

DOI: 10.19797/j.cnki.1000-0852.20220241

梯级水库下游设计洪水地区组成法比较研究

谢雨祚¹, 郭生练¹, 王 俊¹, 钟斯睿¹, 汪 芸¹, 李 帅², 胡 挺²

(1. 武汉大学 水资源与调度全国重点实验室, 湖北 武汉 430072;
2. 中国长江三峡集团有限公司 流域枢纽运行管理中心, 湖北 宜昌 443133)

摘要:水利水电工程兴建运行改变了下游河流的水文情势和来水来沙条件,推求梯级水库下游控制断面设计洪水成了亟需解决的难题。本文综述设计洪水地区组成法的研究进展及问题,推导概率组合离散求和法和改进离散求和法,对比分析两者的离散估计误差;结合清江水布垭~隔河岩~高坝洲梯级水库应用实例,比较讨论各种设计洪水地区组成法的优缺点和实用性。结果表明:(1)基于Copula函数的改进离散求和法可直接推导公式,且无需作独立化处理;离散求和法的总误差随着相关系数 ρ 的增大而显著增加;当 $\rho>0.2$ 时不宜使用。(2)经水布垭~隔河岩水库调蓄后,高坝洲坝址1000年一遇设计洪峰流量削减了50%左右。(3)对于单座或2座水库,离散求和法和同频率地区组成法的适用条件分别为 $0<\rho\leq 0.2$ 和 $\rho>0.7$,改进离散求和法和最可能地区组成法没有限制条件;对于3座及以上的梯级水库群,最可能地区组成法可得到唯一最佳方案解。研究成果可为我国《水利水电工程设计洪水计算规范》的修订提供参考。

关键词:梯级水库;洪水地区组成;Copula函数;离散求和法;统计试验;清江流域

中图分类号:TV122.3

文献标识码:A

文章编号:1000-0852(2023)05-0018-09

0 引言

我国已建9万多座水库并投入运行,形成梯级水库格局并显著地改变了河流水文情势和来水来沙条件。推求受水库调蓄影响后下游控制断面的设计洪水,已经成为流域梯级开发和水工程联合调度过程中亟待解决的科学技术难题。《水利水电工程设计洪水计算规范》(以下简称《规范》)明确指出^[1]：“当设计断面上游有调蓄作用较大的水库或设计水库对下游有防洪任务时,应对大洪水的地区组成进行分析,并拟定设计断面以上或防洪控制断面以上设计洪水的地区组成”、“设计洪水的地区组成可采用典型洪水组成法或同频率洪水组成法拟定。两种洪水组成法的各分区设计洪水过程均应采用同一次洪水过程线为典型”。

国内外学者就梯级水库下游断面设计洪水的问题

开展了一系列研究。王锐琛等^[2]假定水库调蓄后的出库洪峰流量与相应入库洪峰流量的概率相等,首次提出概率组合离散求和法。该法能够考虑各分区的所有可能组合,计算结果比较合理,并于1993年纳入修订的《规范》中。离散求和法受到联合概率分布函数求解的限制,通常需进行独立性处理,这无疑会影响计算结果的精度。第一,独立性处理通常只考虑线性变换,忽略了变量间的非线性关系,数据转化过程中难免出现相关性信息失真;第二,代换变量 E 的物理概念不明确,其拟合频率曲线下端可能出现负值,对于分区较多的梯级水库群,多次独立性变换容易造成误差累积。近年来随着Copula函数在多变量水文分析计算中的广泛应用^[3],可用来推求梯级水库下游设计洪水地区组成。闫宝伟等^[4]应用Copula函数构造了上游断面与区间洪水的联合分布,提出和推导了最可能洪水地区

收稿日期:2022-07-01

网络首发日期:2023-10-23

网络首发地址:https://kns.cnki.net/kcms2/detail/11.1814.P.20231020.1552.019.html

基金项目:国家重点研发计划项目(2021YFC3200305);中国长江三峡集团有限公司(0799254)和国家自然科学基金(51879192)资助项目

作者简介:谢雨祚(1998—),男,湖北宜城人,博士研究生,主要从事水文水资源的研究。E-mail:yuzuo@whu.edu.cn

通信作者:郭生练(1957—),男,福建龙岩人,教授,挪威工程院外籍院士,主要从事水文水资源等方面的研究。E-mail:slguo@whu.edu.cn

组合法。该法考虑了地区间洪水的空间相关性,更加符合客观实际,为设计洪水地区组成分析计算提供了一条新的途径。李天元等^[5]以 Copula 函数理论为基础,构造了水库断面洪量与区间洪量的联合分布,推求了条件概率函数的显式表达式,提出了基于 Copula 函数的改进离散求和法,通过直接对条件概率曲线进行离散,克服了《规范》中离散求和法需要进行变量独立性转换的问题。刘章君等^[6]利用 Copula 函数建立了各分区洪水的联合分布,基于联合概率密度最大原则,推导得到最可能地区组合法的计算通式,并用来推求梯级水库下游断面的设计洪水。Xiong 等^[7]采用理论推导和统计试验法,开展同频率地区组合法和最可能地区组合法的对比研究。发现同频率地区组合法在只有 1 或 2 座水库且分区洪水相关系数 $\rho > 0.7$ 时,其计算结果合理可行;而对于更高维的梯级水库,各分区洪水“同频率”的假设难以成立,且同频率地区组合法需拟定的方案数呈指数增加,带来较大的不确定性。最可能地区组合法的方案数唯一,在梯级水库系统中适用性更优。

现有洪水地区组成方法在工程实践中大都存在较明显的不足^[8]。《规范》推荐的典型年法和同频率地区组合法,虽然简便易行,但人为不确定性较大;概率组合法需要对分区的频率曲线进行离散求和,在独立性转换中难免出现数据失真。对于三座以上的梯级水库群,上述方法不仅分析计算难度大,设计结果具有较大的不确定性。基于 Copula 函数理论的多变量分析方法具有较强的统计基础,计算步骤简单客观,但所得设计结果的合理性、适用性尚有待进一步检验论证。本文基于 Copula 函数和概率组合离散求和的思路,推导单一和梯级水库下游防洪断面年最大流量的概率分布,重点探讨离散求和法的估计误差,结合清江水布垭~隔河岩~高坝洲梯级水库实例,推求受水库调蓄影响的下游断面设计洪水,采用统计试验开展同频率地区组合法、最可能地区组合法、离散求和法和改进离散求和法的比较研究,确定各种地区组合法的适用条件,寻求最佳的设计洪水地区组合法。

1 概率组合离散求和法理论推导

1.1 洪水地区组成

洪水地区组成是各区洪水随机组合的结果,是大中型水利水电工程、防洪系统工程设计洪水计算的重要内容,对流域梯级开发和工程规模的确定具有深远

影响。图 1 为洪水地区组成示意图,展示了 n 个水库 $A_1, A_2, \dots, A_{n-1}, A_n$ 与设计断面 C 的洪水地区组成情况,其中 $B_1, B_2, \dots, B_{n-1}, B_n$ 为梯级水库的 n 个区间流域。随机变量 $X_i, Y_i (i=1, 2, \dots, n)$ 和 Z 分别表示水库 A_i 、区间流域 B_i 和设计断面 C 的天然来水量,依次取值为 x_i, y_i 和 z 。

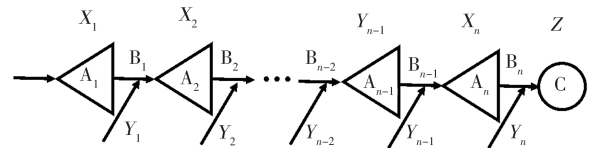


图 1 n 个梯级水库与设计断面的洪水地区组成示意图
Fig.1 Sketch diagram of n -cascade reservoirs and design section

由于水库河网调节等因素的影响,往往难以推求设计洪峰流量的地区组成,且对调洪能力较大的水库,洪量起主要作用。因此,通常将设计断面 C 某一设计频率 p 的时段洪量 z_p 分配给上游 $n+1$ 个组成部分。由水量平衡原理得:

$$x_1 + \sum_{i=1}^n y_i = z_p \quad (1)$$

设计洪水地区组成的本质是给定断面 C 的设计洪量 z_p ,在满足式(1)约束条件下对设计洪量 z_p 进行分配,得到组合 $[x_1, y_1, y_2, \dots, y_{n-1}, y_n]$ 。

1.2 概率组合离散求和法

洪水地区组成依据上游水库和区间洪水的概率组合,推求下游设计断面受调蓄影响的频率曲线。对具有复杂调洪规则的梯级水库,《规范》采用离散求和法推求下游防洪断面最大流量的概率分布,该法可以在分布函数取值区间内考虑各分区的所有可能组合。

设 $P_X = F_X(x)$ 为 X 的概率分布函数, $P_{Y|X} = F_{Y|X}(y|x)$ 为在条件 $X=x$ 下 Y 的条件概率分布函数,由水量平衡 $X + Y = Z$ 、二元分布及离散微积分知, Z 的分布函数为:

$$F_Z(z) = P(X + Y \leq z) = \int_{-\infty}^{\infty} f_X(x) \int_{-\infty}^{z-x} f_{Y|X}(y|x) dy dx = \int_{-\infty}^{\infty} f_X(x) P_{Y|X} dx = \int_0^1 P_{Y|X} dP_X = \sum P_{Y|X} \cdot \Delta P_X = \sum \sum \Delta P_{Y|X} \cdot \Delta P_X \quad (2)$$

因此,求解下游断面天然来水量 Z 的概率问题,可转化为求解在 $X=x$ 条件下 Y 的条件概率函数 $P_{Y|X}$ 的问题。而随机变量 X, Y 相互独立是离散求和法的理论基础,常用相关系数法判断独立性。

1.2.1 单座水库下游断面设计洪水

对于单座水库,根据入库流量和区间流量统计特征,分两种情况讨论。

(1) X 与 Y 相互独立,则式(2)即可表示为:

$$F_Z(z) = \sum \sum \Delta P_Y \cdot \Delta P_X \quad (3)$$

将连续型随机变量 X 与 Y 的频率曲线离散为阶梯状,每一个 X 、 Y 的状态都与一概率区间对应,设 X 取状态 x_i 的概率区间为 $\Delta P_{X,i}$, Y 取状态 y_j 的概率区间为 $\Delta P_{Y,j}$ ($i=1,2,\dots,m_X; j=1,2,\dots,m_Y$)。 Z 相应状态对应的概率区间为 $\Delta P_{Z,ij}$,则 Z 的取值状态数为 $m_Z = m_X \cdot m_Y$,有:

$$\Delta P_{Z,ij} = \Delta P_{X,i} \cdot \Delta P_{Y,j} \quad (4)$$

对 Z 的每一个取值状态,均按照 x_i 及 y_j 控制缩放A断面及区间的典型洪水过程线,将A断面洪水过程线经调洪后得到下泄流量过程线,再与区间过程线组合后得到C断面受A库调洪后最大流量 Q_C 的一个数值 $q_{C,ij}$ 。显然 $q_{C,ij}$ 的出现概率等于 $\Delta P_{Z,ij}$,即:

$$P(Q_C = q_{C,ij}) = \Delta P_{Z,ij} = \Delta P_{X,i} \cdot \Delta P_{Y,j} \quad (5)$$

如果要求经水库A调洪后C断面最大流量等于或大于某一指定流量 q_s 的概率,即:

$$P(Q_C \geq q_s) = P(q_{C,ij} \geq q_s) = 1 - P(q_{C,ij} \leq q_s) = 1 - \sum_{q_{C,ij} \leq q_s} \sum \Delta P_{X,i} \cdot \Delta P_{Y,j} \quad (6)$$

(2) X 与 Y 不相互独立。若经独立性检验表明 X 与 Y 不能视作相互独立的随机变量,则需在计算前将其转换成独立(严格来讲是不相关)随机变量,再进行频率组合计算。如当 X 与 Y 线性相关时,可进行代换如下:

$$E_X = Y - k_1 X \text{ 或 } E_Y = X - k_2 Y \quad (7)$$

式中: E_X 、 E_Y 为新构建的随机变量, k_1 、 k_2 为转换系数。以新变量 E_X 或 E_Y 代替 Y 或 X ,一般对均值较小的变量作代换,而系数 k_1 、 k_2 可由最小二乘法确定。假设采用 E_X 代换变量 Y ,则式(2)可表示为:

$$\begin{aligned} F_Z(z) &= \iint_{X+Y \leq z} f_{X,Y}(x,y) dy dx = \\ &= \iint_{X+k_1 X+E_X \leq z} f_{X,E}(x,e_x) d(kx+e_x) dx = \\ &= \iint_{(k_1+1)X+E_X \leq z} f_{X,E}(x,e_x) de_x dx = \\ &= \sum P_{YE_X} \cdot \Delta P_X = \sum \sum \Delta P_{E_X} \cdot \Delta P_X \end{aligned} \quad (8)$$

若 X 离散成 m_X 种状态, E_X 离散成 m_E 种状态,则 Z

有 $m_Z = m_X \cdot m_E$ 种状态,当 $X=x_i, E_X=e_{x,j}$ 时, Y 的取值 $y_j = k_1 x_i + e_{x,j}$,其出现的概率区间为:

$$\Delta P_{Z,ij} = \Delta P_{X,i} \cdot \Delta P_{E_X,j} \quad (9)$$

经独立性处理后,可以类似独立随机变量进行状态组合求和计算,见式(5)和式(6)。

1.2.2 梯级水库下游断面设计洪水

从单座水库开始,每增加一座水库,离散求和法的全部组合状态将成指数增长,例如有 n 座水库,每一组合变量离散后的状态均取 m 个,则全部组合状态有 m^{n+1} 个。对于两座梯级水库,推求当 $X_1 = x_1$ 时 Y_1 的条件概率函数 $P_{Y_1|X_1}$ 以及当 $X_1 = x_1$ 且 $Y_1 = y_1$ 时 Y_2 的条件概率函数 $P_{Y_2|X_1, Y_1}$,则式(2)可以表示为:

$$F_Z(z) = \sum \sum P_{Y_2|Y_1, X_1} \cdot \Delta P_{Y_1|X_1} \cdot \Delta P_{X_1} = \sum \sum \sum \Delta P_{Y_2|Y_1, X_1} \cdot \Delta P_{Y_1|X_1} \cdot \Delta P_{X_1} \quad (10)$$

以随机变量 X_1 和 Y_1 不相互独立为例,采用独立性检验之后,用随机变量 E_1 代换 Y_1 ,则 X_1 与 E_1 相互独立。同理,用随机变量 E_2 代换 Y_2 ,使 X_1 与 E_2 相互独立。则式(2)可以表示为:

$$F_Z(z) = \sum \sum \sum \Delta P_{E_2|E_1, X_1} \cdot \Delta P_{E_1} \cdot \Delta P_{X_1} \quad (11)$$

若随机变量 E_1 与 E_2 在给定 $X_1=x_1$ 时条件独立,式(11)可以写为:

$$F_Z(z) = \sum \sum \sum \Delta P_{E_2|X_1} \cdot \Delta P_{E_1} \cdot \Delta P_{X_1} = \sum \sum \sum \Delta P_{E_2} \cdot \Delta P_{E_1} \cdot \Delta P_{X_1} \quad (12)$$

离散 X_1 、 E_1 及 E_2 的频率曲线:设 X_1 取 m_{X1} 个状态, E_1 取 m_{E1} 个状态, E_2 取 m_{E2} 个状态,则组合变量 Z 的状态 $m_Z = m_{X1} \cdot m_{E1} \cdot m_{E2}$ 。按照离散后各概率区间对应的洪量放大各分区典型洪水过程,经 A_1 水库调蓄后的过程线再与 $A_1 \sim A_2$ 区间过程线叠加演算到 A_2 水库坝址,再经 A_2 水库调洪并演算至C断面,统计 m_Z 组的不同过程线的C断面最大流量及对应的频率,即可得到经 A_1 、 A_2 两库调洪后C断面最大流量频率曲线。每增加一座水库,组合状态将呈方次增加,例如有 n 座水库,每一组合变量离散后的状态均取 m 个,则全部组合状态有 m^{n+1} 个^[8]。

1.3 改进概率组合离散求和法

受联合概率分布函数求解技术的限制,《规范》中概率组合离散求和法需要将组合变量转换成独立随机变量,再进行频率组合计算,对于分区较多的梯级水库群,多次独立性变换容易造成误差累积。为克服上述问题,基于Copula函数构造上游水库断面来水和区间洪量的联合分布函数。当随机变量 X 、 Y 的联合分布

$C(u, v)$ 确定后, Z 的分布函数式(2)可以表示为:

$$F_Z(z) = P(X + Y \leq z) = \int_{-\infty}^{\infty} \int_{-\infty}^{z-x} f(x, y) dy dx = \int_{-\infty}^{\infty} f_X(x) \int_{-\infty}^{z-x} c(u, v) f_Y(y) dy dx = \int_0^1 \int_0^{F_Y(z-x)} c(u, v) dv du \quad (13)$$

式中: u, v 分别为边缘分布函数,即 X 和 Z 的概率分布函数; $c(u, v) = \partial C(u, v) / \partial u \partial v$ 。

1.3.1 单座水库下游断面设计洪水

单座水库 $n=1$,对式(13)进行离散化处理,将 X 和 Y 的分布函数 u, v 分别离散为 m_X, m_Y 个状态,则 Z 的状态组合情况为 $m_Z = m_X \cdot m_Y$ 。设 $X=x_i$ 时分布函数 u 的概率区间为 (u_i, u_{i+1}) , Y 取状态 y_j 时 v 的概率区间为 (v_j, v_{j+1}) , Z 相应状态对应的概率区间为 $\Delta P_{z,ij}$,则

$$\Delta P_{z,ij} = \int_{u_i}^{u_{i+1}} \int_{v_j}^{v_{j+1}} c(u, v) dv du \quad (14)$$

按照离散后各概率区间对应的洪量放大各分区典型洪水过程,统计 m_Z 组的不同过程线的C断面最大流量及对应的频率,得到经A库调洪后C断面最大流量频率曲线。同式(5), $q_{C,ij}$ 的出现概率等于 $\Delta P_{z,ij}$ 。若要经水库A调洪后C断面最大流量等于或大于某一指定流量 q_s 的概率,即:

$$P(Q_C \geq q_s) = P(q_{C,ij} \geq q_s) = 1 - P(q_{C,ij} \leq q_s) = 1 - \sum_{q_{C,ij} \leq q_s} \sum c(u, v) \cdot \Delta u \cdot \Delta v \quad (15)$$

对 X, Y 的全部取值状态作组合计算,并统计出最大洪峰流量及对应的发生概率,即可得C断面受A库调洪后洪峰的频率曲线。

1.3.2 梯级水库下游断面设计洪水

从单座水库开始,每增加一座水库,离散求和法的全部组合状态将指数增长。例如有 n 座水库,每一组组合变量离散后的状态均取 m 个,则全部组合状态有 m^n 个。当上游有两座梯级水库时,采用Copula函数构造联合分布 $C(u_1, v_1, v_2)$,则式(13)可表示为:

$$F_Z(z) = P(X_1 + Y_1 + Y_2 \leq z) = \int_{-\infty}^{\infty} \int_{-\infty}^{z-x_1} \int_{-\infty}^{z-x_1-y_1} f(x_1, y_1, y_2) dy_2 dy_1 dx_1 = \int_{-\infty}^{\infty} f_{X_1}(x) \int_{-\infty}^{z-x_1} f_{Y_1}(y) \int_{-\infty}^{z-x_1-y_1} f_{Y_2}(y) c(u_1, v_1, v_2) dy_2 dy_1 dx_1 = \int_0^1 \int_0^{F_{Y_1}(z-x_1)} \int_0^{F_{Y_2}(z-x_1-y_1)} c(u_1, v_1, v_2) dv_2 dv_1 du_1 \quad (16)$$

将 X_1, Y_1 及 Y_2 的频率曲线 u_1, v_1, v_2 离散化。设 X_1 取 m_{X1} 个状态, Y_1 取 m_{Y1} 个状态, Y_2 取 m_{Y2} 个状态,则组合变量 Z 的状态 $m_Z = m_{X1} \cdot m_{Y1} \cdot m_{Y2}$ 。C断面最大流量频率曲线计算具体步骤同前。

2 离散求和法的统计试验

概率组合离散求和法的实质是对理论连续分布进行离散化,并用求和代替积分,所以会产生误差。采用统计试验,对离散求和法与改进离散求和法进行对比研究。为便于随机数的生成,假定边缘分布为正态分布,联合分布为二维正态分布。

2.1 理论分布

为简化边缘分布和联合分布计算,设区间流量服从 $Y \sim N(0, 1)$,由于入库流量均值一般大于区间流量均值,因此设入库流量服从 $X \sim N(1, 1)$,二者相关系数 ρ 为^[7]:

$$\rho = \frac{\text{Cov}(X, Y)}{\sqrt{D(X)} \sqrt{D(Y)}} = \text{Cov}(X, Y) \quad (17)$$

式中: $\text{Cov}(X, Y)$ 为随机变量 X 和 Y 的协方差。则 X 和 Y 的联合分布的协方差矩阵为:

$$\Sigma_{XY} = \begin{pmatrix} \text{Cov}(X, X) & \text{Cov}(X, Y) \\ \text{Cov}(Y, X) & \text{Cov}(Y, Y) \end{pmatrix} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \quad (18)$$

根据边缘分布和式(18)的协方差矩阵可以计算二元正态分布 $(X, Y) \sim N(1, 0, 1, 1, \rho)$ 在任何离散点 (x_i, y_j) 上的理论概率分布值 $F_{X,Y}(x_i, y_j)$;由于随机变量 $E \sim N(-r, \sqrt{1-r^2})$ 是独立化处理后的结果,与随机变量 X 的相关系数近似为0,亦可根据协方差公式计算二元正态分布 $(X, E) \sim N(1, -r, 1, \sqrt{1-r^2}, 0)$ 在任何离散点 (x_i, e_j) 上的理论概率分布值 $F_{X,E}(x_i, e_j)$ 。

2.2 离散求和法随机模拟

根据式(18)所示协方差矩阵和两个边缘分布的参数,分别生成2 000个随机变量 X, Y 的随机数,在离散求和法的统计试验中,假定样本仍然服从正态分布,即可估计随机变量 X 和 Y 的分布参数;若构建随机变量 E ,其分布类型亦为正态分布,其转换系数通过式(7)由最小二乘法确定,然后估计正态分布参数。

离散求和法根据随机变量 X 和 Y 的相关系数进行独立性检验,决定是否通过构造与随机变量 X 相互独立的随机变量 E 进行后续的概率离散运算,设置相关系数 ρ 的取值区间为 $[0, 0.98]$ 。当相关系数 ρ 较小时才可以视为相互独立, ρ 较大时构建随机变量 E ,但其阈

值的具体大小尚不明朗。因此本次统计试验分两种情况讨论离散求和法的误差:

(1)离散求和法Y:认为随机变量X和Y相互独立,不再构建随机变量E。在相关系数ρ的取值区间内,直接采用式(3)、式(4)求得Z的概率区间ΔP_{Z,ij},进而计算离散点(x_i, y_j)上的概率分布估计值:

$$\hat{F}_{y,ij}(X = x_i, Y = y_j) = \Delta P_{Z,ij} = \Delta P_{X,i} \cdot \Delta P_{Y,j} \quad (19)$$

(2)离散求和法E:认为随机变量X和Y不相互独立,构建随机变量E。在相关系数ρ的取值区间内,采用式(8)和式(9)计算Z的概率区间ΔP_{Z,ij},并用式(7)将随机变量E转化为Y,进而计算离散点(x_i, y_j)上的概率分布估计值:

$$\hat{F}_{e,ij}(X = x_i, E = e_j) = \hat{F}_{e,ij}(X = x_i, Y = y_j) = \Delta P_{Z,ij} = \Delta P_{X,i} \cdot \Delta P_{E,j} \quad (20)$$

2.3 改进离散求和法随机模拟

改进离散求和法采用与离散求和法相同的随机数,由于选择Normal Copula函数构建随机变量X和Y的联合分布,Normal Copula函数参数即为相关系数ρ。因此在相关系数ρ的取值区间内,采用式(13)和式(14)计算Z的概率区间ΔP_{Z,ij},进而求得离散点(x_i, y_j)上的概率分布估计值:

$$\hat{F}_{e,ij}(X = x_i, Y = y_j) = \Delta P_{Z,ij} = \int_{u_i}^{u_{i+1}} \int_{v_j}^{v_{j+1}} c(u, v) dv du = C(u_{i+1}, v_{j+1}) - C(u_{i+1}, v_j) - C(u_i, v_{j+1}) + C(u_i, v_j) \quad (21)$$

以理论分布计算的各个离散点(x_i, y_j)上的概率分布值F(x_i, y_j)为基准,采用估计值与理论值的绝对误差之和(Sum of Absolute Error, SAE)表征离散求和法与改进离散求和法的总估计误差:

$$SAE = \sum_{i=1}^{m_x} \sum_{j=1}^{m_y} |F(x_i, y_j) - \hat{F}(x_i, y_j)| \quad (22)$$

SAE越小,说明方法的计算结果越精确。需要说明的是,本次统计试验主要研究离散化计算相对连续分布积分的误差,而离散点(x_i, y_j)上的概率值计算之后的水库调洪计算,不再涉及积分的离散且不会产生系统误差。因此,统计试验仅仅涉及离散积分的部分,流程图见图2,将相关系数ρ和随机变量X、Y分别离散成K, m_X, m_Y种状态,并计算所有状态组合相应概率与理论概率间的误差。为避免偶然性,将上述步骤重复L次,再对所有结果进行统计分析。

2.4 统计试验结果

取 m_X = m_Y = 30, K = 20, L = 100, 分析计算了理

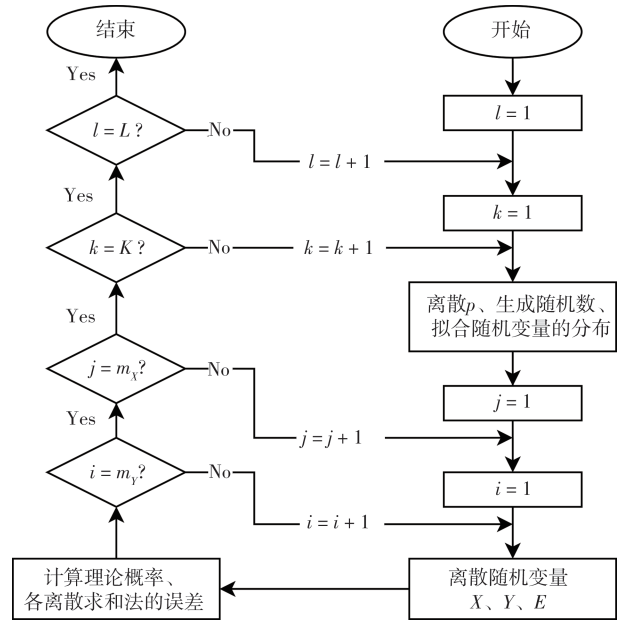


图2 统计试验流程
Fig.2 Statistic experiment flow

论分布概率、离散求和法和改进离散求和法,并对100次循环的结果进行汇总,统计了离散求和法与改进离散求和法在不同相关系数下的估计值与理论值的误差分布,结果见表1和图3。可以看出:①相关系数0<ρ≤0.2时,离散求和法E、离散求和法Y与改进离散求和法的总误差分布较接近,可用于估算各状态的组合概率;②ρ>0.2时,改进离散求和法的总误差分布范围值随相关系数的增大而减小,而离散求和法则随相关系数的增大而加大。从流域产汇流机制上来讲,干流和区间流量一般存在显著的相关关系,离散求和法的独立性假设具有很大的局限性。

2.5 统计试验的维数灾问题

对p维问题,统计实验是需要将各个维度分别离散为m_X, m_{Y1}, m_{Y2}, ..., m_{Yp-1}个状态,假设各个维度离散状态数相同,记为M,则状态组合数达m_X·m_{Y1}·m_{Y2}·...·m_{Yp-1} = M^p个;除此之外,p维正态分布的相关系数矩阵由各维度两两之间的相关系数组成,总共涉及p(p-1)/2个相关系数,为比较各变量不同相关程度的模拟情况,每个相关系数同样需要进行离散,当各系数离散状态个数为K时,状态组合数达K^{p(p-1)/2}个;同时为避免随机抽样的误差,需重复多次上述实验,若进行L次重复实验,则总计涉及M^p·K^{p(p-1)/2}·L次计算。因此随着考虑问题的维度上升,计算量呈指数式上升。以本文统计实验为例(M = 30, K = 20, L = 100),二维只

表1 总误差与相关系数 ρ 关系

Table 1 Relationship between the sum of absolute error and correlation coefficient ρ

ρ	离散求和法E		离散求和法Y		改进离散求和法	
	均值	方差 ($\times 10^{-2}$)	均值	方差 ($\times 10^{-2}$)	均值	方差 ($\times 10^{-2}$)
0.00	0.147	1.09	0.145	0.86	0.147	0.85
0.05	0.147	1.09	0.153	0.57	0.146	0.85
0.10	0.147	0.92	0.171	0.34	0.146	0.69
0.15	0.147	0.95	0.192	0.36	0.145	0.70
0.20	0.151	0.96	0.215	0.49	0.146	0.79
0.25	0.150	0.91	0.240	0.59	0.145	0.77
0.30	0.159	0.97	0.263	0.61	0.144	0.80
0.35	0.166	0.90	0.290	0.70	0.143	0.85
0.40	0.175	0.81	0.321	0.70	0.142	0.80
0.45	0.187	0.78	0.354	0.77	0.142	0.91
0.50	0.202	0.70	0.389	0.82	0.141	0.80
0.55	0.216	0.66	0.428	1.06	0.141	0.93
0.60	0.231	0.48	0.473	0.92	0.137	0.89
0.65	0.244	0.32	0.522	1.04	0.136	0.82
0.70	0.257	0.62	0.580	1.24	0.131	0.81
0.75	0.313	1.43	0.651	1.12	0.128	0.80
0.80	0.401	1.54	0.732	1.06	0.125	0.81
0.85	0.530	1.57	0.838	1.07	0.120	0.96
0.90	0.705	1.67	0.977	1.01	0.115	0.91
0.95	0.927	0.66	1.192	0.97	0.104	0.96
0.98	1.033	0.16	1.424	0.59	0.105	1.30

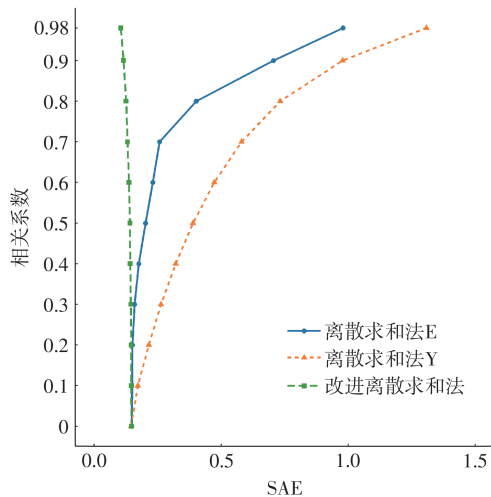


图3 总误差SAE均值和相关系数 ρ 关系
Fig.3 Relational diagram of mean SAE and ρ

需循环180万次计算,三维则需要72亿次,四维高达5184万亿次。随着维数增加,计算量急剧增长,出现明显的“维数灾”难题^[7]。

3 清江梯级水库实例研究

清江流域横贯湖北省西南,是长江出三峡后右岸的第一条大支流,集水面积约为17000 km²。清江中下游干流已经建成水布垭~隔河岩~高坝洲梯级水电站,其中水布垭为多年调节水库,隔河岩为年调节水库,高坝洲水库为隔河岩电站调峰和航运的反调节径流式水库电站。

根据清江梯级水库的运行和观测资料、基于水量平衡原理还原计算得到水布垭~隔河岩~高坝洲水库坝址1951—2020年日尺度流量系列,再根据水文比拟法和上下游流量相减法计算水~隔区间和隔~高区间流量系列。基于清江的场次洪水过程时段,选择最大7d洪量为洪水变量, X_1 、 X_2 、 Z 分别代表水布垭、隔河岩和高坝洲坝址的7d洪量, Y_1 、 Y_2 分别为水~隔区间和隔~高区间的7d洪量。分别采用同频率地区组合法、最可能地区组合法、概率组合离散求和法及改进离散求和法,推求受上游水布垭~隔河岩梯级水库调蓄后,高坝洲断面年最大流量的设计值。

3.1 优选Copula函数

选择常用的Clayton Copula、Gumbel Copula、Frank Copula和t-Copula函数^[9]作为联合分布的备选函数,采用极大似然法估计参数,采用Cramér-Von Mises法假设检验。取检验显著性水平为 $\alpha=0.05$, $p>0.05$ 时通过检验。根据赤池信息准则(AIC)优选Copula函数,AIC值越小则表明Copula函数拟合效果越好。表2为构造洪水地区组成最优的Copula函数和参数。

3.2 离散求和法变量代换

已知 $X_2=X_1+Y_1$,求得与 Y_1 和 Y_2 的相关系数 ρ 分别为0.91和0.94。独立性检验表明 X_1 与 Y_1 和 Y_2 均不独立,且由于 Y_1 和 Y_2 的均值比 X_1 小得多,对 Y_1 和 Y_2 作变量代换,即令 $E_1=Y_1-k_1X_1$, $E_2=Y_2-k_2X_1$,采用最小二乘法计算 $k_1=0.274$ 、 $k_2=0.096$ 。选择P-III分布和线性矩法估计 E_1 和 E_2 的统计参数,频率曲线见图4和图5,可知根据公式(7)得到的代换变量在转换中难免会出现负值,转换过程存在数据失真的可能。

3.3 结果对比分析

以基于改进离散求和法求得高坝洲坝址断面洪峰频率曲线的过程为例,说明离散求和法和改进离散求和法的设计洪水计算过程:将 X_1 、 Y_1 、 Y_2 的概率分布函数 $F(x_1)$ 、 $F(y_1)$ 和 $F(y_2)$ 均离散成30种状态,设变量 X_1

表2 年最大7d洪量系列的联合分布函数和参数

Table 2 Joint distribution function and parameters of annual maximum 7d flood volume

序号	变量	最优Copula函数	Copula参数	洪水地区组成方法
(1)	X_1, X_2	Gumbel Copula	4.81	最可能地区组合法
(2)	X_1, X_2, Z	t-Copula	(0.935 6, 0.948 4, 0.972 7)	最可能地区组合法
(3)	X_1, Y_1	Gumbel Copula	3.93	改进离散求和法
(4)	X_1, Y_1, Y_2	t-Copula	(0.934 3, 0.959 2, 0.987 3)	改进离散求和法

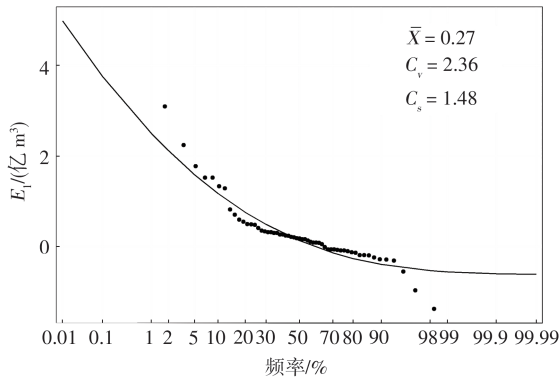


图4 水~隔区间 E_1 的频率曲线

Fig.4 Frequency curve E_1 for Shui-Ge inter-basin

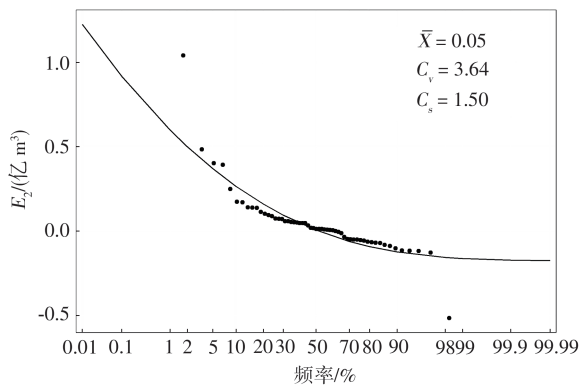


图5 隔~高区间 E_2 的频率曲线

Fig.5 Frequency curve E_2 for Ge-Gao inter-basin

第*i*种状态的取值为 $x_{i,1}$,变量 Y_1 第*j*种状态的取值为 $y_{j,1}$,变量 Y_2 第*k*种状态的取值为 $y_{k,2}$ ($i=1, 2, \dots, 30; j=1, 2, \dots, 30; k=1, 2, \dots, 30$)。选用1998年洪水过程线为典型,按 $x_{i,1}, y_{j,1}$ 和 $y_{k,2}$ 为控制,放大水布垭断面、水~隔区间和隔~高区间的洪水过程线,经水布垭~隔河岩水库调洪后叠加演算得到下游高坝洲坝址断面的洪水过程,并统计出最大流量及对应的发生概率。对 X_1, Y_1, Y_2 的全部取值状态作组合计算,得到高坝洲断面洪峰的频率曲线。其他离散求和与改进离散求和方案的计算过程与上述相似。表3和表4分别列出在不同重现期(1 000年、100年、50年、20年和10年一遇)情况下,隔河岩和高坝洲坝址的设计洪峰流量对比。

如表3和表4所示,受上游水库调蓄影响后,下游断面的设计洪峰均低于原设计值,各种方法所得高坝

表3 水布垭水库调蓄后隔河岩坝址设计洪峰流量对比

Table 3 Comparison of design floods at Geheyan dam-site regulated by Shuibuya reservoir

		重现期/年	1 000	100	50	20	10
设计值	洪峰流量/(m^3/s)	22 100	16 800	15 200	13 000	11 300	
	同频率地区组合法	洪峰流量/(m^3/s)	17 001	14 032	12 674	11 155	10 388
同频率地区组合法	削减量/(m^3/s)	5 099	2 768	2 526	1 845	912	
	削减率/%	23.1%	16.5%	16.6%	14.2%	8.1%	
最可能地区组合法	洪峰流量/(m^3/s)	17 075	14 095	12 727	11 231	10 388	
	削减量/(m^3/s)	50 25	2 705	2 473	1 769	912	
最可能地区组合法	削减率/%	22.7%	16.1%	16.3%	13.6%	8.1%	
	离散求和法	洪峰流量/(m^3/s)	16 100	13 500	11 900	11 100	9 900
离散求和法	削减量/(m^3/s)	6 000	3 300	3 300	1 900	1 400	
	削减率/%	27.1%	19.6%	21.7%	14.6%	12.4%	
改进离散求和法	洪峰流量/(m^3/s)	16 800	14 000	12 300	11 300	10 200	
	削减量/(m^3/s)	5 300	2 800	2 900	1 700	1 100	
改进离散求和法	削减率/%	24.0%	16.7%	19.1%	13.1%	9.7%	

表4 水布垭~隔河岩梯级水库调蓄后高坝洲坝址洪峰流量对比

Table 4 Comparison of design floods at Gaobazhou dam-site regulated by Shuibuya-Geheyan cascade reservoirs

		重现期/年	1 000	100	50	20	10
设计值	洪峰流量/(m^3/s)	24 300	18 800	17 100	14 700	12 300	
	同频率地区组合法	洪峰流量/(m^3/s)	13 300	11 300	11 200	10 300	9 700
同频率地区组合法	削减量/(m^3/s)	11 000	7 500	5 900	4 400	2 600	
	削减率/%	45.3%	39.9%	34.5%	29.9%	21.1%	
最可能地区组合法	洪峰流量/(m^3/s)	13 200	11 300	11 200	10 300	9 700	
	削减量/(m^3/s)	11 100	7 500	5 900	4 400	2 600	
最可能地区组合法	削减率/%	45.7%	39.9%	34.5%	29.9%	21.1%	
	离散求和法	洪峰流量/(m^3/s)	12 500	11 300	10 600	10 100	9 900
离散求和法	削减量/(m^3/s)	12 900	7 300	5 900	3 500	1 500	
	削减率/%	50.8%	39.2%	35.8%	25.7%	13.2%	
改进离散求和法	洪峰流量/(m^3/s)	12 800	11 400	10 900	10 300	10 000	
	削减量/(m^3/s)	12 600	7 200	5 600	3 300	1 400	
改进离散求和法	削减率/%	49.6%	38.7%	33.9%	24.3%	12.3%	

洲洪峰设计值之间差别不大,说明结果较为可靠;洪水的重现期越大,梯级水库对应的削峰率也越大,对于1 000年一遇的设计洪水,削峰率达40%以上,说明梯级水库联合调度对下游防洪断面设计洪水有重要影

响。同频率地区组合法、最可能地区组合法、离散求和法与改进离散求和法的设计结果差别不大,无法区分各种方法的优劣,需要通过统计试验来比较论证。

4 梯级水库下游设计洪水地区组成计算统计试验

由于天然情况下的洪水总体未知,且梯级水库的调洪过程较复杂,难以实现从理论上推求梯级水库下游洪水地区组成的误差公式。因此,本文采用统计试验法,比较清江梯级水库下游设计洪水地区组成计算方法。首先采用蒙特卡洛法生成100 000场包含坝址及区间洪量的洪水总体,从总体数据集中不重复地抽取100组样本,每组样本的理论分布与清江实例的70年连序洪量和233年考证期不连序历史洪水洪量所构成的样本系列相似。对100 000场洪水总体,根据水布垭~隔河岩梯级水库防洪调度规则调洪,统计隔河岩和高坝洲坝址洪峰流量,采用数学期望公式计算隔河岩和高坝洲坝址洪峰流量的经验频率曲线(由于样本数量较大可近似认为是概率分布曲线),统计得出1 000年、100年、50年、20年和10年一遇重现期情况下,隔河岩、高坝洲坝址年最大洪峰流量,作为假设总体以衡量各种设计洪水地区组成计算方法的计算结果。对每一组样本,分别采用同频率、最可能地区组合法、离散求和法及改进离散求和法,计算隔河岩、高坝洲坝址5种设计频率下设计洪峰值,并与假设总体值对比。图6和图7为计算值的箱线图,表5比较了高坝洲坝址洪峰流量计算的均值方差。

从图表中可以看出:①各种计算方法均值与假设总体的相对误差在6%以内,假设总体值基本位于各计算方法100组洪峰箱线图误差棒之间,说明各个梯级水库下游设计洪水地区组成计算方法的结果较可靠;

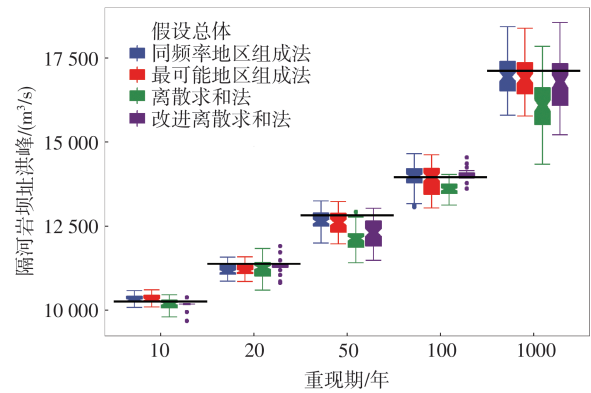


图6 不同重现期隔河岩坝址设计洪峰箱线图
Fig.6 Box-plot of design flood at Geheyan dam-site with different return periods

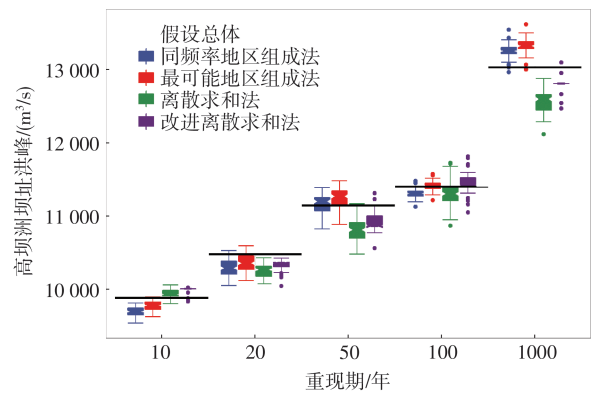


图7 不同重现期高坝洲坝址设计洪峰箱线图
Fig.7 Box-plot of design flood at Gaobazhou dam-site with different return periods

②最可能地区组合法和同频率地区组合法不同重现期隔河岩坝址洪峰方差比较接近,所不同的是最可能地区组合法由于考虑了洪水地区组成的空间相关性,所计算的区间洪量较同频率法偏大,水布垭坝址洪量偏小;③从隔河岩、高坝洲坝址洪峰分布可以看出,随着水库数目增加,改进离散求和法的方差相对小于离散求和法,重现期越小越趋近于假设总体,相对误差较小。

表5 高坝洲坝址设计洪峰流量的均值方差比较
Table 5 Comparison of mean deviation of design floods at Gaobazhou dam-site

重现期/年		1 000	100	50	20	10
假设总体		13 000	11 400	11 100	10 500	9 900
同频率地区组合法	均值	13 200 (1.54%)	11 300 (-0.88%)	11 200 (0.90%)	10 300 (-1.90%)	9 700 (-2.02%)
	方差	98	61	119	117	59
最可能地区组合法	均值	13 300 (2.31%)	11 400 (0.00%)	11 200 (0.90%)	10 400 (-0.95%)	9 800 (-1.01%)
	方差	112	61	127	118	61
离散求和法	均值	12 600 (-3.08%)	11 300 (-0.88%)	10 800 (-2.70%)	10 200 (-2.86%)	9 900 (0.00%)
	方差	142	165	148	88	52
改进离散求和法	均值	12 800 (-1.54%)	11 500 (0.88%)	10 900 (-1.80%)	10 300 (-1.90%)	10 000 (1.01%)
	方差	98	148	124	70	38

5 结论

通过综述设计洪水地区组成法的研究进展及问题,基于Copula函数推导了改进了离散求和法的计算公式,重点探讨概率组合离散求和法和改进离散求和法,比较两者估计的总误差;结合清江水布垭~隔河岩~高坝洲梯级水库应用实例,通过统计试验分析比较了各种设计洪水地区组成法的优缺点和实用性,主要研究结论如下:

(1)离散求和法的独立性转换处理可能使相关性信息失真;改进离散求和法基于Copula函数构造各个水库断面洪量的联合分布,其公式推导更加直接,且无需作独立化处理,既简化了计算、又降低了误差累积。统计试验表明,离散求和法的总误差随着入库和区间流量系列相关系数 ρ 的增加而显著加大,改进离散求和法的总误差则逐渐减小。当 $0 < \rho \leq 0.2$ 时,两者计算的结果差别不大;但当 $\rho > 0.2$ 时,则推荐采用改进离散求和法。

(2)清江梯级水库的计算结果表明,各种方法推求的设计洪峰流量差别较小,受水布垭~隔河岩水库调蓄的影响,高坝洲坝址1000年和100年一遇设计洪峰削减率分别为50%和40%左右,梯级水库调蓄对下游设计断面洪峰的影响不容忽视。

(3)对于单座或2座水库,离散求和法($0 < \rho \leq 0.2$),

改进离散求和法、同频率地区组合法($\rho > 0.7$),最可能地区组成法的计算结果相当;对于3座及以上的梯级水库群,最可能地区组成法可得到唯一方案解,不仅具有统计理论基础,又可获得最优的洪水地区组成方案。研究成果可为我国《水利水电工程设计洪水计算规范》的修订提供科学依据和参考。

参考文献:

- [1] 中华人民共和国水利部. 水利水电工程设计洪水计算规范:SL 44—2006[S]. 北京:中国水利水电出版社,2006.
- [2] 王锐琛,陈源泽,孙汉贤. 梯级水库下游洪水概率分布的计算方法[J]. 水文,1990(1):1-8.
- [3] 刘章君,郭生练,许新发,等. Copula函数在水文水资源中的研究进展与述评[J]. 水科学进展,2021,32(1):148-159.
- [4] 闫宝伟,郭生练,郭靖,等. 基于Copula函数的设计洪水地区组成研究[J]. 水力发电学报,2010,29(6):60-65.
- [5] 李天元,郭生练,刘章君,等. 梯级水库下游设计洪水计算方法研究[J]. 水利学报,2014,45(6):641-648.
- [6] 刘章君,郭生练,李天元,等. 梯级水库设计洪水最可能地区组合法计算通式[J]. 水科学进展,2014,25(4):575-584.
- [7] XIONG F, GUO S, YIN J, et al. Comparative study of flood regional composition methods for design flood estimation in cascade reservoir system[J]. Journal of Hydrology, 2020, 590: 125530.
- [8] 郭生练,熊立华,熊丰,等. 梯级水库运行期设计洪水理论和方法[J]. 水科学进展,2020,31(5):734-745.
- [9] NELSEN R B. An Introduction to Copulas[M]. 2nd ed. Berlin: Springer,2006.

Comparative Study on Design Flood Regional Composition Methods for Downstream Cascade Reservoirs

XIE Yuzuo¹, GUO Shenglian¹, WANG Jun¹, ZHONG Sirui¹, WANG Yun¹, LI Shuai², HU Ting²

(1. State Key Laboratory of Water Resources Engineering and Management, Wuhan University, Wuhan 430072, China;

2. Operation and Administration Center for River Basin Hydro Complex, China Three Gorges Corporation, Yichang 443133, China)

Abstract: The construction and operation of water conservancy and hydropower projects have changed the hydrological situation of the downstream rivers and the conditions of streamflow and sediment transport. Calculate the design flood of the downstream control section of the cascade reservoirs has become an urgent problem. In this study, we reviewed the current design flood regional composition methods and problems, deduced the probabilistic-combined discrete summation (DS) method and the improved discrete summation (IDS) method based on copula function, and compared the sum of absolute error of DS and IDS methods. The Qingjiang cascade reservoirs were selected as case study, the variant design flood regional composition methods and the pros and cons were practically discussed. The results show that: (1) The derivation of the IDS method is more advanced than that of the DS method without independence transformation. The sum of absolute error of the DS method increases significantly with the increase of the correlation coefficient ρ ; when the $\rho > 0.2$, it is recommended to use the IDS method. (2) After the regulation of the Shuibuya-Geheyan reservoirs, the design flood peaks with 1000-year return periods are reduced by about 50%. (3) For single or two reservoirs, the application conditions for the DS and the equivalent frequency regional composition (EFRC) methods are $0 < \rho \leq 0.2$ and $\rho > 0.7$, respectively, while the IDS and the most likely regional composition (MLRC) methods have not limitations. And for three and more cascade reservoirs, the MLRC method can obtain only and optimal scheme solution. The research results could provide a reference for modifying the design flood calculation regulations in China.

Keywords: cascade reservoirs; flood regional composition; Copula Function; discrete summation; statistical experiment; Qingjiang basin