6TH INTERNATIONAL CONFERENCE ON FLOOD MANAGEMENT
September 2014 - São Paulo - Brazil

THE DIVISION OF FLOOD SEASON OF LIXIANJIANG RIVER BASIN BASED ON FISHER OPTIMAL DIVISION METHOD

Gu Shengping[1&2], Xiao Cong[1], Cui Wei[1], Gao Li[1] and Li Zhen[1]
1. *College of Water and Hydropower Engineering, Hohai University, CHINA*
2. *National Engineering Research Center of Water Resources Efficient Utilization and Engineering Safety, CHINA*

ABSTRACT: It may provide a good basis for regulating the reservoir's stage flood control levels to divide flood season into stages properly, for making best use of flood water resources with the fully guaranteed safety of the reservoir. This paper focuses on the application of Fisher optimal division method to division of flood season for Lixianjiang river basin which is located in southwest China. The basic principles and steps of this method are introduced at first. Then, 5 indexes are selected as the characteristic indicators of torrential rain and flood of the basin, and by use of this method, the flood season is divided into four stages, i.e. June 1 to June 30 for pre-flood stage, July 1 to September 10 for main flood stage, September 11 to October 31 for the interim stage, and November 1 to November 30 for post-flood stage. Also, this result is compared with those by the common statistics method and by the fuzzy set theory method in the end, and it is indicated that Fisher optimal division method is of advantages in comprehensively reflecting the characteristics of climate and underlying surface of the basin, so the flood season division result by this method is more objective and more reasonable.

Key Words: Fisher Optimal Division Method, Division of Flood Season, Reservoir Operation, Stage Flood Control Level, Lixianjiang River Basin

1. INTRODUCTION

Traditionally, for the safety of the water and hydropower project, the reservoir is so operated that the water level is kept at the flood control level during the flood season so that it may have a sufficient volume to retain the surplus flood water. When a flood happens, it can be regulated by the reservoir, and the water level may rise gradually with the surplus flood water being retained, and while the flood ends, the flood water retained will be discharged from the reservoir rapidly so that the reservoir may be able to regulate the next flood. It can be seen that in this reservoir operation mode, the water level is relatively low in the whole flood season. As a result, the safety of the water and hydropower project can be well guaranteed. But the flood water resources may not be fully used, while for the hydropower project, water energy resource may not be developed fully due to the low water head in the flood season.

In fact, for most reservoirs, in the whole flood season, the flood characteristics may change with time, which means that different magnitude of the flood may happen in different periods. Then, for improvement of the reservoir operation, the flood season can be divided into several stages with the different flood control levels (Yu, et al. 2006), and specifically, the lower flood control level is for the stages with the greater flood, and the higher flood control level for the stages with the smaller magnitude of flood. In this mode, the safety of the reservoir can be fully guaranteed, and at the same time, more flood water resources can be used, while for the hydropower project, more water energy resource can be developed. However, it is complex to divide the flood season into several stages properly. For doing this, many methods such as cause-effect analysis, statistical analysis, fuzzy set analysis (Jin, et al. 2010), and fractal analysis (Hou, et al. 1999), have been presented. But these methods are of some defects in such aspects as uniqueness in factor being considered, strong subjectivity, and uncertainty in the division result (Liu, et al. 2007). In this paper, Fisher optimal division method, which is of advantages in dealing with the time sequencing of flood season, integrity of factors of precipitation and runoff of the basin, and optimality in the number of stages (Wang, et al. 2011), is studied and applied to the flood season division for Lixianjiang river basin in southwest China.

2. FUNDAMENTALS OF FISHER OPTIMAL DIVISION METHOD

Fisher optimal division method is the method used for dividing a sequence of samples into the number of stages that is the optimal division with the minimum grand total of the sum of square of the deviation of all the stages, and in this division both the minimum difference of the samples within each stage and the maximum difference in between is reached. If P (n, k) represents that a sequence of samples $x_1, x_2, \cdots, x_n$, in which x_i (i=1, 2, … , n), is a vector of m dimensions, corresponding to m characteristic indexes is divided into k stages ($k \leqslant n$), then, totally the number of division results is given as follows:

$$C_{n-1}^{k-1} = \frac{(n-1)!}{(k-1)!\,(n-1)!} \qquad [1]$$

It is certain that of all the division results, one optimal division result or more as stated above can be obtained (Gao, 2005). The steps to find the optimal division of a sequence of samples are as follows.

2.1 Sample Data Standardization

Since in the sequence of samples $x_1, x_2, \cdots, x_n$, x_i (i=1, 2, … , n) is a vector of m dimensions corresponding to m characteristic indexes, a relationship matrix X can be formed:

$$X = (x_{ij})_{n\times m} \qquad [2]$$

Due to the difference in dimension of the indexes in reality, the index value x_{ij} should be standardized as follows, and the matrix X is changed into $X' = (x'_{ij})_{n\times m}$.

$$x'_{ij} = \frac{x_{ij} - x_{min,j}}{x_{max,j} - x_{min,j}} \qquad [3]$$

where, x'_{ij} is the standardized index value; $x_{max,j}$ and $x_{min,j}$ are the maximum and the minimum of the x_{ij} ($1 \leqslant i \leqslant n$) for a given j, respectively.

2.2 Objective Function for Optimal Division

Supposing the standardized sequence of samples $x'_1, x'_2, \cdots, x'_n$, in which x'_i (i=1, 2, … , n) is a vector of m dimensions corresponding to m standardized characteristic indexes, being divided into k stages, one kind of division can be expressed as:

$$P^0(n,k): \{x'_{i_1}, x'_{i_1+1}, \cdots, x'_{i_2-1}\}, \{x'_{i_2}, x'_{i_2+1}, \cdots, x'_{i_3-1}\}, \cdots, \{x'_{i_k}, x'_{i_k+1}, \cdots, x'_{i_{k+1}-1}\}$$

or for short,

$$P^0(n,k): \{ i_1, i_1+1, \cdots, i_2-1 \}, \{ i_2, i_2+1, \cdots, i_3-1 \}, \cdots, \{ i_k, i_k+1, \cdots, i_{k+1}-1 \}$$

where, $1=i_1< i_2< \cdots < i_k < i_{k+1}-1 = n$.

In the division above, for a given stage, expressed by $G_{i,j}$={ i, i+1, ⋯, j }, j>i, the difference between the corresponding samples within this stage can be represented by the stage diameter defined as Equation 4, and the smaller the stage diameter value, the less the difference.

$$D(i, j) = \sum_{r=i}^{j} (x'_r - \bar{x}'_{ij})^T (x'_r - \bar{x}'_{ij}) \qquad [4]$$

where, $\bar{x}'_{ij} = \frac{1}{j - i + 1} \sum_{r=i}^{j} x'_r$.

Then, for obtaining the optimal division, the objective function can be defined below:

$$e[P(n, k)] = \min \sum_{r=1}^{k} D(i_r, i_{r+1} - 1) \qquad [5]$$

For the given n and k, the optimal division result, with the minimum sum of diameters of all the stages, $P^*(n, k)$ can be achieved.

2.3 Solution to the Optimal Division

From Equation 5 the following recursive equations can be given below:

$$e[P(n,2)] = \min_{2 \le j \le n} \{D(1, j - 1) + D(j, n)\} \qquad [6]$$

$$e[P(n, k)] = \min_{k \le j \le n} \{e[P(j - 1, k - 1)] + D(j, n)\} \qquad [7]$$

When the sequence of samples is divided into k stages, j_k corresponding to the optimal function value can be determined by solving the following equation.

$$e[P(n, k)] = e[P(j_k - 1, k - 1)] + D(j_k, n) \qquad [8]$$

Then, j_{k-1} can be so determined that it is satisfied with the following equation:

$$e[P(j_k - 1, k - 1)] = e[P(j_{k-1} - 1, k - 2)] + D(j_{k-1}, j_k - 1) \qquad [9]$$

If the stage j_k is expressed as $G_k = \{j_k, j_k + 1, \cdots, n\}$, then, the stage j_{k-1} can be obtained as $G_{k-1} = \{j_{k-1}, j_{k-1} + 1, \cdots, j_k - 1\}$, and by deduction, all the stages G_1, G_2, …, G_k can be achieved, which is the optimal division of k stages.

2.4 Preferable Number of Stages

For determining the preferable number of stages, the curve $e[P(n, k)] \sim k$ indicating the relationship between the value of the objective function and the number of the stages may be created. Generally, the value of the objective function $e[P(n, k)]$ may decrease with the number of stages k being increased. Then for each section of the curve, the slope $\beta'(k)$ may be calculated by Equation 10, and also the slope difference between two adjacent sections $\beta(k)$ can be calculated by Equation 11. Then the turning point of the curve $e[P(n, k)] \sim k$ can be found at k=k* where β(k)=β(k*) reaches the biggest value. Thus k=k* can be taken as the preferable number of stages.

$$\beta'(k) = \frac{e[P(n,k)] - e[P(n,k-1)]}{k-(k-1)} \quad [10]$$

$$\beta(k) = \beta'(k+1) - \beta'(k) \quad [11]$$

3. DIVISION OF THE FLOOD SEASON OF LIXIANJIANG RIVER BASIN

Lixianjiang river basin, located in southwest China, is of the catchment area about 19300km^2. This basin is characterized by south subtropical plateau monsoon climate. The distribution of precipitation varies significantly from time to time, about 85% of annual rainfall occurring in the period from May to October, especially 55% in the period from June to August, while less rainfall in November to next April. Correspondingly, the runoff of this river varies at the same trend as the precipitation of the basin.

In this paper, it is considered that the flood season is from June 1 to November 30, which can be divided into stages by Fisher optimal division method.

3.1 Characteristic Index Data and the Standardization

For flood season division, the factors to be considered should be the features of precipitation and the runoff, the former mainly indicating the influence of the weather and the latter reflecting the integrated effect of the weather, the underlying surface of the basin, and other factors. Specifically, in this paper, such 5 indexes as the mean of the annual average ten days precipitation (P), the mean of the annual maximum 1 day precipitation of ten days precipitation (P_1), the mean of the annual maximum 3 day precipitation of ten days precipitation (P_3), the mean of the annual maximum 7 day precipitation of ten days precipitation (P_7), and the mean of the annual average flow rate of ten days (Q), all of which are calculated for the flood season of the basin totally involving 18 ten days from the first ten days of June to the last ten days of November, on the measurement data of the basin, are put into analysis. Then, the sequence of index samples of the data vectors associated with these 18 ten days could be formed as listed in Table 1, with each vector being of 5 index values.

Table 1 Characteristic Index Data Samples

No	Ten-day	P/mm	P_1/mm	P_3/mm	P_7/mm	Q/$m^3 \cdot s^{-1}$
1	First ten-day, Jun	8.81	36.84	54.62	76.27	257.20
2	Second ten-day, Jun	8.01	31.51	51.27	69.22	420.84
3	Last ten-day, Jun	10.52	39.33	62.40	92.37	502.78
4	First ten-day, Jul	11.21	38.97	63.03	94.55	678.42
5	Second ten-day, Jul	11.11	41.22	65.27	97.19	827.20
6	Last ten-day, Jul	9.91	39.73	64.73	91.75	932.81
7	First ten-day, Aug	9.19	37.33	57.39	79.27	1035.89
8	Second ten-day, Aug	9.18	34.90	55.17	80.28	960.86
9	Last ten-day, Aug	7.66	32.49	51.17	70.48	866.41
10	First ten-day, Sept	7.49	31.83	48.69	65.56	750.05
11	Second ten-day, Sept	6.23	28.69	42.64	56.92	657.63
12	Last ten-day, Sept	5.59	25.24	37.45	51.36	606.59
13	First ten-day, Oct	5.72	24.77	39.09	52.75	585.21
14	Second ten-day, Oct	4.17	18.69	28.75	37.50	564.39
15	Last ten-day, Oct	4.74	24.04	37.11	48.38	522.52

16	First ten-day, Nov	3.01	15.33	22.03	26.95	407.60
17	Second ten-day, Nov	2.02	11.16	15.93	18.45	364.08
18	Last ten-day, Nov	1.55	9.84	13.52	15.38	307.16

By use of Equation 3, these samples can be standardized as listed in Table 2.

Table 2 Standardized Index Data Samples

No	P'	P'_1	P'_3	P'_7	Q'
1	0.75	0.86	0.79	0.74	0.00
2	0.67	0.69	0.73	0.66	0.21
3	0.93	0.94	0.94	0.94	0.32
4	1.00	0.93	0.96	0.97	0.54
5	0.99	1.00	1.00	1.00	0.73
6	0.87	0.95	0.99	0.93	0.87
7	0.79	0.88	0.85	0.78	1.00
8	0.79	0.80	0.80	0.79	0.90
9	0.63	0.72	0.73	0.67	0.78
10	0.61	0.70	0.68	0.61	0.63
11	0.48	0.60	0.56	0.51	0.51
12	0.42	0.49	0.46	0.44	0.45
13	0.43	0.48	0.49	0.46	0.42
14	0.27	0.28	0.29	0.27	0.39
15	0.33	0.45	0.46	0.40	0.34
16	0.15	0.18	0.16	0.14	0.19
17	0.05	0.04	0.05	0.04	0.14
18	0.00	0.00	0.00	0.00	0.06

3.2 Stage Diameter

The stage diameters D (i, j) (i=1, 2, …, 17; j=i+1, …, 18), are calculated by use of Equation 4, as listed in Table 3.

Table 3 Stage Diameter D (i, j)

j \ i	1	2	3	4	5	6	7	8	…	17
2	0.046									
3	0.186	0.133								
4	0.368	0.249	0.028							
5	0.617	0.410	0.096	0.022						
6	0.849	0.561	0.192	0.071	0.020				…	
7	1.129	0.768	0.374	0.197	0.104	0.036				
8	1.278	0.872	0.473	0.265	0.151	0.058	0.009			
9	1.432	1.014	0.644	0.422	0.286	0.153	0.069	0.033		
10	1.571	1.162	0.834	0.608	0.458	0.287	0.162	0.085		

11	1.882	1.493	1.218	0.983	0.802	0.568	0.373	0.227		
12	2.345	1.975	1.747	1.490	1.262	0.946	0.661	0.431		
13	2.717	2.359	2.165	1.884	1.612	1.225	0.865	0.565		
14	3.561	3.203	3.038	2.703	2.351	1.847	1.367	0.959		
15	3.953	3.599	3.452	3.086	2.684	2.110	1.558	1.085		
16	5.146	4.787	4.655	4.221	3.721	3.010	2.316	1.710		
17	6.720	6.342	6.211	5.688	5.066	4.194	3.332	2.570		
18	8.372	7.967	7.826	7.207	6.456	5.417	4.383	3.458		0.006

3.3 Objective Function Value

On the basis of the expression of the objective function, Equation 5, and the recursive equations, Equation 6 and Equation 7, the values of objective function $e\left[P(n,k)\right]$ (n=3, 4, …, 18; k=2, 3, …, 17) can be calculated, and the results are listed in Table 4. In this table, the numbers in brackets mean that at these numbers, the minimum value of objective function can be reached, corresponding to the optimal division for the given n and k.

Table 4 Objective Function Values $e\left[P(n,k)\right]$

n \ k	2	3	4	5	6	7	8	…	17
3	0.046(3)								
4	0.074(3)	0.028(3)							
5	0.142(3)	0.068(4)	0.022(4)						
6	0.238(3)	0.095(5)	0.049(5)	0.020(5)					
7	0.382(4)	0.178(5)	0.095(7)	0.049(7)	0.020(7)				
8	0.451(4)	0.200(6)	0.103(7)	0.057(7)	0.029(7)	0.009(7)			
9	0.607(4)	0.296(6)	0.163(7)	0.103(9)	0.057(9)	0.029(9)	0.009(9)		
10	0.794(4)	0.400(7)	0.215(9)	0.118(9)	0.072(9)	0.043(9)	0.023(9)		
11	1.168(4)	0.537(9)	0.286(9)	0.189(9)	0.118(11)	0.072(11)	0.043(11)	…	
12	1.484(9)	0.657(9)	0.394(10)	0.232(11)	0.136(11)	0.090(11)	0.061(11)		
13	1.554(9)	0.727(9)	0.424(11)	0.239(11)	0.142(11)	0.096(11)	0.068(11)		
14	1.727(11)	0.945(10)	0.557(11)	0.371(11)	0.239(14)	0.142(14)	0.096(14)		
15	1.740(11)	0.963(11)	0.569(11)	0.383(12)	0.278(14)	0.182(14)	0.136(14)		
16	2.032(11)	1.254(11)	0.860(12)	0.569(16)	0.383(16)	0.278(16)	0.182(16)		
17	2.485(11)	1.707(11)	0.991(16)	0.598(16)	0.411(16)	0.307(16)	0.210(16)		
18	2.938(12)	1.802(16)	1.025(16)	0.632(16)	0.445(16)	0.341(16)	0.244(16)		0.001(18)

3.4 Optimal Flood Season Division

On the basis of Table 4, for the whole sequence of samples, n=18, the optimal division results for the different number of stages are listed in Table 5. Then, the curve $e\left[P(n,k)\right] \sim k$ is drawn as shown in Figure 1. Besides, the slope of the curve for each section β'(k) and the slope difference between two adjacent sections β(k), calculated on Equation 10 and Equation 11, are listed in Table 5.

Table 5 Optimal Division Results for the Different Number of Stages

k	e[P(n, k)]	β'(k)	β(k)	Optimal Division Result
2	2.938			{1～11} {12～18}
3	1.802	-1.136	0.358	{1～10} {11～15} {16～18}
4	1.025	-0.777	0.384	{1～3} {4～10} {11～15} {16～18}
5	0.632	-0.393	0.207	{1～3} {4～6} {7～10} {11～15} {16～18}
6	0.445	-0.187	0.082	{1～3} {4～6} {7～10} {11} {12～15} {16～18}
7	0.341	-0.104	0.008	{1～3} {4～6} {7～10} {11} {12～13} {14～15} {16～18}
8	0.244	-0.097	0.041	{1～3} {4～5} {6} {7～8} {9～10} {11～13} {14～15} {16～18}
9	0.188	-0.056	0.010	{1～3} {4～5} {6} {7～8} {9～10} {11～13} {14～15} {16}{17～18}
10	0.142	-0.046	0.006	{1}{2} {3～4} {5～6} {7～8} {9～10} {11～13} {14～15} {16}{17～18}
11	0.103	-0.039	0.011	{1}{2} {3～4} {5～6} {7～8} {9～10} {11～13} {14}{15} {16}{17～18}
12	0.074	-0.029	0.006	{1}{2} {3}{4} {5～6} {7～8} {9～10} {11～13} {14}{15} {16}{17～18}
13	0.051	-0.023	0.003	{1}{2} {3}{4} {5～6} {7～8} {9～10} {11}{12～13} {14}{15} {16}{17～18}
14	0.031	-0.02	0.006	{1}{2} {3}{4} {5}{6} {7～8} {9～10} {11}{12～13} {14}{15} {16}{17～18}
15	0.016	-0.015	0.006	{1}{2} {3}{4} {5}{6} {7～8} {9}{10} {11}{12～13} {14}{15} {16}{17～18}
16	0.008	-0.008	0.002	{1}{2} {3}{4} {5}{6} {7}{8} {9}{10} {11}{12～13} {14}{15} {16}{17～18}
17	0.001	-0.007		{1}{2} {3}{4} {5}{6} {7}{8} {9}{10} {11}{12～13} {14}{15} {16}{17}{18}

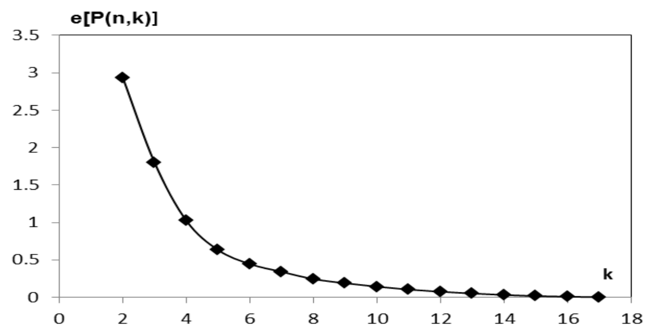

Figure 1 Relationship of $e\left[P(n,k)\right] \sim k$

From Table 5 and Figure 1, β(k) takes the biggest value at k*=4, and β(k*)=0.384, correspondingly, the turning point of the curve $e\left[P(n,k)\right] \sim k$, is found here, and then k*=4 can be taken as the preferable number of stages. As a result, the flood season of Lixianjiang river basin is divided into 4 stages as listed in Table 6.

Table 6 Result of the Flood Season Division of Lixianjiang River Basin

Division	Pre-flood Stage	Main Flood Stage	Interim Stage	Post-flood Stage
Period	1 Jun – 30 Jun	1 Jul – 10 Sept	11 Sept – 31 Oct	1 Nov – 30 Nov

3.5 Discussion

In this paper, the comparison of Fisher optimal division method in flood season division with the common statistics method and the fuzzy set theory method is made, and the flood season division results of the latter two methods are listed in Table 7.

Table 7 Flood Season Division Results by Common Statistics Method & Fuzzy Set Theory Method

Flood Season Division Method	Flood Season Division Result		
	Pre-flood Stage	Main Flood Stage	Post-flood Stage
Common Statistics Method	1 Jun - 30 Jun	1 Jul - 20 Sept	21 Sept – 30 Nov
Fuzzy Set Theory Method	1 Jun - 30 Jun	1 Jul - 19 Oct	20 Oct - 30 Nov

From the results listed in Table 6 and Table 7, it can be seen that all these three methods are the same in definition of the pre-flood stage. Also, they are the same in that the period from 1, July to 10, September, is involved in the main flood season, although they are different in lengths of the main flood stage and the post-flood stage. Besides, by Fisher optimal division method, the interim stage is added between the main flood stage and post-flood stage, while the post-flood stage is shorter than those by the other two methods.

It is normal that there are some differences of the flood season division results above. Relatively speaking, the common statistics method is so simple but quite subjective, and the fuzzy set theory method is of uncertainty in threshold for the starting and ending of the flood season, resulting in the uncertainty in flood season division. Moreover, because in flood season division by these two methods, only one index, e.g. the flood peak occurring time, can be considered, and the flood season can be normally divided into 3 stages, these two methods are not enough in objectivity, all-sidedness, and adaptability. However, by Fisher optimal division method, the time of sequencing of the samples, and more characteristic indexes of the precipitation and runoff, may be fully considered, and the number of stages can be optimized. Therefore, the flood season division result by Fisher optimal division method is more reasonable and reliable.

4. CONCLUSION

1) By the application of Fisher optimal division method to the flood season division of Lixianjiang river basin, it can be seen that this method is not only applicable, but also very effective in finding the optimal flood season division from mathematical viewpoint.

2) In comparison with the common statistics method and fuzzy set theory method, more characteristic indexes of the precipitation and runoff, which have influence on flood season division, can be properly considered by use of Fisher optimal division method, the flood season division result is of more objectivity and more adaptability.

3) In the use of Fisher optimal division method, an interval of certain days, e.g. five days, ten days, and one month, is normally taken as the basic unit in flood season division, then, the continuity of the natural hydrologic event, an entire flood process, may be broken. So the basic time unit should be defined carefully. Furthermore, the optimal flood season division result should be finally determined, not only on the mathematical analysis, but also on the real hydrologic situation of the water project.

5. ACKNOWLEDGEMENTS

This work was supported by the National Twelfth Five-Year Key Technology R & D Program of China (Grant No 2013BAB06B01).

6. REFERENCES

Gao Bo, 2005: *Theory and application of safe utilization of flood water resources* Hohai University, CHINA. (in Chinese)

Hou Yu, Wu Boxian, and Zheng Guoquan, 1999: “Preliminary study on the seasonal period classification of flood by using fractal theory” *Advances in Water Science* 10:2, 140-143. (in Chinese)

Jin Baoming and Fang Guohua, 2010: “Application of Fuzzy set analysis method to flood season division for Nanping city” *Water Power* 36:3, 20-22. (in Chinese)

Liu Kelin, Wang Yintang, Hu Siyi, etc. 2007: “Application of fisher optimal dissection method to flood season division” *Advances in Science and Technology of Water Resources* 27:3, 14-16. (in Chinese)

WANG Shuangyin，YANG Xiaoxiao, 2011: “Flood season division based on correlation coefficient and optimum partition method of Fisher” *Journal of China Hydrology* 31:4, 1-5. (in Chinese)

Yu Ting, Guo Shenglian, Liu Pan, etc. 2006: “Study and application of method of flood division of reservoir” *China Rural Water and Hydropower* 56:8, 24-26. (in Chinese)