Chapter 4
Copula-Based Seasonal Design Flood Estimation

4.1 Introduction

Since the rain-producing systems vary with season, the river flood is usually characterized as seasonality. Seasonal fluctuations are a significant source of variability in runoff records. However, seasonality is often overlooked when evaluating flood risk due to the use of annual value for defining extreme values. The phrase "1 in 100 years" flood does not inform whether a given extreme value is more likely to come from one season over another. The oversight of seasonality is also common to the peak-over-threshold method, even though this method is capable of obtaining more than one extreme value per year (Michael et al. 2007).

In the reservoir operation, the water level of the reservoir should be limited below the flood control water level (FCWL) during flood season to offer adequate storage for flood control. The current FCWL, which plays a key role in the flood prevention and floodwater utilization, is mainly determined according to the design floods estimated from annual maximum flood series while neglecting the seasonal information. This results in over-standard for flood prevention and a waste of floodwater in most of the years. Therefore, the design flood guideline in China stresses the importance of classifying the annual maximum floods caused by different generating mechanisms (MWR 1993). For floodwater utilization, it's very valuable to use the seasonal flood information in flood frequency analysis o operate the reservoir more effectively during flood seasons without enhancing the flood prevention risk. How to reasonably and optimally design seasonal floods that reflects seasonal variations poses a challenge to hydrologists and engineers nowadays. It is a very important and urgent issue in the management of reservoirs in China (Guo et al. 2004; Fang et al. 2007).

The conventional flood frequency analysis methods are based on univariate distributions, mainly concentrated on the analysis of peak discharge or flood volume series. For seasonal design flood, statistical analysis of the flood occurrence dates is also very useful and important. Generally, the annual maximum flood often

L. Chen and S. Guo, *Copulas and Its Application in Hydrology and Water Resources*, Springer Water, https://doi.org/10.1007/978-981-13-0574-0_4

occurs in main flood season, and median or small floods occur in other ones. The flood occurrence date is also a random variable and follows a particular distribution, which is different from that of flood magnitude. Thus, seasonal design flood should consider both the dates and magnitudes of flood events that may be described by a bivariate joint distribution. Chen et al. (2010) proposed a new seasonal design flood method, which considers dates of flood occurrence and magnitudes of the peaks (runoff) based on copula function. Their results show that the proposed method can satisfy the flood prevention standard, and provide more information about the flood occurrence probabilities in each sub-season. Yin et al. (2017) used three bivariate flood quantile selection methods, namely equivalent frequency combination (EFC) method, conditional expectation combination (CEC) method and conditional most likely combination (CMLC) method, to estimate unique seasonal design flood to meet the needs in engineering. Results showed that the CMLC method is more rational in physical realism and recommended for estimating the seasonal design floods, which can provide rich information as the references for flood risk assessment, reservoir scheduling, and management.

4.2 Review of Seasonal Design Flood Methods

The issue of seasonal flood frequency analysis was identified by Creager et al. as early as in 1951. The aim of the seasonal design flood is to determine the relationship between hydrograph and return period in each season. Two current seasonal design flood methods: one was suggested by Chinese design flood guideline (MWR 1993), other was proposed by Singh et al. (2005). These two methods are referred as Chinese method and Singh's method in this study and reviewed as follows.

4.2.1 Chinese Method

The seasonal maximum (SM) flood series extract the maximum peak discharge (or runoff volumes) from each season during each year of record. The seasonal T-year design flood is obtained by fitting a particular distribution, such as P-III distribution used in China (MWR 1993).

In this method, the annual maximum values Y can be described as:

$$Y = \{Y_1, Y_2, \ldots, Y_s\} \tag{4.1}$$

where $Y_1, Y_2, \ldots, Y_s$ are the seasonal maximum flood series; and s is the number of the sub-seasons.

According to the Eq. (4.1), the extreme value distribution of the annual maximum flood series can be defined as:

$$F_T(y) = F_1(y)F_2(y)\cdots F_S(y) \tag{4.2}$$

where $F_T(y)$ is the distribution of the annual maximum flood series; and $F_1(y),\ldots,F_s(y)$ are the distributions of seasonal maximum flood series (Waylen and Woo 1982). For a fixed value y, Eq. 4.2 shows that $F_T(y)$ will always be less than or equal to the smallest of the $F_i(y)$, since each of the latter values must always be in the range [0, 1]. In other words, the annual frequency curve must always lie on or above the highest of the seasonal frequency curves on a common probability paper, i.e., T-year seasonal design floods are always less than the annual design floods (Durrans et al. 2003).

The Chinese flood prevention standard is defined by annual return period $T, T = 1/(1 - F_T(y))$, while the Chinese design flood guideline assumes that the seasonal design frequency is equal to the annual design frequency, namely $F_T(y) = F_1(y_1) = F_2(y_2) = 1 - 1/T$(MWR 1993). Assuming two sub-seasons, take the 100-year design flood for example. If $F_1(y) = F_2(y)$, i.e., in the case of identical distribution, Eq. 4.2 leads to $F_T(y) = (F_1(y))^2 = 0.98$. This means that when the seasonal design method is used, the combined frequency of them cannot reach the annual prevention standard. If the combination frequency must reach to the annual design frequency, at least one of the seasonal frequencies must exceed the annual frequency analysis. Therefore, the current Chinese seasonal design flood method cannot satisfy the flood prevention standard.

4.2.2 Singh's Method

Annual maximum flood series are formed by extracting the annual maximum peak discharge (or runoff volumes) from each year of record. If n is the number of recorded years and n_i is the number of annual maxima that occur in the ith season, then $n = \sum_{i=1}^{s} n_i$ (Durrans et al. 2003).

This method can be described as follows: considering that the occurrence of a flood event $B = \{Y > y\}$ must be associated with one of the events $\{A_i\}$, i = 1, …, s. $\{A_i\}$ means the annual maximum flood that occurs during the ith season.

The exceedance frequency $P(y, A_i)$ of seasonal design flood is defined as:

$$P(y, A_i) = P(y|A_i)P(A_i) \tag{4.3}$$

where $P(y|A_i)$ is the exceedance probability that an annual flood maximum occurring in the ith season. $P(A_i)$ is the probability of an annual maximum occurring in the ith season, i = 1, …, s.

This method has been described by Thomas et al. (1998), who pointed out that its use is valid for both independent and dependent seasonal flood distributions. Singh et al. (2005) applied this method to estimate design flood from a nonidentically distributed series and provided an estimation procedure for practical use.

The sum of the probabilities of seasonal design flood is given by:

$$P(Y \geq y) = \sum_{i=1}^{s} P(B \cap A_i) = \sum_{i=1}^{s} P(A_i)P(y|A_i) \tag{4.4}$$

Equation 4.4 is the total probability law and expresses the frequency distribution of the annual maximum flood as the sum of the frequency distribution of those annual maximum floods that are conditioned on the maxima occurring in the ith season with the probability weight $P(A_i)$ (Singh et al. 2005).

Assuming the annual maxima occurring in different seasons are identically distributed, the conditional frequency distribution $P(y|A_i)$ is free of A_i, then Eq. 4.4 leads to

$$P(y) = P(y_0)\sum_{i=1}^{s} P(A_i) = P(y_0) \tag{4.5}$$

where $P(y_0)$ is a fixed frequency distribution indicating that the overall annual maxima are identically distributed. Equation 4.5 shows the validity of Eq. 4.4 which can satisfy the flood prevention standard (Singh et al. 2005).

The flood frequency distribution $P(y|A_i)$ should be estimated from those observed values of the Y_i flood series that are picked as the annual maximum floods. For some drier season, there may be few or even no samples to be drawn. It is not accurate and reliable to use these data series for calculation. Equation 4.4 suffers in practice from the fact that n_i for one season will usually be considerably smaller than n_i for another season. Because of this, the reliability with which each conditional distribution in Eq. 4.4 may be estimated will vary from season to season. Furthermore, since n_i for any season will always be less than or equal to n, this approach essentially limits the lengths of the record samples (Durrans et al. 2003).

4.3 A New Seasonal Design Flood Method

The sampling methods, flood seasonality identification methods, and the copula functions are introduced and discussed. The von Mises distribution is used to describe the flood occurrence dates, while the P-III distribution or exponential distribution (Ex) is selected as marginal distribution for annual maximum flood series or peak-over-threshold samples, respectively. A new seasonal design flood method is described as follows.

4.3.1 Sampling Method

Sampling methods play an important role in flood frequency analysis. The annual maximum (AM), seasonal maximum (SM) and peaks-over-threshold (POT) sampling methods are used and compared in this section.

The POT sampling method is also widely used in flood frequency analysis because more information can be obtained compared with that of the AM or SM sampling method. To guarantee the independence of the samples, the flood peaks are selected by two criteria suggested by Institute of Hydrology of UK (IH 1999): (1) two peaks have to be separated by at least three times the average time to rise, in which the average time to rise is determined from the synthetic records as 2 days and is kept constant throughout the study; and (2) The minimum discharge between two peaks has to be less than two-thirds of the discharge of the first of the two peaks.

4.3.2 Identification of Seasonality

The whole flood season is usually divided into three sub-seasons, and these sub-seasons are defined as the pre-flood season, main flood season and post-flood season (MWR 1993; Ngo et al. 2007).

Several types of approaches for detecting flood seasonality have been proposed. One type of approaches is to segment flood season regarding climatological and river basin physiographic characteristics by analyzing the rain-producing system (Black and Werrity 1997; Singh et al. 2005). The other type is to segment flood season by using visual identification based on some measurements of flood seasonality. Ouarda et al. (1993) proposed two variations of a graphical method for identification of river flood season from peaks-over-threshold (POT) data. Cunderlik et al. (2004) used the relative frequency (RF) method and directional statistics (DS) method to identify the seasonality. The RF method is based on counting the number of events in each season, to allow comparisons between records, expressing these counts as a percentage of the total number of events in each record (Black and Werritty 1997; Cunderlik et al. 2004; Ouarda et al. 2006). The DS method describes the seasonality by defining the mean day of the flood (directional mean) and the flood variability measure. The DS, RF, and POT methods are compared in this study.

4.3.3 Seasonal Design Flood Estimation

The season design flood can be characterized by flood occurrence dates and flood magnitudes. In this section, first, the marginal distribution of flood occurrence dates and flood magnitudes are established.

4.3.3.1 Margin Distribution of Flood Occurrence Dates

The von Mises distribution introduced in Chap. 3. Three can only be used for unimodal distribution. Since the annual maximum floods may be generated by different mechanisms, the flood occurrence data series often obey a multimodal distribution. Thus, a mixed von Mises distribution which can describe the multimodal character is comprised of a finite mixture of von Mises distributions. The probability density function for a mixture of N von Mises distributions (vM-pdf) takes the following form:

$$f_X(x) = \sum_{i=1}^{N} \frac{p_i}{2\pi I_0(\kappa_i)} \exp[\kappa_i \cos(x - \mu_i)]$$
$$0 \leq x \leq 2\pi,\ 0 \leq \mu_i \leq 2\pi,\ \kappa_i \geq 0 \tag{4.6}$$

where p_i is the mixing proportion, μ_i is the mean direction, and κ_i is the concentration parameter.

Various methods can be used to estimate the 3 N parameters on which the mixture of N vM-pdfs depends (Carta and Ramírez 2007). The least squares (LS) method is used in this book, in which the 3 N unknown parameter values can be estimated by minimizing the sum of the squares of the deviations between the experimental data and the calculated value (Carta et al. 2008).

4.3.3.2 Margin Distribution of Flood Magnitudes

For the AM flood series, the P-III distribution has been recommended by MWR (1993) as a uniform procedure for flood frequency analysis in China. The formula of P-III distribution is given in Table 1.1.

The classical use of the POT sampling method comprises the assumptions of a Poisson-distributed number of threshold exceedances and exponentially distributed peak exceedances (Lang et al. 1999). The probability density function of 2-parameter Ex distribution is given in Table 1.1 as well. For POT flood series, the 1-parameter Ex distribution is used by setting the parameter $\gamma = 0$

$$f_{YPOT}(y) = \frac{1}{\alpha} \mathrm{e}^{(-y/\alpha_0)} \tag{4.7}$$

where α_0 is a parameter of the Ex distribution.

The parameters of P-III and Ex distributions are estimated by L-moments method (Hosking and Wallis 1997).

4.3.3.3 Bivariate Distribution of Flood Occurrence Dates and Magnitudes

For estimating seasonal design flood, the bivariate joint distributions of flood occurrence dates and magnitudes need to be built. Every joint distribution can be written regarding a copula and its univariate marginal distributions. The copula is a function that links univariate marginal distribution functions to construct a multivariate distribution function. The definition and establishment of copulas can be seen in Chap. 2. The Gumbel–Hougaard, Frank, Clayton, and Ali-Mikhail–Haq copulas are used to establish the joint distribution.

4.3.3.4 Seasonal Design Flood Estimation

Seasonal design flood is related to the flood dates X and magnitudes Y and follows a two-dimensional distribution $F(x, y)$. Assuming all floods occur during whole flood season, the annual exceedance probability can be defined as:

$$P(X \leq t, Y > q) = F_X(t) - F(t, q) \tag{4.8}$$

where t is the last day of the flood season, and q is a specific discharge value. $F_X(t)$ is the marginal distribution function of t.

$F(t, q)$ is the joint distribution of the flood peak which occurs before the date t with the value less than or equal to the discharge q, and can be described by

$$\begin{aligned}
F(t, q) &= \int_{-\infty}^{q} \int_{-\infty}^{t} f(x, y) dxdy \\
&= \int_{-\infty}^{q} \int_{-\infty}^{t} f_X(x) f_Y(y|x) dxdy \\
&= \int_{-\infty}^{t} f_X(x) \int_{-\infty}^{q} f_Y(y|x) dydx \\
&= \int_{-\infty}^{t} f_X(x) F(q|x) dx \\
&= \int_{0}^{t} F(q|x) dF_X(x) \\
&= \sum_{i=1}^{s} P(Y \leq q | x_i < x < x_{i+1}) F_X(x_i < x < x_{i+1})
\end{aligned} \tag{4.9}$$

$$
\begin{aligned}
P(X \le t, Y > q) &= F_X(t) - \sum_{i=1}^{s} P(Y \le q | \Delta x_i) F_X(x_i < x < x_{i+1}) \\
&= \sum_{i=1}^{s} F_X(x_i < x < x_{i+1}) \\
&\quad - \sum_{i=1}^{s} P(Y \le q | \Delta x_i) F_X(x_i < x < x_{i+1}) \\
&= \sum_{i=1}^{s} F_X(x_i < x < x_{i+1})(1 - P(Y \le q | x_i < x < x_{i+1})) \\
&= \sum_{i=1}^{s} F_X(x_i < x < x_{i+1}) P(Y \ge q | x_i < x < x_{i+1})
\end{aligned}
\tag{4.10}
$$

where $f_X(x)$ is the marginal density function of variable x; $f(x, y)$ is two-dimensional density function; $f_Y(y|x)$ and $F_Y(y|x)$ are the conditional probability density and distribution function of y; and x_i represents a segmentation point. If s equals the number of the sub-seasons, then Eq. 4.10 is as the same as that of Eq. 4.4. It is also indicated from Eq. 4.10 that the seasonal design flood frequency curves are located below the annual one.

If $F_X(x_{i-1} < x < x_i)$ is replaced with $P(A_i)$, the exceedance probability for the seasonal design flood frequency $P(q, A_i)$ is defined as:

$$
\begin{aligned}
P(q, A_i) &= P_i(Y_i > q | A_i) P(A_i) \\
&= P(x_{i-1} \le X \le x_i, Y_i > q) \\
&= F(x_i) - F(x_{i-1}) - F(x_i, q) + F(x_{i-1}, q)
\end{aligned}
\tag{4.11}
$$

where x_{i-1} and x_i are the segmentation points. Equation 4.11 indicates that seasonal design flood is related to bivariate joint distributions. The seasonal design flood frequency $P(q, A_i)$ can be described by the probability weight $P(A_i)$ and the conditional frequency distribution $P(q|A_i)$. Since the range of the $P(q|A_i)$ is from 0 to 1, the value of $P(q, A_i)$ is restricted within $P(A_i)$.

4.4 Case Study

4.4.1 Identification of Flood Seasonality

The Geheyan reservoir is selected as a case study. Fifty-four year (1954–2004) discharge records are used to analyze seasonal design flood. For the POT sampling method, the threshold value with 3500 m^3/s is selected, which corresponds to a mean of 2.57 exceedances per year.

In the Qingjiang basin, floods frequently occur in summer from June to early August when the monsoon fronts advance from south to north or in the fall from late August to early October when the fronts withdraw from north to south. Although both summer and fall floods result from frontal rains, their hydrological characteristics are distinctly different because the intensity of the rain-producing system is varied with seasons (Singh et al. 2005). The statistical analysis results of 10-day rainfall data are listed in Table 4.1. It can be seen from Table 4.1 that most of the rainfall occurs from late June to middle July, whereas in other time of the flood season a relatively small amount of rainfall is received. Seasonal variation of trends is that flood events begin to increase from late June and decrease in late July. Therefore, those two periods might be the segmentation points.

The DS, RF and POT methods are used to describe the flood seasonality, and the results of these methods are listed in Table 4.2. It can be seen that the RF method has the shortest main flood season (from June 21 to July 20) because of the clustering of dates of flood occurrences into ten days. Ouarda et al. (2006) pointed out that the seasonality method based on the peaks-over-threshold (POT) approach lead to the best results. However, the result of POT method also has the shortest main flood season (from June 26 to July 26) at the Qingjiang basin as shown in Table 4.2. Compared with POT method, the result of DS approach has a 5-day difference for each sub-season.

In order to ensure the flood control safety, the results of DS method are chosen, since it has the longest main flood season from June 21 to July 31. The mean day of the flood (directional mean) is on July 2. The flood occurrence dates sampled by

Table 4.1 Statistical analysis results of 10-day rainfall data of the Geheyan basin

Date		≥ 60 mm	50–59 mm	40–49 mm	30–39 mm	Mean values (mm)	Percentage (%)
May	Early	1	2	3	9	58	6.26
	Mid.	1	1	4	11	58	6.19
	Late	0	2	2	13	61	6.51
June	Early	3	3	2	18	61	6.56
	Mid.	4	0	1	9	60	6.40
	Late	4	2	18	13	85	9.14
July	Early	3	6	14	11	86	9.20
	Mid.	7	7	11	9	84	9.03
	Late	1	2	6	7	60	6.42
August	Early	2	2	4	6	60	6.42
	Mid.	1	1	4	9	59	6.27
	Late	3	1	5	8	56	5.96
Sept.	Early	0	2	7	11	49	5.30
	Mid.	4	3	5	9	58	6.25
	Late	0	2	2	3	38	4.08

Table 4.2 Results of three methods for identification of the seasonality

Methods	The pre-flood season	The main flood season	The post-flood season
DS	May 1–June 20	June 21–July 31	Aug. 1–Sept. 30
RF	May 1–June 20	June 21–July 20	July 21–Sept. 30
POT	May 1–June 25	June 26–July 26	July 27–Sept. 30

POT method are translated into a location on the circumference of a circular drawn in Fig. 4.1. It can be seen from Fig. 4.1 that the flood events are mainly centered June 20 to July 31, and the interval time of the adjacent flood events is obviously shorter in this period.

In summary, the flood season of the Qingjiang basin can be divided into three sub-seasons, i.e., the pre-flood season, main flood season and post-flood season. Based on the analysis results above, the pre-flood season is from May 1 to June 20; the main flood season is from June 21 to July 31, and the post-flood season is from August 1 to Sept. 30.

4.4.2 Computation of Empirical Frequency

The empirical probabilities can be computed by Eqs. 3.8 and 3.9.

Fig. 4.1 Application of DS method for flood occurrence dates based on POT samples

4.4.3 Bivariate Distribution

The joint distribution is established for the AM and POT samples respectively, and the estimated parameters of the margin distribution and joint distribution are listed in Table 4.3. Some statistical tests are used for margin and joint distributions. A chi-square goodness-of-fit test is performed to test the assumption H_0 that the flood magnitude follows P-III distribution or Ex distribution. A Kolmogorov–Smirnov (K-S) test is used to test the assumption H_0 that the flood occurrence dates follow mixed von Mises distribution. The results shown in Table 4.4 indicate that these assumptions cannot be rejected at the 5% significant level. The fitted frequency histograms of the flood occurrence date by the mixed von Mises distribution for POT sample series are drawn in Fig. 4.2. The margin distribution frequency curves of flood occurrence dates and magnitudes are shown in Figs. 4.3 and 4.4, respectively, of which the line represents the theoretical distribution, and the crossing means empirical probability of observations. Figures 4.3 and 4.4 indicate that all the theoretical distributions can fit the observed data reasonably well, although there are some uncertainties in the dataset itself.

Four widely used copulas, namely the Gumbel-Hougaard, the Ali-Mikhail-Haq, the Frank and the Clayton are compared and discussed. The root mean square error (RMSE) and Akaike's information criterion (AIC) are used to identify the most appropriate copula distribution (Zhang and Singh 2006). The equation for RMSE can be expressed by

Table 4.3 Estimated parameters of the marginal and joint distributions for peak discharges

Sampling method	P-III or Ex distribution			Mixed von Mises distribution			Joint distribution
				μ_i	κ_i	pi	θ
	α	β	δ	1.03	4.80	0.17	
AM	2.520	0.001	2 467	2.81	3.28	0.70	1.82
				5.66	27.43	0.13	
	λ	α_0		0.62	16.52	0.11	
POT	2.58	2.361		2.72	1.61	0.72	0.83
				5.60	3.76	0.17	

Table 4.4 Hypothesis test for margin distributions

Samples	P-III or Ex distribution		Mixed von Mises distribution	
	$\chi^2_{0.95}$	Chi-squared statistics χ^2	$D_{n,0.95}$	K-S statistics D_n
AM	7.815	1.800	0.180	0.089
POT	12.592	2.758	0.115	0.047

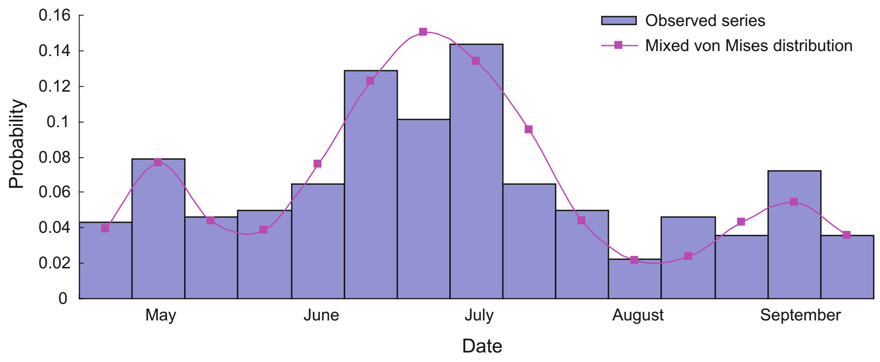


Fig. 4.2 Fitted frequency histograms of flood occurrence dates by the mixed von Mises distribution for POT samples

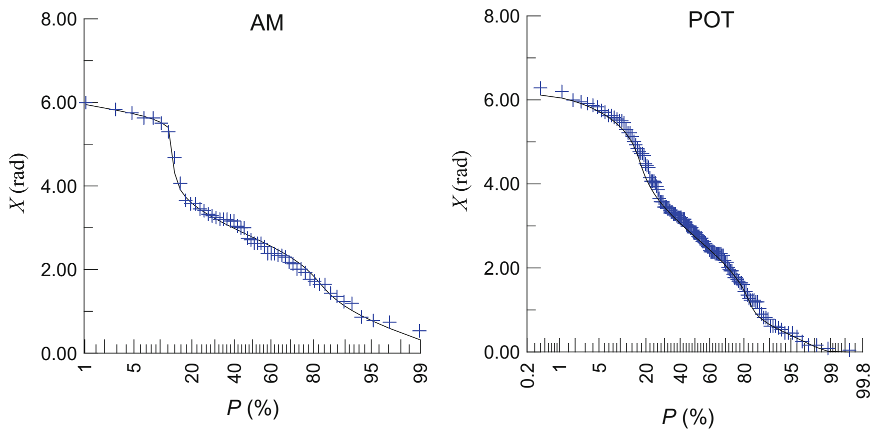


Fig. 4.3 Frequency curves of flood occurrence dates based on AM and POT samples

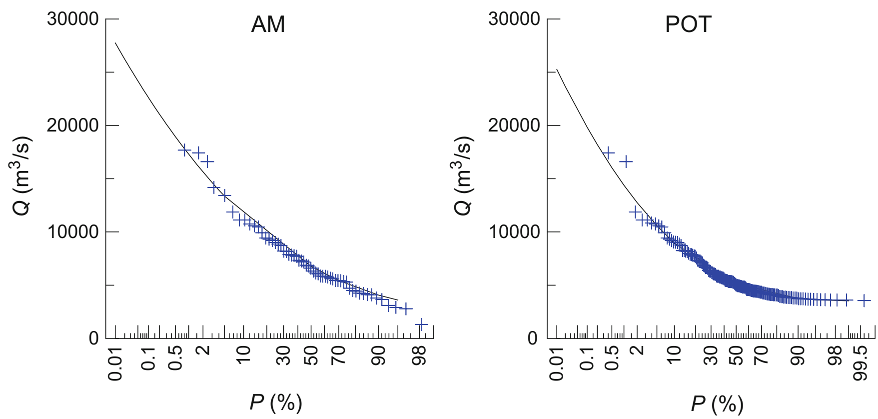


Fig. 4.4 Frequency curves of flood magnitudes based on AM and POT samples

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(x_{the}(i) - x_{emp}(i)\right)^2} \tag{4.12}$$

where N represents the number of observations; and $x_{the}(i)$ and $x_{emp}(i)$ denote the ith calculated and observed values, respectively.

The Akaike information criterion (AIC), developed by Akaike (1974), is used to identify the appropriate probability distribution. The AIC can be obtained either by calculating the maximum likelihood or by calculating the mean square error of the model (Zhang and Singh 2006). The AIC values related to maximum likelihood values can be expressed by

$$AIC = 2k - 2\ \ln(L) \tag{4.13a}$$

The AIC values related to mean square error can be expressed by

$$AIC = 2k + N\ \ln(MSE) \tag{4.13b}$$

where k is the number of parameters in the statistical model; L is the maximized value of the likelihood function for the estimated model; and $MSE = RMSE^2$.

The *RMSE* and *AIC* values related to *MSE* (Eq. 4.13b) for different copulas are listed in Table 4.5. The best family is the one which has the minimum *RMSE* and *AIC* values. It can be seen that Frank and Clayton family fit the empirical joint probabilities better than Gumbel and Ali-Mikhail–Haq. No obvious difference exists between Frank family and Clayton family.

The empirical joint probabilities of the combinations of flood occurrence dates and flood peak magnitudes are plotted versus theoretical probabilities as shown in Fig. 4.5, which shows that no significant difference between empirical and theoretical joint probabilities can be detected.

It may be concluded that the proposed bivariate joint distribution is suitable to represent the flood occurrence dates and magnitudes at the Geheyan reservoir basin. The joint distribution of the AM flood series is shown in Fig. 4.6.

Table 4.5 The RMSE and AIC values for different copula functions

Family		AM			POT	
	θ	*RMSE*	*AIC*	θ	*RMSE*	*AIC*
Gumbel–Hougaard	1.24	0.042	−318	1.10	0.031	−942
Ali-Mikhail–Haq	0.70	0.080	−249	0.37	0.045	−838
Frank	1.82	0.038	−329	0.83	0.028	−970
Clayton	0.49	0.039	−326	0.20	0.028	−970

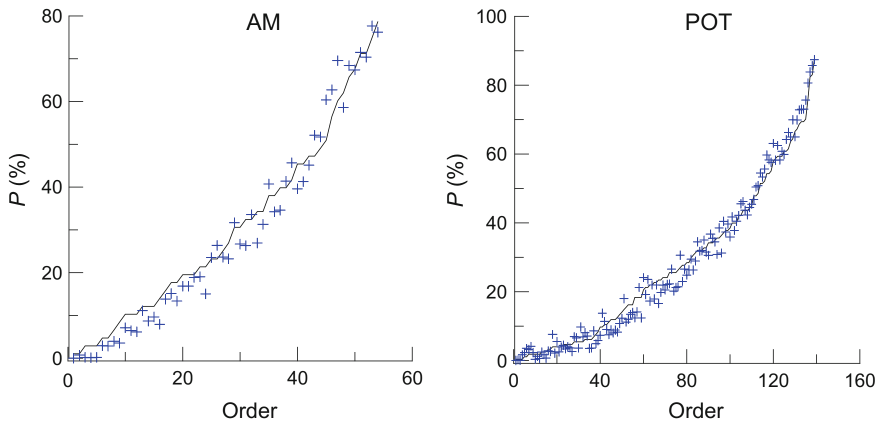


Fig. 4.5 Joint distribution and empirical probabilities of the observed combinations based on AM and POT samples

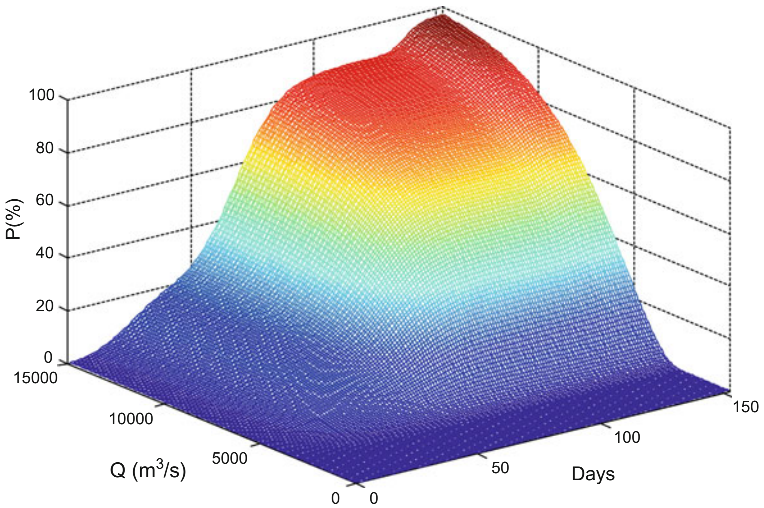


Fig. 4.6 Joint distribution of the flood occurrence dates and magnitudes based on AM samples

4.4.4 Seasonal Design Flood Estimation

The seasonal design floods in the pre-flood, main flood and post-flood sub-seasons are calculated by Eq. 4.11. The curve fitting method that based on minimizing the sum of the squares of the deviations between the observed values obtained from a plotting position formula and theoretical values calculated by Eq. 4.11 for each sub-season is used. An objective function of curve fitting method is given by

$$\text{Min } G(q_j) = \sum_{i=1}^{s}\sum_{j=1}^{N_i}\left(P(j) - P(q_j, A_i)\right)^2 \tag{4.14}$$

where N_i is the number of the observed data in the ith sub-season; q_j is the observed data in the ith sub-season. $P(j)$ is the cumulative frequency calculated by Eq. 3.9.

The Quasi-Newton method is used to optimize above objective function, and the estimated parameters of von Mises distribution and the seasonal design flood values for AM and POT samples are listed in Tables 4.6 and 4.7 respectively. Figure 4.7 shows that the theoretical curve of seasonal design floods can fit the observational data well.

The relations between the seasonal and annual frequency curves are shown in Fig. 4.8. The seasonal design flood frequency curves are rational, from the point of view that they are lower than the annual design flood frequency curve. Furthermore, the relations between the annual and seasonal design flood frequency curves must be obeyed the Eq. 4.4 or 4.10, which is also taken as a criterion to test the rationality of the seasonal design flood. A goodness-of-fit test for observed and

Table 4.6 Estimated parameters of the Mixed von Mises distribution for AM and POT samples

AM			POT		
μ_i	κ_i	p_i	μ_i	κ_i	p_i
1.03	4.80	0.19	0.61	14.11	0.12
2.81	3.28	0.71	2.77	1.96	0.73
5.66	27.43	0.10	5.61	5.17	0.15

Table 4.7 Comparisons of annual maximum design flood with seasonal design floods estimated by different methods (m^3/s)

Methods	Return period (year)	Annual design	Design values Pre-flood season		Main flood season		Post-flood season	
Chinese method	1,000	22,800	18,700	(−17.98%)	22,200	(−2.63%)	20,500	(−10.09%)
	200	18900	15,090	(−20.31%)	18,383	(−2.92%)	15,890	(−16.08%)
	100	17,400	13,500	(−22.41%)	16,800	(−3.45%)	14,500	(−16.67%)
Singh's method	1,000	22,800	18,384	(−19.37%)	27,298	(19.73%)	15,018	(−34.13%)
	200	18,900	15,206	(−19.69%)	22,883	(20.85%)	13,395	(−29.26%)
	100	17,400	14,282	(−17.92%)	20,626	(18.54%)	12,545	(−27.90%)
Proposed method AM	1,000	22,800	20,300	(−10.96%)	24,000	(5.26%)	22,200	(−0.03%)
	200	18,900	15,400	(−18.52%)	19,200	(1.59%)	17,300	(−0.08%)
	100	17,400	14,200	(−18.39%)	18,100	(4.02%)	16,200	(−0.07%)
POT	1,000	22,044	21,040	(−4.55%)	22,750	(3.20%)	22,042	(−0.01%)
	200	18,270	17,013	(−6.88%)	19,216	(5.18%)	17,986	(−1.55%)
	100	16,606	15,608	(−6.01%)	17,315	(4.27%)	16,604	(−0.01%)

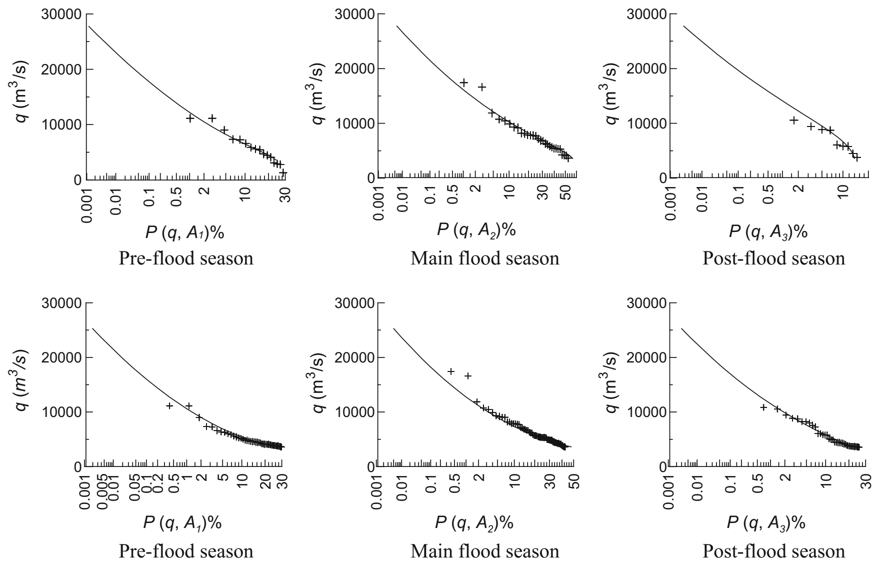


Fig. 4.7 Frequency curves of sub-season design floods based on AM and POT samples

calculated data are shown in Fig. 4.9, in which the line represents the annual theoretical probabilities derived by summing up the seasonal probabilities calculated by Eqs. 4.10 and 4.11, and the crossings represent the empirical probabilities.

1000-year seasonal design floods are estimated by Eq. 4.11, and the results based on AM series are shown in Fig. 4.10. It shows that a surface formed in a three-dimensional Cartesian coordinate system, indicating that various combinations of seasonal design floods can be obtained for a given return period T. As the increase of either or both of two seasonal design flood values, another design flood values will be decreased.

4.4.5 Comparisons of Different Methods

The current seasonal design flood method used in China assumes that the design frequency in each sub-season is identical. In accordance with this hypothesis and the demand of satisfying the flood prevention standards, the seasonal design frequencies must obey the following rules

$$P_X = P_Y = P_Z = P' \tag{4.15}$$

where P_X, P_Y, and P_Z are the design frequencies of the pre-flood season, main flood season and post-flood season, respectively. If the annual maximum flood series is

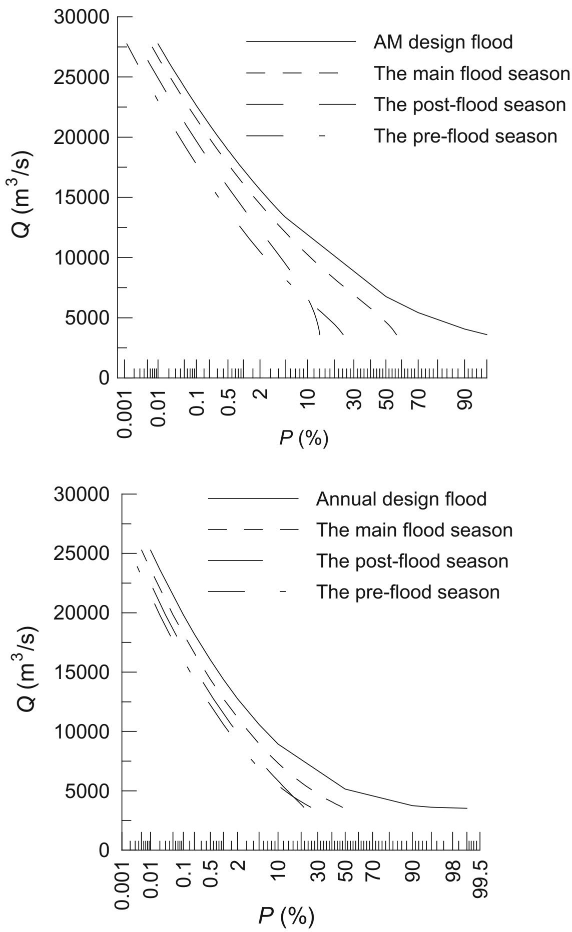


Fig. 4.8 The relations between seasonal and annual design flood frequency curves based on AM and POT samples

used, then P' equals $1/(3T)$. If the POT samples are used, the annual return period needs to be converted to the exceedance probability of the POT method as follows (Rosbjerg 1993):

$$P(Q \geq Q_T) = \frac{1}{\lambda T} \tag{4.16}$$

The comparisons of the annual maximum and seasonal design floods estimated by different methods are given in Table 4.7, where the relative error describes the deviation between the annual design values and seasonal design values. Table 4.7 implies that the seasonal design flood values based on the seasonal maximum series are underestimated in all sub-seasons. Since all of the seasonal design values are less than the annual ones, the current seasonal design flood method used in China is unable to satisfy the flood prevention standards.

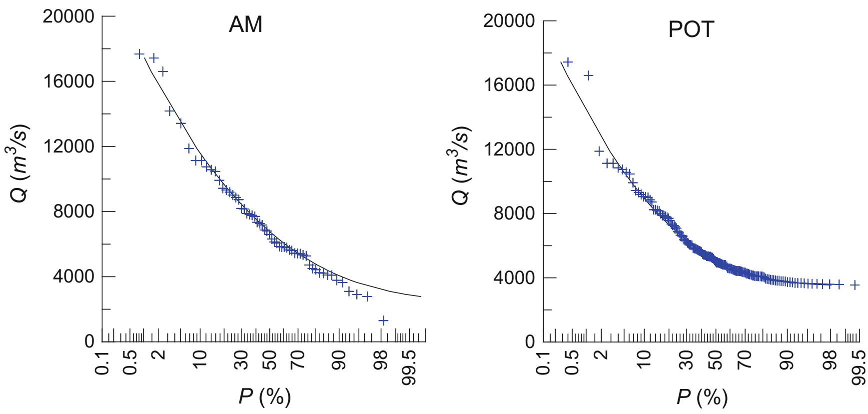


Fig. 4.9 Rational tests of the seasonal design floods based on AM and POT samples

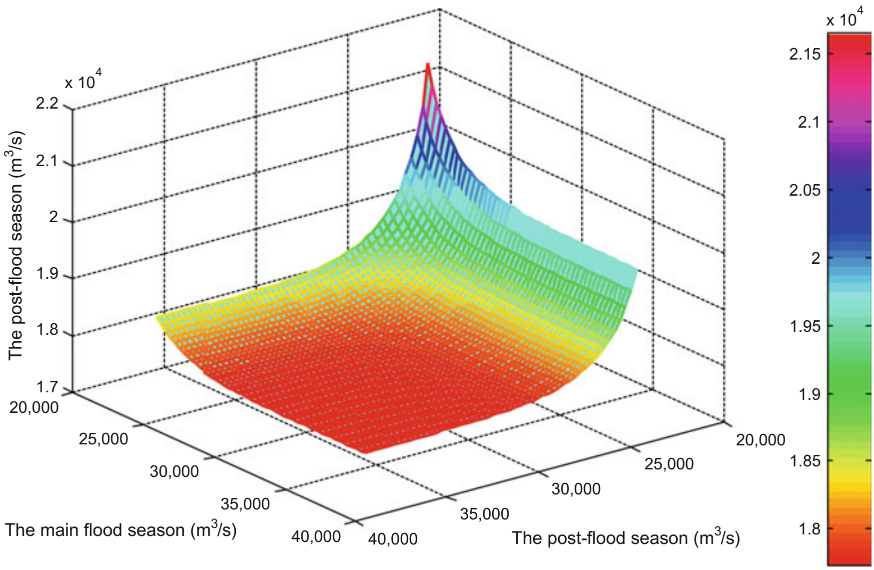


Fig. 4.10 1000-year seasonal design flood peak discharges with different combinations

For the seasonal design flood method suggested by Singh et al. (2005), the sample size of the pre-flood season, main flood season and post-flood season are 16, 20 and 19, respectively. The design values in the pre-flood and post-flood season are lower than those calculated by the Chinese seasonal design flood method. On the other hand, the design values in the main flood season are much higher than that of the annual maximum design floods. The annual and seasonal frequency curves based on Singh's method are drawn in Fig. 4.11. It is shown that the seasonal

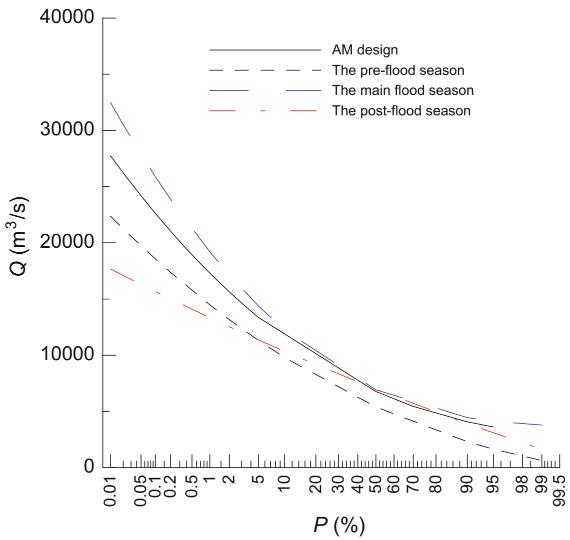


Fig. 4.11 Comparisons of frequency curves calculated by Singh's method with the annual maximum design method

frequency curve of the main flood season is above the annual maximum one. The seasonal frequency curve in the pre-flood season is higher than that in the post-flood season. Actually, the flood in the post-flood season is much larger than that in the pre-flood season. The mean values of the annual maximum flood peak in the pre-flood and post-flood season are 5792 m^3/s and 7060 m^3/s, respectively. The reason for these unreasonable results may be mainly due to that the sample series for some sub-seasons are too short for flood frequency analysis.

The seasonal design floods calculated by the proposed method are also listed in Table 4.7. It indicates that the seasonal design flood values based on AM samples are much higher than the annual maximum design flood values in the main flood season. The design values in the other sub-seasons are less than their corresponding annual maximum ones but greater than those calculated by current seasonal design flood methods. The T-year design flood values calculated by the POT samples are also listed in Table 4.7. It is shown that the seasonal design values of the POT samples in the main flood season exceed the T-year annual design values and less than them in other sub-seasons. The results based on two sampling methods demonstrate that the seasonal design floods estimated by the proposed method are slightly greater than annual design floods in the main flood season and less than them in other flood sub-seasons. Furthermore, the seasonal design floods calculated by Eqs. 4.10 and 4.11 can meet the flood prevention standards.

The design values of the POT samples is less than that of the AM flood series in the main flood season, whereas the values of POT series are larger than AM series in the pre-flood season. The reason for this is mainly due to that the discrepancy exists between the P-III distribution and Ex distribution. For example, 1000-year design values based on the POT and AM samples are equal to 22,044 m^3/s and 22,800 m^3/s, respectively. Compared with the 24% flood occurrence probability for

AM samples, it is about 30% for POT samples in the pre-flood season. No significant difference exists between the results of the POT and AM series in the post-flood season.

The design values of the 1-day maximum runoff volume W_{1d}, 3-day maximum runoff volume W_{3d}, and 7-day maximum runoff volume W_{7d} are also calculated by the proposed method. The parameters of the margin and joint distributions are estimated and listed in Table 4.8. 1000-year and 100-year design flood runoff volumes for each sub-season are calculated by Eq. 4.11 and listed in Table 4.9.

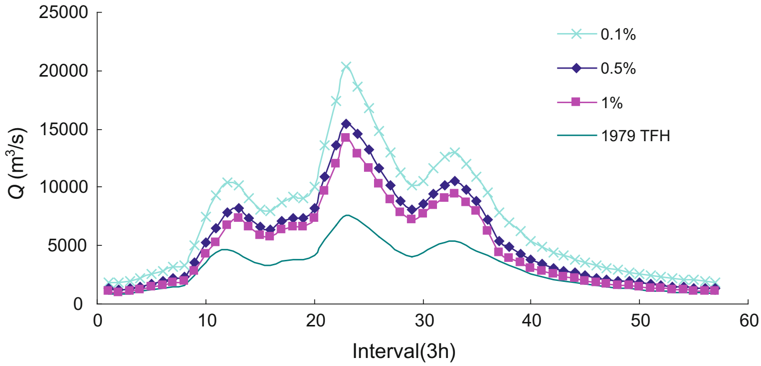
The seasonal FCWL is obtained by the flood hydrograph routing method based on the design flood hydrograph (DFH). One of the methods to derive the DFH is the typical flood hydrograph (TFH) method which has been widely used by practitioners (Nezhikhovsky 1971; Yue et al. 2002). The flood hydrograph with the highest peak or biggest volume is usually selected as a TFH. The DFH is constructed by multiplying each discharge ordinate of the TFH by an amplifier. The TFH of 1979, 1997 and 1998 were selected for the pre-flood season, the main flood season and the post-flood season, respectively. The peak and volume-amplitude (PVA) method is used to derive DFH (MWR 1993; Xiao et al. 2009), and the results are shown in Fig. 4.12. The design flood hydrographs are routed through the reservoir, and the seasonal design FCWL values are determined. They are 201.2, 192.1 and 200.1 m in the pre-flood season, main flood season and post-flood season, respectively.

Table 4.8 Estimated parameters of the marginal and joint distributions for runoff volumes

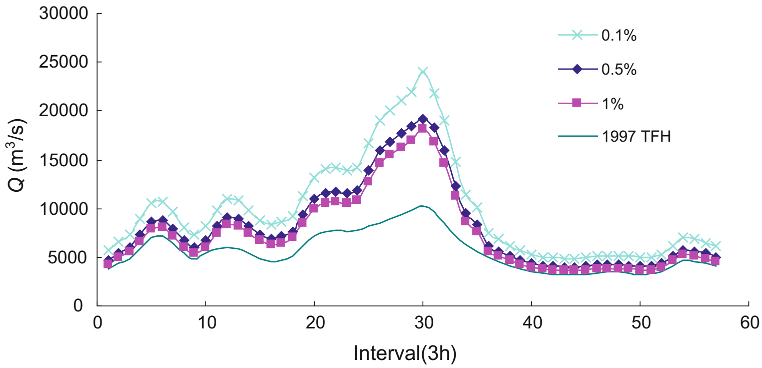
	P-III distribution			Mixed von Mises distribution			Joint distribution
	α	β	δ	μ_i	κ_i	p_i	θ
W_1				1.03	7.22	0.10	
	2.195	0.646	1.700	2.83	3.55	0.78	2.28
				5.67	35.59	0.11	
W_3				0.93	7.00	0.10	
	1.582	0.231	3.423	2.71	2.95	0.79	1.39
				5.50	8.88	0.11	
W_7				0.60	1.45	0.04	
	1.644	0.156	5.267	2.70	3.22	0.85	2.26
				5.55	4.35	0.12	

Table 4.9 Estimated design runoff volumes by the proposed method (billion m^3)

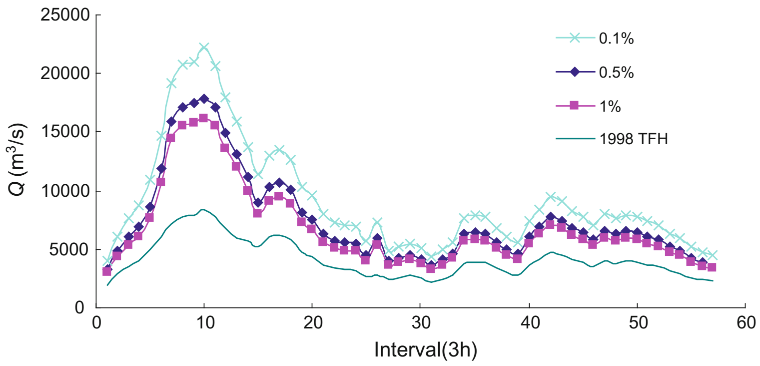
T	1000			200			100		
Days	Pre-flood	Main flood	Post-flood	Pre-flood	Main flood	Post-flood	Pre-flood	Main flood	Post-flood
W_1	13.55	17.71	16.61	10.64	14.9	13.77	9.38	13.67	12.51
W_3	31.27	37.45	32.93	24.99	31.3	26.68	22.23	28.62	23.93
W_7	44.78	60.15	55.89	34.38	49.97	45.60	29.89	45.53	41.08



(a) The pre-flood season



(b) The main flood season



(c) The post-flood season

Fig. 4.12 The derived DFH of the Geheyan reservoir by PVA method

Table 4.10 Comparisons of energy output and flood water utilization rates based on different seasonal design FWCL

Index	Comparison of methods	Year		
		Wet	Normal	Dry
Annual electricity generation	Current (10^8 kW h)	35.23	27.44	23.21
	Proposed (10^8 kW h)	35.89	28.04	23.68
	Increment of generation (%)	1.87	2.19	2.02
Annual spill release	Current ($10^8 m^3$)	25.57	6.85	3.83
	Proposed ($10^8 m^3$)	24.01	5.34	1.75
	Reduced spill release (%)	6.04	10.14	55.24
Flood water resources utilization rate	Current (%)	82.57	93.02	95.70
	Proposed (%)	83.29	94.56	98.05

The daily discharge data set from 1951 to 2004 is used to analyze and compare the benefit of seasonal design FCWL with the current scheme. Three representative years, wet year (1964), normal year (1985) and dry year (2001) are selected for the analysis. Annual electricity generation, spill release and flood water resources utilization at the Geheyan reservoir are calculated and listed in Table 4.10. It can be seen that compared with the current scheme, the annual electricity generations based on the proposed FCWL is increased 1.87, 2.19 and 2.02% in the wet year, normal year and dry year respectively. The annual spill release is reduced. The flood water utilization rate is increased from 82.57 to 83.29% for the wet year, and from 95.70 to 98.05% for the dry year. Therefore, the proposed FCWL can increase energy output and flood water utilization rate.

4.5 Conclusion

Seasonal design floods, which reflect the seasonal flood variation, are very important for reservoir operation and management. A bivariate joint distribution based on copula function, which considers the flood occurrence dates and magnitudes is proposed and established. The main conclusions of this chapter are summarized as follows:

(1) The current seasonal design flood method used in China cannot satisfy the flood prevention standards. Although the Singh's method based on the annual maximum series can meet these standards, the estimated design floods have large errors due to the short length of sample series.
(2) Compared with two current seasonal design flood methods, the proposed method that considers both flood occurrence dates and flood magnitudes is

much more rational in the physical mechanism and can satisfy flood prevention standards in China.

(3) The proposed method can increase energy output and flood water utilization rate and provides a new way for seasonal flood estimation.

References

Akaike H (1974) A new look at the statistical model identification. IEEE Trans Autom Control AC 19(6):716–722

Black AR, Werritty A (1997) Seasonality of flooding: a case study of north Britain. J Hydrol 195:1–25

Carta JA, Ramírez P (2007) Analysis of two-component mixture Weibull statistics for estimation of wind speeds using mixture of von Mises distribution. Renew Energy 32:518–531

Carta JA, Bueno C, Ramírez P (2008) Statistical modeling of directional wind speeds using mixtures of von Mises distributions: case study. Energy Convers Manage 49:897–907

Chen L, Guo S, Yan B, Liu P, Fang B (2010) A new seasonal design flood method based on bivariate joint distribution of flood magnitude and date of occurrence. Hydrol Sci J 55(8):1264–1280

Creager WP, Kinnison HB, Shifrin H, Snyder FF, Williams GR, Gumbel EJ, Matthes GH (1951) Review of flood frequency methods: final report of the subcommittee of the joint division committee on floods. ASCE Transactions 116:1220–1230

Cunderlik JM, Ouarda TBMJ, Bobée B (2004) Determination of flood seasonality from hydrological records. Hydrol Sci J 49(3):511–526

Durrans SR, Eiffe MA, Thomas WO, Goranflo HM (2003) Joint seasonal/annual flood frequency analysis. J Hydrol Eng ASCE 8(4):181–189

Fang B, Guo SL, Wang SX, Liu P, Xiao Y (2007) Non-identical models for seasonal flood. Hydrol Sci J 52(5):974–991

Guo SL, Zhang HG, Chen H, Peng DZ, Liu P, Pang B (2004) Reservoir flood forecasting and control system in China. Hydrol Sci J 49(6):959–972

Hosking JRM, Wallis JR (1997) Regional frequency analysis: an approach based on L-moments. Cambridge University Press, London

Institute of Hydrology (IH) (1999) Flood estimation handbook, vol 3. Inst Hydro. Wallingford, UK

Lang M, Ouarda TBMJ, Bobée B (1999) Towards operational guidelines for over-threshold modeling. J Hydrol 225(3–4):103–117

Michael L, Andrew M, Martin L (2007) Frequency analysis of rainfall and streamflow extremes accounting for seasonal and climatic partitions. J Hydrol 348:135–147

Ministry of Water Resources (MWR) (1993) Regulation for calculating design flood of water resources and hydropower projects. Chinese Water Resour. Hydrop. Press, Beijing, China (in Chinese)

Nezhikhovsky RA (1971) Channel network of the basin information. Hydrometeorological, Leningrad, Russia

Ngo LL, Madsen H, Rosbjerg D (2007) Simulation and optimization modelling approach for operation of the Hoa Binh reservoir. Vietnam J Hydrol 336:269–281

Ouarda TBMJ, Ashkar F, EJabi N (1993) Peaks over threshold model for seasonal flood variations. In: Kuo CY (ed) Engineering hydrology, Proceedings of international symposium, pp 341–346, ASCE, Reston, VA

Ouarda TBMJ, Cunderlik JM, St-hilaire A, Barbet M, Brunear P, Bobée B (2006) Data-based comparison of seasonality-based regional flood frequency methods. J Hydro 330:329–339

Rosbjerg D (1993) Partial duration series in water resources. Technical University of Denmark, Denmark

Singh VP, Wang SX, Zhang L (2005) Frequency analysis of nonidentically distributed hydrologic flood data. J Hydrol 307:175–195

Thomas W, O. Jr, Crockett KL, Johnson A(1998) Discussion of techniques for analysis of ice-jam flooding. In: Proceedings of Association of State Floodplain Managers

Waylen P, Woo MK (1982) Prediction of annual floods generated by mixed processes. Water Resour Res 18(4):1283–1286

Xiao Y, Guo SL, Liu P, Yan BW, Chen L (2009) Design flood hydrograph based on multi-characteristic synthesis index method. J Hydrol Eng 14(12):1359–1364

Yin JB, Guo SL, Liu ZJ, Xiong F (2017) Bivariate seasonal design flood estimation based on copulas. J Hydrol Eng 22(12). https://doi.org/10.1061/(asce)he.19435584.0001594

Yue S, Quarda TBMJ, Bobée B, Legendre P, Bruneau P (2002) Approach for describing statistical properties for flood hydrograph. J Hydrol Eng 7:147–153

Zhang L, Singh VP (2006) Bivariate flood frequency analysis using the copula method. J Hydrol Eng 11(2):150–164