{H}_i$, then the element of *R* should be $r_{ik} = \mu_k(\overline{H}_i)$.

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:

$$Q = H \circ R \quad (4)$$

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 variable in the fuzzy mathematics. Then, the following expression can be constructed:

$$Q = (0,\cdots,0,1,0,\cdots,0) \circ \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mn} \end{bmatrix} = \frac{r_{i1}}{Q_1} + \frac{r_{i2}}{Q_2} + \cdots + \frac{r_{in}}{Q_n} \qquad □5□$$

The above equation represents a fuzzy subset on the universe of the peak discharge. It is just 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 flood influencing factors. Therefore, different peak discharges may occur under the same peak stage, while same peak discharge may also occur under the different peak stages. Table 1 lists some data of the annual maximum peak stage *H* and annual maximum peak discharge *Q* surveyed in Hankou hydrologic station of Changjiang River of China.

Table 1 Some data of the annual maximum peak stage *H* and annual maximum peak discharge *Q* in the Hankou hydrologic 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

From the table it can be seen that the peak stage values *H* are about the same in 1948 and 1949 in the Hankou hydrologic station and, 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 and, conversely, the two corresponding peak stage values *H* are very different. Furthermore, a situation that the peak discharge value *Q* became even larger at a small peak stage value *H* was also observed.

In fact, a definite relationship can only exist in a certain range between the annual maximum peak stage *H* and annual maximum peak discharge *Q*, owing to the action of the flood influencing factors. Therefore, flood data in a range from 1865 to 1970 in the Hankou hydrologic station of Changjiang River, which was scarcely influenced by human activity, was selected in the present discussion. The annual maximum peak stage *H* in the range was divided into nine classes and then, the universe of *H* should be $U_H = \{H_1, H_2, \cdots, H_9\}$ as shown in Table 3. The annual maximum peak discharge *Q* in the range was also divided into ten classes and then, the universe of *Q* should be $U_Q = \{Q_1, Q_2, \cdots, Q_{10}\}$ as shown in Table 2. According to the calculation from the equation (2) and (3), 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 and Q_{10}, several constants can be simply determined as $b_8^2 = b_9^2 = b_{10}^2 = 0.2$ at present, which is just the average value of b_1^2 □ b_7^2. Afterwards, the mid-value $\overline{H}_i$ of the universe of the annual maximum peak stage was substituted into the equation (1) and then, 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 the equation (4). The fit rate of the calculated result for the flood data from 1865 to 1970 in the Hankou hydrologic station can satisfactorily achieve to 73%.

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

U_Q	Range of *Q* (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	36-27	0	0	0	0	0.45	0.90	0	0	0	0
H_7	27-28	0	0	0	0	0	0	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

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

Year	H(m)	U_H	U_Q	Range of Q (m^3/s)	Q (m^3/s)	Fit
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	√

Thus, the annual maximum peak discharge in the Hankou hydrologic station can be predicted according to the present fuzzy relational matrix R. For example, the annual maximum peak stage in 1971 in the Hankou hydrologic station is 24.21m, at the 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 the equation (4) and the result can be represented as follows:

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

If taking the value of Q with the highest probability as the forecasted value, then the annual maximum peak discharge in 1971 could 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 hydrologic station were also forecasted as listed in Table 4. The forecasted values within those five years are in accord with the actual situations besides 1972.

Of course, the effect by human activities on the flood peak level must also be considered in the forecast. Table 5 shows the two largest floods in the 20th century at Hankou hydrologic station. We can see that the forecast for the flood peak discharge of 1954 was accurate, while that for 1998 had an error of one scale. The reason: since the 1980s, large scales of forest felling happened in the middle to upper streams of the Yangtze River, causing soil erosion and riverbed silting. Therefore, compared with the period before the 1980s, the flood peak level has been obviously increased at Hankou hydrologic station at the same discharge. This teaches us that to achieve better forecast accuracy, reduction calculation must be conducted when predicting the flood peak level and the flood peak discharge after the 1980s.

Table 5 The two largest floods in the 20th century at Hankou hydrologic station

Year	H(m)	U_H	U_Q	Range of Q (m^3/s)	Q (m^3/s)	Fit
1954	29.73	H_9	Q_{10}	74000-79000	76100	√
1998	29.43	H_9	Q_{10}	74000-79000	71100	×

4. Conclusion

The relationship between the peak stage and peak discharge is rather undetermined due to the complexity of the flood influencing factors. In fact, there exists a fuzzy relation 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 the forecast of corresponding peak stage or peak discharge and precipitation runoff between the hydrologic stations in the upper and lower reaches of the same river. Although the result of the flood forecast is always in a range, it can already satisfy the demands in most situations. To conclude, flood forecast using the fuzzy information analysis technology should be effective.

References

[1] L. H. Feng, L. R. Chen. Three large floods along the Yangtze River in the 20th century. *Journal of Natural Disasters*, 10(1), 8-11, 2001.

[2] Z. P. Fan, J. Ma, Q. Zhang. An approach to multiple attribute decision making based on fuzzy preference information on alternatives. *Fuzzy Sets and Systems*, 131(1), 101-106, 2002.

[3] O. Cordon, F. Moya, C. Zarco. A new evolutionary algorithm combining simulated annealing and genetic programming for relevance feedback in fuzzy information retrieval systems. *Soft Computing*, 6(5), 308-319, 2002.

[4] C. H. Huang, J. D. Wang. *Analysis of fuzzy information and its application*. Beijing Normal University Publisher, pp. 177-182, 1992.

[5] H. Nils, S. Peter. Precautionary saving and fuzzy information. *Economics Letters*, 70(1), 107-114, 2001.

[6] B. Isabelle, G. Thierry. Representation and fusion of heterogeneous fuzzy information in the 3D space for model-based structural recognition--Application to 3D brain imaging. *Artificial Intelligence*, 148(1-2), 141-175, 2003.